Pol. J. Environ. Stud. Vol. 30, No. 3 (2021), 2755-2767
DOI: *10.15244/pjoes/129917* ONLINE PUBLICATION DATE: 2021-02-22

Original Research

Prediction and Analysis of CO_2 Emissions Based on Regularized Extreme Learning Machine Optimized by Adaptive Whale Optimization Algorithm

Wei Sun, Yawen Wang*

Department of Business Administration, North China Electric Power University, Baoding 071000, China

Received: 23 September 2020
Accepted: 2 November 2020

Abstract

With the rapid increase in CO_2 emissions, there is a profound impact on global climate change, seriously hindering the sustainable development of the low-carbon economy. Therefore, it is particularly significant to predict CO_2 emissions. To improve the accuracy and robustness, a new hybrid model (KPCA-CEEMDAN-AWOA-RELM) is proposed when considering multiple factors about historical CO2 emissions, energy consumption, economic and social. First, factors are determined by Pearson doubly significant test, and the principal component is extracted by kernel principal component analysis (KPCA) to realize nonlinear dimension reduction. Then, the CO_2 emissions sequence is decomposed by the ensemble empirical model with adaptive noise (CEEMDAN) model to reduce noise interference and abate reconstruction error. Finally, CO_2 emissions can be predicted via the extreme learning machine with regularization parameter modification optimized by an improved input weight matrix and the deviation matrix of the adaptive optimization whale algorithm (AWOA-RELM). Taking Hebei and China as examples, it is found that the selected model is better than other comparative models.

Keywords: CO_2 emissions, kernel principal component analysis, adaptive whale optimization algorithm, regularized extreme learning machine, ensemble empirical model with adaptive noise

Introduction

The 2019 Taiyuan energy low CO_2 development forum pointed out that: since the adoption of the 2030 agenda for sustainable development and the Paris Agreement, the world has promoted sustainable economic development and accelerated the transformation of global energy [1]. To deal with weather problems, some countries and international organizations need to do their best. According to BP world energy statistics review, China has become the world's largest energy consumer, accounting for 34% of global primary energy consumption in 2018 [2]. In the 13th Five-Year Plan, compared to 2015, the Chinese government proposed that the unit's gross domestic product (GDP) CO_2 emissions will be decreased to 18% by 2020, and the total amount of carbon emissions

*e-mail: 2394037961@qq.com

will be effectively controlled [3]. Therefore, it is very important to the formulation and implementation of emissions reduction policies and the realization of low-carbon economic development that CO_2 emission is predicted by scientific and reasonable methods when several factors comprehensively analyzed.

Many scholars have carried out in-depth research on CO_2 emissions. One of the important research directions is to explore various factors of CO_2 emissions by traditional statistical methods. Based on four main factors that is economic activity, employment, energy intensity and CO_2 intensity, the energy-related CO_2 emissions were analyzed by the refined Laspeyres index decomposition [4]. The determinants of carbon emissions have been proved to exist through the Environmental Kuznets Curve (EKC) and the STIRPAT model [5]. Between total energy consumption and CO_2 emission was found that had a strong relationship through regression analysis [6]. There was a certain relationship between regional economic development and industrial CO_2 emissions when main factors affecting the change of CO_2 emissions were taken into account [7]. A dynamic comprehensive input-output simulation model was established to study greenhouse gas emissions [8]. These are only based on traditional statistical methods to analyze various affecting factors of CO_2 emissions.

Another important direction is that many scholars have also researched on CO_2 emissions prediction combining with many factors. Combined with a variety of influencing factors, the CO_2 emissions are predicted through the system dynamics method [9]. Household CO_2 emissions were explored and predicted with butterfly optimization algorithm optimized the least square support vector machine (BOA-LSSVM) [10]. CO_2 emission was predicted with the extreme learning machine improved by particle swarm optimization (PSO-ELM) [11]. CO_2 emission related to energy consumption was predicted in the extreme learning machine (ELM) optimized by grey prediction theory and support vector machine algorithm when considering coal, gasoline, natural gas, and coal power generation [12]. CO_2 emission was predicted based on principal component analysis with regularized extreme learning machine [13]. CO_2 emission was predicted by using random forest and ELM when the economy and factors of energy were considered [14]. Compared with the traditional neural network, ELM has a faster convergence rate and less human interference in the forecasting aspect.

ELM has been used in various prediction fields, such as short-term load forecasting [15], carbon price forecasting [16], the scale of electric vehicles [17] and so on. However, the generalization ability of ELM may be affected by the input weight matrix and the hidden layer threshold with randomly assigned. So it is necessary that ELM is optimized with some optimization methods which mainly contain particle swarm optimization (PSO), moth algorithm and whale optimization algorithm (WOA) [18]. Among, WOA was verified that its ability of search was better than other traditional optimization algorithms [19]. However, WOA is easy to fall into the local area and affect the optimal solution. Therefore, adaptive whale optimization algorithm (AWOA) is introduced to weight of adaptive based of WOA which can enhance local optimization ability and improve convergence accuracy [20]. AWOA is used to optimize the input weight and hidden layer threshold of ELM to maximize realize the global and local search capability, improve the training rate, and overcome the inherent instability.

When multiple influencing factors are used as independent variables in the prediction model, the operation time and complexity will be increased, and the prediction accuracy will be affected. Therefore, it is necessary to reduce the dimension variables. The methods of dimension reduction are mainly concluded regression analysis [21], factor analysis [22]. In addition, the entropy data were input into nuclear principal component analysis (KPCA) for feature fusion and dimension reduction, which is usually used for rolling bearing fault diagnosis [23]. The performance of SVM using principal component analysis (PCA), KPCA or independent component correlation algorithm (ICA) for feature extraction has been proved to be superior to that of SVM without feature extraction, and KPCA is the best feature extraction effect [24]. However, there are few studies on KPCA for CO_2 emissions forecasting. So this paper will use KPCA to realize the feature dimension reduction of CO_2 emissions.

Because of the instability and unique complexity of CO_2 emissions, the prediction accuracy will be greatly reduced. To reduce the influence of noise, the original sequence is usually decomposed by empirical mode decomposition (EMD) [25], ensemble empirical mode decomposition (FEEMD) [26]. However, these decomposition methods have some problems of pattern mixing and reconstruction error. Therefore, the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) is an adaptive noise optimization improvement in EEMD [27]. CEEMDAN was employed to decompose the original data and improve the prediction accuracy [28]. Financial time series is decomposed by CEEMDAN into multiple intrinsic mode functions of different time scales [29].

Based on the study of the above literature, a hybrid prediction model is proposed when various factors are considered from the perspective of energy, economy and society. Firstly, the main component of factors is extracted by KPCA to achieve nonlinear dimension reduction, as input variables. Secondly, CEEMDAN is used to decompose CO_2 emissions series to reduce the reconstruction error and improve the calculation speed. Lastly, taking Hebei and China as examples, the regularized extreme learning machine optimized by the adaptive whale algorithm with adaptive weights of the input weight matrix and deviation matrix is used to predict CO_2 emissions to improve the prediction accuracy and robustness.

