Acta Technica *61*, No. 4B/2016, 175–184

Study on soft sensor model of boiler thermal efficiency based on support vector regression and grid search algorithm[1]

Bian Heying[2], Zhang Xiaoli[3]

Abstract. A novel soft sensor model of boiler thermal efficiency is established by applying support vector regression(SVR) and grid search algorithm (GSA). This method divides data collected into training data and testing data by researching on the 1000 MW unit of Sanbaimen Power Plant in Datang Chaozhou, uses grid search algorithm to optimize the parameters C and g of the soft sensor model, and applies test data and random data to test the model accuracy and generalization capability. The simulation results show that the built soft sensor model has higher prediction accuracy where relative error is controlled within 1 % and better generalization ability, which provides a feasible scheme for online measurement of boiler thermal efficiency.

Key words. Support vector regression, grid search algorithm, boiler thermal efficiency, soft sensor, model.

1. Introduction

Boiler thermal efficiency is a key indicator for measuring the economic operation of the boiler. To improve its thermal efficiency, combustion should not only save energy, but also reduce pollution and greenhouse gas emissions [1]. Therefore, it is very important to measure accurately boiler thermal efficiency, which is helpful to adjust timely the boiler combustion optimization control strategy, improve the energy efficiency and reduce the cost of power generation. The boiler thermal efficiency model is a key factor to measure accurately the boiler thermal efficiency. There are

[1]The authors would like to thank for the support by Natural Science Foundation of the education department of Henan Province under the Grant 14B470009.

[2]College of Electrical and Mechanical Engineering, Pingdingshan University, Henan, Pingdingshan, 467000, China

[3]Food and Chemical Engineering Department, College of Henan Quality Project, Pingdingshan, 467000, China

two kinds of methods to establish the model at present: mechanism modeling and data modeling [2]. As the boiler combustion process has multivariable, nonlinear, strong coupling and large delay characteristics, it is very difficult to establish boiler thermal efficiency model by mechanism modeling. Data modeling method has attracted more attention In recent years [3], [4] the soft sensor model of boiler thermal efficiency was built according to the test data of boiler combustion by applying neural network and genetic algorithm; research results show that the model has better measurement accuracy and generalization ability. Literature [5] proposed the least squares support vector machines model between extracted feature and boiler efficiency by analyzing the boiler combustion historical data. Research results show that the prediction accuracy of the model was also obviously improved.

2. Support vector regression and modeling of grid search algorithm

2.1. Theory of modeling of support vector regression

The basic idea of modeling of support vector regression is that non-linear problems of a low-dimensional space are transformed into linear problems of a high-dimensional space by introducing the kernel function $K(x_i, x_j)$ [6].

If $\{(x_1, x_1), \ldots, (x_n, x_n)\}$ is the training sample set, then the regression function F is given by the formula

$$F = \left\{ f | f(x) = w^T \phi(x) + b, \quad w \in R^n \right\}, \tag{1}$$

where w stands for weight vector, T denotes the sample points sets, $\phi(x)$ is nonlinear mapping from input to output space and b is the threshold. Then

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

where $\|w_i\|$ describes the model complexity, the role of the penalty coefficient C is a compromise between the experience risk and the complexity of the model, n stands for the number of training samples, R_{emp} is the experience risk function and R_{reg} is the structure risk function. The value of w can be determined from the formula

$$w = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) \phi(x_i), \tag{3}$$

where α_i, α_i^* represents the solution of R_{reg} minimized, x_i is the support vector, n stands for the number of training samples, and $f(x)$ can be represented by the formula

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) [\phi(x_i) \cdot \phi(x)] + b. \tag{4}$$

Non-linear support vector regression function $f(x)$ can be defined according to the kernel function $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ using the formula

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x_i, x) + b = \sum_{i=1}^{n} w_i \cdot K(x_i, x) + b, \tag{5}$$

where w_i is the coefficient of support vector, $K(x_i, x)$ is the kernel function, and b is defined by the formula

$$b = y_i - w_i \cdot \phi(x_i) . \tag{6}$$

2.2. Parameter selection of SVR model

The radial basis kernel function $K(x_i, x_j) = \exp\left(-g \cdot |x_i - x_j|^2\right)$ is selected as kernel function of support vector regression model in this research. The accuracy of SVR model generally depends mainly on four parameters, particularly on the kernel function parameter g, penalty coefficient C, maximum allowable error e, and insensitive loss function ε. Parameters e and ε are, however, usually ignored because they are controlled by the human's behavior and exhibit only a little impact on the model of soft sensor ability [7]. On the other hand, parameters C and g affect directly the model soft sensor precision and generalization ability [8]. This research applies the grid search algorithm to optimize parameters C and g in order to obtain parameters that are as good as possible.

