

## ORIGINAL RESEARCH ARTICLE

# Modeling of boiler variable load combustion system based on gradient lifting decision tree and improved bidirectional threshold cycle unit

Guotian Yang, Yuchen He\*, Xin Li, Xinli Li

School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China. E-mail: he775916@163.com

## ABSTRACT

Boiler combustion system is a typical dynamic system with many variables, strong coupling, large-lag, and multiple input/output. It is very difficult to build a combustion system model that conforms to the actual working conditions. This paper presents a new modeling method of boiler combustion system based on bidirectional threshold cycle unit (Bi-GRU), and establishes the training model of combustion system under variable load (low, medium and high load) conditions. At the same time, gradient lifting decision tree (GBDT) is used to reduce the dimension of input characteristic matrix. GBDT model can evaluate the weight of input features under different loads and outputs, and can identify the feature with the largest weight proportion on the basis of retaining the original physical meaning of the feature. The feature selection model based on GBDT can not only reduce the original input dimension, but also provide theoretical guidance for the subsequent combustion control strategy. The calculation results of actual operation data show that the new combustion system model established by Bi-GRU and GBDT can accurately reflect the dynamic changes of main steam flow, main steam pressure and NO<sub>x</sub> emission under different loads. Compared with the traditional recurrent neural network (RNN) model, the accuracy and performance of the new model in this paper are significantly improved, and the structure is simple and the amount of calculation is small.

**Keywords:** Boiler Combustion System; Bidirectional Threshold Circulation Unit; Gradient Lifting Decision Tree; Output Characteristics

## ARTICLE INFO

Received: 3 March 2022  
Accepted: 5 April 2022  
Available online: 17 April 2022

## COPYRIGHT

Copyright © 2022 Guotian Yang, *et al.*  
EnPress Publisher LLC. This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).  
<https://creativecommons.org/licenses/by-nc/4.0/>

## 1. Introduction

The boiler combustion system is a typical multi-input and multi-output system with complex variables, which has the characteristics of nonlinearity, strong coupling, large lag, and strong timing and correlation of input and output signals<sup>[1,2]</sup>. Large qualitative changes, unit load fluctuations caused by frequent grid adjustments, and fewer measurement points of relevant parameters in the boiler all add difficulties to the modeling of boiler combustion system. Therefore, establishing a combustion system model closer to the actual working conditions has always been a hot spot in the field of boiler combustion<sup>[3-6]</sup>.

The existing boiler combustion system modeling is mainly divided into white box model based on combustion mechanism and black box model based on data drive. The white box model is established based on the dynamic physical characteristics of thermodynamic variables<sup>[7-12]</sup>, and its model is complex and difficult to be applied to the actual boiler combustion control site. Black box model is based on data modeling, and its calculation cost is low. It is generally modeled by data mining, data fitting and other methods. Common black box models are based

on neural network<sup>[13]</sup>, support vector machine<sup>[14]</sup>, Gaussian process<sup>[15]</sup>, etc., which are established on the basis of analyzing the correlation between known operating parameters and experimental data. Compared with the white box model, the black box model is more suitable for complex industrial field control. With the continuous advancement of intelligent control in power plants, the traditional rough boiler combustion control has gradually changed to fine combustion control, and the amount of data collected by boiler side measurement points and distributed control system (DCS) has increased; at the same time, industrial process modeling is not a static model, but a dynamic process in which the variables of the system change with time. Traditional data-driven methods cannot support large amount of data operations, and the implementation of dynamic models often increases with the geometric multiple of the input model complexity. Redundant information input seriously affects the accuracy and generalization ability of the model. As a new machine learning algorithm, deep learning shows more and more powerful application value in big data mining, but it is less applied in boiler combustion system modeling.

