Fuel Consumption Prediction Based on Whale Algorithm Optimized Support Vector Machine

Andi Liu

School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo, Shandong, 255000, China

Abstract: In order to improve the prediction accuracy of vehicle fuel consumption, this paper proposes a support vector machine model based on whale optimization algorithm for prediction. First, SPSS software is used to analyze the relationship between the influencing factors of fuel consumption. On this basis, the correlation between various influencing factors and fuel consumption is analyzed. Then, the whale optimization algorithm is used to optimize the penalty parameter C and the kernel parameter g in the support vector machine kernel function, and it is brought into the model for simulation. Taking the determination coefficient and mean square error as the evaluation indexes, the prediction results show that the model has high accuracy.

Keywords: correlation analysis, WOA, SVM, fuel consumption

1. Introduction

The fuel economy of automobile refers to the ability of automobiles to drive economically with as little fuel consumption as possible under the condition of ensuring power performance. Automobiles have a good fuel economy, which can reduce the use cost of automobiles and reduce the dependence of countries on oil imports. Therefore, the study of automobile fuel economy can effectively analyze the use of oil in the automotive industry in China, and it can provide a basis for buyers. Analyzing the relationship between fuel consumption and various influential factors is helpful to analyze how to reduce fuel consumption.

Since the 1973 world oil crisis, world oil prices have escalated. What's more, the gradual depletion of oil as a non-renewable energy will inevitably affect the future production and life of mankind. In addition, since the 1980s, countries around the world have begun to pay attention to global warming caused by the greenhouse effect, among which carbon dioxide is the main factor causing the greenhouse effect. Therefore, countries have formulated laws and regulations to control fuel consumption. In China 's laws and regulations, cars in the simulation of urban and suburban operating cycle, by measuring carbon dioxide, carbon monoxide and hydrocarbons emissions, using carbon balance method to calculate fuel consumption. However, due to the short time of carbon balance automobile fuel consumption detector in China and it is used in a bad environment, so the accuracy and stability of actual working conditions are still insufficient, and there is measurement uncertainty[1]. In this regard, Zhou et al. proposed a new method of vehicle fuel consumption detection based on the carbon balance method and the Vmas system in the paper. The content of carbon dioxide, carbon monoxide and hydrocarbons in exhaust gas was measured by the Vmas system, and then the fuel consumption was calculated by the carbon balance method. According to this idea, a fuel consumption test system was built for in-use cars[2]. Zhang et al. used principal component analysis to simplify the variables of automobile fuel consumption in the paper, electing some main variables into the BP neural network to predict fuel consumption provides a new idea [3].

Therefore, this paper will use correlation analysis and support vector machine based on whale optimization algorithm to predict fuel consumption of fuel vehicles. Firstly, SPSS software is used to analyze the correlation of influencing factors of fuel consumption. Then the whale optimization algorithm is used to optimize the penalty parameter C and the kernel parameter g in the kernel function of the support vector machine. Finally, the model is used to simulate and analyze the data.

2. Basic principles

2.1. Whale Optimization Algorithm (WOA)

The whale optimization algorithm[4] is based on the behavior of whales in capturing prey. Whales usually create some unique bubbles on circular or '9' -shaped paths when hunting. This predatory strategy is called upward-spirals and double-loops. The whale optimization algorithm simulates the predatory behavior of this kind of bubble network.

Assuming that the current best candidate solution is the target prey, other search agents will find the location of the best search agent according to the following equation:

$$\vec{D} = \left| \vec{C} \overrightarrow{X^*}(t) - \vec{X}(t) \right| \tag{1}$$

$$\vec{X}(t+1) = \overrightarrow{X^*}(t) - \vec{A} \cdot \vec{D} \tag{2}$$

In the form, t represents the number of current iterations. $\vec{A}$ and $\vec{C}$ are all coefficient vectors. $\overrightarrow{X^*}(t)$ is the vector of the optimal solution to the current whale position. $\vec{X}(t)$ is s the current whale position vector. $\vec{A}$ and $\vec{C}$ are calculated as follows:

