Study on Prediction Model for Boiler Thermal Efficiency Based on Support Vector Regression and Particle Swarm Optimization

Bian He-Ying[1,*], Dai Ke-Jie[1] and Fang Yan-Jun[2]

[1]*College of Electrical and Information Engineering, Pingdingshan University , Henan, Pingdingshan, 467000, China*

[2]*Department of Automation, Wuhan University, Hubei, Wuhan, 430072, China*

Abstract: Aiming at the problem that the boiler thermal efficiency can't be on-line calculated in the field, data collected are divided into training data and testing data by researching on the 1000MW unit of Sanbaimen Power Plant in Datang Chaozhou, and then prediction model for boiler thermal efficiency is established by applying support vector regression and particle swarm optimization. This method uses optimization function of particle swarm algorithm to optimize the parameters C and g of the prediction model, and applys test data and the randon data to test the model's accuracy and generalization capability. The simulation results show that prediction model for boiler thermal efficiency has higher prediction accuracy that relative error is controlled within 1% and has better generalization ability, which provides feasible scheme for online measurement of boiler thermal efficiency.

1. INTRODUCTION

Boiler thermal efficiency is a key indicator to measure the economic operation of the boiler and be able to embody the effect of the boiler combustion optimization control. The improvement of thermal efficiency of the boiler combustion not only saves energy, but also reduces pollution and greenhouse gas emissions [1]. Therefore, accurately measuring boiler thermal efficiency and adjusting the boiler combustion optimization control strategy according to it, is helpful in the improvement of the energy efficiency and reduction in the cost of power generation. The accurate boiler thermal efficiency model is a key factor to measure. Typical model of boiler thermal efficiency has mechanism modeling and data modeling [2]. Because the boiler combustion process is a very complex physical and chemical process, it has multi-variable, nonlinear, strong coupling and large delay characteristics, it is very difficult to establish boiler thermal efficiency model by mechanism modeling. In recently years, data modeling method has attracted more attention. For example, Zhang Yunfeng and Wang Jingcheng proposed the least squares support vector machines model between extracted feature and boiler efficiency by analyzing the boiler combustion historical data, research results show the model has better prediction accuracy and generalization ability [3]; Lu Yukun and Peng Xing established the boiler thermal efficiency model by applying the back propagation neural network according to the field data in power factory, research results show the prediction accuracy of the model has also improved [4].

In this research, the boiler thermal efficiency model was established by applying the support vector machine and particle swarm optimization, and verified the model has higher prediction accuracy and stronger generalization ability by detecting the test data from the power factory.

2. CALCULATION AND PREDICTION MODEL STRUCTURE OF THE BOILER THERMAL EFFICIENCY

Simplified Model of AMSE Boiler Thermal Efficiency Calculation

Presently, there are two calculation standards for the boiler thermal efficiency, one is the people's republic of China national standard of power performance test procedures and the other is American society of mechanical engineers power performance test procedures, referred to as ASME. The standard model must be simplified for improving the calculation efficiency. In this research [5], the calculation values, by applying the boiler thermal efficiency simplified model based on ASME standard, are as the real values of boiler thermal efficiency. Model simplification principle is as follows.

1) Dry flue gas heat loss is shown in formula (1).

$$l_G = \frac{C_{py}}{Q_d^y} \cdot \left(k_1 + k_2 \cdot \alpha_{py} \right) \left(t_{py} - t_R \right) \times 100\% \tag{1}$$

