Application Research of PSO-LSSVM in Carbon Emission Prediction in Hebei Province

Wei Sun¹, Xuyang Liu^{2,*}

¹Department of Business Administration, North China Electric Power University, Baoding 071000, China

²Department of Business Administration, North China Electric Power University, Baoding 071000, China

Abstract. This paper firstly introduces the background and significance of carbon emission prediction in Hebei Province, then collects the data and uses the PSO-LSSVM prediction model to train and predict the data collected through MATLAB software, finally obtains the prediction results. The results show that the prediction of PSO-LSSVM are accurate, and it can be used to predict carbon emissions in Hebei Province, Providing a new solution to the carbon emission forecasting of Hebei Province.

1. Introduction

With the development of the economy, China's annual energy consumption is also increasing. At the same time, due to economic development, the annual carbon emissions are also increasing, and China's per capita carbon emissions have leapt to the forefront of the world. The developing economy has brought severe environmental pressure, and the global warming caused by greenhouse gases has seriously affected human survival. Reducing energy consumption and carbon dioxide emissions while maintaining economic development is a topic studied by many scholars at home and abroad. Accurately predicting the carbon emissions of a region has certain practical significance to formulate emission reduction measures to ensure the economic developing stably.

2. PSO-LSSVM

2.1 Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO), an intelligent algorithm which was originally inspired by the foraging behavior of group birds, was firstly proposed by J. Kennedy and RC Eberhart. PSO uses the sharing of information by individuals in the group to solve the problem of group motion and obtain the optimal solution, making evolution process from disorder to order in vector space [1]. In this paper, particle swarm optimization is used to find the optimal parameters in *LSSVM*. The basic steps of the particle swarm algorithm are as follows [2]:

(1) Initialize particle and velocity. The fitness value which indicates the quality of the particle is calculated according to the velocity and position of the particle by the fitness function. Each particle represents a potentially optimal solution to the extreme value optimization

problem. We usually use position, velocity, and fitness values to represent the characteristics of a particle.

(2) Find individual extremum and group extremum. The individual extremum refers to the optimal position of the fitness value calculated in the position experienced by the individual. And the group extremum means the optimal position of the fitness searched by all the particles of the population. The individual position is then updated by tracking individual extremum and population extremum.

(3) Update speed and location

$$V_{id}^{k+1} = \omega V_{id}^k + c_1 r_1 (P_{id}^k - X_{id}^k) + c_2 r_2 (P_{gd}^k - X_{id}^k) \quad (1)$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \quad (2)$$

(4) Calculate particle fitness value. The particle fitness value is calculated according to the fitness function which is generally the optimization function of the problem under study. The optimization function of this paper is:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (3)$$

The values of sig2 and gam are the optimal parameter values when the RMSE reaches a minimum.

(5) Update the individual extremum and the group extremum and we should stop the cycle if it is judged that the termination condition is satisfied. Otherwise, we need return to the third step to continue the loop calculation until the termination condition is satisfied or the highest number of iterations is reached.

2.2 LSSVM

The least squares support vector machine which puts the least squares linear system as the loss function optimizes and extends the model of support vector machine [3]. It solves problems by replacing the inequality constraint with the equality constraint, using the principle of structural risk minimization. In this way, we can obtain

* Xuyang Liu:15733220487@163.com

the result just by solving only a linear equation system, which reduced computational complexity and improved solution efficiency [4]. The modeling process is based on equation (4):

$$y = \omega^T \varphi(x) + b \quad (4)$$

In equation (4), $x \in R^n$ is the input vector of a given set of n -dimensional training samples; $y \in R$ is the corresponding output vector; R is the sample space; $\omega \in R^{nh}$ is the weight vector, ω^T is The transposed matrix of ω , $b \in R$ is the offset amount, and $\varphi(x)$ is the nonlinear mapping function.

LSSVM regression can be expressed as a constrained optimization problem. And the optimization objective function is as follows:

$$\begin{cases} \min_{\omega, e, b} J(\omega, e) = \frac{1}{2} \|\omega\|^2 + \frac{C}{2} \sum_{i=1}^N e_i^2 \\ s. t. y_i = \omega^T \varphi(x_i) + b + e, i = 1, \dots, N \end{cases} \quad (5)$$

The operation process of *PSO-LSSVM* is as following steps [5]: firstly the parameters such as position, velocity and population size of the particles are initialized; calculate the fitness value of each particle with the fitness value function; update the individual position by tracking individual extremum and population extremum; the fitness value is calculated after updating particle position which is concluded by comparing the new fitness value of the particle with the fitness value of individual extremum and the group extremum; we stop the loop and output two optimal parameters when the number of cycles reaches the threshold or when the error precision reaches the set value; then we put the two optimal parameters into the *LSSVM* to deal with the data that need to be predicted ,finally obtain the prediction result.