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \tag{3}$$

$$\vec{C} = 2 \cdot \vec{r} \tag{4}$$

In the form, $\vec{a}$ is the parameter that decays linearly during iteration. $\vec{r}$ is a random vector. $\vec{a}$ is calculated as follows:

$$\vec{a} = 2 - \frac{2t}{T_{\max}} \tag{5}$$

In the form, t represents the number of current iterations. $T_{\max}$ is the largest number of iterations. The humpback whale captures prey by spiraling. The mathematical model is established as follows:

$$\vec{X}(t+1) = \left| \overrightarrow{X^*}(t) - \vec{X}(t) \right| \cdot e^{bl} \cdot \cos(2\pi l) + \overrightarrow{X^*}(t) \tag{6}$$

In the form, e^{bl} is the distance coefficient. b is a constant, defining the spiral shape. l is a random vector.

The swimming path of humpback whales when capturing prey is a shrinking circle, swimming in a spiral manner. It is assumed that the probability of the humpback whale to choose contraction is p_i, the probability for spirally updating whale position is $1 - p_i$.

$$\vec{X}(t+1) = \begin{cases} \overrightarrow{X^*}(t) - \vec{A} \cdot \vec{D}, & \text{if } p < p_i \\ \left| \overrightarrow{X^*}(t) - \vec{X}(t) \right| \cdot e^{bl} \cdot \cos(2\pi l) + \overrightarrow{X^*}(t), & \text{if } p \geq p_i \end{cases} \tag{7}$$

In the form, $p \in [0,1]$.

Moreover, due to linear attenuation of $\vec{a}$, the values of $\vec{A}$ fluctuate accordingly. Whales also search for prey at random. When $\left| \vec{A} \right| \geq 1$, select a search agent to keep other whales away from prey. This process is as follows:

$$\vec{D} = \left| \vec{C}\overrightarrow{X_{rand}} - \vec{X}(t) \right| \tag{8}$$

$$\vec{X}(t+1) = \overrightarrow{X_{rand}} - \vec{A} \cdot \vec{D} \tag{9}$$

In the form, $\overrightarrow{X_{rand}}$ is a randomly selected whale position vector.

2.2. Support vector machine (SVM)

Support Vector Machine (SVM) is a kind of linear classifier, which can classify linear and nonlinear data according to statistical learning theory. Sample data can be mapped to high-dimensional feature space by kernel function and then processed by linear regression. It is suitable for solving nonlinear, small sample and high dimension problems.

2.3. SVM Parameter Optimization Based on WOA

When the support vector machine is used to establish the model, for the nonlinear regression prediction in the sample data, the kernel function is usually used to map the data to the high-dimensional space for processing. In the kernel function, penalty parameter C and kernel parameter g have important influences on the prediction performance of the model. Therefore, the whale optimization algorithm (WOA) is used to optimize the parameters C and g.

Firstly, the data is imported and normalized. Secondly, set the whale parameters and define the number and upper and lower bounds of the optimization parameters. Then set the number of whales and the maximum number of iterations and the probability of whales choosing spiral bounding. Then, calculate the fitness value and update the position. Finally, the optimal penalty parameter C and kernel parameter g are obtained to establish the WOA-SVM model.

3. Experimental calculation

3.1. Correlation analysis

The experimental data are from the car house. To facilitate data analysis, set data labels for the following indicators. With the wide application of automobile electronic control technology. At present, more and more electronic control systems appear in automobiles, which have an impact on fuel consumption. The factors affecting the fuel consumption rate are analyzed from four aspects, namely, the total size and the quality of the vehicle, the engine, the transmission system and the vehicle shape. 17 factors including vehicle length (mm), vehicle width (mm), vehicle height (mm), axle distance (mm), complete vehicle curb mass (kg) and displacement (L) were selected for analysis. As shown in Table 1.

