Introduction

The 2019 Taiyuan energy low CO₂ 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) CO₂ emissions will be decreased to 18% by 2020, and the total amount of carbon emissions...
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 CO₂ emission is predicted by scientific and reasonable methods when several factors comprehensively analyzed.

Many scholars have carried out in-depth research on CO₂ emissions. One of the important research directions is to explore various factors of CO₂ emissions by traditional statistical methods. Based on four main factors that is economic activity, employment, energy intensity and CO₂ intensity, the energy-related CO₂ 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 CO₂ emission was found that had a strong relationship through regression analysis [6]. There was a certain relationship between regional economic development and industrial CO₂ emissions when main factors affecting the change of CO₂ 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 CO₂ emissions.

Another important direction is that many scholars have also researched on CO₂ emissions prediction combining with many factors. Combined with a variety of influencing factors, the CO₂ emissions are predicted through the system dynamics method [9]. Household CO₂ emissions were explored and predicted with butterfly optimization algorithm optimized the least square support vector machine (BOA-LSSVM) [10]. CO₂ emission was predicted with the extreme learning machine improved by particle swarm optimization (PSO-ELM) [11]. CO₂ 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]. CO₂ emission was predicted based on principal component analysis with regularized extreme learning machine [13]. CO₂ 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 CO₂ emissions forecasting. So this paper will use KPCA to realize the feature dimension reduction of CO₂ emissions.

Because of the instability and unique complexity of CO₂ 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 (EEemd) [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 CO₂ 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.
Prediction and Analysis of CO₂ Emissions...

to predict CO₂ 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 \( \delta(n) \) is added, and the \( t \)-th decomposition can be expressed as \( x(t)(n) = x(n) + \delta(n) \). The i-th mode component generated by EMD and CEEMDAN decomposition is decomposed: \( E_i(\cdot) \) and IMF1. The steps of CEEMDAN are as follows:

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

\[
\text{IMF}_1(n) = \frac{1}{T} \sum_{t=1}^{T} E_1(r_1(n)) + \varepsilon_1 E_1(\delta^t(n))
\]  

(5)

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

\[
r_1(n) = x(n) - \text{IMF}_1(n)
\]

(2)

3. Repeat the test \( t \) time. In each test, the signal \( r_1(n) + \varepsilon_2 E_1(\delta^t(n)) \) is decomposed and stops when the first EMD mode component is obtained. The second mode component is obtained:

\[
\text{IMF}_2(n) = \frac{1}{T} \sum_{t=1}^{T} E_1(r_2(n)) + \varepsilon_1 E_1(\delta^t(n))
\]

(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:

\[
r_i(n) = r_{i-1}(n) - \text{IMF}_i(n)
\]

(4)

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:

\[
R(n) = x(n) - \sum_{i=1}^{l} \text{IMF}_i
\]

(6)

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

\[
x(n) = \sum_{i=1}^{l} \text{IMF}_i + R(n)
\]

(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:

\[
D = |C \ast X_L(t) - X(t)|
\]

(8)

\[
X(t + 1) = X_L(t) - A \ast D
\]

(9)

...where \( t \) represents the current number of iterations, \( X(t) \) 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.

\[
A = 2a \ast r - a
\]

(10)

\[
C = 2 \ast r
\]

(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:

\[
X(t + 1) = D' \ast e^{b1} \ast \cos(2 \pi l) + X_L(t)
\]

(12)

...where \( D' \) is the distance between whales and prey \( X(t) \), local optimization \( X_L(t) = |X(t) - X(t)| \). b defined Constant of helix shape. l 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:

\[ X(t+1) = f(x) = \begin{cases} \ X(t) - A \ast D & \text{if } p < 0.5 \\ \ D \ast e^{bl \ast \cos(2\pi t)} + X(t) & \text{if } p \geq 0.5 \end{cases} \]  

(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:

\[ D = |C \ast X_{rand} - X(t)| \]  