This paper presents a boiler combustion system model based on data-driven bidirectional threshold cycle unit. Because the boiler combustion process is sensitive to load changes, the training model is divided into three parts: low load, medium load and high load. 40 features are selected as input features, such as coal feed volume, primary air volume, secondary air volume, pulverized coal fineness, oxygen content, etc. of each combustion layer; the main steam pressure, main steam temperature and NO<sub>x</sub> emission are taken as outputs to establish the boiler combustion system model. At the same time, the gradient boosting decision tree (GBDT) is used to reduce the dimension of the input features on the basis of retaining the original physical meaning of the features, and the most relevant features under different loads are selected to facilitate the modeling and optimal control of the subsequent combustion system. Compared with the traditional recurrent neural network (RNN) method, the boiler combustion system model based on GBDT-BiGRU proposed in this paper has higher

accuracy and less calculation time.

## 2. GBDT-BiGRU algorithm

### 2.1 Gradient lifting decision tree

Gradient boosting decision tree is an iterative decision tree algorithm composed of multiple high-dimensional decision trees. It uses computational features to select features of the relative importance of a single tree to achieve the purpose of dimensionality reduction of high-dimensional data. GBDT algorithm has the advantages of decision tree algorithm, that is, it automatically combines multiple features without standardizing or normalizing the features, does not consider whether the data is linearly separable, and the algorithm is highly interpretable; at the same time, it helps to suppress the complexity of the decision tree, reduce the fitting ability of a single decision tree, and eliminate the over fitting problem. Due to the excessive model input of the boiler combustion system itself, in order to achieve combustion refinement and refine some feature quantities, feature selection using GBDT algorithm can not only reduce the dimension of model input, but also globally explain the overall role of features in the model<sup>[16,17]</sup>.

The core of GBDT algorithm is to combine multiple weak classifiers cart tree into a strong classifier, follow the downward direction of negative gradient to ensure the convergence of the algorithm and realize the global convergence of the model<sup>[18,19]</sup>. The implementation steps of GBDT algorithm are as follows:

(1) the model input data set is  $\{x_i, y_i\}$ , where  $i, j = 1, 2, \dots, n$ ; GBDT model loss function  $L()$  is softmax function, namely:

$$L(x_i) = -x_i + \log \sum_{j=1}^n e^{x_j} \quad (1)$$

(2)  $F(x)$  is the GBDT classifier function, where  $F_0(x)$  is the initial classifier. Let the partial derivative of  $F_0(x)$  be 0, and the initial weight  $y_0$  of the model can be obtained from equation (2).

$$F_0(x) = \arg \min_{\beta} \sum_{i=1}^n L(y_i, \gamma_0) \quad (2)$$

(3)  $t = 1, 2 \dots, T$  is the number of iterations,

and the gradient descent direction  $g_{i,t}$ :

$$g_{i,t}(x_i) = - \left[ \frac{\partial L(y_i, F(x_i))}{\partial F(x_i)} \right]_{F(x)=F_{t-1}(x)} \quad (3)$$

(4) Use the sample data of  $\{x_i, g_{i,t}\}$  to fit the  $m$ -th CART regression tree. According to the least square method, calculate the best fitting parameters  $a_{m,r}$  of the  $r$ -th leaf node of the  $m$ -th CART regression tree and the fitted CART regression tree model  $h(x_i, a_{m,r})$ :

$$a_{m,r} = \arg \min_{a_{m,r}, \gamma_0} \sum_{i=1}^n [-g_{i,t} - \gamma_0 h(x_i; a_{m,r})]^2 \quad (4)$$

(5) Minimize the loss function  $L()$ , and calculate the new step size of the model according to equation (5), that is, the new weight of the model  $y_i$ :

$$\gamma_t = \arg \min_{a_{m,r}, \gamma_t} \sum_{i=1}^n L(y_i, F_{t-1}(x) + \gamma_t h(x_i; a_{m,r})) \quad (5)$$

(6) Update the model until the model meets the maximum number of iterations  $T$  or convergence, and get the strongest CART regression tree  $F_t$ :

