Prediction of carbon dioxide emissions based on principal component analysis with regularized extreme learning machine: The case of China

Wei Sun[*], **Jingyi Sun**[*†]

Department of Business Administration, North China Electric Power University, Baoding 071000, China
*These authors contributed equally to this work

Abstract

Nowadays, with the burgeoning development of economy, CO_2 emissions increase rapidly in China. It has become a common concern to seek effective methods to forecast CO_2 emissions and put forward the targeted reduction measures. This paper proposes a novel hybrid model combined principal component analysis (PCA) with regularized extreme learning machine (RELM) to make CO_2 emissions prediction based on the data from 1978 to 2014 in China. First eleven variables are selected on the basis of Pearson coefficient test. Partial autocorrelation function (PACF) is utilized to determine the lag phases of historical CO_2 emissions so as to improve the rationality of input selection. Then PCA is employed to reduce the dimensionality of the influential factors. Finally RELM is applied to forecast CO_2 emissions. According to the modeling results, the proposed model outperforms a single RELM model, extreme learning machine (ELM), back propagation neural network (BPNN), GM(1,1) and Logistic model in terms of errors. Moreover, it can be clearly seen that ELM-based approaches save more computing time than BPNN. Therefore the developed model is a promising technique in terms of forecasting accuracy and computing efficiency for CO_2 emission prediction.

Keywords: CO_2 emissions prediction, Influential factors, Pearson coefficient test, PACF, PCA, RELM

http://eeer.org

1. Introduction

After over 30 y of the economic reforms in China, there emerges a remarkable rise with average annual GDP growth rate at nearly 10% [1]. However, the burgeoning development of economy inevitably results in the large increase in energy consumption and hence CO_2 emissions. China overtook the United States as the world's leading emitter of CO_2 in the year of 2006 [2]. To respond to the cause of serious global warming, China commits to continue taking effective measures for CO_2 emission control during the 13th Five-Year Plan, thus the carbon emission peaking can be reached by 2030. Accordingly, it's of great significance to focus on CO_2 emissions prediction research, which provides a valuable reference for practical measures of CO_2 emission reduction.

A report published by the National Petroleum Council in the United States predicted a 50%-60% growth in total global demand for energy by 2030 [3]. Energy consumption is the main source of CO_2 emissions [4], thus a lot of researchers have paid attention to this area. Say and Yucel [5] studied the relationship between total energy consumption and total CO_2 emissions through regression analysis, which displayed a strong relationship between these two factors. In [6], the energy consumptions were modeled using artificial neural network (ANN) based on economic and demographic variables. The results showed that the correlation coefficients between ANN predictions and actual energy consumptions were higher than 90%, which indicated a high reliability of ANN for forecasting future energy consumption. Azadeh and Tarverdian [7] presented an integrated algorithm based on genetic algorithm, computer simulation and design of experiments using stochastic procedures for

monthly electrical energy consumption prediction. In [8], Utgikar and Scott explored the possible causes of inaccuracy in energy forecasting which could provide a better understanding of prediction process and design a strategy for reducing the errors in energy prediction. Aydin [9] utilized multiple linear regression analysis to study the relationship between CO_2 emissions and energy-related factors where correlation analysis were employed to determine the influential factors. In [10], an approach was proposed for coal-related CO_2 projections in future planning, wherein coal-related CO_2 emissions were modeled by trend analysis. Feng and Zhang [11] conducted a case study to predict the effects of different development alternatives on future energy consumption and carbon emission, namely under three scenarios: business-as usual, basic-policy and low-carbon. The results provided insights into the energy future and highlighted possible steps to develop a sustainable low-carbon city.

At this stage, researches on carbon emissions can be mainly divided into two parts: discussion on influential factors and study on prediction models. For the influential factors, existing studies related to this part include the methods such as index decomposition means [12-13] and input-output structural analysis [14-15]. Andres and Rustemoglu [16] introduced refined laspeyres index method into the research of relationships between CO_2 emissions and four identified factors in Brazil and Russia for the period 1992-2011 to explore the determinants of accelerating CO_2 emissions. Li et al. [17] applied logarithmic mean divisia index method (LMDI) to decompose the change in carbon emissions into some influencing factors caused by urbanization. The results revealed that energy intensity contributed largely to carbon emission reduction in Hubei Province. Li et al. [18] estimated the agricultural CO_2

emissions in China during the period of 1994 to 2011 and applied LMDI as the decomposition technique. The results illustrated that agricultural subsidy acts to reduce CO_2 emissions effectively and has increased in recent years. Wang et al. [19] proved that economic development, energy structure and low energy efficiency are three main driving factors of increasing CO_2 emissions in China based on a modified production-theoretical decomposition analysis approach. Ahmed [20] studied the relationship between CO_2 emissions, economic growth, urbanization and trade openness by two steps: (a) Autoregressive distributed lag bounds test was carried out to explore whether there existed cointegration between the variables. (b) The relationship between the factors was analyzed according to the long-run and short-run dynamics. Based on the last research, Ali et al. [21] added the factor of energy consumption to examine its dynamic impact on CO_2 emissions. Deng et al. [22] combined structural decomposition analysis and logarithmic mean divisia index method to study the drivers behind CO_2 emissions in Yunnan province. This technique could take both production and final demand into account in less-developed regions. Cointegration and Granger causality were adopted by Tang and Tan [23] to examine the relationship among CO_2 emissions, energy consumption, foreign direct investment and economic growth in Vietnam. They pointed out that there existed long-run equilibrium among these variables. Lin et al. [24] evaluated the relation between CO_2 emissions and industrial growth through an autoregressive distributed lag bounds testing and cointegration analysis. The results suggested that there was a reduction potential of CO_2 emissions in the Chinese manufacturing sector without intimidating industrial growth. Wang et al. [25] applied a two-level decomposition model based on Kaya identity to

uncover the main influential factor for CO_2 emissions. The results indicated that energy intensity reduction was conducive to low-carbon economic development. Based on combining correlation analysis, gray correlation analysis and principal component regression analysis, Bian et al. [26] integrated data envelopment analysis with energy structure adjustment to measure CO_2 emission reduction in China. The findings showed that it was a practical way to decrease CO_2 emissions through the abatement of coal consumption and development of non-fossil energy.