(14)

\[ X(t+1) = X_{rand} - A \ast D \]  

(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:

\[ \omega = (\omega_{max} - \omega_{min}) \ast i / \text{Max}_{\text{iteration}} \]  

(16)

...where \( i \) is the current number of iterations, and \( \text{Max}_{\text{iteration}} \) 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:

\[ \sum_{l=1}^{L} \beta_l g(\omega_l \cdot x_j + b_l) = y_j, j = 1, 2, ..., N \]  

(18)

It can be abbreviated as:

\[ H \beta = y \]  

(19)

...where

\[ H(\omega_1, ..., \omega_L, x_1, ..., x_N, b_1, ..., b_L) = \begin{bmatrix} g(\omega_1 \cdot x_1 + b_1) & ... & g(\omega_L \cdot x_1 + b_L) \\ ... & ... & ... \\ g(\omega_1 \cdot x_N + b_1) & ... & g(\omega_L \cdot x_N + b_L) \end{bmatrix}_{N \times L} \]  

(20)

The output weight can be obtained by linear least squares:

\[ ||H\beta - y|| = ||H H^+ y - y|| = \min_\beta ||H\beta - y|| \]  

(22)

The least square solution obtained is as follows:

\[ \beta = H^+ y \]  

(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:

\[ \min_\beta C \|y - H\beta\|^2 + \|\beta\|^2 \]  

(24)

The constraints are as follows:

\[ \min_\beta C \|e\|^2 + \|\beta\|^2 \]  

(25)

s. t. \( y - H\beta = e \)  

(26)

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

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

\[ L(\beta, e, \delta) = C \|e\|^2 + \|\beta\|^2 + \delta^T(y - H\beta - e) \]  

(27)

...where non-negative \( \delta \) is the Lagrangian multiplier. The relevant optimization conditions are as follows:

\[ \frac{\partial L}{\partial \beta} = 0 \Rightarrow 2\beta - H^T \delta = 0 \]  

(28)

\[ \frac{\partial L}{\partial e} = 0 \Rightarrow 2e - \delta = 0 \]  

The output weight matrix \( \beta \) is as follows:

\[ \beta = \begin{cases} (H^T H + (1/C))^{-1} H^T y & N > L \\ (H^T H + (1/C))^{-1} y & N < L \end{cases} \]  

(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 CO2 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.

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

Experimental

Data Selection

To verify the effectiveness of the selected model, this paper selects Hebei Province and China’s CO2 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. CO2 emissions can’t be directly obtained, but it can be obtained through the conversion to energy CO2 emissions coefficient from the IPCC guidelines on national greenhouse gas inventories [33].

The historical data is a key to CO2 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 CO2 emissions. CO2 emission is non-linear and non-stationary. Results of PACF analysis were shown that CO2 emissions with a confidence level of 90% and lag 1 have a strong correlation. So the first order lag of CO2 emissions was also selected as one of influencing factors.
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>
<thead>
<tr>
<th>Index</th>
<th>Pearson correlation</th>
<th>Index</th>
<th>Pearson correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>First order lag of CO$_2$ emissions</td>
<td>0.997</td>
<td>Industrial added value</td>
<td>0.966</td>
</tr>
<tr>
<td>Coal</td>
<td>0.999</td>
<td>GDP per capita</td>
<td>0.963</td>
</tr>
<tr>
<td>Crude oil</td>
<td>0.982</td>
<td>Total import and export amount</td>
<td>0.949</td>
</tr>
<tr>
<td>Natural gas</td>
<td>0.789</td>
<td>Fixed asset investment of the whole society</td>
<td>0.872</td>
</tr>
<tr>
<td>Primary electricity</td>
<td>1.000</td>
<td>Consumption level of residents</td>
<td>0.939</td>
</tr>
<tr>
<td>GDP</td>
<td>0.895</td>
<td>Population</td>
<td>0.913</td>
</tr>
<tr>
<td>Value of primary industry</td>
<td>0.981</td>
<td>City rate</td>
<td>0.946</td>
</tr>
<tr>
<td>Value of secondary industry</td>
<td>0.964</td>
<td>Power generation</td>
<td>0.991</td>
</tr>
<tr>
<td>Value of third industry</td>
<td>0.923</td>
<td>Steel production</td>
<td>0.922</td>
</tr>
</tbody>
</table>

Table 2. Results of KPCA in Hebei.

