SVR-based prediction of carbon emissions from energy consumption in Henan Province

Guohua GOU
Department of Modern Management, Zhengzhou Technical College, Zhengzhou, China
E-mail:gouguohua0930@126.com

Abstract. This paper analyzes the advantage of support vector regression (SVR) in the prediction of carbon emission and establishes the SVR-based carbon emission prediction model. The model is established using the data of Henan's carbon emissions and influence factors from the 1991 to 2016 to train and test and then predict the carbon emissions from 2017 to 2021. The results show that: from the perspective of carbon emission from energy consumption, it raised 224.876 million tons of carbon dioxide from 1991 to 2016, and the predicted increment from 2017 to 2021 is 30.5563 million tons with an average annual growth rate at 3%. From the perspective of growth rate among the six factors related to carbon emissions it is proved that population urbanization rate per capita GDP and energy consumption per unit of GDP influences the growth rate of carbon emissions less than the proportion of secondary industry and coal consumption ratio of carbon. Finally some suggestions are proposed for the carbon emission reduction of Henan Province.

1. Introduction
As one of the traditional industrial bases in Henan Province, its total energy consumption in 2016 had exceeded 30 million t standard coal. Therefore, the carbon reduction work in Henan Province has an important impact on the national emission reduction results. Building a scientific model of carbon emissions in Henan Province and forecasting carbon emissions in the next few years will help to better grasp the trend of carbon emissions, so as to provide effective basis for the development of carbon reduction measures.

At present, the carbon emission forecasting model can be divided into two categories: one is the mixed construction mode, and the other is the direct construction mode. In this paper, the direct construction model is used to forecast the carbon emissions in Henan Province. In the direct construction model, the least commonly used least squares regression method and the ridge regression method have made some contributions to the carbon emission system modelling theory, the actual carbon emission forecast and planning, but they still exist stability and The explanation is weak, the parameter is difficult to determine and so on. Based on this, this paper introduces a new method of machine learning, which is a support vector machine(SVU), which specializes in the problem of nonparametric estimation under limited samples. It not only solves the small sample in the prediction model, but also shows some advantages in the high-dimensional pattern recognition. Therefore, this paper constructs a support vector regression model to predict the carbon emissions of Henan Province, so as to provide data support for carbon emission reduction.

2. Formatting the title, authors and affiliations
China's carbon emissions mainly from energy consumption, so all the factors related to energy consumption will also affect the carbon emissions. According to the influencing factors of China's energy consumption carbon emissions determined by the team, combined with the research results of other scholars in China and the actual situation of energy consumption in Henan Province, under the principle of comprehensive, comparable and data availability, , Urbanization rate, per capita GDP, the proportion of secondary industry, unit GDP energy consumption and coal consumption ratio as the main influencing factors of carbon emissions. Among them, the total population and urbanization rate, per capita GDP and the proportion of secondary industry, unit GDP energy consumption and coal consumption ratio, respectively, reflects the population factors, wealth factors, technical factors on the impact of carbon emissions.

3. SVR model of carbon emission forecast in Henan province

This paper uses the SVM algorithm proposed by C Jin and S W Jin in 2014. The algorithm draws the merits of other methods, and gives a method to determine the working set. Compared with the ordinary optimization problem, the numerical solution method takes up less memory and has more superiority in precision and speed.

Step 1, classify the samples. In this paper, the carbon emissions and their influencing factors in Henan Province from 1991 to 2011 and 2011 to 2015 are respectively used as training samples and test samples.

Step 2, All sample data are normalized. The independent and dependent variables in all samples are normalized by the following formula, so that all data are in [0, 1].

\[
x_{il} = \frac{x_{il}^0 - \min x_{il}^0}{\max x_{il}^0 - \min x_{il}^0} \quad i = 1,2,\cdots,n, \quad l = 1,2,\cdots,6
\]

\[
y_i = \frac{y_i^0 - \min y_i^0}{\max y_i^0 - \min y_i^0} \quad i = 1,2,\cdots,n
\]

Step 3, select the best parameters \( C, \gamma \). First, the parameters \( C, \gamma \) a rough selection, the range of values: \([-10,10]\), and then into the model to get a rough view of Figure; According to the rough map to adjust the parameters, the parameter range: \([-8,8]\), and then into the model was fine map; In which the grid width \( m \), the cross validation fold \( k \) and the return bandwidth \( \varepsilon \) remain unchanged, respectively, 1, 5 and 0.05. Finally, the optimal values \( C \) and \( \gamma \) of the parameters \( C^* \) and \( \gamma^* \) are obtained, respectively, 9.18959, 0.0358968.