Where, Q_d^y is low heat value boiler efficiency, α_{py} is excess air coefficient of exhaust, t_{py} is exhaust flue gas temperature, t_R is cool air temperature, C_{py} is mean specific heat of exhaust, k_1 and k_2 are the coefficient of Q_d^y.
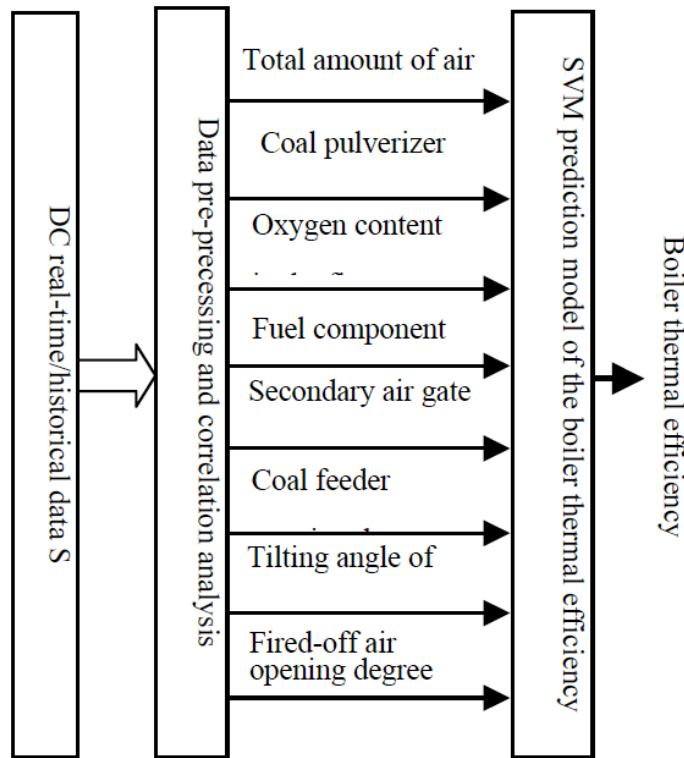
Fig. (1). Prediction model structure of the boiler thermal efficiency.

2) Water content heat loss is shown in formula (2).

$$l_m = \frac{C_{pH_2O}}{Q_d^y} \cdot \left(k_3 + 0.01\left(k_4 + k_2 \cdot \alpha_{py}\right)\right) \cdot \left(t_{py} - t_R\right) \times 100\% \qquad (2)$$

Where, C_{pH_2O} is mean specific heat of water vapor, k_3 and k_4 are the coefficient of Q_d^y.

3) The heat loss of incomplete combustion of carbon is shown in formula (3).

$$l_{uc} = \frac{33730}{Q_d^y} \cdot A^y \cdot \left(0.9 \cdot \frac{C_{hf}}{100 - C_{hf}} + 0.1 \frac{C_{hz}}{100 - C_{hz}}\right) \times 100\% \quad (3)$$

4) Radiation and convection heat loss is $l_{un} = 0.5\%$.

5) Other heat losses are $l_{un} = 0.5\%$.

6) The boiler thermal efficiency is shown in formula (4).

$$\eta = 100 - l \qquad (4)$$

The thermal efficiency of the boiler is calculated according to the simplified model of the thermal efficiency of the boiler, but the calculation speed is slow, is not real-time and very difficult to realize for on-line measurement of the boiler thermal efficiency. Therefore, the prediction model of the boiler thermal efficiency is established by applying SVR and PSO algorithm in this research. The key parameters that influence the boiler thermal efficiency are chosen as prediction model input parameters. The part boiler thermal efficiency values calculated according to the formula (4) are the model output parameter for building the prediction model.

Model Structure of the Boiler Thermal Efficiency

The boiler thermal efficiency is mainly influenced by many factors such as the total amount of the fuel, the total amount of air and so on, its model input parameters are chosen by mechanism analysis and correlation analysis, which are shown as follows: total amount of the fuel, total amount of air, five coal feeders opening degree, six secondary air dampers opening degree, a fired-off air damper opening degree, three oxygen content in the flue gas, six coal pulverizers ventilation rate and the tilting angle of the burners.

Effects of load on the boiler thermal efficiency is described by the total amount of the fuel and air; coal feeder opening degree describes effect of the pulverized coal; effect of oxygen content is described by several parameters such as secondary air damper opening degree, coal feeders ventilation rate and oxygen content in the flue gas; the influence of other factors are described by the fired-off air damper opening degree and tilting angle of the burners. The prediction model structure of the boiler thermal efficiency is established and shown in Fig. (**1**).