2.3. Method of GSA-SVR modeling

In this research, the method supporting vector machine combined with grid search algorithm is known as GSA-SVR modeling. The GSA-SVR training steps are as follows:

1. To determine training data and test data according to the data collected from the field.

2. To deal with data such as noise reduction and data normalization.

3. To optimize parameters C and g that are obtained by applying grid search algorithm.

4. To establish soft sensor SVR model of the boiler thermal efficiency by using the optimum parameters and training data.

5. To verify the accuracy and generalization ability of the model built by applying the test data.

6. To use the model to predict the data measured if the accuracy and generalization ability of the model meet the requirements; otherwise to build the model again.

3. Soft sensor model structure of the boiler thermal efficiency

3.1. Simplified model of boiler thermal efficiency calculation based on ASME

There are two calculation standards of the boiler thermal efficiency at present, one of them being the national standard of power performance test procedures of people's republic of China [9]; the other being the standard of power performance test procedures of the American society of mechanical engineers, referred to as ASME. The standard model must be simplified in order to improve the calculation efficiency because some parameters are not able to be measured online under the boiler operation condition. In this research, the model simplification principle based on ASME is as follows.

1. Water content heat loss is given by the formula

$$l_{\mathrm{m}} = \frac{C_{\mathrm{pH_2O}}}{Q_{\mathrm{d}}^{\mathrm{y}}} \cdot \left(k_3 + 0.01\left(k_4 + k_2 \cdot \alpha_{\mathrm{py}}\right)\right) \cdot \left(t_{\mathrm{py}} - t_{\mathrm{R}}\right) \times 100\%\,, \tag{7}$$

   where $C_{\mathrm{pH_2O}}$ is mean specific heat of water vapor, $Q_{\mathrm{d}}^{\mathrm{y}}$ is the low-heat value of boiler efficiency, K_2, k_3, k_4 are the coefficients of $Q_{\mathrm{d}}^{\mathrm{y}}$, α_{py} is the excessive air coefficient of exhaust, t_{py} is the exhaust flue gas temperature and t_{R} is the temperature of cool air.

2. Dry flue gas heat loss is given as

$$l_{\mathrm{G}} = \frac{C_{\mathrm{py}}}{Q_{\mathrm{d}}^{\mathrm{y}}} \cdot \left(k_1 + k_2 \cdot \alpha_{\mathrm{py}}\right)\left(t_{\mathrm{py}} - t_{\mathrm{R}}\right) \times 100\%\,, \tag{8}$$

   where C_{py} is the mean specific heat of exhaust and k_1 is the coefficient.

3. The heat loss caused by incomplete combustion of carbon is given as

$$l_{\mathrm{uc}} = \frac{33730}{Q_{\mathrm{d}}^{\mathrm{y}}} \cdot A^{\mathrm{y}} \cdot \left(0.9 \cdot \frac{C_{\mathrm{hf}}}{100 - C_{\mathrm{hf}}} + 0.1\frac{C_{\mathrm{hz}}}{100 - C_{\mathrm{hz}}}\right) \times 100\%\,, \tag{9}$$

   where A^{y} is the application base ash, C_{hf} is is the fuel percentage in fly ash and C_{hz} is the fuel percentage in ash slag.

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

5. Other heat loss is $l_{\mathrm{o}} = 0.5\,\%$.

6. Then the boiler thermal efficiency is

$$\eta = 100 - l\,, \tag{10}$$

   where $l = l_{\mathrm{G}} + l_{\mathrm{uc}} + l_{\mathrm{un}} + l_{\mathrm{o}}$.

The thermal efficiency of the boiler can be calculated according to the simplified model of the thermal efficiency of the boiler, but the calculation speed is slow and it is very difficult to realize on-line measurement of the boiler thermal efficiency. Therefore, the soft sensor model of the boiler thermal efficiency is built by applying SVR and GSA algorithm in this research. The key parameters influenced by the boiler thermal efficiency are selected as soft sensor model input parameters, the partial boiler thermal efficiency values calculated according to the formulae (6)-(9) are taken as model output parameters during establishing soft sensor model.