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

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

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:

$$\text{MAPE} = \left(\frac{1}{n}\sum_{t=1}^{n} \left|\hat{y}_t - y_t\right| \right) \times 100\% \text{ (31)}$$

$$\text{MAE} = \left(\frac{1}{n}\sum_{t=1}^{n} \left|\hat{y}_t - y_t\right| \right) \text{ (32)}$$

Fig. 2. Result of CEEMDAN in Hebei.
In order to more intuitively 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:

\[
P_{\text{MAPE}} = \frac{\text{MAPE}_1 - \text{MAPE}_2}{\text{MAPE}_1} \times 100\% \tag{34}
\]

\[
P_{\text{MAE}} = \frac{\text{MAE}_1 - \text{MAE}_2}{\text{MAE}_1} \times 100\% \tag{35}
\]

\[
P_{\text{RMSE}} = \frac{\text{RMSE}_1 - \text{RMSE}_2}{\text{RMSE}_1} \times 100\% \tag{36}
\]

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.
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-ELM are 0.6243\%, 130.9433 and 166.3573.

### Table 3. The prediction results in Hebei.

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

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

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

<table>
<thead>
<tr>
<th>Model</th>
<th>MAPE (%)</th>
<th>MAE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>KPCA-CEEMDAN-AWOA-RELM</td>
<td>0.3474</td>
<td>123018.54</td>
<td>111467.50</td>
</tr>
<tr>
<td>PCA-CEEMDAN-AWOA-RELM</td>
<td>0.3985</td>
<td>127922.67</td>
<td>139873.50</td>
</tr>
<tr>
<td>KPCA-CEEMDAN-RELM</td>
<td>0.5634</td>
<td>184318.44</td>
<td>227738.61</td>
</tr>
<tr>
<td>KPCA-CEEMDAN-AWOA-ELM</td>
<td>0.4079</td>
<td>129281.47</td>
<td>156222.19</td>
</tr>
<tr>
<td>KPCA-CEEMDAN-WOA-ELM</td>
<td>0.4952</td>
<td>158297.67</td>
<td>177815.30</td>
</tr>
<tr>
<td>KPCA-CEEMDAN-AWOA-ELM</td>
<td>0.5411</td>
<td>172717.50</td>
<td>180320.59</td>
</tr>
<tr>
<td>KPCA-EMD-AWOA-ELM</td>
<td>0.4514</td>
<td>146448.02</td>
<td>162829.16</td>
</tr>
<tr>
<td>KPCA-AWOA-ELM</td>
<td>0.5557</td>
<td>179411.75</td>
<td>194018.41</td>
</tr>
<tr>
<td>KPCA-WOA-ELM</td>
<td>0.6293</td>
<td>200592.50</td>
<td>219846.10</td>
</tr>
<tr>
<td>KPCA-PSO-ELM</td>
<td>0.9673</td>
<td>297710.32</td>
<td>368587.15</td>
</tr>
<tr>
<td>KPCA-ELM</td>
<td>1.0656</td>
<td>334529.93</td>
<td>389682.31</td>
</tr>
<tr>
<td>KPCA-BP</td>
<td>1.9995</td>
<td>640618.14</td>
<td>651116.11</td>
</tr>
</tbody>
</table>
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 CO₂ 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 CO₂ emissions cause large prediction errors.
without data processing. Therefore, the application of CEEMDAN decomposition to CO₂ 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 CO₂ emissions, energy, economy and society, are tested by a double significance test, to verify that the selected factors have a significant impact on CO₂ 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 CO₂ 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, CO₂ 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