For the forecasting techniques, CO_2 emissions are predicted mainly through the relationship models between carbon emission and its influencing factors based on different scenarios. Kang et al. [27] employed STIPRAT model to examine the impact of energy- related factors on CO_2 emissions and tested the spillover effects of per capita CO_2 emissions through a spatial panel data technique. This study provided some policy advice on reduction of China's CO_2 emissions. Based on STIPRAT model, Sheng and Guo [28] extended this basic method to be a panel error-correction one which can dynamically take the influence of urbanization changes on total CO_2 emissions into consideration. Their findings indicated that the rapid urbanization augmented CO_2 emissions both in the short-run and long-run. Wu et al. [29] utilized a multi-variable grey model to forecast CO_2 emissions on the basis of energy consumption, urban population and economic growth. Pérez-Suárez et al. [30] compared environmental kuznets curve with logistic growth model in CO_2 emission prediction considering a sample of 175 countries. The results showed that extended environmental kuznets curve tended to outperform the forecasting accuracy of the latter one. Vector autoregressive

model was adopted by Xu and Lin [31] to identify the drivers of CO_2 emissions in China's iron and steel industry. The findings revealed that energy efficiency played a significant part in CO_2 emission reduction. Baareh [32] introduced four input data including oil, natural gas, coal and primary energy consumption to build ANN for CO_2 emission prediction. The results proved ANN was a powerful and efficient tool in forecasting CO_2 emissions. A hybrid model that combined ANN with bees algorithm for analyzing CO_2 emissions in the world was presented by Behrang et al. [33]. Two steps were carried out: (a) The bees algorithm was applied to determine the indicators. (b) World CO_2 emissions were forecasted up to the year of 2040 based on ANN.

With the propositions and prosperities of artificial intelligent algorithms, traditional neural networks offer a new way of CO_2 emission prediction. Despite their strong nonlinear mapping ability and parallel processing capability, the drawbacks of these methods are the slow learning speed, complex training parameters and easily tapping into the local minimum. Huang et al. [34] introduced extreme learning machine (ELM) to solve the stated issue of conventional training methods. With the advantages such as fast convergence speed, high training accuracy and no manual tuning, the ELM model has been successfully applied to forecasting problems in many fields, such as wind speed [35], electricity load [36], oil price [37] and so on. However, ELM is based on empirical risk minimization principle which easily causes over-fitting phenomenon. Therefore, in order to guarantee the global optimization and generalization ability, RELM model, in which the calculation process of Moore-Penrose generalized inverse

and the introduction of the regularization factor are added to ELM, is used for CO_2 emission prediction in this paper.

In general, based on the aforementioned studies, it can be found that the appropriate selection of influential factors has momentous influence on the prediction results of CO_2 emissions. However, most studies only put emphasis on the impact of these factors on the total CO_2 emissions and ignore the correlation to each other. In reality, there exist overlaps of information contained in the data, thus, the computational efficiency is greatly depressed due to the complex network. Therefore, principal component analysis (PCA) is employed in this paper to reduce the dimension of pre-select influential factors with retention of information to the utmost so that the network structure can be simplified and operation efficiency and prediction accuracy can be significantly improved.

Therefore, compared with past works, there exist two main differences: (i) RELM, a new kind of neural networks, is firstly introduced into CO_2 emission prediction, which overcomes the disadvantages of slow learning speed, the need of numerous training samples, over-fitting and so on in the previous researches. (ii) The correlations among influential factors are paid close attention in this paper, thus PCA is utilized to manipulate them for dimension reduction to improve the computational efficiency and forecasting precision. The rest of this paper is organized as follows: Section 2 presents a brief description of PCA, ELM and RELM; Section 3 displays the framework of the proposed novel approach in this study; Section 4 elaborates the selection of input; Section 5 validates the established model through a case study; Finally, the

paper is concluded and several concrete mitigation measures have been further put forward in Section 6.

2. Methodology

2.1. Principal Component Analysis

PCA was initially introduced in the discussion of non-random variables by Pearson [38] and extended to random one by Hotelling [39]. This method can effectively reduce the dimensionality of a data set on the premise of retaining main variance. It is achieved by applying orthogonal transformation to convert the data into a new set of indexes, also called PCs, that meet: (i) Each PC is a linear combination of original variables. (ii) PCs are uncorrelated to each other. The first PC accounts for the most information of original index and the largest proportion of variability which its predecessors have not explained is interpreted by each subsequent one. In this paper, the PCA calculation was performed on SPSS v.19.0 and the accumulative explained variation of the selected PCs should be more than 0.85.

2.2 Extreme Learning Machine

The ELM is a novel machine learning algorithm for single layer feed-forward neural networks. The main nature of ELM is the random initialization of the input weights and hidden biases without iterative adjustments during the learning process, thus the optimal output weights can be quickly obtained based on the predefined network structure [40]. Besides its fast learning speed, ELM also avoids numerous problems such as local minima and learning rate faced by

other ANNs [41]. The topological structure of ELM network is illustrated in Fig. 1. The specific procedures of ELM are described as follows:

Given a training data set with N samples $\{(x_i, y_i)\}_{i=1}^{N}$, the ELM model with L hidden nodes are expressed as

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

where x_j is the input pattern, y_j is the desired output, $w_i \in R$ is the randomly assigned input weight vector between the ith hidden node and input nodes. b_i is the randomly selected bias of the ith hidden node. $g(\cdot)$ is an activation function. β_i represents the weight connecting the ith hidden node and output nodes.