Materials and Methods

Nuclear Principal Component Method

The principal component method (PCA) is to find the best linear combination of variables by calculating the characteristic covariance matrix to judge the variance consistency between variables [30]. However, it is necessary to consider the non-linear characteristics of some research objects in reality. Kernel principal component analysis (KPCA) can effectively extract the non-linear characteristics of kernel function calculation [31]. KPCA is a nonlinear mapping realized by kernel function calculation. Then, the samples are mapped from low latitude space to higher dimensional feature space. Finally, PCA is used to reduce dimension in this space. Thus, any vector can be expressed linearly by the samples in this space.

Ensemble Empirical Mode Decomposition with Adaptive Noise

To effectively eliminate the mode mixing problem, and to reduce the reconstruction error, CEEMDAN is proposed to optimize based on EEMD. If the original time series x(n), then the t-th white noise sequence $\delta^t(n)$ is added, and the t-th decomposition can be expressed as $x^t(n) = x(n) + \delta^t(n)$. The i-th mode component generated by EMD and CEEMDAN decomposition is decomposed: $E_i(\cdot)$ and $\widetilde{IMF}_1$. The steps of CEEMDAN are as follows:

1. Firstly, the original sequence is decomposed by EMD to obtain the first model component:

$$\widetilde{IMF}_1(n) = \frac{1}{T}\sum_{t=1}^{T} IMF_1^t(n) = \overline{IMF_1}(n) \tag{1}$$

2. In the first stage (i = 1), the first unique residual signal is calculated:

$$r_1(n) = x(n) - \overline{IMF_1}(n) \tag{2}$$

3. Repeat the test t time. In each test, the signal $r_1(n) + \varepsilon_1 E_1(\delta^t(n))$ is decomposed and stops when the first EMD mode component is obtained. The second mode component is obtained:

$$\widetilde{IMF}_2(n) = \frac{1}{T}\sum_{t=1}^{T} E_1(r_1(n)) + \varepsilon_1 E_1(\delta^t(n)) \tag{3}$$

4. In each remaining stage, the i-th residual signal is calculated in the same way as in the third step. In this case, the i+1-th mode component can be obtained:

$$r_i(n) = r_{i-1}(n) - \widetilde{IMF}_1(n) \tag{4}$$

$$\widetilde{IMF_{1+1}} = \frac{1}{T}\sum_{t-1}^{T} E_1(r_i(n)) + \varepsilon_i E_i(\delta^t(n)) \tag{5}$$

5. Judge whether the number of extreme points of the residual signal is no more than two. If it is, the algorithm stops, that is, the residual sequence can't be further decomposed. The final residual signal can be expressed as:

$$R(n) = x(n) - \sum_{i=1}^{I} \widetilde{IMF}_1 \tag{6}$$

Therefore, the original sequence x (n) can be decomposed into:

$$x(n) = \sum_{i=1}^{I} \widetilde{IMF}_1 + R(n) \tag{7}$$

Whale Optimization Algorithm

WOA is a new heuristic optimization algorithm proposed, inspired by the foraging behavior of the humpback whale bubble-net. In WOA, each whale can be regarded as a particle, and the position of each particle represents a decision variable. In the process of whaling, whales hunt not along a straight line, but in a spiral way. The algorithm flow is as follows:

1. Surround prey:

Whales usually surround their prey first when hunting, and their mathematical model is as follows:

$$D = |C * X_L(t) - X(t)| \tag{8}$$

$$X(t+1) = X_L(t) - A * D \tag{9}$$

...where t represents the current number of iterations, $X_L(t)$ is the optimal whale positions vector so far, and X(t) represents the current whale position vector.

A and C are learning factors, but they are calculated by the following formula.

$$A = 2a * r - a \tag{10}$$

$$C = 2 * r \tag{11}$$

...where the value of a decreases linearly in (0, 2), and r is the random number between [0, 1].

2. Spiral Hunt:

In the process of hunting, whales usually encircle their prey and hunt in a spiral motion. The mathematical model is as follows:

$$X(t+1) = D' * e^{bl} * \cos(2\pi l) + X_L(t) \tag{12}$$

...where D' is the distance between whales and prey X(t), local optimization $X_L(t)(D = |X_L(t) - X(t)|)$. b defined Constant of helix shape. I is the random number in (-1,1). To ensure that whales swim to prey in a spiral shape and at the same time shrink the enclosure, WOA uses the random probability p selection spiral model to update the whale's position. Where p is a random number between [0, 1] and the formula is as follows:

$$X(t+1) = f(x) = \begin{cases} X_L(t) - A * D & \text{if } p < 0.5 \\ D' * e^{bl} * \cos(2\pi l) + X_L(t) & \text{if } p \geq 0.5 \end{cases} \quad (13)$$

3. Search for prey:

The individual random location is used to search for prey, called a global search, and its formula is as follows:

$$D = |C * X_{rand} - X(t)| \quad (14)$$

$$X(t+1) = X_{rand} - A * D \quad (15)$$

...where X_{rand} is a randomly selected position vector. To enhance global search ability of WOA algorithm, it is required that randomly selected solution should be used to update the location accordingly.

Adaptive Weight Whale Optimization Algorithm

Similar to other heuristic optimization algorithms, local optimization ability of WOA is enhanced and the convergence accuracy is improved by introducing adaptive weight. To improve the ability of local search, the larger weight is introduced to change the position of the whale when the whale is close to the food and then the adaptive weight AWOA is obtained. The adaptive weight formula and the improved formula are as follows:

$$\omega = (\omega_{max} - \omega_{min}) * i/Max\_iteration \quad (16)$$

...where i is the current number of iterations, and Max_iteration is the maximum number of iterations.

Regularized Extreme Learning Machine

The extreme learning machine (ELM) is an improved algorithm based on single hidden layer feed forward BP neural network (BPNN) [32]. Compared with the traditional BPNN, the input weight and the threshold value of hidden neurons in the learning process of ELM is given randomly.

Give N training samples $\{(x_i, y_i)\}_{i=1}^{N}$, the ELM regression model with L hidden layer neuron functions can be expressed as:

$$\sum_{i=1}^{L} \beta_i \, g(\omega_i \cdot x_j + b_i) = y_j, j = 1, 2 \ldots, N \quad (18)$$

It can be abbreviated as:

$$H\beta = y \quad (19)$$

where

$$H(\omega_1, \ldots, \omega_L, x_1, \ldots, x_N, b_1, \ldots, b_L) = \begin{bmatrix} g(\omega_1 \cdot x_1 + b_1 & \cdots & g(\omega_L \cdot x_L + b_L) \\ \cdots & \cdots & \cdots \\ g(\omega_1 \cdot x_N + b_1 & \cdots & g(\omega_L \cdot x_N + b_L \end{bmatrix}_{N \times L} \quad (20)$$

$$\begin{cases} \beta = [\beta_1^T, \ldots, \beta_L^T] \\ y = [y_1^T, \ldots, y_N^T] \end{cases} \quad (21)$$

The output weight can be obtained by linear least squares:

$$||H\beta - y|| = ||HH^{+}y - y|| = \min_\beta ||H\beta - y|| \quad (22)$$

The least square solution obtained is as follows:

$$\beta = H^{+}y \quad (23)$$

...where H^+ represents the Moore-Penrose generalized inverse of the hidden layer output matrix H.