3.2. Soft sensor model structure of the boiler thermal efficiency

The boiler thermal efficiency is influenced by many factors [10]. The model input parameters are selected by mechanism analysis and correlation analysis as follows: total amount of air, ventilation rates of six coal pulverizes, three values of the contents of oxygen in the flue gas, fuel component, six opening degrees of secondary air dampers, five opening degrees of the coal feeders, a tilting angle of the burners and the opening degree of fired-off air damper.

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

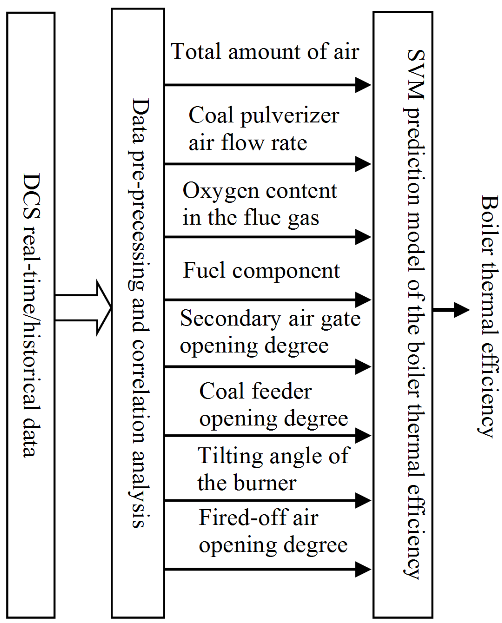

Fig. 1. Prediction model structure of boiler thermal efficiency

4. Soft sensor modeling of the boiler thermal efficiency based on GSA-SVR

The task was to establish a soft sensor model according to the method of GSA-SVR modeling introduced in the previous section. The simulation data come from the historical data of a 1000 MW ultra-supercritical unit of Datang Chaozhou power station.

Now how to determine the training sample data and testing sample data. First, the data coming from the distributed control system were preprocessed and selected randomly 200 groups data among them. 40 groups data were obtained by choosing a group data among every 5 groups, 40 thermal efficiency values were calculated according to the ASME model by applying the 40 groups, which served as the training sample output. Second, the data groups of 8 multiples were selected among the 200 groups data, 25 groups data obtained served as the test sample input, the thermal efficiency values match the 25 groups data served as the test sample output

To determine the variation range of the penalty parameter C and radial basis core parameter g according to the relevant theories, the variation range of the former parameter is [-8, +8], while the variation range of the latter parameter is [0, 300]. To optimize the parameters C and g by using the grid search algorithm, the fitness curve of the training sample set acquired by applying GSA-SVR algorithm is shown in Fig. 2.

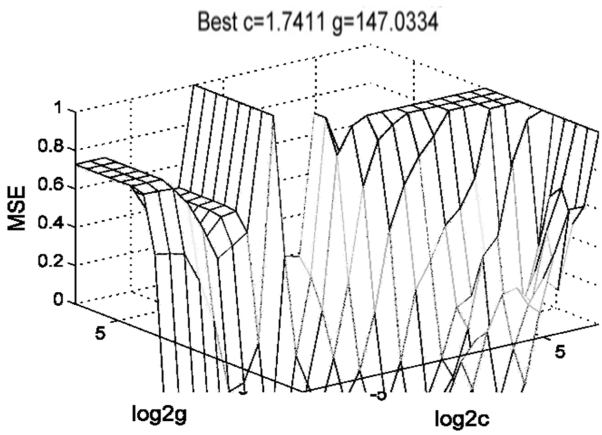

Fig. 2. Fitness curve of SVR optimized by GSA

Fig. 2 shows that the optimal penalty factor C equals to 1.7411, the kernel function parameter g equals to 147.0334 by applying the GSA-SVR model, and that the soft sensor model built has 17 support vectors. The decision function of the SVR model obtained by Matlab is

$$f(x) = \sum_{i=1}^{n} w_i \exp\left(-g\left|x_i - x\right|^2\right) + b\,. \tag{11}$$

5. Verification of validity of soft sensor model of boiler thermal efficiency

The final prediction effect of the soft sensor model is evaluated by using mean square error (MSE) and squared correlation coefficient (r^2). With the square correlation coefficient approaching more close to 1, the result of regression fitting is better, the value of MSE is smaller and prediction precision of the model is also better.