The Eq. (1) can be simply written as:

$$H\beta = y \tag{2}$$

where

$$H(w_1, ..., w_L, x_1, ..., x_N, b_1, ..., b_L) = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & ... & g(w_L \cdot x_1 + b_L) \\ ... & ... & ... \\ g(w_1 \cdot x_N + b_1) & ... & g(w_L \cdot x_N + b_L) \end{bmatrix}_{N \times L} \tag{3}$$

$$\begin{cases} \beta = \left[\beta_1^T, ..., \beta_L^T \right] \\ y = \left[y_1^T, ..., y_N^T \right] \end{cases} \tag{4}$$
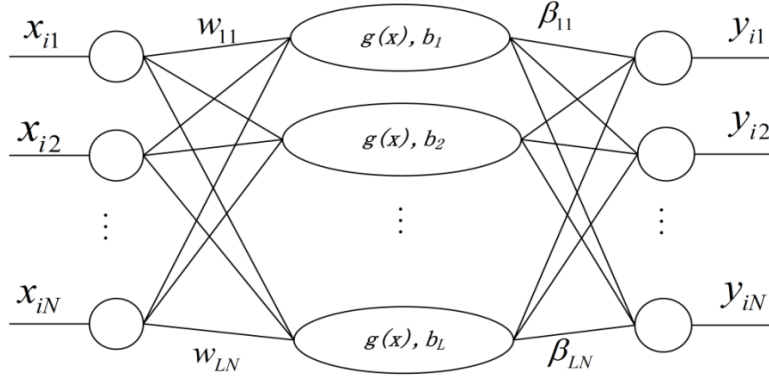
The output weights can be derived by finding the least square solutions to the linear Eq. (5):

$$\left\| H\beta - y \right\| = \left\| HH^+ y - y \right\| = \min_{\beta} \left\| H\beta - y \right\| \tag{5}$$

Here the least square solutions are obtained as follows:

$$\beta = H^+ y \tag{6}$$

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

8

Fig. 1. The ELM network.

2.3. Regularized Extreme Learning Machine

The drawback of standard ELM algorithm is the single consideration of empirical error minimization which gives rise to overfitting and depresses the generalization ability [42]. To solve this problem, both empirical error minimization and structural risk minimization are simultaneously taken into account to achieve the best tradeoff with a regularization parameter C in RELM model [43]. The formula can be described as follows:

$$\min_{\beta} C \|y - H\beta\|_2^2 + \|\beta\|_2^2 \tag{7}$$

Eq. (7) can be also expressed as the following optimization problem with a constraint condition:

$$\min_{\beta} C \|e\|_2^2 + \|\beta\|_2^2 \tag{8}$$

$$s.t. \quad y - H\beta = e \tag{9}$$

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

According to Karush-Kuhn-Tucker (KKT) condition, the corresponding Lagrange function is given by:

$$L(\beta, e, \lambda) = C \|e\|_2^2 + \|\beta\|_2^2 + \lambda^T (y - H\beta - e) \tag{10}$$

where the nonnegative λ is the Lagrangian multiplier. The relevant optimization conditions are shown as follows:

$$\begin{cases} \dfrac{\partial L}{\partial \beta} = 0 & \Rightarrow \quad 2\beta - H^T \lambda = 0 \\[2mm] \dfrac{\partial L}{\partial e} = 0 & \Rightarrow \quad 2Ce - \lambda = 0 \\[2mm] \dfrac{\partial L}{\partial \lambda} = 0 & \Rightarrow \quad y - H\beta - e = 0 \end{cases} \tag{11}$$

For N is the number of training samples, L is the number of hidden nodes, the final output weight matrix β can be derived as follows:

$$\beta = \begin{cases} (H^T H + \dfrac{I}{C})^{-1} H^T y & N > L \\[3mm] H^T (H^T H + \dfrac{I}{C})^{-1} y & N < L \end{cases} \tag{12}$$

3. Approaches of PC-RELM Model

The framework of the proposed model for carbon dioxide emission prediction is displayed in Fig. 2. This novel approach can be explained in detail as follows:

In part I, the Pearson coefficient analysis and bilateral significance test are carried out to study the relationships between the impact factors and carbon dioxide emissions. The partial autocorrelation analysis is adopted to select historical carbon dioxide emissions with highest correlation on the target emission. This section contributes to the pre-selection of input for research. In part II, PCA is employed for feature extraction and dimension reduction of the pre-selected data, which can improve the computational efficiency. Part III aims at realizing carbon dioxide emission prediction through RELM model.

10
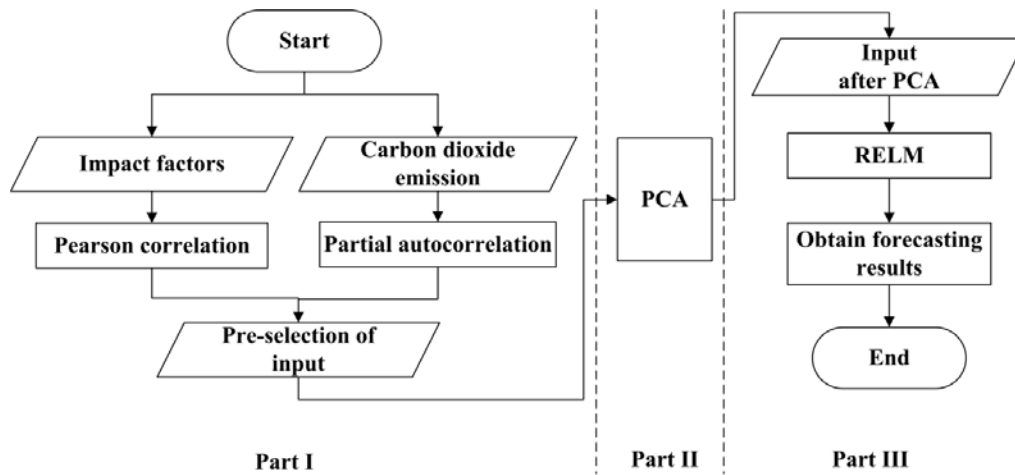
Fig. 2. Framework for carbon dioxide emission prediction based on PC-RELM.

4. Input Selection

4.1. Data Source and Conversion

The research is made based on energy consumption as well as other related data in China from the year of 1978 to 2014. The consumption of total energy and percentage composition of four kinds of energies that contain raw coal, crude oil, natural gas and primary electricity are recorded in China Statistical Yearbook on the basis of standard coal. Considering there is no direct promulgation of CO_2 emissions, conversion coefficients listed in Table 1 are employed to convert the standard coal data to corresponding values of CO_2 emissions, according to the comprehensive report of "China sustainable development of energy and carbon emission scenarios analysis" [44]. The yearly CO_2 emissions in China during the period 1978-2014 are shown in Table 2.