However, the standard ELM may have the problem of over fitting and reducing the generalization ability, so ELM is needed to modify: both empirical error minimization and risk need to be considered and use the regularization parameter C to modify to achieve the best compromise. The modified extreme learning machine is called regularized extreme learning machine (RELM). The formula is as follows:

$$\min_\beta C\,||y - H\beta||_2^2 + ||\beta||_2^2 \quad (24)$$

The constraints are as follows:

$$\min_\beta C\,||e||_2^2 + ||\beta||_2^2 \quad (25)$$

$$\text{s.t. } y - H\beta = e \quad (26)$$

...where $e = [e_1, e_2, \ldots, e_N]$ is the output error of the training sample x_i.

According to Karush-Kuhn-Tucker (KKT) condition, the corresponding Lagrange functions:

$$L(\beta, e, \delta) = C||e||_2^2 + ||\beta||_2^2 + \delta^T(y - HB - e) \quad (27)$$

...where non-negative δ is the Lagrangian multiplier. The relevant optimization conditions are as follows:

$$\begin{aligned} \partial L/\partial\beta = 0 &\Rightarrow 2\beta - H^T\delta = 0 \\ \partial L/\partial e = 0 &\Rightarrow 2cE - \delta = 0 \\ \partial L/\partial\delta = 0 &\Rightarrow y - H\beta - E = 0 \end{aligned} \quad (28)$$

The output weight matrix β is as follows:

$$\beta = \begin{cases} (H^TH + (I/C))^{-1}H^Ty & N > L \\ H^T(H^TH + (I/C))^{-1}y & N < L \end{cases} \quad (29)$$

In addition, in order to avoid non optimal or unnecessary weights and thresholds and improve the performance of RELM, this paper proposes an improved whale optimization algorithm to adjust the weights and deviations from input layer and hidden layer, and selects an optimal group to predict energy consumption.

Hybrid Models

Fig. 1 is the flow chart of the KPCA-CEEMDAN-AWOA-RELM model. In part 1, the principal component is extracted by KPCA from 17 factors as input variables of the prediction model. The second part is the adaptive whale optimization algorithm. It can be seen from this that if P<0.5 and |A<1|, the whales will transfer the adaptive weight. In the third part, it can be seen that AWOA is used to optimize the weight of the RELM input layer and the threshold of the hidden layer to obtain the optimal network. In the fourth part, original CO_2 emissions series was decomposed to obtain a more stable series with CEEMDAN as output variables, and used AWOA-RELM model to predict, and finally add each prediction series to get the final prediction results.

Experimental

Data Selection

To verify the effectiveness of the selected model, this paper selects Hebei Province and China's CO_2 emissions as cases. So, data is selected from 1980 to 2017 in the Economic Yearbook of Hebei, from 1980 to 2018 in China Statistical Yearbook. Specimens are divided into the training set and testing set according to 8:2 scales. CO_2 emissions can't be directly obtained, but it can be obtained through the conversion to energy CO_2 emissions coefficient from the IPCC guidelines on national greenhouse gas inventories [33].

The historical data is a key to CO_2 emissions prediction as one of input variables. To eliminate the internal correlation, the partial autocorrelation function (PACF) is used to explore the correlation between historical and predicted of CO_2 emissions. CO_2 emission is non-linear and non-stationary. Results of PCAF analysis were shown that CO_2 emissions with a confidence level of 90% and lag 1 have a strong correlation. So the first order lag of CO_2 emissions was also selected as one of influencing factors.

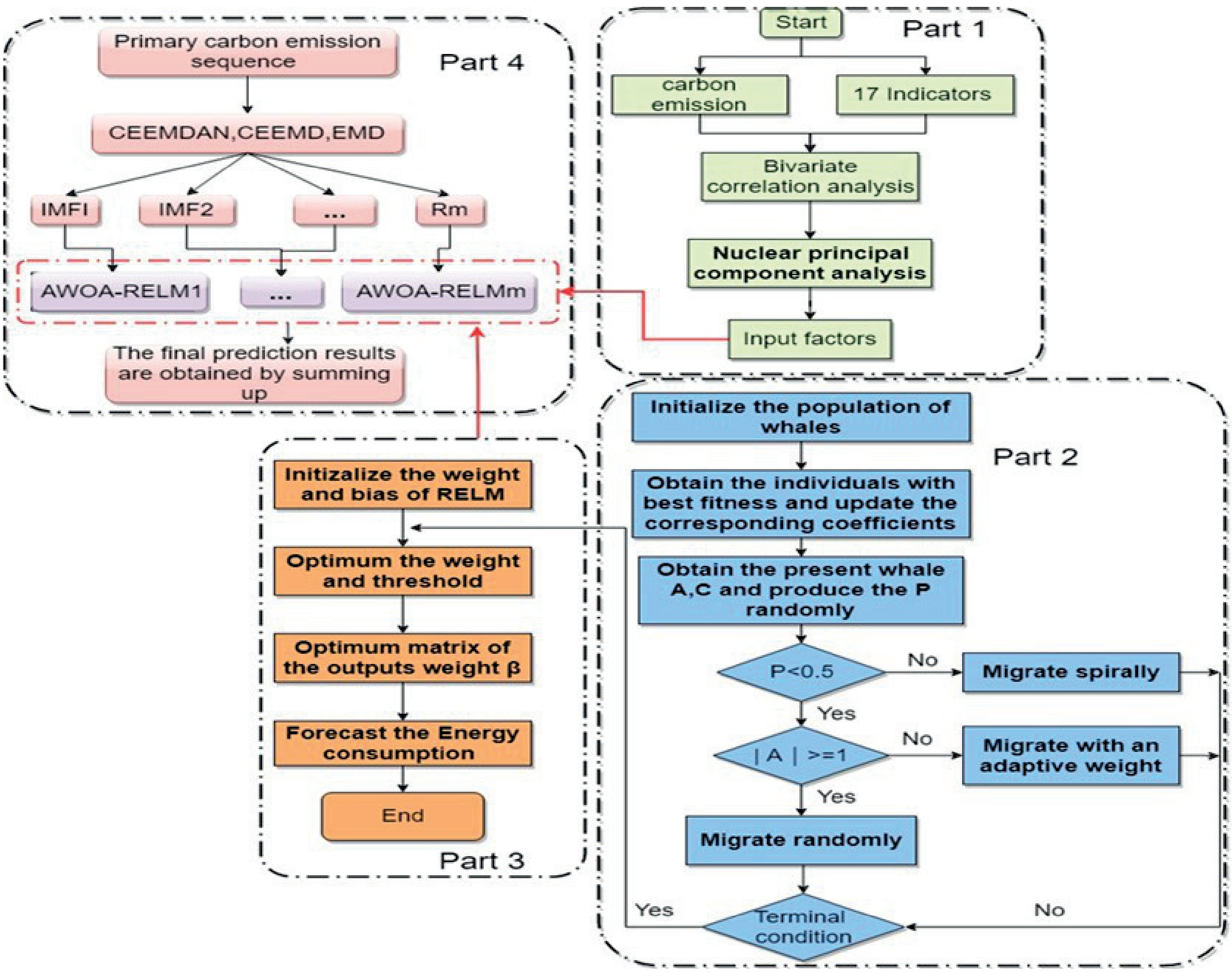

Fig. 1. Flow chart of KPCA-AWOA-RELM model.