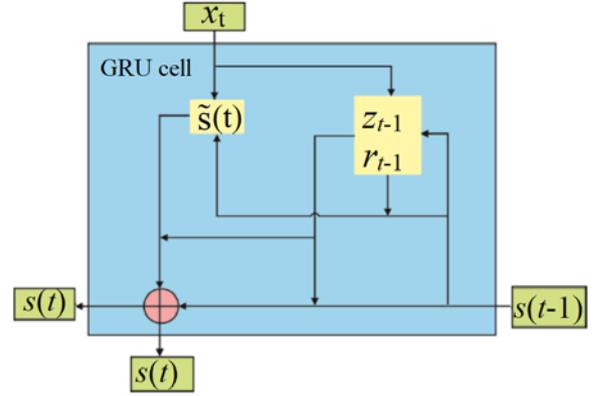
$$F_t(x) = F_{t-1}(x) + \gamma_t h(x_i; a) \quad (6)$$

(7) Output  $F_t$ .

## 2.2 GRU neural network model

RNN model has obvious advantages in dealing with short-term ordinal series, but it is easy to disappear gradient when dealing with time series with high analysis dimension. In this regard, Grave *et al.* proposed the long short-term memory (LSTM) neural network to improve the RNN model structure, set hidden layer memory units (i.e. Input gate, forgetting gate and output gate) to realize the memory control of time sequence, and solve the problem of gradient disappearance of RNN model. However, the hidden layer structure of LSTM neural network is complex and the training sample time is too long<sup>[20]</sup>. Therefore, Cho *et al.*<sup>[21]</sup> proposed the structure of gated recurrent unit (GRU) neural network to simplify the gate setting of LSTM neural network and reduce the model training time. Compared with the three gating units of LSTM neural network, GRU neural network has only two gating

units, namely reset gate ( $z$ ) and update gate ( $r$ ); at the same time, GRU neural network does not have a separate storage unit, which simplifies the number of model parameters, improves the convergence speed of the algorithm, retains the advantages of LSTM neural network, and improves the efficiency of sample training. **Figure 1** shows the structure of GRU neural network.



**Figure 1.** Schematic diagram of the GRU neural network structure.

The gating update formula of GRU neural network model is:

$$\begin{aligned} r(t) &= f(W_r \cdot [s(t-1), x(t)]) \\ z(t) &= f(W_z \cdot [s(t-1), x(t)]) \\ \tilde{s}(t) &= \tanh(W_s \cdot [r(t) \circ s(t-1), x(t)]) \\ s(t) &= (1 - z(t)) \circ s(t-1) + z(t) \circ \tilde{s}(t) \end{aligned} \quad (7)$$

Where:  $r(t)$  and  $z(t)$  are the status of update door and reset door at time  $t$  respectively;  $W_r$ ,  $W_z$ ,  $W_s$  are the weight matrix of update gate, reset gate and hidden layer network state respectively;  $s(t)$ ,  $\tilde{s}(t)$  are the network states of hidden layer and candidate hidden layer respectively;  $\circ$  indicates Hadamard product.

The output of GRU neural network is:

$$y(t) = W \cdot s(t) \quad (8)$$

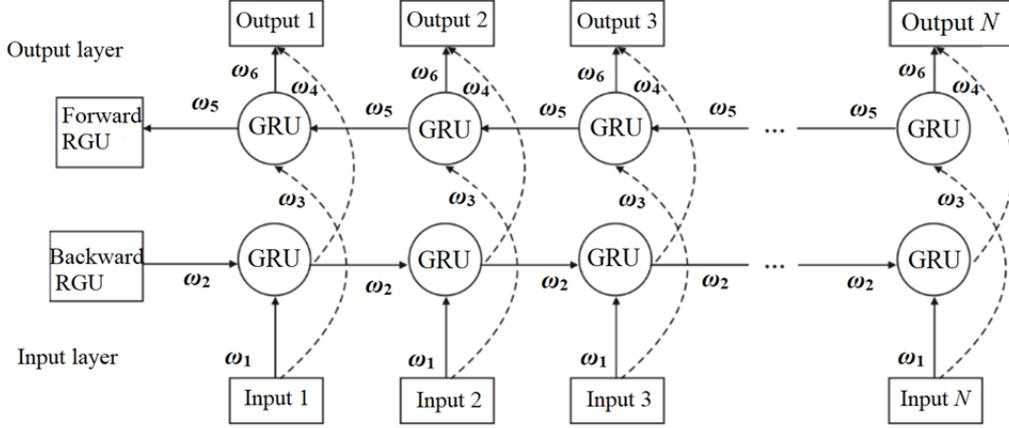
Where:  $W$  is the output matrix.