Table 1: Pearson

	NEDC comprehensive fuel consumption (L)
vehicle length (mm)	0.417**
vehicle width (mm)	0.711**
vehicle height (mm)	0.601**
axle distance (mm)	0.465**
complete vehicle curb mass (kg)	0.751**
maximum horsepower (Ps)	0.691**
maximum power (kW)	0.690**
maximum power speed (rpm)	-0.368**
maximum torque (N·m)	0.771**
maximum torque speed (rpm)	-0.286*
maximum net power (kW)	0.657**
gear (or simulated gear)	-0.072
shift form	0.393**

* $p<0.05$ ** $p<0.01$

From the table above, correlation analysis was used to study the correlation between NEDC

comprehensive fuel consumption (L) and a total of 13 items, including vehicle length (mm), vehicle width (mm), vehicle height (mm, axle distance (mm), preparation quality (kg), maximum horsepower (Ps), maximum power (kW), maximum power speed (rpm), maximum torque (N m), shift form, maximum torque speed (rpm), maximum net power (kW), and gear (or simulated gear). The Pearson correlation coefficient is used to indicate the strength of the correlation.

The correlation coefficients between NEDC comprehensive fuel consumption (L) and vehicle length (mm), vehicle width (mm), vehicle height (mm) and vehicle preparation quality (kg) are 0.417, 0.711, 0.601 and 0.751. And they all showed the significance at 0.01 level. Therefore, there is a significant positive correlation between NEDC comprehensive fuel consumption (L) and vehicle length (mm), vehicle width (mm), vehicle height (mm) and overall quality (kg). This is because large cars rolling resistance, air resistance, slope resistance and acceleration resistance increased significantly. In addition, in order to ensure the power performance, the load rate of the vehicle decreases while the fuel consumption rate increases. Therefore, reducing vehicle size and mass is conducive to improving fuel economy.

In addition, we can see that the correlation coefficient between the comprehensive fuel consumption (L) of NEDC and the gear (or the simulated gear) is-0.072. At present, CVT is widely used in automobiles, and the working diameter of the master-slave driving wheel transmitting power is variable. The conveyor belt works with it, and the speed output is linear, and the transmission ratio is continuously variable. In addition, the transmission system is matched with the engine working condition at any time, so there is no specific stop. In this experiment, the simulated gear is used for analysis, so the results are not obvious. But qualitative point of view, still has a certain negative correlation. That is to say, the increase of transmission gear increases the opportunity for the engine to work in economic conditions, which is conducive to improving fuel economy and reducing fuel consumption rate. This corresponds to reality.

There are three fuel supply modes in the samples analyzed, which are multi-point electric injection, direct injection and mixed injection. Multi-point EFI is to install the injector on the intake manifold of each cylinder. The gasoline is injected into the intake manifold to form a mixture outside the cylinder. The advantage of this injection method is that the required pressure is not high, and the inlet pipe can prolong the fuel evaporation time. The mixing time of the mixture is long. However, when the nozzle is injected, some fuel is injected into the intake pipe, so it is impossible to accurately control the amount of gasoline entering the cylinder. In-cylinder direct injection technology installs the injector on the engine cylinder head and directly injects fuel into the cylinder. This injection mode has high pressure and good atomization performance, which can realize accurate control of the amount of gasoline injected into the cylinder. Since the high-pressure fuel injected enters in the form of fog, it is conducive to mixing with air to form a combustible mixture[5]. The third is the mixed injection fuel supply. This fuel supply method is a combination of multi-point electric injection and direct injection in cylinder, and an injector is installed in the cylinder and the intake manifold. It is the most advanced injection method, not only high efficiency, but also high accuracy. This way is beneficial to reduce fuel consumption. However, the structure is complex and the cost is high.

In addition, other auxiliary equipment in the car will have an impact on fuel consumption. Therefore, it is necessary to obtain the actual value based on the external characteristic curve provided by the manufacturer.

3.2. Experimental environment

First, normalize the data and set the number of optimized parameters. Then, set the upper and lower bounds of the optimization parameters. The ranges of the two parameters are $C \in [0.01,1]$, $g \in [2^\wedge -5, 2^\wedge 5]$. Set the number of whales to 20, $T_{\max} = 30$. Finally, the model is built on MATLAB.