3. SUPPORT VECTOR REGRESSION MODEL

Support Vector Regression Theory

The basic idea of support vector regression is that the non-linear problems of low dimensional space are transformed into linear problems of a high dimensional space by introducing the kernel function $K(x_i, x_j)$.

If $\{(x_1,x_1),\ldots,(x_n,x_n)\}$ is the training sample set, then the regression function is formula (5) [6].

Boiler Thermal Efficiency

$$F = \left\{ f \,\middle|\, f(x) = w^T\phi(x) + b, w \in R^n \right\} \tag{5}$$

Where, *w* stands for weight vector, b is threshold. R_{emp} is structure risk function.

$$R_{reg}(w) = \frac{1}{2}\|w\|^2 + C \cdot R_{emp}\left[f(x)\right]$$
$$= \frac{1}{2}\|w\|^2 + C \cdot \frac{1}{n}\sum_{i=1}^{n}|y - f(x)| \tag{6}$$

Where, $\|w\|^2$ describes the model complexity, the role of the penalty coefficient C is a compromise between experience risk and the complexity of the model, n stands for the number of training samples. *w* can be counted according to formula (7).

$$w = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\phi(x_i) \tag{7}$$

α_i, α_i^* is solution of R_{reg} minimized, x_i is support vector, n stands for the number of training samples. *f(x)* can be represented as formula (8) according to the formula (7).

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\left[\phi(x_i)\cdot\phi(x)\right] + b \tag{8}$$

Non-linear support vector regression function *f(x)* can be defined according to the kernel function $K(x_i,x_j) = \phi(x_i)\cdot\phi(x_j)$, it is shown in formula (9) [7].

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x_i,x) + b = \sum_{i=1}^{n}w_i \cdot K(x_i,x) + b \tag{9}$$

Where, w_i is the coefficient of support vector, x_i is support vector, $K(x_i,x)$ is kernel function, b is defined as formula (10).

$$b = y_i - w \cdot \phi(x_i) \tag{10}$$

Introduction of slack variable ε can increase the robustness of vector regression. Concrete expression is shown in equation (11).

$$|y - f(x)|_\varepsilon = \begin{cases} 0 & |y - f(x)| \le \varepsilon \\ |y - f(x)| - \varepsilon & a \end{cases} \tag{11}$$

Parameter Selection

In this work, the radial basis kernel function $(\exp(-g \times |x_i - x_j|^2))$ is selected as Kernel function of support vector regression model. Training parameters of the support vector regression model have four parameters; they are kernel function parameter g, penalty coefficient *C*, maximum allowable error e, and insensitive loss function ε. Parameters ε and e are controlled by the human's behavior, they can often be ignored because of their little impact on model prediction ability [8]. Parameters *C* and g directly affect the modeling calculation process, the properties of the model, and great impact on the prediction precision and generalization ability [9]. This research selects the particle swarm optimization algorithm to optimize parameters *C* and g in order to obtain good parameters.

Method of PSO-SVR Model

In this research, the method that support vector machine is combined with optimization algorithm is known as PSO-SVR model method.

Final prediction effect is evaluated by using mean square error (MSE) and squared correlation coefficient r^2, expression formulas of MSE and r^2 are (12) and (13). The value of MSE is smaller and prediction precision of the model is the better; Square correlation coefficient is more close to 1, the result of regression fitting is better [10].

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(f(x_i) - y_i\right)^2 \tag{12}$$

$$r^2 = \frac{\left(n\sum_{i=1}^{n}f(x_i)y_i - \sum_{i=1}^{n}f(x_i)\sum_{i=1}^{n}y_i\right)^2}{\left(n\sum_{i=1}^{n}f(x_i)^2 - \left(\sum_{i=1}^{n}f(x_i)\right)^2\right)\left(n\sum_{i=1}^{n}y_i^2 - \left(\sum_{i=1}^{n}y_i\right)^2\right)} \tag{13}$$

Where, $f(x_i)$ stands for prediction value of the output, y_i stands for actual value of the output.