Table 1. Pearson analysis results of various factors.

Index	Pearson correlation	Index	Pearson correlation
First order lag of CO_2 emissions	0.997	Industrial added value	0.966
Coal	0.999	GDP per capita	0.963
Crude oil	0.982	Total import and export amount	0.949
Natural gas	0.789	Fixed asset investment of the whole society	0.872
Primary electricity	1.000	Consumption level of residents	0.939
GDP	0.895	Population	0.913
Value of primary industry	0.981	City rate	0.946
Value of secondary industry	0.964	Power generation	0.991
Value of third industry	0.923	Steel production	0.922

The prediction effect of CO_2 emissions is affected as various factors, among which historical data and energy sources (coal, crude oil, natural gas and primary electricity) have a greater impact on CO_2 emissions. Also from the perspective of economic and social development: those are GDP, the value of primary, secondary and tertiary industries, industrial added value, total imports and exports, the whole society's fixed-asset investment, the level of consumption of residents, population, urbanization rate, power

Table 2. Results of KPCA in Hebei.

Number	Characteristic value	Contribution rate /%	Cumulative contribution rate /%
1	0.624626	93.4120	93.4120
2	0.026430	3.9525	97.3645
3	0.006604	0.9876	98.3520
4	0.003938	0.5889	98.9409
8	0.002869	0.4291	99.3700
6	0.002154	0.3221	99.6921
7	0.000873	0.1305	99.8226
8	0.000506	0.0756	99.8982
9	0.000263	0.0393	99.9375
10	0.000165	0.0246	99.9621
11	0.000106	0.0159	99.9780
12	0.000061	0.0091	99.9872
13	0.000026	0.0038	99.9910
14	0.000015	0.0023	99.9933
15	0.000015	0.0022	99.9955
16	0.000012	0.0018	99.9974
17	0.000008	0.0013	99.9986
18	0.000003	0.0005	99.9991
19	0.000002	0.0003	99.9995
12	0.000001	0.0002	99.9996
…	…	…	…
37	0.0000	0.0000	100.0000

generation and steel production are selected. Therefore, 18 indexes are taken as the preliminary selection factors.

Correlation Analysis

To avoid the repetition of content description, this paper will introduce CO_2 emissions prediction and analysis of Hebei Province in detail, and only analyze the prediction results of China's CO_2 emissions.

According to the above data, CO_2 emissions and influencing factors were tested bilaterally, the results of correlation analysis as showed in Table 1. The Pearson correlation coefficient of natural gas is 0.789, less than 0.87, which is not considered. Pearson correlation coefficients of other factors were greater than 0.87, and the remaining 17 factors were selected. The probability value of the bilateral significance test is 0.000 less than 0.01, so the preliminary selection of factors is reasonable.

KPCA Analysis

KPCA is a nonlinear dimension reduction method based on the kernel function. In order to eliminate the calculation errors caused by different dimensions, it is necessary to standardize the sample data. The kernel function chosen in this paper is the Gaussian kernel function. The principal component contribution rates and cumulative contribution rate obtained by KPCA is shown in Table 2.

When the cumulative contribution to principal component is more than 85%, the principal component can be extracted. From Table 2, when the dimension of the extracted input variable is 1, the cumulative contribution rate has reached 93.4120%, which can replace the original 17 dimensions as the input variable. Under the premise of ensuring effective information, the network model structure is simplified.

CO_2 Emissions Decomposition

The original CO_2 emissions sequence is not stable enough. Therefore, to reduce the interference of white noise and further excavate the intrinsic characteristics, original CO_2 emissions series are decomposed by CEEEMDAN. Four IMF Eigen modes and one residual term are obtained, as showed in Fig. 2. CEEEMDAN adding a finite number of adaptive white noise sequences in each stage can reduce the reconstruction error and white noise interference.

The principal component extracted by KPCA is used as the input of each sequence decomposed by CEEMDAN. The input of each sequence is consistent, and the final predicted value is obtained by adding the predicted output value of each sequence.

Parameter Setting and Prediction Evaluation Standard

In this paper, KPCA-CEEMDAN-AWOA-RELM is used to predict CO_2 emissions. In this paper, compared with the traditional WOA, the effectiveness of AWOA is verified according to the reference function F1, and the smaller inertia does enhance the local search ability of WOA.

$$F1(x) = \sum_{i=1}^{n} x_i^2 \tag{30}$$

MAPE, MAE and RMSE are selected to verify the rationality of the model. As a general error index, the smaller the value, the accuracy is the better. The formula is defined as:

$$MAPE = (1/n) \sum_{t=1}^{n} |(\hat{y}_t - y_t)/y_t| * 100\% \tag{31}$$

$$MAE = (1/n) \sum_{t=1}^{n} |\hat{y}_t - y_t| \tag{32}$$

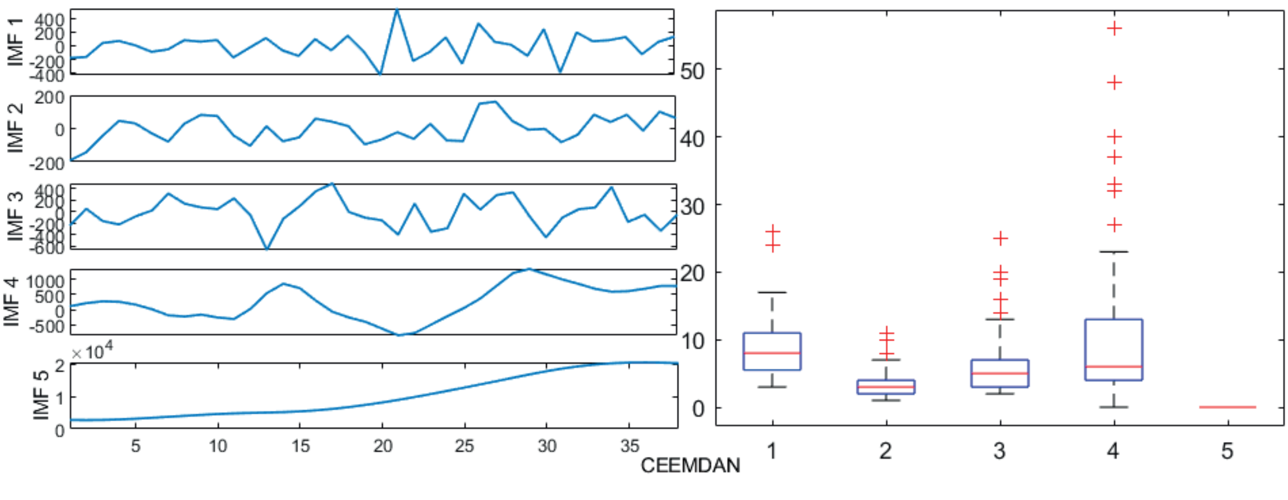

Fig. 2. Result of CEEMDAN in Hebei.