## 2.3 Bi-GRU neural network model

Due to the particularity of the boiler combustion system, the current time output of the model is the result of the accumulation of all historical inputs and outputs<sup>[22]</sup>. Therefore, the model needs to be able to learn the complete before and after infor-

mation of time series. In order to improve the training accuracy of the model, the hidden layer of GRU neural network in this paper selects the bidirectional gated recurrent unit (Bi-GRU) structure, and adds the full connection layer of linear correction unit function between the feature output layer and the

input layer, so that the features extracted during network training are more effective. In order to prevent training from over fitting, the hidden layer loss rate of GRU neural network is set to 0.2. The structure of Bi-GRU neural network is shown in **Figure 2**.



**Figure 2.** Schematic diagram of the Bi-GRU neural network structure.

The update formulas of the forward and backward candidate hidden layer network states  $\tilde{s}^f(t)$ ,  $\tilde{s}^b(t)$  of Bi-GRU neural network are shown in equations (9) and (10):

$$\tilde{s}^f(t) = \tanh(\mathbf{W}_s^f \cdot [r(t) \circ s^f(t-1), x(t)]) \quad (9)$$

$$\tilde{s}^b(t) = \tanh(\mathbf{W}_s^b \cdot [r(t) \circ s^b(t+1), x(t)]) \quad (10)$$

The final output formula of Bi-GRU neural network is:

$$\mathbf{y}(t) = \mathbf{W}_s^f \cdot s^f(t) + \mathbf{W}_s^b \cdot s^b(t) \quad (11)$$

Where,  $\mathbf{W}_s^f, \mathbf{W}_s^b$  are the weight matrices of the forward and backward hidden layer network states, and  $s^f(t), s^b(t)$  are the forward and backward hidden layer network states.

### 3. Boiler combustion system model based on GBDT-BiGRU

#### 3.1 Evaluation function

In this paper, three criterion functions are selected to evaluate the accuracy of the model: root mean square error  $\delta_{\text{RMSE}}$ , mean absolute error  $\delta_{\text{MAE}}$  and absolute error  $\delta_{\text{AE}}$ . The root mean square error

is sensitive to the main error of prediction. The average absolute error reveals the average distribution of the overall error. The absolute error represents the deviation degree between the identification value and the target value, and the expressions are:

$$\delta_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_{\text{target}} - y_{\text{ident}})^2} \quad (12)$$

$$\delta_{\text{MAE}} = \frac{1}{n} \sum_{t=1}^n |y_{\text{target}} - y_{\text{ident}}| \quad (13)$$

$$\delta_{\text{AE}} = y_{\text{ident}} - y_{\text{target}} \quad (14)$$

Where:  $y_{\text{target}}$  is the measured value of the target;  $y_{\text{ident}}$  is an analog value.

#### 3.2 Boiler combustion system modeling

Based on GBDT and Bi-GRU neural network, the boiler combustion system model is established. The main steps are as follows:

(1) Data preparation. Experimental input from 10,500 sample data collected from the on-site DCS, including training data set (8,000 samples) and test data set (2,500 samples). One sample data is collected every 30 s, and a total of 66 hours of continuous and stable operation of the boiler under normal

working conditions is collected. 40 features are selected as inputs (input coal volume, pulverized coal fineness, primary air volume, secondary air volume, etc. of each combustion layer), and 3 features are selected as outputs. The output characteristic variables are main steam flow, main steam pressure and  $\text{NO}_x$  emission.