Step 4, build and test the model. The SVR model is constructed using the best parameter values \( C^* \), \( \gamma^* \) and training samples obtained above. The regression function:

\[
f(x) = 2.5460K(x_2,x) + 1.3375K(x_3,x) - 9.1896K(x_4,x) + 4.274K(x_5,x) + 9.1896K(x_6,x) - 9.1896K(x_7,x) - 7.3493K(x_{10},x) + 5.0624K(x_{12},x) - 9.1896K(x_{13},x) - 9.1896K(x_{14},x) - 9.1896K(x_{15},x) + 9.1896K(x_{16},x) + 9.1896K(x_{17},x) - 0.4126K(x_{19},x) + 1.7260K(x_{21},x) + 9.1896K(x_{22},x) - 1.7442K(x_{24},x) + 0.9176
\]

Where \( x \) is the input sample whose model parameter is not zero, that is, the support vector; \( f(x) \) is the set of output vectors. The regression function \( f(x) \) is followed by the training sample and the test sample data into \( f(x) \) to persuade the correlation coefficient, respectively, \( R_1 = 0.997601 \), \( R_2 = 0.976871 \). At the same time, the training samples and test samples together into the model, the original data and fitting data comparison chart, as shown in Fig.1. It can be seen that the model has good learning and generalization ability, so the regression function \( f(x) \) can be used as energy consumption forecast model of Henan Province.
Step 5, forecast data. The results of carbon sequestration and related influencing factors in Henan Province from 1991 to 2016 were put into the carbon emission forecasting model to obtain the output results. Then the results were subjected to the anti-normalization treatment, that is, from 1991 to 2016, Henan Province Emissions, see Table 1. At the same time combined with the carbon emissions in 1991-2016, the carbon emissions trends in Henan Province from 2017 to 2021 are shown in Fig 2.

**Fig.1** Comparison of raw and regression data

4. data sources

4.1 sample data

The total population, the proportion of secondary industry, coal consumption ratio of 1991-2016 Henan Province three factors from the "Henan Economic Yearbook 2016"; 1991-2016 urbanization rate of Henan Province, according to the Henan Provincial Bureau of Statistics published in the calendar year, "Henan Province, the development of urbanization statistical monitoring news release"; In order to eliminate the influence of price factors, the GDP data of Henan Province from 1990 to 2015 were recalculated according to the comparable prices in 2000, and the per capita GDP and unit GDP energy consumption were measured accordingly. In this paper, the "IPCC Guide", combined with other scholars of the relevant research, the energy consumption of Henan Province, the calculation of carbon emissions, using the following formula:
\[ C = \sum_{i=1}^{3} E_i F_i \]

Where \( C \) is the energy consumption caused by energy consumption; \( E_i \) is the energy consumption of the \( i \)-th class of energy; \( F_i \) is the energy-emitting element of the \( i \)-th class; \( i = 1, 2, 3 \) respectively represent coal, oil and natural gas. 1990 - 2015 Henan Province carbon emissions, see Table 1.

| Years | Carbon emissions (10000t) | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 |
|-------|--------------------------|------|------|------|------|------|------|------|
| 1998  |                          | 4544.22 | 4712.38 | 5001.78 | 5716.86 | 5956.97 | 6484.54 | 6519.71 |
| 1999  |                          | 4590.58 | 6668.4 | 6839.76 | 8186.11 | 8878.83 | 9809.32 | 11234.08 |
| 2000  |                          | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
| 2001  |                          | 12687.26 | 14530.81 | 15957.66 | 17300.09 | 18635.64 | 19974.39 | 21296.55 |
| 2002  |                          | 21699.24 | 22301.81 | 22500.08 | 23189.14 | 18635.64 | 21296.55 |

4.2. influencing factors predictive value

According to the "13th Five-Year Plan" of Henan Province, the average annual growth rate of each influencing factor was calculated as follows: 7.1‰, 3.3%, 5.8%, -1.4%, -5%, -6‰, and calculated from 2017 to 2021 in Henan Province, the impact of carbon emission factors indicators, see Table 2.