$$\mathrm{RMSE} = \sqrt{(1/n)\sum_{t=1}^{n}[\hat{y}_t - y_t]^2} \quad (33)$$

In order to more intuitively to observe whether the selected model can improve the prediction accuracy, three corresponding performance improvement evaluation indexes are added based on three error evaluation indexes. They are MAPE improvement percentage (PMAPE), MAE improvement percentage (PMAE) and RMSE improvement percentage (PRMSE). The formula of the three performance improvement evaluation indicators is as follows:

$$P_{MAPE} = \frac{\mathrm{MAPE}_1 - \mathrm{MAPE}_2}{\mathrm{MAPE}_1} \times 100\% \quad (34)$$

$$P_{MAE} = \frac{\mathrm{MAE}_1 - \mathrm{MAE}_2}{\mathrm{MAE}_1} \times 100\% \quad (35)$$

$$P_{RMSE} = \frac{\mathrm{RMSE}_1 - \mathrm{RMSE}_2}{\mathrm{RMSE}_1} \times 100\% \quad (36)$$

MAPE_1, MAE_1, RMSE_1 represent MAPE, MAE, RMSE, MAPE of the benchmark model. MAPE_2, MAE_2, RMSE_2 represent MAPE, MAE, RMSE of the comparison model.

To display the estimated performance of the hybrid model, comparative models are established in Fig. 3. In the blue part, PCA and KPCA were used to process the influencing factors, and the principal component components were used as input variables. In the yellow part, the model components and the residual quantity of CO_2 emissions series are obtained by decomposition of CEEMDAN, CEEMD and EMD. The green part is prediction models, which mainly include AWOA-RELM, AWOA-ELM, WOA-ELM, PSO-ELM, ELM and BP. The pink part is combination contrast models to verify the prediction performance of the selected model.

Results and Discussion

Case One

In order to verify the prediction performance and robustness of the proposed hybrid model, CO_2 emissions prediction is carried out based on comparison models. Taking Hebei case, the prediction results of each model are shown in Fig. 4. The calculation results of MAPE, MAE and RMSE are shown in Table 3 and Table 4 contains the percentage improvement results of the models. The conclusions are shown as follows:

1. Compared with other models in this paper, KPCA-CEEMDAN-AWOA-RELM has perfect prediction performance and robustness. The evaluation index of the model is the best, which is MAPE of 0.1349%, MAE of 28.2314 and an RMSE of 36.2668, respectively. Moreover, the curve of prediction results is closest to the actual CO_2 emissions curve in Fig. 4, so it can fully prove that KPCA-CEEMDAN-AWOA-RELM has a better effective than the traditional ELM and WOA, and improves the global ability of search.

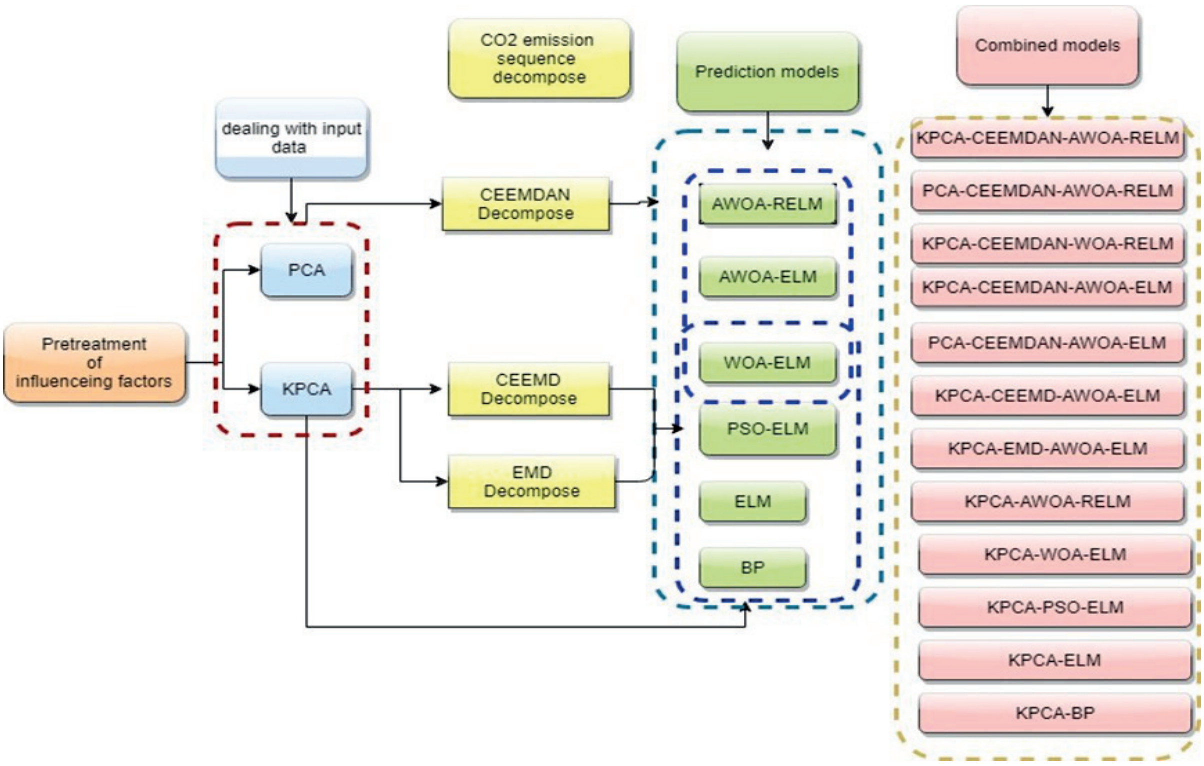

Fig. 3. Framework of comparative prediction models.

Table 3. The prediction results in Hebei.

	MAPE (%)	MAE	RMSE
KPCA-CEEMDAN-AWOA-RELM	0.1349	28.2314	36.2668
PCA-CEEMDAN-AWOA-RELM	0.6243	130.9433	166.3573
KPCA-CEEMDAN-WOA-RELM	0.6963	146.6076	156.9578
KPCA-CEEMDAN-AWOA-ELM	0.3607	75.3439	92.9914
PCA-CEEMDAN-AWOA-ELM	0.7383	154.9099	178.2320
KPCA-CEEMD-AWOA-ELM	0.5302	110.0257	184.8101
KPCA-EMD-AWOA-ELM	1.1184	209.2306	342.2678
KPCA-AWOA-RELM	0.6463	135.6178	173.0556
KPCA-WOA-ELM	0.7133	149.3975	186.5788
KPCA-PSO-ELM	0.9600	200.1688	286.5797
KPCA-ELM	0.9621	199.9011	327.0336
KPCA-BP	1.1758	234.4838	303.5495

Table 4. The percentage improvement results of the models in Hebei.