(2) Feature selection the research object of the boiler in this paper is DG-2060/26.15-II 2660MW ultra supercritical variable pressure once through boiler manufactured by Dongfang Boiler Factory, which adopts the opposed swirl combus-

tion mode of front and rear walls. The boiler has 6 layers of burners, 3 layers are arranged on the front and rear walls, and each layer is equipped with 6 sets of burners, a total of 36 burners. Therefore, the boiler output below 200 MW is low load, 200–400 MW is medium load, and more than 400 MW is high load. The pulverizing system adopts MPS212HP-II medium speed coal mills, and each boiler is equipped with 6 medium speed coal mills and 6 EG-2690 electronic weighing coal feeders. See **Table 1** and **Table 2** for boiler parameters and pulverizer parameters.

**Table 1.** Main design parameters of the ultra-supercritical 660 MW unit boiler

Project	BMCR value
Superheated steam flow/( $\text{t}\cdot\text{h}^{-1}$ )	2,060
Superheater outlet steam pressure/MPa	26.15
Superheater outlet steam temperature/ $^{\circ}\text{C}$	605
Reheat steam flow/( $\text{t}\cdot\text{h}^{-1}$ )	1,676.9
Reheater inlet steam pressure /MPa	5.33
Reheater outlet steam pressure /MPa	5.13
Reheater inlet steam temperature/ $^{\circ}\text{C}$	362
Reheater outlet steam temperature/ $^{\circ}\text{C}$	603
Feed water temperature at economizer inlet/ $^{\circ}\text{C}$	297
Primary air outlet temperature/ $^{\circ}\text{C}$	339
Secondary air outlet temperature/ $^{\circ}\text{C}$	345
Exhaust outlet temperature (correction)/ $^{\circ}\text{C}$	122
Actual fuel consumption/( $\text{t}\cdot\text{h}^{-1}$ )	297.2
Calculate boiler thermal efficiency/%	93.10
Guaranteed thermal efficiency/%	93.00

**Table 2.** Main design parameters of the ultra-supercritical 660 MW unit boiler equipped with medium speed coal mill

Project	Numerical value
Maximum output/( $\text{t}\cdot\text{h}^{-1}$ )	76.1
Minimum output/( $\text{t}\cdot\text{h}^{-1}$ )	19.0
Guaranteed output/( $\text{t}\cdot\text{h}^{-1}$ )	71.4
Maximum ventilation/( $\text{t}\cdot\text{h}^{-1}$ )	104.3
Minimum ventilation volume/( $\text{t}\cdot\text{h}^{-1}$ )	78.2
Speed of coal mill/( $\text{r}\cdot\text{min}^{-1}$ )	31.6
Ventilation resistance/Pa	7,190
Unit power consumption of coal mill/(( $\text{kW}\cdot\text{h}$ ) $\cdot\text{t}^{-1}$ )	7.89

According to different load requirements, GBDT is used to calculate the characteristic weight matrix under high, medium and low loads respectively. The high, medium and low loads mentioned in this paper take more than 60% of the rated output of the unit for safe operation as the high load, 30%–60% as the medium load, and less than 30% as the low load.

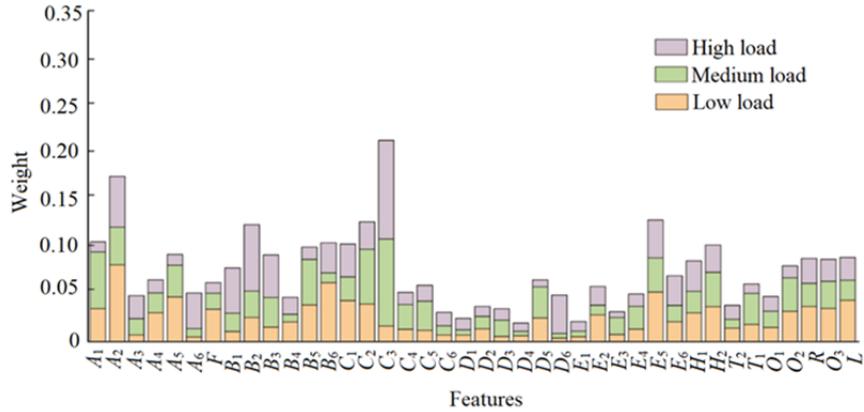
The model in this paper has three outputs, each of which has an appropriate feature selection model. However, the boiler combustion system model only needs one feature selection model under different loads. Therefore, under different outputs, the same input characteristic weights under different working