To verify the validity of the built soft sensor model, the model is tested by applying test data. The test results show that $r^2 = 0.99365$, MSE is 0.031, the regression curve fitting degree and the accuracy of the model built can meet the practical requirements. The simulation output curve of the actual value and soft sensor value of the boiler thermal efficiency by using Matlab is shown in Fig. 3.

The solid line represents the actual values of the boiler thermal efficiency; the dotted line represents the prediction values of the boiler thermal efficiency by applying the soft sensor model built.

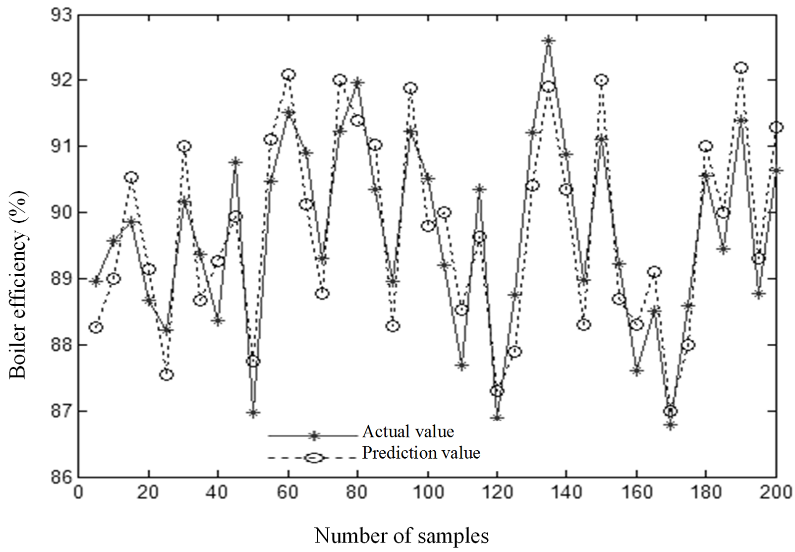

Fig. 3. Output curve of actual value and soft sensor value of boiler thermal efficiency

Fig. 4 shows the same trend of the rise and fall both of the soft sensor values and actual values, only the amplitude of variation is slightly different in the local range, which confirms that the soft sensor model built of the boiler thermal efficiency has higher prediction accuracy and better generalization ability. Fig. 4 contains the curve of the relative error and corresponds to Fig. 3.

The relative error is controlled basically within 1% and variance is 0.0043, as is shown in Fig. 4.