Benchmark model		Comparative model	PMAPE (%)	PMAE (%)	PRMSE (%)
KPCA-BP	VS.	KPCA-ELM	18.18	14.75	-7.74
KPCA-PSO-ELM	VS.	KPCA-WOA-ELM	25.70	25.36	34.89
KPCA-WOA-ELM	VS.	KPCA-AWOA-RELM	9.39	9.22	7.25
KPCA-AWOA-RELM	VS.	KPCA-CEEMD-AWOA-ELM	17.97	18.87	-6.79
KPCA-CEEMD-AWOA-ELM	VS.	KPCA-CEEMDAN-AWOA-ELM	31.95	31.52	49.68
KPCA-CEEMDAN-AWOA-ELM	VS.	KPCA-CEEMDAN-AWOA-RELM	62.61	62.53	61.00
KPCA-CEEMDAN-WOA-RELM	VS.	KPCA-CEEMDAN-AWOA-RELM	80.63	80.74	76.89
PCA-CEEMDAN-AWOA-ELM	VS.	KPCA-CEEMDAN-AWOA-ELM	51.14	51.36	47.83
PCA-CEEMDAN-AWOA-RELM	VS.	KPCA-CEEMDAN-AWOA-RELM	78.39	78.44	78.20

2. The prediction accuracy of the regularized extreme learning machine optimized by the adaptive whale algorithm is rather higher. The MAPE, MAE and RMSE of KPCA-AWOA-RELM are 0.6463%, 135.6178 and 173.0556, respectively. But three indicators of KPCA-BP model are 1.1758%, 234.4838 and 303.5495. According to the results in Table 4, comparing with KPCA-WOA-ELM and KPCA-AWOA-RELM, MAPE decreases by 9.39%, MAE increases by 9.22%, and RMSE decreases by 7.25%, and it is shown that RELM optimized by AWOA works well.
3. When comparing with CEEMD, EMD and not decomposed models, CEEMDAN model is rather excellent. With regard to KPCA-CEEMDAN-AWOA-ELM, MAPE, MAE and RMSE is 0.3607%, 75.3439 and 92.9914, respectively. Compared with KPCA-AWOA-RELM and KPCA-CEEMD-AWOA-ELM, the improvement rates of the three indicators are 17.97%, 18.87% and -6.79%. For KPCA-CEEMD-AWOA-ELM and KPCA-CEEMDAN-AWOA-ELM, the improvement percentages are 31.95%, 31.52% and 49.68%, respectively. This shows that CEEMDAN decomposed the original data and transformed into more regular subsequences and eliminates reconstruction errors which make prediction easier.
4. Applying the principal component extracted by kernel principal component analysis as input, the prediction effect of the former is better than that of PCA. For KPCA-AWOA-RELM, the values of MAPE, RMSE, and MAE are 0.1349%, 28.2314 and 36.2668, respectively. Three indexes of PCA-CEEMDAN-AWOA-RELM are 0.6243%, 130.9433

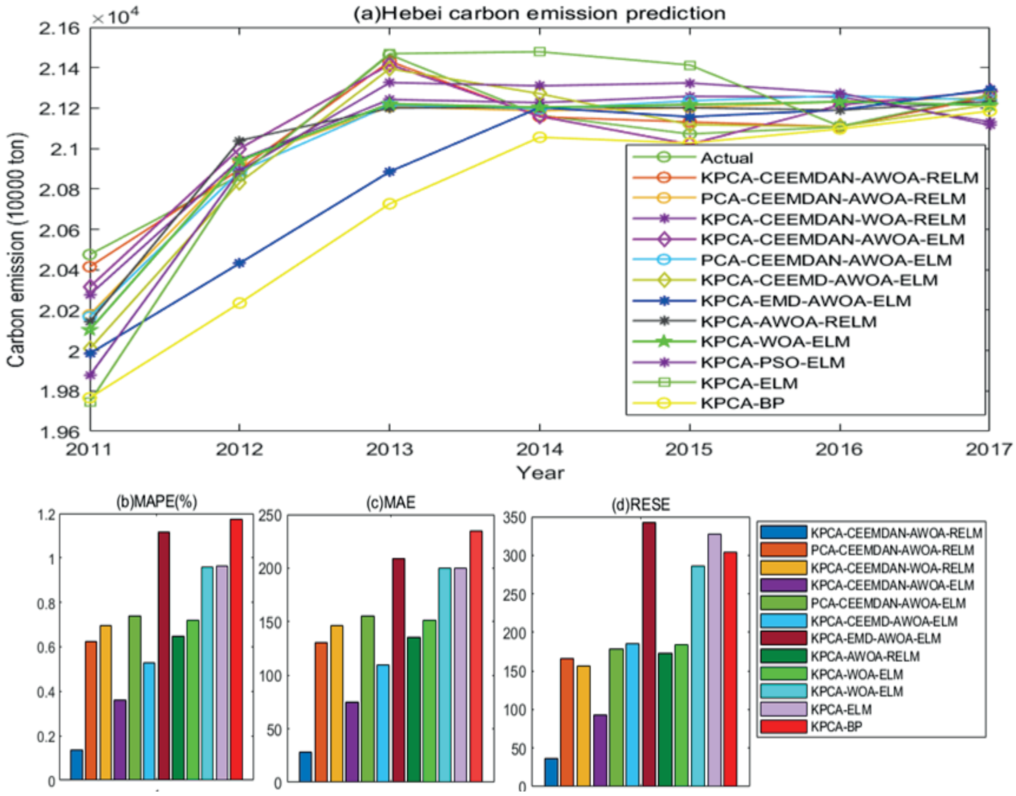

Fig. 4. a) The CO_2 emission prediction fitting curve in Hebei; b) MAPE; c) MAE; d) RMSE.

and 166.3573. For PCA-CEEMDAN-AWOA-ELM and KPCA-CEEMDAN-AWOA-ELM, the improvement rates of the three indicators are 51.14%, 51.36% and 47.83%, respectively. The improvement percentages of PCA-CEEMDAN-AWOA-RELM and KPCA-CEEMDAN-AWOA-RELM are 78.39%, 78.44% and 78.20%, respectively. This further demonstrates that KPCA is superior to PCA.

Case Two

Taking the actual data of China's CO_2 emissions as an example, the prediction results of each model are shown in Fig. 5. Table 5 shows the performance comparison results of each model. Table 6 represents the percentile of improving results. From these charts, we can get similar analysis results as case 1.

Table 5. The prediction results in China.

	MAPE (%)	MAE	RMSE
KPCA-CEEMDAN-AWOA-RELM	0.3474	123018.54	111467.50
PCA-CEEMDAN-AWOA-RELM	0.3985	127922.67	139873.50
KPCA-CEEMDAN-WOA-RELM	0.5634	184318.44	227738.61
KPCA-CEEMDAN-AWOA-ELM	0.4079	129281.47	156222.19
PCA-CEEMDAN-AWOA-ELM	0.4952	158297.67	177815.30
KPCA-CEEMD-AWOA-ELM	0.5411	172717.50	180320.59
KPCA-EMD-AWOA-ELM	0.4514	146448.02	162829.16
KPCA-AWOA-RELM	0.5557	179411.75	194018.41
KPCA-WOA-ELM	0.6293	200592.50	219846.10
KPCA-PSO-ELM	0.9673	297710.32	368587.15
KPCA-ELM	1.0656	334529.93	389682.31
KPCA-BP	1.9995	640618.14	651116.11

Table 6. The percentage improvement results of the models in China.