conditions are linearly accumulated, and the weights are averaged to obtain the input characteristic weight models under three different outputs, and the sum of weights greater than 85% is selected as the input characteristics of the boiler combustion process model.

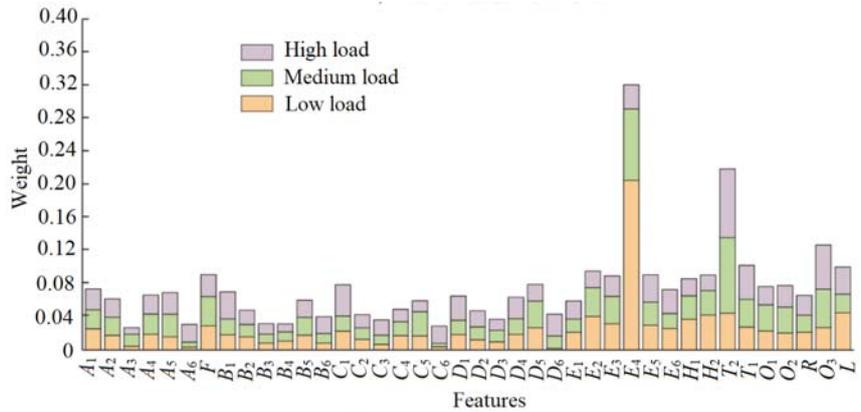
(3) Model construction and training to avoid over fitting, the loss rate of the hidden layer of Bi-GRU neural network is set to 0.2. The results of GBDT model are selected as the model input, and its output is the main steam flow, main steam pressure and  $\text{NO}_x$  emission, with mean square error  $\delta_{\text{MSE}}$  is used as the model loss function, and the initial learning rate is set to 0.05 training model.

(4) Model validation input 2,500 characteristic data in the test data set into the trained Bi-GRU neural network model, and output the main steam flow, main steam pressure and  $\text{NO}_x$  emission. Three

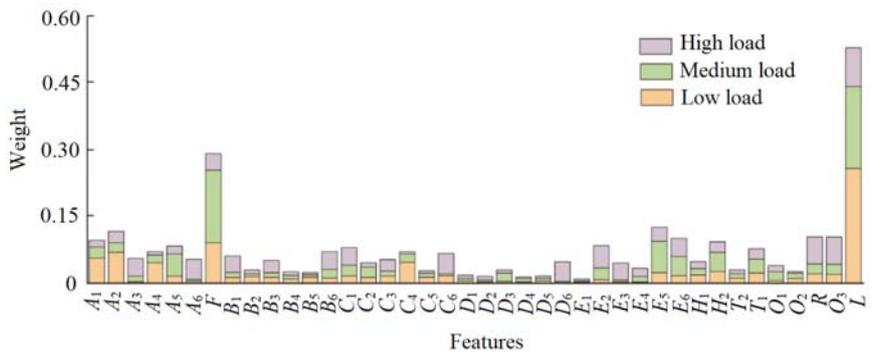
criterion functions are introduced to evaluate the generalization performance and accuracy of the model.



(a) Main steam flow input characteristic weight



(b)  $\text{NO}_x$  emission input characteristic weight



(c) Main steam pressure input characteristic weight

**Figure 3.** The weight cumulative bars of GBDT-based input to output features at variable load.

Note:  $A_n, B_n, C_n, D_n, E_n$  ( $n = 1, 2, \dots, 6$ ) are the coal feeding volume, pulverized coal concentration, grinding fineness, primary air volume and secondary air volume of burner layers A, B, C, D, E respectively;  $H_1, H_2$  are the burnout air volume of the front and rear walls respectively; F is feedwater flow; R is the actual water coal ratio;  $O_1, O_2, O_3$  are the oxygen content of flue gas at the outlet of front and rear walls and the oxygen content at the inlet of the reactor respectively;  $T_1, T_2$  are the main steam temperature and flue gas temperature of the air preheater respectively; L is the load.

