Coal consumption prediction based on least squares support vector machine

Li Zhang1,4, Liansheng Zhou1, Yingtian Zhang2, Kun Wang1, Yu Zhang1, Zhijun E3, Zhiyong Gan1, Ziyue Wang1, Bin Qu1 and Guohao Li1

1 Tianjin Electric Power Science & Research Institute, Tianjin, 300384;
2 Tianjin power technology development co. LTD;
3 State Grid Tianjin Electric Power Company.
4 Email: zhanglihust@163.com

Abstract. At present, China's economic construction continues to move forward, and the demand for energy is increasing day by day, and the problem of energy shortage is showing. All industries are actively conducting research on energy saving and emission reduction. However, thermal power plants have large demand for coal and heavy pollution from flue gas. Therefore, technological upgrading, energy saving and operation control of coal-fired power plants need to be optimized. Support vector machine (SVM) algorithm is applied in the aspect of function fitting based on the strict statistics basis, and the coal consumption prediction model based on the least squares support vector machine is the product of the informatization application of power plant and the competition requirement of power market. Firstly, the research status of the optimization algorithm of thermal power unit is described. Secondly, the principle of support vector machine and least squares support vector machine is introduced, so as to do the basic work for establishing the prediction model of coal consumption later. Finally, the key problems to be further studied and solved in the field of coal consumption prediction in thermal power plants are discussed.

1. Introduction
At present, thermal power is still dominant in China, power generation coal accounts for more than 50% of China's coal industry in 2014. With the implementation of energy conservation and emissions reduction, the coal-fired power plant must improve efficiency, reduce cost and optimize the allocation of resources to adapt to the development of the times. Therefore, under the premise of ensuring the safe operation of the unit, the power generation enterprise urgently needs to understand the thermal economy of the unit in time, and analyzes and adjusts the factors affecting the normal operation of the unit, which can keep the unit in the optimal operating condition at all times and minimize the coal consumption of power supply. In order to improve the economic benefit of power plant, it is necessary to deeply realize the operation of the unit coal consumption trends. Thus, it can form a set of scientific and quantitative coal consumption prediction model for thermal power unit and reduce the actual coal consumption.

In this paper, the practicability research of least squares support vector machine (LS-SVM) theory in coal consumption prediction is proposed. It can use the measured coal consumption data of a certain thermal power plant for a period of time and find out the function relation of each indicator to coal consumption in different working conditions. Thus, the coal consumption is predicted in advance...
based on the LS-SVM theory. Then, it can keep the unit in the optimal operating condition, reduce the coal consumption of the unit and improve the efficiency of the unit.

2. Least squares support vector machine theory

2.1. The basic theory of the LS-SVM
Support Vector Machine (SVM) is a new method based on statistical theory. It is a theory of small sample training and classification machine learning proposed by V. Vapnik et al. of AT&T Bell LABS. It contains minimum experience risk and confidence range, and the core problem of realizing the minimum risk is to realize structural risk minimization to achieve the best generalization ability.

After a lot of research, SVM has a variety of improved algorithms. For example, Zhang Fen proposed an improved support vector machine estimation algorithm, SORR, which can measure cholesterol levels in three plasma protein samples [1]. He Fan proposed the BA+SVM algorithm, which uses the bat algorithm (BA) to synchronize the parameter optimization of the SVM and the selection of feature attributes of the input data, and improves the classification ability of the SVM [2]. Zhu Yong Sheng combined feature, vector selection method and linear SVM to form a new type of SVM named PSVM, which is faster than SVM in running [3]. Nevertheless, the LS-SVM used in this paper has the characteristics of generalization and global optimization. Then, the time of extracting data is shorter, and the results of computing are more accurate [4-7]. It is widely used in signal processing, system identification and modeling, advanced control and soft measurement [8]. Furthermore, it has a strong advantage in solving the problem of regression. The algorithm applied to the regression problem is given as follows:

Assuming that the training sample set \( T \) is consist of \( i \) sample points:
\[
T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_i, y_i)\}
\]  
(1)

Where \( x_i \in \mathbb{R}^n \) is the input vector; \( y_i \in \mathbb{R} \) is the output vector corresponding to the input vector. The optimization problem of the LS-SVM is:
\[
\min_{w,b,\xi} R = \frac{1}{2} \|w\|^2 + \frac{1}{2} c \sum_{i=1}^{n} \xi_i^2 \\
s.t. \quad y_i = w^T \varphi(x_i) + b + \xi_i, \quad i = 1, 2, \ldots, n
\]  
(2)

The Lagrange function corresponding to the optimization problem is:
\[
L(w, b, \xi, \alpha) = R - \sum_{i=1}^{n} \alpha_i \left( w^T \varphi(x_i) + b + \xi_i - y_i \right)
\]  
(3)

\[\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_n] \]
(4)

Where \( \alpha \) is the Lagrange multiplier; \( w \) and \( b \) are the model parameters; \( c \) is the regularization parameter, and \( c > 0 \); \( \xi \) is the training set predicts error vector.