Benchmark model		Comparative model	PMAPE (%)	PMAE (%)	PRMSE (%)
KPCA-BP	VS.	KPCA-ELM	46.71	47.78	40.15
KPCA-PSO-ELM	VS.	KPCA-WOA-ELM	34.94	32.62	40.35
KPCA-WOA-ELM	VS.	KPCA-AWOA-RELM	11.69	10.56	11.75
KPCA-AWOA-RELM	VS.	KPCA-CEEMD-AWOA-ELM	2.64	3.73	7.06
KPCA-CEEMD-AWOA-ELM	VS.	KPCA-CEEMDAN-AWOA-ELM	24.61	25.15	13.36
KPCA-CEEMDAN-AWOA-ELM	VS.	KPCA-CEEMDAN-AWOA-RELM	14.82	4.84	28.65
KPCA-CEEMDAN-WOA-RELM	VS.	KPCA-CEEMDAN-AWOA-RELM	38.33	33.26	51.05
PCA-CEEMDAN-AWOA-ELM	VS.	KPCA-CEEMDAN-AWOA-ELM	17.63	18.33	12.14
PCA-CEEMDAN-AWOA-RELM	VS.	KPCA-CEEMDAN-AWOA-RELM	12.82	3.83	20.31

All in all, conclusions are drawn as follows in light of the above case analysis.

1. In the above two cases, the KPCA-CEEMDAN-AWOA-RELM model proposed is the best in this paper, compared with other comparative models. According to the evaluation criteria, it shows that the model performs the optimum and robustness in CO_2 emissions prediction.
2. It is necessary to optimize the regularized ELM by an adaptive whale optimization algorithm. According to the experimental results, the prediction results of AWOA-RELM are more satisfactory than conventional WOA. At the same time, it is also verified that the prediction effect of RELM is better than the traditional ELM.
3. CEEMDAN can effectively improve the prediction accuracy model. The prediction effect of the model by CEEMDAN decomposed is better than other models with CEEMD, EMD decomposed and not decomposed. The possible reasons are that there are the high nonlinearity, complexity and chaotic of CO_2 emissions cause large prediction errors

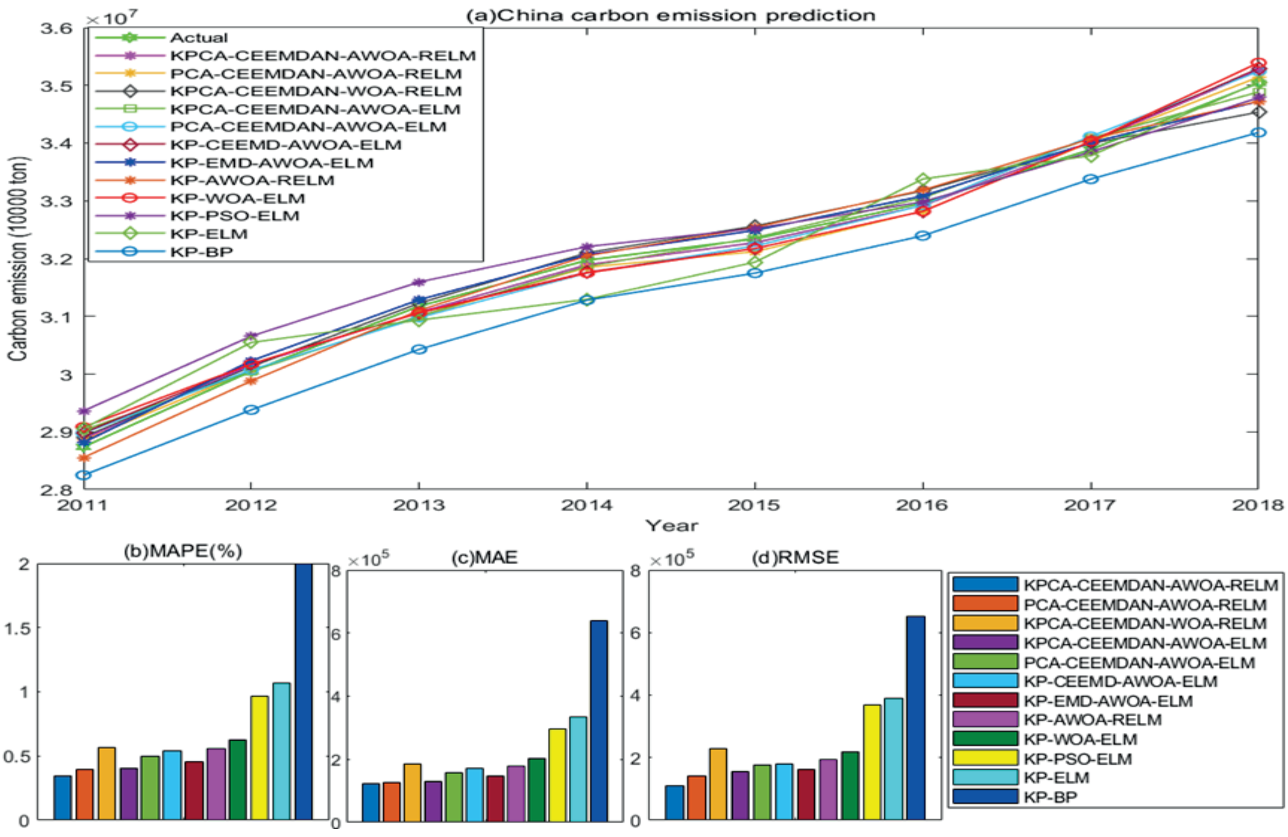

Fig. 5. a) The CO_2 emission prediction fitting curve in China; b) MAPE; c) MAE; d) RMSE.

without data processing. Therefore, the application of CEEMDAN decomposition to CO_2 emissions prediction is reasonably and effectively, eliminating reconstruction errors as much as possible.
4. As an input variable, the principal component extracted by KPCA is better than PCA. On the basis of results, it is found that KPCA is reasonable and effective for nonlinear dimension reduction.

Conclusions

In this paper, a new hybrid model (KPCA-CEEMDAN-AWOA-RELM) which mainly involve KPCA, CEEMDAN, and RELM optimized by AWOA. Firstly, these influencing factors which mainly contain historical CO_2 emissions, energy, economy and society, are tested by a double significance test, to verify that the selected factors have a significant impact on CO_2 emissions. Then, to reduce the input variables, those factors are nonlinear dimension reduced by KPCA. Taking Hebei and China as examples, KPCA-CEEMDAN-AWOA-RELM is compared with other models to verify the validity and applicability of the model. The results show that the local optimization defects can be avoided through AWOA of input weight matrix and deviation matrix, and the prediction effect of RELM is better than the traditional ELM. In addition, KPCA is preferable to PCA in nonlinear dimension reduction. To reduce energy consumption, promote CO_2 emissions reduction and low-carbon energy development, some policy recommendations is as follows: (1) The development of clean energy is energetically promoted to advance the renewal energy to replace traditional energy; (2) The realization of "coal to electricity" can be faster promoted, such as central heating, geothermal energy and other clean heating methods; (3) Reducing the use of primary energy, especially coal, is the most critical measure to improve air quality and promote the sustainable development of energy.