PSO-SVR training steps are as follows:

1) To choose the training data collected from field, reduce noise and normalize data.

2) Optimal parameters *C* and g are obtained by applying particle swarm optimization algorithm to identify training data. Detailed process is as follows:

① Initialize the particle population: Set the number of iterations to 200, particle number is 20, and randomly generate a set of initial positions and velocity of the particles;

②Evaluation fitness by using the fitness function MSE;

③Replace the previous position with the current best position 'pbest' if the current fitness value of particles is better than the previous best position 'pbest'; Replace the previous global best position with the current best global position 'gbest' if the current fitness value of particles is superior to the global best position experienced; Where, 'pbest' stands for the current best position of particle, 'gbest' stands for the current best position of group;

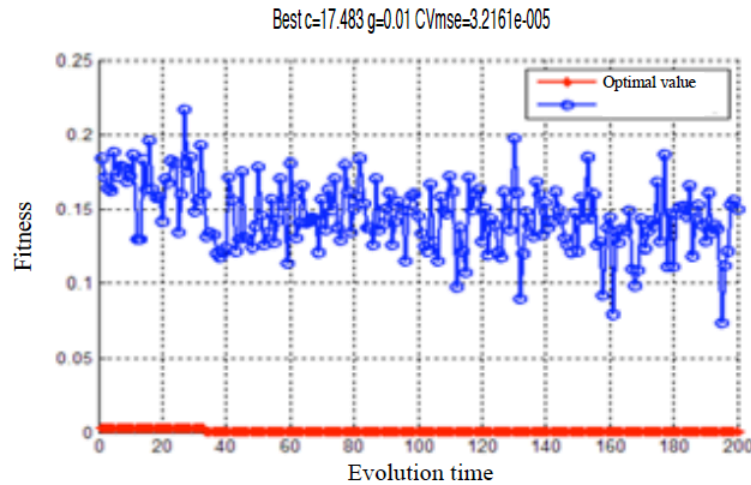④Update the particle's speed and position according to the expression (14) and (15).

Fig. (2). Fitness curve of SVR optimized by PSO.

$$V_i = w_v * V_i + c_1 * Rand(\ )*(pBest[i] - X_i)$$
$$+ c_2 * Rand(\ )*(pBest[g] - X_i) \tag{14}$$

$$X_i = X_i + V_i \tag{15}$$

Where, the constant c_1 and c_2 are called learning factors, *Rand* () and *Rand* () are random numbers from [0, 1], w_v is inertia weight that controls the degree by which the past speed effects the current speed, particle position, and particle speed.

⑤The iteration stop and output optimal solution when the number of iterations achieves target value, otherwise, jump to step ②.

3)To perform the SVR algorithm, train training data and get the SVR model.

4) To verify model's accuracy and generalization capability by using the test data.

4. PREDICTION MODEL OF THE BOILER THER-MAL EFFICIENCY BASED ON PSO-SVR

The Establishment of SVR Model

To establish prediction model according to PSO-SVR algorithm introduced in the previous section. Simulation data is from the historical data of a 1000MW ultra-supercritical unit of Datang Chaozhou power station.

Determine training sample data and testing sample data. First, preprocess the data from the distributed control system and choose 200 sets of data at random. 25 sets of data were selected randomly as the training sample input from the 200 group of data, the thermal efficiency value which are calculated by ASME match the 25 sets of data selected, which are as the training sample output.

Initialization parameters of the particle swarm algorithm are as follows: learning factor $c_1 = 1.5$, $c_2 = 1.7$, population quantity $pop = 20$, maximum generation $\max gen = 200$. The

fitness curve of the training sample set acquired by applying PSO-SVR is shown in Fig. (**2**).

Fig. (**2**) shows that the prediction model has 17 support vectors, the kernel function parameter g=0.01 and constant b=0.0592. Based on these parameters, the decision function of the SVR model is obtained by using matlab, which is shown in formula (16).

$$f(x) = \sum_{i=1}^{n} w_i \exp\left(-g|x_i - x|^2\right) + b \tag{16}$$

Where, w_i is support vector coefficient, x_i is support vector of the model, x is the data for forecasting, g is the kernel function parameter, b is constant and n is the number of support vectors.