According to KKT conditions (which is Karush-Kuhn-Tucker Optimization Conditions, this condition is needed in the SVM to transform inequality constraints into a set of equality constraints):
\[
\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^{n} \alpha_i \varphi(x_i)
\]  
(5)

\[
\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{n} \alpha_i = 0
\]  
(6)

\[
\frac{\partial L}{\partial \xi_i} = 0 \rightarrow c \xi_i = \alpha_i
\]  
(7)

\[
\frac{\partial L}{\partial \alpha_i} = 0 \rightarrow y_i \left( w \cdot \varphi(x_i) + b \right) - 1 + \xi_i = 0
\]  
(8)
Another form as follows:

\[
\begin{pmatrix}
0 \\
1 \\
\end{pmatrix} + \frac{c}{l}
\begin{pmatrix}
\beta \\
\alpha \\
\end{pmatrix} =
\begin{pmatrix}
0 \\
Y \\
\end{pmatrix}
\]

(10)

\[\Omega_{ij} = \varphi(x_i)^T \varphi(x_j) = K(x_i, x_j)\]  

(11)

\[Y = \begin{pmatrix} y_1, \ldots, y_n \end{pmatrix}^T\]  

(12)

\[1 = (1, \ldots, 1)^T\]  

(13)

\[\alpha = (\alpha_1, \ldots, \alpha_i)^T\]  

(14)

Where, \(K(x_i, x_j)\) is a kernel function that satisfies the Mercer condition. Thus, the decision function expression of LS-SVM is as follows:

\[y(x) = \sum_{i=1}^{n} \alpha_i K(x, x_i) + b\]  

(15)

2.2. Common kernel function of the LS-SVM

The polynomial kernel function is shown as follows:

\[k(x, x_i) = \left((x^T x_i) + 1\right)^d, d = 1, 2\]  

(16)

The radial basis function (RBF) kernel function is shown as follows:

\[k(x, x_i) = \exp\left(-\frac{x - x_i}{2\sigma^2}\right)\]  

(17)

The sigmoid kernel function is shown as follows:

\[k(x, x_i) = \tanh\left[v(x^T x_i) + \delta\right]\]  

(18)

The liner kernel function is shown as follows:

\[k(x, y) = x \cdot y\]  

(19)

3. The coal consumption prediction model based on the LS-SVM

3.1. Model input and output variable selection

Table 1. Major thermal economic indicators of thermal power plants

| Index             | Name                                      | Name                                      | Name                                      | Name                                      |
|-------------------|-------------------------------------------|-------------------------------------------|-------------------------------------------|-------------------------------------------|
| Turbine index     | Vacuum degree                             | Efficiency                                | Steam rate                                | Feed water consumption rate              |
|                   | the superheated steam temperature         | Consumption rate of fan                   | Blowdown rate                             | Exhaust gas temperature                  |
| Boiler index      | Loss rate                                 | Rate of railway car inspection            | Power consumption rate of coal handling   | Qualified rate of coal blending           |
|                   |                                           |                                            | Make-up water percentage                 | Auxiliary water ratio                    |
| Fuel index        | Alkali consumption                        | Acid consumption                          | Generating capacity                       | Auxiliary power ratio                    |
|                   | coal consumption of electricity supply     |                                            |                                            |                                            |
| Chemical index    |                                           |                                            |                                            |                                            |
| Overall index     |                                           |                                            |                                            |                                            |

3
The thermal economic indicators of thermal power plants are related to each other. These economic indicators reflect the operating standards of all thermal power units. By identifying the main aspects of the problem, the complexity of the system is simplified [9]. Table 1 shows the major thermal economic indicators of thermal power plants.

In addition to the main thermal economic indicators shown in Table 1, the factors affecting coal consumption are varied. In this paper, the main steam temperature, main steam pressure, reheating steam temperature, vacuum degree and flue gas exhaust temperature are selected as the influencing factors of coal consumption prediction.