**Table 1.** CO_2 Emission Conversion Coefficients for Different Energy Species

Energy species	Coal	Crude oil	Natural gas	Power
C/(t/t)	0.7467	0.5825	0.4435	0

11

Table 2. CO_2 Emissions over the Period of 1978-2014 in China   (10,000 tons)

Year	Total	Year	Total	Year	Total
1978	38,534.27924	1991	70,231.63738	2004	149,752.0567
1979	39,563.73908	1992	73,756.9878	2005	171,180.9606
1980	40,591.80094	1993	77,973.55682	2006	187,499.1829
1981	42,344.55343	1994	82,210.04039	2007	203,585.7424
1982	44,734.0561	1995	87,488.97443	2008	207,193.8969
1983	47,669.68216	1996	90,002.04612	2009	217,033.0961
1984	52,387.94459	1997	89,694.06873	2010	229,304.1098
1985	51,788.19323	1998	89,684.42261	2011	248,653.6079
1986	54,686.05874	1999	92,955.14501	2012	254,071.7937
1987	58,677.9241	2000	95,437.90723	2013	261,149.6338
1988	62,947.84656	2001	99,844.06383	2014	263,144.034
1989	65,524.42751	2002	109,210.1317		
1990	66,623.92291	2003	128,392.9525		

In Fig. 3, it can be clearly found out that there exists a continuous rise in CO_2 emissions of the total one, which displays the same trend with raw coal. This is due to the fact that coal is the main fuel in China which accounts for nearly 65% in primary energy consumption structure. However, crude oil and natural gas contribute a relatively small proportion of CO_2 emissions. Fig. 4 displays the growth multiple of CO_2 emissions of different energies from 1979 to 2014 with the data in 1978 as the base. The trend can be divided into two stages: there was a gentle rise with slow growth rate before the year of 2002. After this turning point, the growth rate increased rapidly and the growth multiple reached about 6.8 in 2014. Moreover, in response to the national call for energy conservation and emission reduction nowadays, the utilization of natural gas is vigorously promoted. This is the main reason why natural gas presents a significant increase in CO_2 emissions in recent years.
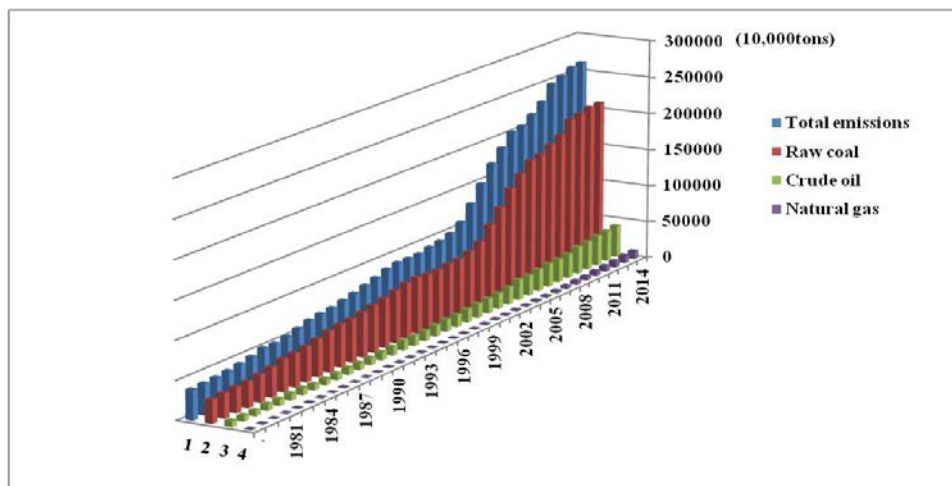
Fig. 3. CO_2 emissions of total energies and main sources during 1978-2014 in China.
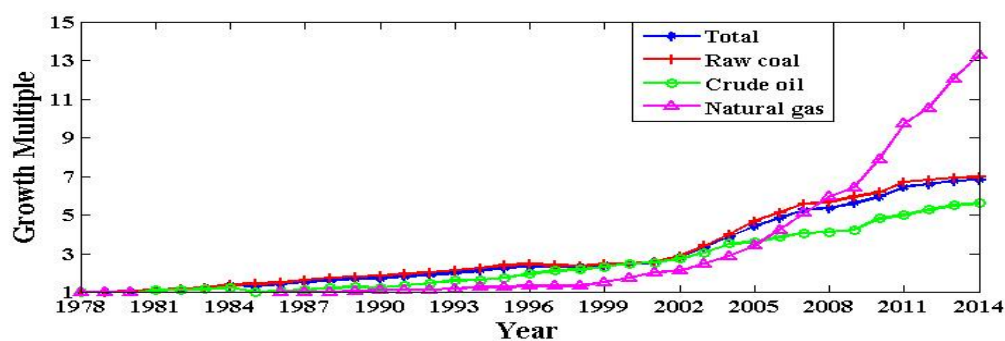
Fig. 4. CO2 emission growth multiple of of total energies and main sources.

4.2. SPSS Analysis

In the previous studies on carbon dioxide emission forecasting, influential factors mainly contain energy consumption, GDP, population, urbanization rate, service industry and so on [45-48]. In this paper, eleven variables are pre-selected from China Statistical Yearbook for CO_2 emissions prediction including coal consumption, GDP of primary industry, GDP of secondary industry, GDP of tertiary industry, population, urbanization level, transportation possession quantity, power generation, steel production, total investment in fixed assets of the whole society and area final consumption.

13

Mining the relationships between CO_2 emissions and the eleven pre-selected variables are essential for the establishment of a good prediction model. Pearson coefficient and bilateral significance test are selected for correlation analysis in this paper. Table 3 presents the values of correlation coefficients. It can be found that all the correlation coefficients are more than 0.8 and the concomitant probability value of bilateral significance test is 0.000 less than 0.05, which reveals that there exists positive and significant correlation between CO_2 emissions and the eleven above-mentioned indicators. Thus, the eleven pre-selected variables all should be taken into account in the CO_2 emission prediction.