3.3. Evaluation indicators

In this experiment, R^2 and MSE are used as evaluation indexes, and the calculation formula is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\overline{y}_i - y_i)^2} \qquad (10)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{m} w_i (y_i - \hat{y}_i)^2 \qquad (11)$$

3.4. Experimental results

The WOA-SVM model is established on MATLAB software. The simulation results are as follows:
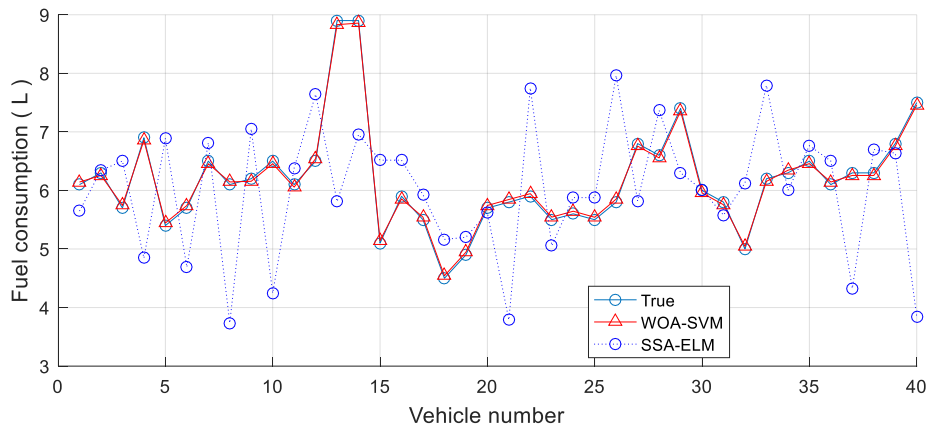

Figure 1: Forecast results

3.5. Analysis of experimental results

At present, with the development of electronic control technology in automobiles and the improvement of people ' s living standards, various auxiliary equipment is continuously added to the automobile. For example, the configuration of multimedia systems, voice recognition systems, seat electric adjustment and other functions will also have a certain impact on fuel consumption in the use process. Forecasting results of this model are $R^2 = 99.8\%$, $MSE = 0.0017$. The prediction accuracy is high. In addition, the prediction accuracy of ELM model is $R^2 = 98.4\%$, which is lower than the first one.

4. Conclusions

This paper uses SPSS software for correlation analysis to get the model as Table 1. The influencing factors of each fuel consumption in the table are analyzed, and then the support vector machine based on whale optimization algorithm is used to optimize. The results show that the prediction results of the model are basically consistent with the actual value, and the prediction accuracy is high, which can be effectively used to predict the use of petroleum energy in the automotive industry in China. At the same time, it provides some basis for future enterprises to improve technology in the direction of energy saving and emission reduction. When consumers buy products, they can get some reference through this model.

References

[1] Tianlong Zhou, Jian Li, Gang Yan, Xin Zhang, Lei Luo, Junjun Guo. Evaluation of Uncertainty in Measurement of Indication Error for Automobile Fuel Consumption Detector by Carbon Balance Method. China Inspection Body & Laboratory, 2021,29(05):56-57.DOI:10.16428/j.cnki.cn10-1469/tb.2021.05.017.
[2] Ju Zhou, Yongjun Min. Design and Development of Fuel Consumption Test System for Light Vehicle.

Highways & Automotive Applications, 2010(03):15-18.

[3] Deng Zhang, Chunsheng Zu, Chaochao Zhao. Fuel Consumption Prediction Based on Principal Component Analysis and Neural Network. Agricultural Equipment & Vehicle Engineering, 2015,53(06):47-52.

[4] Peng Xu, Yeying Zhang. Support Vector Machine Transformer Fault Diagnosis Method Based on Whale Optimization Algorithm. Lamps & Lighting, 2021(09):85-88.

[5] Shufang Gao. Application of In-cylinder Direct Injection Technology in Small and Medium Gasoline Engines. Internal Combustion Engine & Parts, 2022(04):94-96. DOI: 10.19475/j.cnki.issn1674-957x.2022.04.030.