## 5. VERIFY THE VALIDITY OF THE PREDICTION MODEL OF THE BOILER THERMAL EFFICIENCY

Final prediction effect of the prediction model is evaluated by using mean square error (MSE) and squared correlation coefficient r^2. The value of MSE is smaller, prediction precision of the model is better; Square correlation coefficient is closer to 1 and the result of regression fitting is better.

To verify the validity of the prediction model built in the last section, the model is tested by applying test data. The test results show that MSE is 0.019 and r^2 is 0.99107. The accuracy of the model and the regression fitting curve can meet the practical requirements. Simulation output curve of the boiler thermal efficiency by using matlab is shown in Fig. (**3**).

The dotted line represents the predicted value of the boiler thermal efficiency by applying the prediction model; the solid line represents actual values of the boiler thermal efficiency. Fig. (**3**) shows the same trend of the rise and fall of both the predicted value and the actual value, only amplitude of variation is slightly different in the local range, which verifies the build prediction model. The boiler thermal effi-
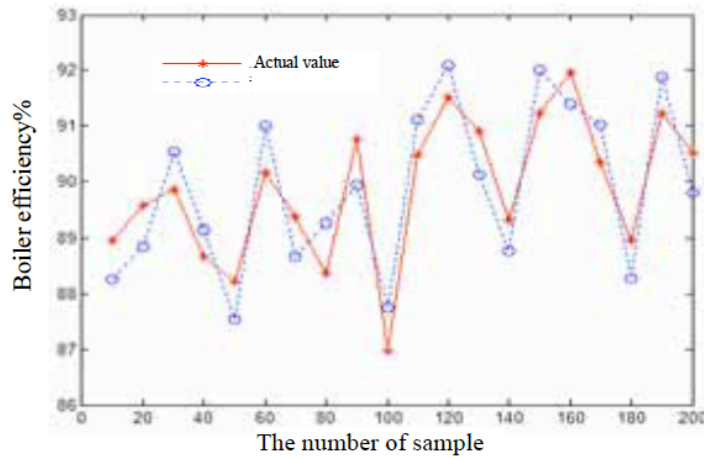
Fig. (3). Prediction model output curve of the boiler thermal efficiency.
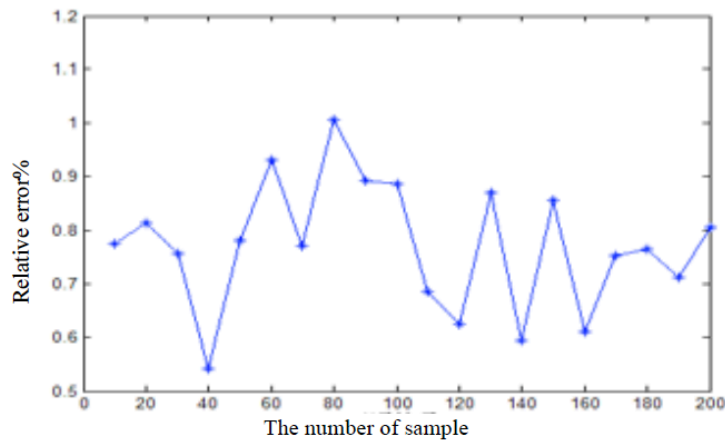
Fig. (4). Relative error curve of the boiler thermal efficiency.
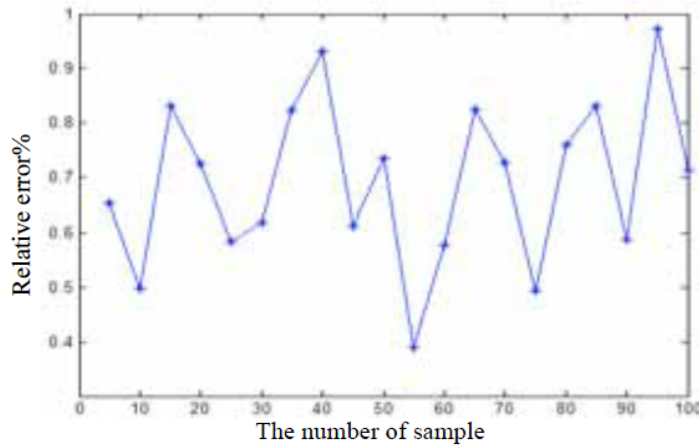
Fig. (5). Relative error curve of the boiler thermal efficiency.