| Years | Population (million people) | Urbanization rate (%) | Per capita GDP (Yuan/person) | Secondary industry Proportion(%) | Unit GDP energy consumption (Ton of standard coal) | Coal consumption ratio (%) | Carbon emission (10000t) |
|-------|-----------------------------|-----------------------|-------------------------------|---------------------------------|---------------------------------------------|---------------------------|------------------------|
| 2017  | 9478                        | 52.58                 | 30956.8                       | 50.08                           | 1.34                                        | 86.46                      | 23886.21               |
| 2018  | 9532                        | 54.52                 | 32736.4                       | 49.33                           | 1.28                                        | 85.95                      | 24612.32               |
| 2019  | 9585                        | 56.54                 | 34681.7                       | 48.59                           | 1.20                                        | 85.40                      | 25377.74               |
| 2020  | 9594                        | 58.63                 | 36704.6                       | 47.86                           | 1.16                                        | 84.88                      | 26228.45               |
| 2021  | 10103                       | 60.01                 | 38852.5                       | 47.14                           | 1.10                                        | 84.30                      | 26941.84               |

5. Conclusions

From the growth ratio. Analysis of Table 1,2 and Figure3 shows that in 1991-2021, energy consumption in Henan Province, the overall trend of carbon emissions, an increase of 22487.62 million tons, the forecast interval 2017-2021 growth of 3055.63 million tons, The average annual growth rate remained at around 3%, so the next few years in Henan Province, the road is still very serious carbon emission reduction.

From the growth rate. The growth rate from 2001 to 2011 was 1178.88 million tons, respectively, higher than 373.9 million tons in 1991-2001 and 696.75 million tons in 2011-2021. However, through comprehensive analysis of Table1, Table2 and Figure3 shows that the six carbon influencing factors in the population, urbanization rate and per capita GDP from 1991 to 2021 is increasing year by year, unit GDP energy consumption is gradually reduced The energy utilization rate increased year by year, the change was 14.44 million, 45.64%, 36398.75 (yuan/person), -2.95 tons of standard coal. It seems that these four factors and carbon emissions growth rate of change in the pace is inconsistent, indicating that these four factors contribute to the growth rate of carbon emissions is small. While the proportion of secondary industry and coal consumption ratio was 46.75% and 90.81% in 1991-2001, respectively, and the average proportion of 2011-2021 was 50.23% and 87.9% The average share of 2001-2011 was 52.85% and 91.93% respectively. It can be seen that these two factors are consistent
with the changes in the growth rate of carbon emissions, indicating that the proportion of secondary industry and coal consumption than carbon emissions growth rate of contribution.

6. Recommendations
In view of the above analysis, this paper puts forward the following suggestions for the carbon reduction work in Henan Province in the next few years: First, to gradually compress the secondary industry and vigorously develop the tertiary industry as the guide, increase the adjustment of industrial structure, and promote the optimization and upgrading of industrial structure; Second, to ensure the premise of energy supply, and gradually reduce the dependence on coal, increase the natural gas, wind and solar energy as the representative of the proportion of clean energy, optimize the energy consumption structure; Third, the government should increase investment in science and technology and talent, and then improve the environmental protection products and low-carbon technology research and development capabilities and research and development efforts; Fourth, learn from and learn from the experience of the development of carbon emissions trading system in Europe, and actively cooperate with China's carbon trading market construction and accelerate the development of carbon finance products; Fifth, to strengthen the ecological environment of the restoration, continue to implement forestation, grassland restoration, wetland protection and other measures to create a good living environment.

References
[1] Moody J E. The Effective Number of Parameters: An Analysis of Generalization and Regularization in Nonlinear Learning Systems. NIPS, 2012, 4:847-854.
[2] Vapnik V N, Golowich S L, Smola A. Support Vector Method for function Approximation, Regression Estimation Processing and Signal Proceeding. Proceeding of NIPS, 2006:281-287.
[3] J K Song. China Carbon Emission Forecasting Model Based on Support Vector Regression. Journal of China University of Petroleum, 2016,36 (2) 183-187.
[4] H Wu, S Z Gu, Zhou Hong, et al. Relationship between energy consumption, carbon emissions and economic growth in Henan Province. Resources Science, 2015,3 (10): 1898 -1905
[5] J K Song. China’s carbon emissions prediction model based on support vector regression. Journal of China University of Petroleum,2014,6(10):23-29
[6] C Jin, S W Jin. Software reliability prediction model based on support vector regression with improved estimation of distribution algorithms. Applied Soft Computing,2014,6:56-61.