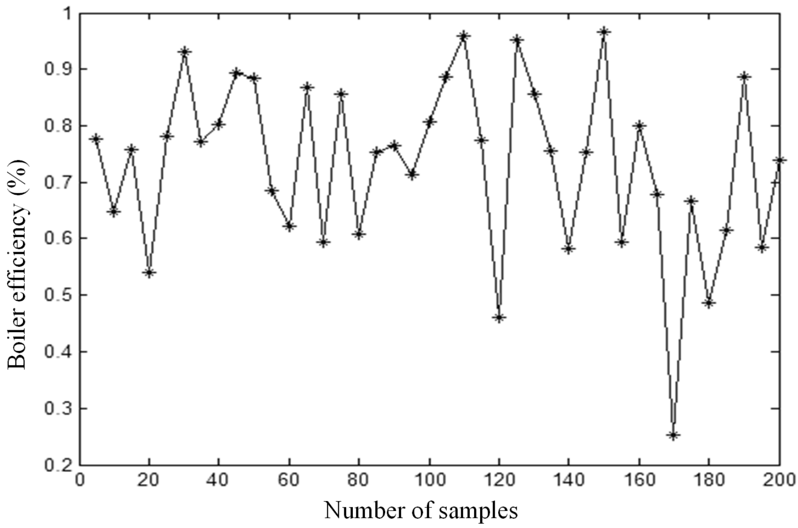

Fig. 4. Relative error curve of the boiler thermal efficiency

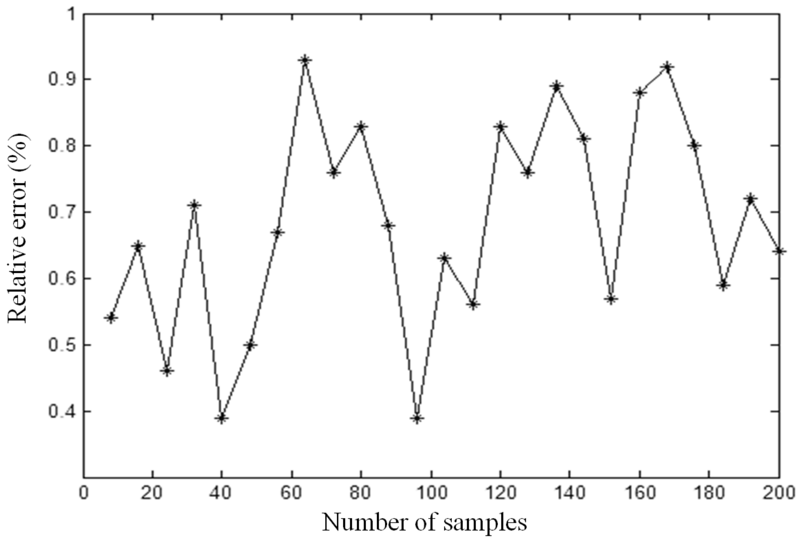

Fig. 5. Relative error of the boiler thermal efficiency

25 groups of data are selected randomly in order to further verify the generalization of the model established. The soft sensor results of the model show that the relative error can be limited in the range of accuracy and the soft sensor model has better generalization ability, as is shown in Fig. 5.

6. Conclusion

A novel soft sensor model of the boiler thermal efficiency is established by applying support vector regression and grid search algorithm about the problem that the boiler thermal efficiency is difficult to measure accurately. To optimize the parameters C and g of the prediction model are found using adaptive grid search algorithm. Soft sensor accuracy and generalization capability of the model are improved by preprocessing of acquisition data and verified by using test data. The simulation results show that the relative error of the soft-sensing model established is controlled within 1%, the variance is only 0.0043. In summary, soft sensor accuracy and generalization capability of the soft sensor model established can meet the practical requirements and provide an effective way of measuring the boiler thermal efficiency in power plant.

References

[1] Y. Torii, S. Abe: *Decomposition techniques for training linear programming support vector machines*. Neurocomputing *72* (2009), Nos. 4–6, 937–984.

[2] A. Borin, M. F. Ferrao, C. Mello, D. A. Maretto, R. J. Poppi: *Least-squares support vector machines and near infrared spectroscopy for quantification of common adulterants in powdered milk*. Analytica Chimica Acta *579* (2006), No. 1, 25–32.

[3] P. Han, F. Gao, Y. J. Zhai, Y. Lu : *Combustion modeling for utility boiler based on multi-output least squares support vector machine*. IJ Advancements Comput. Technol. *4* (2012), No. 14, 60–68.

[4] Y. He: *Combustion system model based prediction of boiler efficiency and NOx emission*. Thermal Power Generation *42* (2013), No. 4, 56–61.

[5] Y. Gu, W. Zhao, Z. Wu: *Online adaptive least squares support vector machine and its application in utility boiler combustion optimization systems*. J Process Control *21* (2011), No. 7, 1040–1048.

[6] L. X. Yin, M. C. Wang, Q. Shang: *The combustion optimization of a coal-fired boiler based on support vector machine and particle swarm algorithm*. Boiler Technol. *45* (2014) No. 4, 123–128.

[7] B. Ren, Z. P. Feng, X. Jiaotong: *Improved genetic algorithm and particle swarm optimization as well as comparison between them*. J Nanjing Normal University (Engineering and Technology Edition) (2002), No. 2, 14–20.

[8] M. A. Mohandes, T. O. Halawani, S. Rehman, A. A. Hussain: *Support vector machine for wind speed prediction*. Renewable Energy *29* (2004), No. 6, 939–947.

[9] F. Si, C. E. Romero, Z. Yao, E. Schuster, Z. Xu, R. L. Morey, B. N. Liebowitz: *Optimization of coal-fired boiler SCRs based on modified support vector machine models and genetic algorithms*. Fuel *88* (2009), No. 5, 806–816.

[10] H. Li, M. D. Li, H. J. Chen: *Soft sensors for thermal efficiency in industry oil boiler based on BPNN*. Energy Research & Utilization *3*, (2010), No. 4, 135–139.

Received November 16, 2016