Table 3. Correlation Analysis at the Bilateral Significance Level of 0.01

Factors	Coefficient	Significance	Factors	Coefficient	Significance
Coal consumption	1.000	0.000	Transportation possession quantity	0.931	0.000
GDP of primary industry	0.974	0.000	Power generation	0.990	0.000
GDP of secondary industry	0.974	0.000	Steel production	0.988	0.000
GDP of tertiary industry	0.956	0.000	Total investment in fixed assets of the whole society	0.918	0.000
Population	0.866	0.000	Area final consumption	0.966	0.000
Urbanization level	0.973	0.000			

4.3. PACF Analysis

The influence of historical CO_2 emissions on the target one is taken into consideration in this part. PACF is employed to find out the inherent relationship of the dataset. The partial autocorrelograms is illustrated in Fig. 5, where the confidence level is 90%. The results indicated that carbon emssion data in lag 1 and lag 2 showed a strong correlation, thus these two variables are also selected as influential factors.
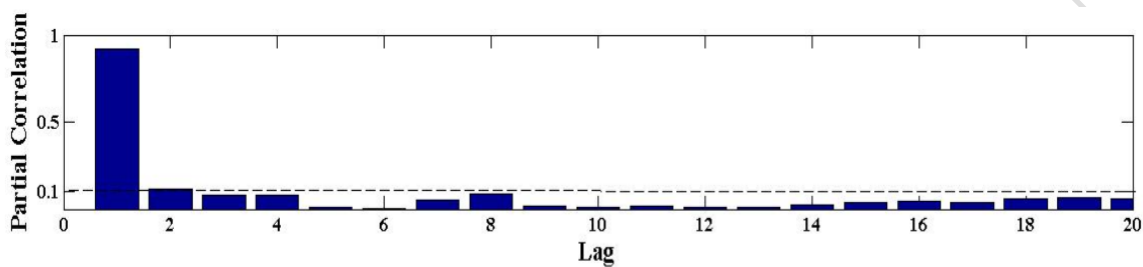

Fig. 5. PACF of total CO_2 emissions dataset.

4.4. PCA Process

Based on the pre-selected thirteen variables in the section of 4.1 and 4.2, PCA is utilized to remove the multicollinearity presented in the predictors. We mine the major information containing in the data through this method. The PCA process result is shown in Table 4 and Fig. 6. It can be seen that the first principal component explain more than 95% of the factors, so this principal component is utilized to replace the predictors as the input.
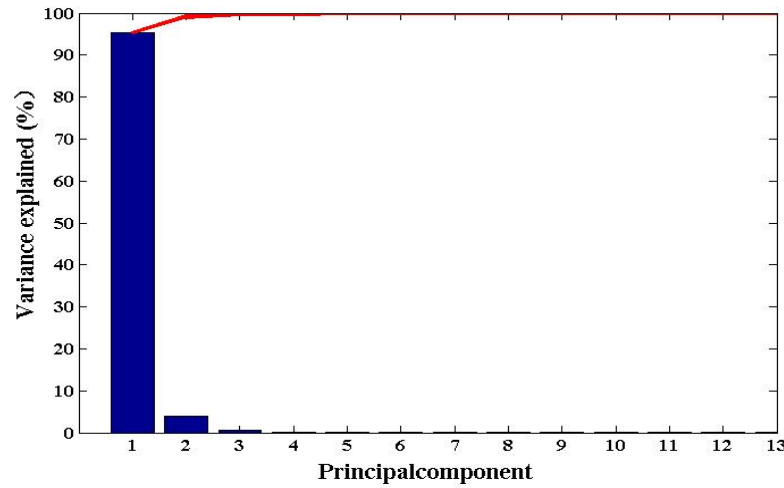
Fig. 6. Scree plot in PCA analysis.

Table 4. Component Matrix

Component	PC1
Coal consumption	0.984
GDP of primary industry	0.996
GDP of secondary industry	0.995
GDP of tertiary industry	0.987
Population	0.845
Urbanization level	0.961
Transportation possession quantity	0.973
Power generation	0.999
Steel production	0.992
Total investment in fixed assets of the whole society	0.962
Area final consumption	0.994
Historical CO_2 emissions(Lag 1)	0.995
Historical CO_2 emissions(Lag 2)	0.993

5. Experiment of CO_2 Emission Prediction in China Based on RELM Model

5.1. Comparative Framework

The experiment of CO_2 emission prediction in China is carried out based on the aforementioned

related data from 1978 to 2014, totally 37 data points. Wherein, the data from 1980 to 2009 are

selected as training set and the remaining 5 data are utilized as test set.

16

As shown in Fig. 7, three comparative parts are contained in the framework. In part I, four basic models including ELM, BPNN, GM(1,1) and Logistic model are introduced to forecast CO_2 emissions. Part II utilizes RELM to test whether the regularization parameter donates to the prediction accuracy and the effectiveness of PCA is explored in Part III.
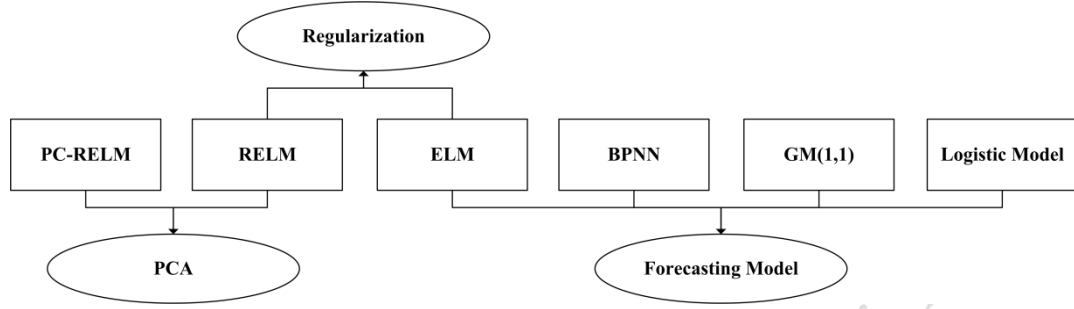
Fig. 7. Comparison framework for CO_2 emission forecasting.