## 4. Experimental analysis and result discussion

### 4.1 GBDT feature weight result analysis

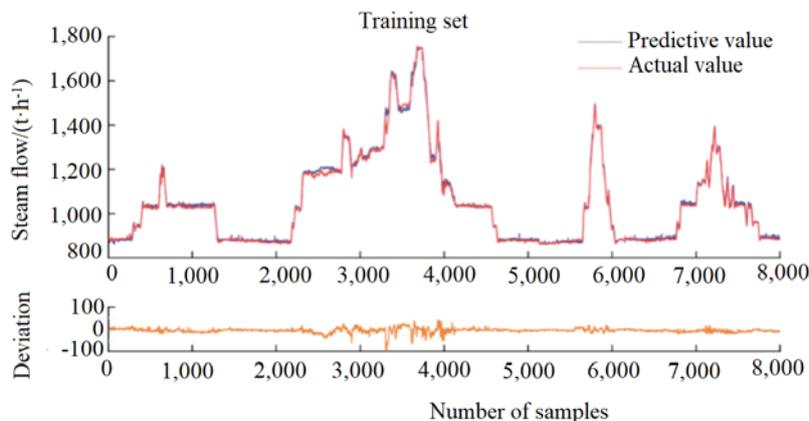
**Figure 3** is the cumulative bar graph of the weight of GBDT input characteristics to output under variable load. **Figure 3** shows the change and relationship of different load characteristic weights. The refined input characteristic division can make the selection and control of corresponding characteristic parameters easier in the process of boiler load change. It can be seen from **Figure 3(a)** that for the main steam flow, the coal volume, pulverized coal concentration and pulverized coal fineness are the main influence characteristics. Except for the primary air volume of each layer, the weights of other characteristics are almost equal. It can be seen from **Figure 3(b)** that for  $\text{NO}_x$  emission, the main influence characteristics are secondary air volume, temperature and oxygen content, while other factors have less influence. It can be seen from **Figure 3(c)** that load and feedwater are the main characteristics affecting steam pressure, while the effects of other characteristics can be ignored. At the same time, it is further found that for different load characteristic models, the influence of primary air volume accounts for a small proportion. This is because relevant characteristics such as coal feeding volume already exist, and primary air volume can be ignored to a certain extent.

As shown in **Figure 3**, when the load changes

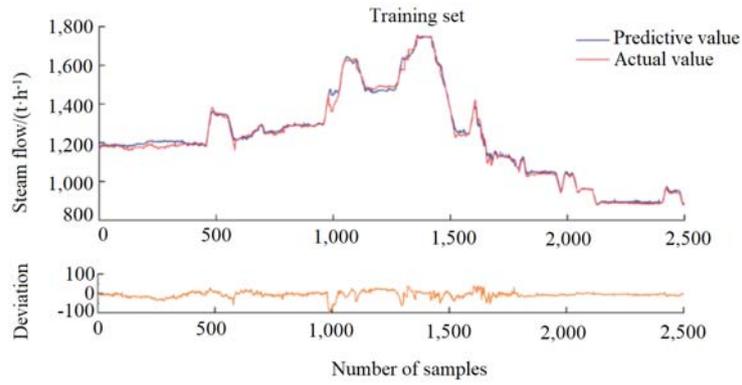
from low load to medium load, the operator can give priority to controlling the amount of pulverized coal, pulverized coal concentration and pulverized coal fineness of the burners on layer B and E to ensure the stability of steam flow and pressure, and then control the change of secondary air volume to estimate the emission of  $\text{NO}_x$ . Therefore, the feature selection model of GBDT is not only to select the feature with the largest amount of information, realize the dimensionality reduction of the input feature matrix, but also provide a theoretical basis for the priority of operator control parameters.