3.2. Pre-processing of sample data
The corresponding pre-treatment of the original sample data is helpful to accelerate the training speed and convergence speed of the model, and improve the prediction accuracy. There are many methods to deal with sample data. This paper adopts the normalization method. The normalized processing of the sample data is mapped to the [1, 2] interval. Suppose the maximum value is Max Value and minimum value is Min Value in the current sample data. And x and y represent the sample values before and after normalization. Then, the linear transformation is as follows:

\[
y = \frac{x - \text{min value}}{\text{max value} - \text{min value}} + 1
\]

(20)

3.3. Selection of LS-SVM model parameters
In this paper, the RBF (Radial Basis Function) is selected as the kernel function of LS-SVM model. The RBF function has relatively few parameters and numerical restrictions, which can reduce the complexity of the model and improve the training speed. Therefore, only two parameters that influence the performance of the model are selected, which named the kernel parameter g and the penalty factor C, and then, the determination of these two parameters is the best model selection problem. In this paper, the cross validation method is adopted to select the combination of the minimum cross validation error as the optimal value of the parameter by adjusting the values of two parameters. After calculation, the kernel parameter of LS-SVM model was determined as 0.4, and the penalty factor was 60.

3.4. Results and discussion
In this paper, a total of 100 groups of historical operation data of Heze power plant were selected. Due to the large amount of data and the same data type, several sets of data sources were selected as listed in Table 2. In the previous 85 groups, the coal consumption historical value and other influencing factors were used as training samples of the model, and the historical value of coal consumption in the next 15 groups was used as the test sample of the model. Then, the LS-SVM coal consumption prediction model is established with the data sampling interval of 1 h. It is given identical as previously that the kernel parameter was determined as 0.4, and the penalty factor was 60. Figure 1 and Figure 3 shows the comparison between actual coal consumption and forecast coal consumption, and Figure 2 shows the relative error of the coal consumption.

Programs for developing LS-SVM method are written in C++ and compiled using Microsoft’s Visual C++ 6.0.

| Group | 1    | 2    | 3    | … | 85   | 86   | … | 100 |
|-------|------|------|------|---|------|------|---|-----|
| Coal Consumption(g/kwh) | 318.07 | 317.09 | 315.60 | … | 307.14 | 317.06 | … | 310.98 |
Figure 1. The comparison between actual coal consumption and forecast coal consumption.

Figure 2. The relative error of the coal consumption.

Figure 3. The comparison between actual coal consumption and forecast coal consumption.

It can be seen that all the forecast coal consumption were very close to the actual value. And all the relative error of the coal consumption were less than 0.1%, which indicated the feasibility of the least squares support vector machine theory in coal consumption prediction.

4. Conclusions
(1) Based on the development and application of the least square support vector machine theory, a coal consumption prediction model based on LS-SVM was established.

(2) In this paper, the coal consumption prediction model of the LS-SVM is proposed, and the relative error of the prediction is all less than 0.1 % which shows that the prediction results are ideal.

(3) The LS-SVM has overcome the disadvantages of long training time, poor generalization ability and easy to get into local minima. Furthermore, it also has strong small sample learning and generalization ability.
References

[1] Zhang F 2006 Improved Support Vector Machine Regression Estimation Algorithm and Its Application [D]. Anhui University

[2] He F, Lu C J 2017 A Synchronous Optimization Algorithm for Improving SVM Classification Ability [J]. Advances in Applied Mathematics 06(09)

[3] Zhu Y, Zhang Y 2004 A new type SVM—projected SVM [J]. Science in China 47(S1) 21-28

[4] Yan W W, Zhu H D, Shao H H 2003 Based on least squares support vector machine soft sensor modeling [J]. Journal of System Simulation 15(10) 1494-1496

[5] Suykens J A K, Vandewalle J 2000 Recurrent least squares support vector machines[J]. IEEE transactions on circuits and systems-I: Fundamental theory and applications 47(7) 1109-1114

[6] Xu Z M, Wen X Q, Sun Y Y, et al 2009 Prediction of Slagging Characteristics of Coal-fired Boilers Based on Least Squares Support Vector Regression [J]. Chinese Journal of Electrical Engineering 29(17) 8-13

[7] Chen M S, Liu D P 2006 Multi-objective Optimization of Utility Boiler Combustion Based on LS-SVM and SPEA2 [J]. East China Electric Power 34(3) 50-54

[8] Wang X P, Cao L M 2002. Genetic Algorithm Theory, Application and Software Implementation [M]. Xi'an: Xi'an Jiaotong University Press 18-50

[9] Keerthi S S, Lin C J 2003 Asymptotic behaviors of support vector machines with Gaussian kernel[J]. Neural Computation 15(7) 1667-1689