5.2. Evaluation Criteria of Model Performance

In order to determine which forecasting model outperforms the other models, the performance of the prediction models is usually assessed by statistical criteria: relative error (RE), mean absolute percentage error (MAPE), maximum absolute percentage error (MaxAPE), median absolute percentage error (MdAPE) and root mean square error (RMSE). The smaller the values are, the better the forecasting performance is. These five error indexes are defined as follows:

$$RE = \left| \frac{y_t - y_t^*}{y_t} \right| \times 100\% \tag{13}$$

$$MAPE = \frac{1}{N} \sum_{t=1}^{N} \left| \frac{y_t - y_t^*}{y_t} \right| \times 100\% \tag{14}$$

$$MaxAPE = \max(\left| \frac{y_t - y_t^*}{y_t} \right| \times 100\%) \tag{15}$$

$$MdAPE = \text{median}\left(\left|\frac{y_t - y_t^*}{y_t}\right| \times 100\%\right) \tag{16}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(y_t - y_t^*\right)^2} \tag{17}$$

where y_t and y_t^* are the actual and forecast CO_2 emissions at time period t, respectively. N represents the number of CO_2 emissions to be predicted.

5.3. Parameter Setting

As above mentioned, only two parameters need to be pre-set in RELM model. The regularization parameter C and the number of node in hidden layer are set as 2^{10} and 100, respectively. As compared models, the selection of parameters in ELM and BPNN is listed in Table 5.

Table 5. Parameters for ELM and BPNN Model

| Forecasting model	Parameters	Value
ELM	Number of node in hidden layer	100
BPNN	Number of node in hidden layer	50
	Maximum number of convergence	100
	Learning rate	0.1
	Error	0.00004

5.4. Results and Discussion

In Fig. 8, CO_2 emission prediction curves from 2010 to 2014 are obtained by six different models. It can be obviously found out that: (a) in contrast with other five models, the goodness of fit between the forecasted value by PC-RELM and actual value reaches the highest degree; (b)

the fitting condition of ELM-based models is generally better than other techniques mainly due to the strong generalization ability; (c) the hybrid model PC-RELM presents higher predicted precision than RELM which indicates that the PCA part can effectively improve the prediction performance of the single RELM.
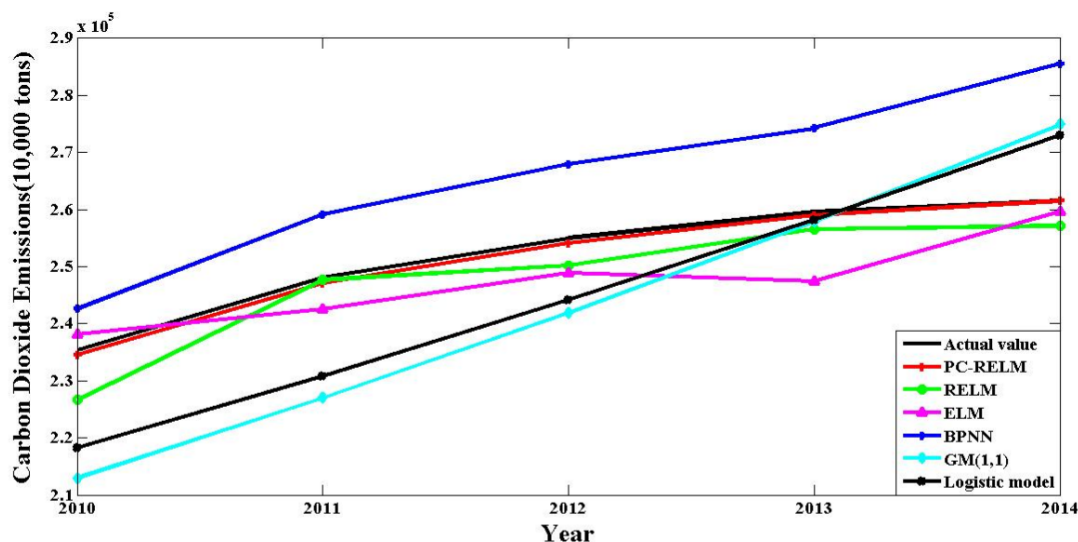
Fig. 8. CO_2 emission forecasting results from2010 to 2014.

Fig. 9 displays the relative errors achieved by the six prediction models. The relative errors obtained by PC-RELM are all under 0.5% which outperforms other methods except in the year of 2011. The single RELM model exhibits a slightly lower error in the second point than PCA-RELM. In addition, there emerges large deviation between the actual values and predicted ones in BPNN and GM(1,1) model where the maximum relative errors are both over 9%.
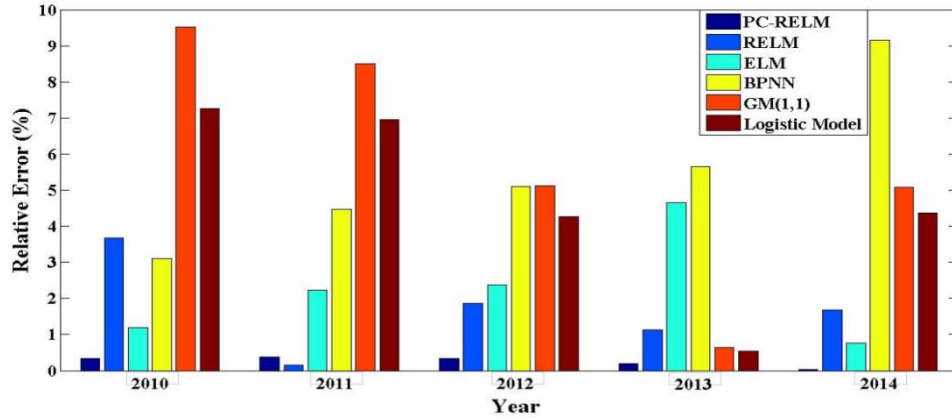
Fig. 9. Comparison of relative errors for different forecasting models.