3. Data sources and influencing factors analysis

The carbon emissions data of Hebei Province are all collected from Hebei Economic Yearbook and China Energy Statistics Yearbook. The carbon emission data of Hebei is calculated with the method of carbon emission coefficient, and the calculation formula is as follows:

$$CO_2 = k * E_T \quad (6)$$

Among them, CO_2 is the annual carbon emission of Hebei Province, and k is the carbon emission coefficient of standard coal, which indicates the amount of carbon dioxide released per ton of standard coal. E_T is the total annual energy consumption of Hebei Province. Total annual energy consumption is calculated from several kinds of energies using standard coal coefficient method. Analysis of influencing factors

Hebei Province is a traditional industrial province who makes coal as the main source of energy supply for industrial development. Due to its large population and developed industries, this paper selects the following seven indicators as the influencing factors of its high carbon emissions: the total GDP of Hebei Province, total population, industrial GDP, urbanization rate, per capita GDP, energy efficiency, and fiscal revenue. The gray correlation analysis shows that the correlation between

total GDP and carbon emissions is the largest, and there is a negative correlation between energy efficiency and carbon emissions, which means the higher the energy efficiency is, the less carbon emissions produces.

4. Application of PSO-LSSVM

Firstly, the carbon emission data and the impact factor data are divided into two parts: training set and prediction set. The author got the function between the impact factor and carbon emission by putting the training set into the model. Then, put the prediction set data into the trained model to obtain the prediction result of prediction set. Finally, the author compared the prediction result with the original carbon emission data in the prediction set and calculated the error index to obtain the result and conclusion. The following are the prediction results:

Table 1. Comparison table of forecast result and original value

year	2011	2012	2013
Raw data	68972	70660	72876
Predict value	70819	70785	70917
year	2014	2015	2016
Raw data	72031	72216	73196
Predict value	70987	70879	70863

The error indicators are shown in the following table:

Table 2. Values of prediction result error indicator

index	<i>MAPE</i>	<i>RMSE</i> (10,000 tons)	<i>SD</i> (10,000 tons)
Value	1.65%	1376	452

As is shown in the table 2, *MAPE* expresses the mean absolute percentage error, which represents the mean percent of the error between the predicted result and the original data. *RMSE*, root-mean-square error, is the extraction of square root of specific value between the summation and the number of samples. The summation is the square error of forecast error which is the error of prediction of outcome and original data. *SD* signifies the standard deviation, indicating the evolution of the sum of error square.

The prediction results are obtained by running the *PSO-LSSVM* algorithm program in MATLAB. It can be seen in the table that *MAPE*=1.66%, *RMSE*=1376, *SD*=452, and the prediction error is relatively small.

5. Conclusion

This paper collects data concerning energy consumption and factors affecting carbon emissions from 1995 to 2016 of Hebei Province, and calculate the annual carbon emissions based on the energy consumption method, then use the *PSO-LSSVM* prediction model to predict the results of test set. As we can see that *MAPE*=1.65%, indicating that the prediction accuracy is high and the *SD*

is low relatively. According to the results, the algorithm can be used for carbon emission prediction in Hebei Province. At the same time, the *PSO-LSSVM* algorithm provides a solution for carbon emission prediction in Hebei Province.

References

1. Hossain, MA ; Pota, HR ; Squartini, S; Abdou, AF. Modified PSO algorithm for real-time energy management in grid-connected microgrids[J] *RENEWABLE ENERGY*, **136**: 746-757,2019.
2. Han, QT ; Gao, XY. Research on Cooperative Task Allocation and Route Plan Based on PSO[M] *BASIC & CLINICAL PHARMACOLOGY & TOXICOLOGY*, **124**: 242-242,2019.
3. Guo,T; He, W; Jiang, ZL; Chu, XM; Malekian, R ; Li, ZX . An Improved LSSVM Model for Intelligent Prediction of the Daily Water Level[J] *ENERGIES*, **12** (1),2019.
4. Liu, JX ; Jia, ZH. Telecommunication Traffic Prediction Based on Improved LSSVM[J] *INTERNATIONAL JOURNAL OF PATTERN RECOGNITION AND ARTIFICIAL INTELLIGENCE*,**32** (3),2018.
5. Roushangar, K; Saghebian, SM ; Mouaze, D. Predicting characteristics of dune bedforms using PSO-LSSVM[J] *INTERNATIONAL JOURNAL OF SEDIMENT RESEARCH*,**32**(4):515-526,2017.