ciency has high prediction accuracy and good generalization ability. Fig. (**4**) is the relative error curve corresponding to Fig. (**3**).

The relative error is controlled within 1%, variance is 0.0028, as are shown in Fig. (**4**).

20 groups of data sets are selected randomly in order to further verify the generalization of the model established. The prediction results of the model show that the relative

error can be limited in the range of accuracy and the prediction model has good generalization ability, as is shown in Fig. (**5**).

CONCLUSION

A prediction model of the boiler thermal efficiency is established by applying support vector regression and article swarm algorithm for the problem that the boiler thermal

efficiency is difficult to measure accurately. To optimize the parameters C and g of the prediction model by adapting article swarm algorithm. Prediction accuracy and generalization capability of the model are improved by preprocessing acquisition data and verifying them by using test data. The simulation results show that the prediction value relative error of the soft-sensing model established is controlled within 1% and variance is only 0.019. In summary, prediction accuracy and generalization capability of the prediction model established can meet the practical requirements and provide an effective way for the measurement of the boiler thermal efficiency in a power plant.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Y. P. Gu, and W. J. Zhao, "Combustion optimization for utility boiler based on least square-support vector machine," In: *Proceedings of the CSEE*, vol. 30, pp. 91-97, Sep. 2012.

[2]   J. L. Gao, and J. Ling, "Test study on combustion adjustment for 1000MW ultra supercritical boilers," *Journal of Chinese Society of Power Engineering*, vol. 27, pp. 741-747, May 2012.

[3]   Y. F. Zhang, and J. C. Wang, "Coal-fired boiler efficiency modeling based on partial least squares support vector machines," *Control and Instruments in Chemical Industry*, vol. 39, pp. 1432-1437, Nov. 2012.

[4]   Y. K. Lu, and X. Peng, "Hybrid modeling optimization of thermal efficiency and NOx emission of utility boiler," In: *Proceedings of the CSEE*, vol. 31, pp. 16-24, June 2011.

[5]   Z. H. Hu, and W. D. Hao, "Experimental study on combustion and Nox emission characteristic of 1000MW ultra-supercritical coal-fired boiler," *Journal of mechanical engineering*, vol. 46, pp. 105-112, Apr. 2010.

[6]   G. Zhang, and L. Bao, "Soft-sensing modeling of the carbon content in fly ash based on information fusion for thermal power plant," In: *Proceedings of the IEEE International Conference on Mechatronics and Automation,* Changchun, China, pp. 3860-3865, Aug. 2009.

[7]   M. A. Mohanders, S. R. Halawani, and A.A. Hussain, "Support vector machines for wind speed prediction," *Renewable Energy*, vol. 29, pp. 939-947, Feb. 2004.

[8]   B. Ren, and Z. P. Feng, "Improved genetic algorithm and particle swarm optimization as well as comparison between them," *Journal of Nanjing Normal University (Engineering and Technology),* vol. 2, pp. 14-20, Aug. 2002.

[9]   Y. Sheng, B. Guo, and T.X. Gu, "Particle swarm optimization algorithm and comparison with genetic algorithm," *Journal of University of Electronic Science and Technology of China*, vol. 34, pp. 696-699, May 2005.

[10] R. Li, H. N. Wang, H. He, Y. M. Cui, and Z. L. Du, "Support vector machine combined with k-nearest neighbors for solar flare forecasting," *Chinese Journal of Astronomy and Astrophysics*, vol. 7, pp. 1088-1095, Mar. 2007.