The prediction results and conclusions in this paper lay a solid foundation for our future research, especially the research focus of energy consumption, CO_2 emissions and low-carbon energy development.

Acknowledgements

Thanks to Wei Sun for checking the language logic of this article and Zhaoqi Li's English grammar check. And we appreciate sincerely the electronic library and kinds of databases purchased by North China Electric Power University for reliable empirical data.

Conflict of Interest

The authors declare no conflict of interest.

References

1. TAIYUAN energy low carbon development forum. Available online: http://www.taiyuanforum.com (accessed on 29 October 2019).
2. BP statistical reviews of world energy beyond Petroleum. Available online: http://www.bp.com/statisticalreview (accessed on 29 October 2019).
3. HUANG Y.S., SHEN L., LIU H. Grey relational analysis, principal component analysis and forecasting of carbon emissions based on long short-term memory in China. Journal of Cleaner Production, **209**, 415, **2019**.
4. RUSTEMOGLU H., ANDRES A.R. Determinants of CO_2 emissions in Brazil and Russia between 1992 and 2011: a decomposition analysis. Environmental Science & Policy, **58**, 95, **2016**.
5. SHAFIEI S., SALIM R.A. Non-renewable and renewable energy consumption and CO_2 emissions in OECD countries: a comparative analysis. Energy Policy, **66**, 547, **2014**.
6. SAY N.P., YUCEL M. Energy consumption and CO_2 emissions in Turkey: Empirical analysis and future projection based on an economic growth. Energy Policy, **34**, 3870, **2006**.
7. ZHOU X., ZHANG M., ZHOU M.H., ZHOU M. A comparative study on decoupling relationship and influence factors between China's regional economic development and industrial energy-related carbon emissions. Journal of Cleaner Production, **142**, 783, **2017**.
8. SONG J.N., YANG W., WANG S., WANG X.E., HIGANO Y., FANG K. Exploring potential pathways towards fossil energy-related GHG emission peak prior to 2030 for China: an integrated input-output simulation model. Journal of Cleaner Production, **178**, 688, **2018**.
9. XIE Z.Q., GAO X.N., YUAN W.H., FANG J.D., JIANG Z.B. Decomposition and prediction of direct residential carbon emission indicators in Guangdong Province of China, Ecological Indicators, **115**, 106344, **2020**.
10. WEN L., CAO Y. A hybrid intelligent predicting model for exploring household CO_2 emissions mitigation strategies derived from butterfly optimization algorithm. Science of the Total Environment, **727**, 138572, **2020**.
11. SUN W., WANG C.F., ZHANG C.C. Factor analysis and forecasting of CO_2 emissions in Hebei, using extreme learning machine based on particle swarm optimization. Journal of Cleaner Production, **162**, 1095-1101, **2017**.
12. LI M.L., WANG W., DE G., JI X.H., TAN Z.F. Forecasting carbon emissions related to energy consumption in Beijing-Tianjin-Hebei region based on grey prediction theory and extreme learning machine optimized by support vector machine algorithm, Energies, **11** (9), 2475, **2018**.
13. SUN W., SUN J.Y. Prediction of carbon dioxide emissions based on principal component analysis with regularized extreme learning machine: The case of China. Environmental Engineering Research, **22** (3), 302, **2017**.
14. SUN W., WANG Y.W., ZHANG C.C. Forecasting CO_2 emissions in Hebei, China, through moth-flame optimization based on the random forest and extreme learning machine. Environmental Science and Pollution Research, **25** (29), 28985, **2018**.
15. LI S., WANG P., GOEL L. Short-term load forecasting by wavelet transform and evolutionary extreme learning machine. Electric Power Systems Research, **122**, 96, **2015**.
16. HAO Y., TIAN C.S. A hybrid framework for carbon trading price forecasting: the role of multiple influence factors. Journal of Cleaner Production, **262**, 120378, **2020**.

17. WANG W.J., ZHAO D., MI Z.Q., FAN L.G. Prediction and Analysis of the Relationship between Energy Mix Structure and Electric Vehicles Holdings Based on Carbon Emission Reduction Constraint: A Case in the Beijing-Tianjin-Hebei Region, China. Sustainability, **11** (10), 2928, **2019**.
18. MO X.Y., ZHANG L., LI, H., QU, Z.X. A Novel Air Quality Early-Warning System Based on Artificial Intelligence. International Journal of Environmental Research and Public Health, **16.**19, 3505, **2019**.
19. MIRJALILI S., LEWIS A. The whale optimization algorithm. Advances in engineering software, **95**, 51, **2016**.
20. SUN W., ZHANG C.C. Analysis and forecasting of the carbon price using multi-resolution singular value decomposition and extreme learning machine optimized by adaptive whale optimization algorithm. Apply Energy, **231**, 1354, **2018**.
21. WEI S.W., WANG T., LI Y.B. Influencing factors and prediction of carbon dioxide emissions using factor analysis and optimized least squares support vector machine. Environmental Engineering Research, **22** (2), 175, **2017**.
22. SUN W., ZHANG C. A hybrid ba-elm model based on factor analysis and similar-day approach for short-term load forecasting. Energies, **11** (5), 1282, **2018**.
23. GAO S.Z., LI T.C., ZHANG Y.M. Rolling bearing fault diagnosis of PSO-LSSVM based on CEEMD entropy fusion. Transactions of the Canadian Society for Mechanical Engineering, **44** (3), 405, **2020**.
24. CAO L.J., CHUA K.S., CHONG W.K., LEE H.P., GU Q.M. A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. Neurocomputing, **55** (1-2), 321, **2003**.
25. SUN W., WANG Y.W. Factor analysis and carbon price prediction based on empirical mode decomposition and least squares support vector machine optimized by improved particle swarm optimization. Carbon Management, **11** (3), 315, **2020**.
26. SUN W., DUAN M. Analysis and forecasting of the carbon price in China's regional carbon markets based on fast ensemble empirical mode decomposition, phase space reconstruction, and an improved extreme learning machine. Energies, **12** (2), 277, **2019**.
27. TORRES M.E., COLOMINAS M.A., SCHLOTTHAUER G., FLANDRIN P. A complete ensemble empirical mode decomposition with adaptive noise, 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4144, **2011**
28. LU H.F., MA X., HUANG K., AZIMI M. Carbon trading volume and price forecasting in China using multiple machine learning models. Journal of Cleaner Production, **249**, 119386, **2020**.
29. CAO J., LI Z., LI J. Financial time series forecasting model based on CEEMDAN and LSTM, Physica A-Statistical Mechanics and its Applications, **519**, 127, **2019**.
30. CALVINO A. A simple method for limiting disclosure in continuous microdata based on principal component analysis. Journal of Official Statistics, **33** (1), 15, **2017**.
31. ALAM S., KWON G.R. Alzheimer's Disease Neuroimaging Initiative. Alzheimer disease classification using KPCA, LDA, and multi-kernel learning SVM. International Journal of Imaging Systems and Technology, **27** (2), 133, **2017.**
32. HUANG G.B., ZHU Q.Y., SIEW C.K. Extreme learning machine: a new learning scheme of feed forward neural networks. 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541), **2**, 985, **2004**.
33. IPCC Guidelines for National Greenhouse Gas Inventories, **2006**.