The statistical errors of the six forecasting techniques are clearly shown in Table 6. The analysis manifests that: (a) PC-RELM model provides the best prediction results in terms of MAPE, MaxAPE, MdAPE and RMSE. (b) Compared with RELM model, the PCA part in PC-RELM removes the multicollinearity in pre-selected influential factors and simplifies the network structure which contributes to the operation efficiency and the improvement of forecasting performance. (c) The errors of RELM is lower than ELM mainly due to the fact that the introduction of the regularization parameter in RELM enhances the global optimization and generalization ability of ELM model in CO_2 emission prediction. (d) Considering the significant influence of representative samples on BPNN, the MAPE, MaxAPE, MdAPE and RMSE values are higher than ELM-based algorithms. (e) The errors obtained by GM(1,1) are largest among the six models mainly because the smoothness degree of the original data has an impact on the prediction accuracy. (f) The prediction precision of Logistic model is higher than BPNN and GM(1,1) while it is remarkably lower than ELM-based models. The MAPE value of Logistic model is 18.5 times larger than PC-RELM.

20

Table 6. Statistical Error Measures of Prediction Methods

Forecasting Methods	Indexes			
	MAPE	MaxAPE	MdAPE	RMSE
PC-RELM	0.252%	0.378%	0.330%	0.003
RELM	1.704%	3.681%	1.690%	0.021
ELM	2.239%	4.651%	2.223%	0.026
BPNN	5.498%	9.164%	5.097%	0.059
GM(1,1)	5.774%	9.517%	5.123%	0.066
Logistic Model	4.683%	7.270%	4.377%	0.053

The computing time of PC-RELM, RELM, ELM and BPNN for continuously running for 100 times in MATLAB 2014a on a Windows 7 system is shown in Table 7. It can be clearly seen that ELM-based models save more computing time than BPNN model mainly because there is no need to update the randomly selected parameters in the learning process. In contrast with ELM, RELM takes 2.44 s for computing, which is slightly longer than ELM. Therefore, the regularization part has little impact on the running speed of ELM while improving the forecasting accuracy. Notably, PC-RELM is 0.3 s shorter than RELM, thus the PCA part can upgrade the running speed to some degree with the improvement of prediction precision.

Table 7. Computing Time of PC-RELM, RELM, ELM and BPNN Models

	PC-RELM	RELM	ELM	BPNN
$t(s)$ (100 times)	2.14	2.44	1.52	65.87

6. Conclusions

This paper chooses eleven influential factors including the lag phases of historical CO_2 emissions. After reducing the dimensionality of the influential factors through PCA process, RELM is introduced to forecast the CO_2 emissions. Several conclusions can be obtained as

follows: (a) the PCA process is conducive to improving the operation speed and forecasting accuracy; (b) the high prediction precision of RELM model is attributed to the introduction of regularization part which enhances the global optimization and generalization ability with little time cost. (c) RELM combined with PCA outperforms other models with the lowest MAPE, MaxAPE, MdAPE and RMSE, indicating that PC-RELM model is a promising technique for CO_2 emission prediction.

Based on the findings in this paper, some suggestions for CO_2 emission reduction have been proposed with the consideration of selected influential factors: (a) According to the correlation analysis, coal consumption is completely correlated with CO_2 emission, thus it's necessary to substitute fossil energy with renewable and clean energy so as to achieve the diversification of energy consumption. (b) The economic growth should rely on innovative talents and technological advancements to improve resource allocation efficiency. The proportion of primary, secondary and tertiary industry ought to be reasonably adjusted thereby reducing energy consumption of GDP per capita. (c) People should enhance their low carbon awareness and cut down energy consumption during their life and labor. (d) To control vehicle exhaust emissions, traffic restrictions based on even-numbered and odd-numbered license plates can be implemented to reduce pollutant emissions and traffic pressure.

Acknowledgments

References

1. Jiang X, Zhu K, Wang S. The potential for reducing China's carbon dioxide emissions: Role of foreign-invested enterprises. *Global Environ. Chang*. 2015;35:22-30.

2. Gurney KR. Global change: China at the carbon crossroads. *Nature* 2009;458:977-979.

3. Holditch SA, Chianelli RR. Factors that will influence oil and gas supply and demand in the 21st century. *MRS Bull*. 2008;33:317-323.

4. Longwell HJ. The future of the oil and gas industry: Past approaches, new challenges. *World Energ*. 1998;5:100-104.

5. Say NP, Yücel M. Energy consumption and CO_2 emissions in Turkey: Empirical analysis and future projection based on an economic growth. *Energ. Policy* 2006;34:3870-3876.

6. Aydin G, Jang H, Topal E. Energy consumption modeling using artificial neural networks: The case of the world's highest consumers. *Energ. Source. Part B*. 2016;11:212-219.

7. Azadeh A, Tarverdian S. Integration of genetic algorithm, computer simulation and design of experiments for forecasting electrical energy consumption. *Energ. Policy* 2007;35:5229-5241.

8. Utgikar PV, Scott PJ. Energy forecasting: Predictions, reality and analysis of cause of error. *Energ. Policy* 2006;34:3087-3092.

9. Aydin G. The development and validation of regression models to predict energy-related CO_2 emissions in Turkey. *Energ. Source. Part B*. 2016;10:176-182.

10. Aydin G. The modeling of coal-related CO_2 emissions and projections into future planning. *Energ. Source. Part A*. 2014;36:191-201.

11. Feng YY, Zhang LX. Scenario analysis of urban energy saving and carbon abatement policies: A case study of Beijing city, China. *Procedia Environ. Sci.* 2012;13:632-644.

12. Lin BQ, Moubarak M. Decomposition analysis: Change of carbon dioxide emissions in the Chinese textile industry. *Renew. Sust. Energ. Rev.*2013;26:389-396.

13. González PF, Landajo M, Presno MJ. The driving forces behind changes in CO_2 emission levels in EU-27.Differences between member states. *Environ. Sci. Policy* 2013;38:11-16.

14. Zhang Y, Wang HK, Liang S, et al. Temporal and spatial variations in consumption-based carbon dioxide emissions in China. *Renew. Sust. Energ. Rev.* 2014;40:60-68.

15. Wang YF, Zhao HY, Li LY, et al. Carbon dioxide emission drivers for a typical metropolis using input–output structural decomposition analysis. *Energ. Policy* 2013;58:312-318.

16. Andres AR, Rustemoglu H. Determinants of CO_2 emissions in Brazil and Russia between 1992 and 2011: A decomposition analysis. *Environ. Sci. Policy* 2016;58:95-106.

17. Li Q, Wei YN, Dong YF. Coupling analysis of China's urbanization and carbon emissions: example from Hubei Province. *Nat. Hazards* 2016;81:1333-1348.

18. Li W, Ou Q, Chen Y. Decomposition of China's CO_2 emissions from agriculture utilizing an improved Kaya identity. *Environ. Sci. Pollut. Res.* 2014;21:13000-13006.

19. Wang QW, Chiu YH, Chiu CR. Driving factors behind carbon dioxide emissions in China: A modified production-theoretical decomposition analysis. *Energ. Econs.* 2015;51:252-260.

20. Ahmed K. The sheer scale of China's urban renewal and CO_2 emissions: multiple structural breaks, long-run relationship, and short-run dynamics. *Environ. Sci. Pollut. Res.* 2016;23:16115-16126.

21. Ali HS, Law SH, Zannah TI. Dynamic impact of urbanization, economic growth, energy consumption, and trade openness on CO_2 emissions in Nigeria. *Environ. Sci. Pollut. Res.* 2016;23:12435-12443.

22. Deng MX, Li W, Hu Y. Decomposing industrial energy-related CO_2 emissions in Yunnan Province, China: Switching to low-carbon economic growth. *Energies* 2016; 9.

23. Tang CF, Tan BW. The impact of energy consumption, income and foreign direct investment on carbon dioxide emissions in Vietnam. *Energy* 2015;79:447-454.

24. Lin BQ, Moubarak M, Ouyang XL. Carbon dioxide emissions and growth of the manufacturing sector: Evidence for China. *Energy* 2014;76:830-837.

25. Wang CJ, Wang F, Zhang HG, et al. Carbon emissions decomposition and environmental mitigation policy recommendations for sustainable development in Shandong Province. *Sustainability* 2014;6:8164-8179.

26. Bian YW, He P, Xu H. Estimation of potential energy saving and carbon dioxide emission reduction in China based on an extended non-radial DEA approach. *Energ, Policy* 2013;63:962-971.

27. Kang YQ, Zhao T, Wu P. Impacts of energy-related CO_2 emissions in China: A spatial panel data technique. *Nat. Hazards* 2016;81:405-421.

28. Sheng PF, Guo XH. The long-run and short-run impacts of urbanization on carbon dioxide emissions. *Econ. Model* 2016;53:208-215.

29. Wu LF, Liu SF, Liu DL et al. Modelling and forecasting CO_2 emissions in the BRICS (Brazil, Russia, India, China, and South Africa) countries using a novel multi-variable grey model. *Energy* 2015;79:489-495.

30. Pérez-Suárez R, López-Menéndez AJ. Growing green? Forecasting CO_2 emissions with environmental Kuznets curves and logistic growth models. *Environ. Sci. Policy* 2015;54:428-437.

31. Xu B, Lin BQ. Assessing CO_2 emissions in China's iron and steel industry: A dynamic vector autoregression model. *Appl. Energ.* 2016;161:375-386.

32. Baareh AK. Solving the carbon dioxide emission estimation problem: An artificial neural network model. *J. Software Eng. Applications* 2013;6:338-342.

33. Behrang MA, Assareh E, Assari MR et al. Using bees algorithm and artificial neural network to forecast world carbon dioxide emission. *Energ. Source* 2011;33:1747-1759.

34. Huang GB, Zhu QY, Siew CK. Extreme learning machine: Theory and applications. *Neurocomputing* 2006;70:489-501.

35. Liu H, Tian HQ, Li YF. Four wind speed multi-step forecasting models using extreme learning machines and signal decomposing algorithms. *Energ. Convers. Manage.* 2015; 100:16-22.

36. Ömer FE. Forecasting electricity load by a novel recurrent extreme learning machines approach. *Int. J .Elec. Power.* 2016; 78: 429-435.

37. Yu L, Dai W, Tang L. A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting. *Eng. Appl. of Artif. Intel.* 2015;47:110-121

38. Pearson K. On lines and planes of closest fit to systems of points in space. *Philos Mag.* 1901;2:559-572.

39. Hotelling H. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 1933;24:417-441.

40. Li S, Goel L, Wang P. An ensemble approach for short-term load forecasting by extreme learning machine. *Appl. Energ.* 2016;170:22-29.

41. Deo RC, Şahin M. Application of the extreme learning machine algorithm for the prediction of monthly Effective Drought Index in eastern Australia. *Atmos. Res.* 2015;153:512-525.

42. Lombardi AM. Some reasoning on the RELM-CSEP likelihood-based tests. *Earth Planets Space* 2014;66:286-301.

43. Zhang K, Luo MX. Outlier-robust extreme learning machine for regression problems. *Neurocomputing* 2015;151:1519-1527.

44. Energy Research Institute National Development And Reform Commission. The comprehensive report of China sustainable development of energy and carbon emission

scenarios analysis [cited May, 2003]. Available form: http://www.docin.com/p-419481261.html. Accessed May 2003.

45. Asongu S, Montasser GE, Toumi H. Testing the relationships between energy consumption, CO_2 emissions, and economic growth in 24 African countries: a panel ARDL approach. *Environ. Sci. Pollut. Res.* 2015;23:6563-6573.

46 Auffhammer M, Carson RT. Forecasting the path of China's CO_2 emissions using province-level information. *J. Environ. Econ. Manage.* 2008;55:229-247.

47. Farhani S, Ozturk I. Causal relationship between CO_2 emissions, real GDP, energy consumption, financial development, trade openness, and urbanization in Tunisia. *Environ. Sci. Pollut. Res.* 2015;22:15663-15676.

48. Song JK. China's carbon emissions prediction model based on support vector regression. *J. China Univ. Pet.* 2012;36:182-187.