### 4.2 Experimental results of Bi-GRU neural network identification

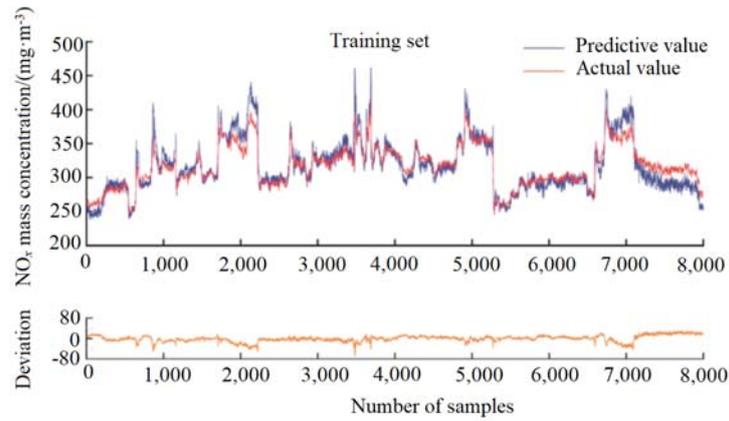
Using the feature selection results, the Bi-GRU neural network boiler combustion system model is constructed. The training and testing identification results and errors of main steam flow,  $\text{NO}_x$  emission and main steam pressure are shown in **Figures 4-9**. It can be seen from **Figure 4** to **Figure 9** that the deviation between the identified value of main steam flow,  $\text{NO}_x$  emission and main steam pressure and the actual value is small. Compared with the root mean square error and average absolute error of main steam flow and pressure, the accuracy of  $\text{NO}_x$  emission identification results is much lower. This is because the model in this paper must first ensure the stability of main steam flow and pressure.



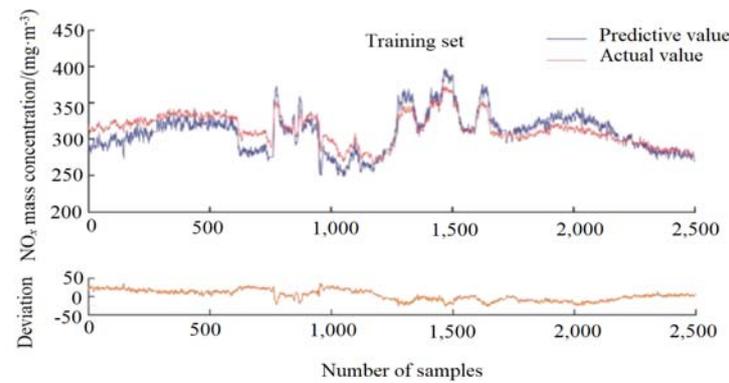
**Figure 4.** The identification results and errors of main steam flow training set.



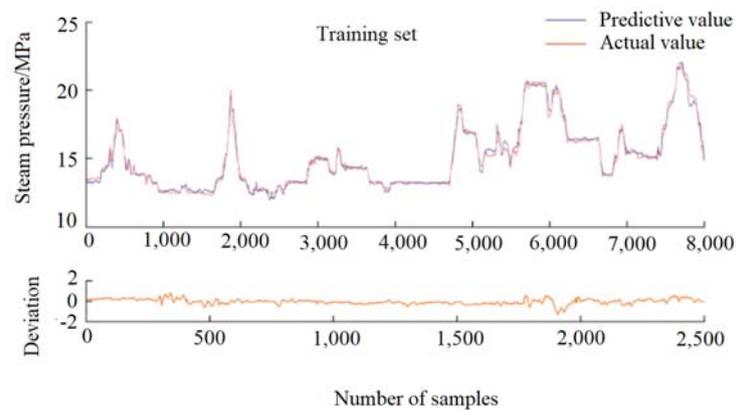
**Figure 5.** The identification results and errors of main steam flow testing set.



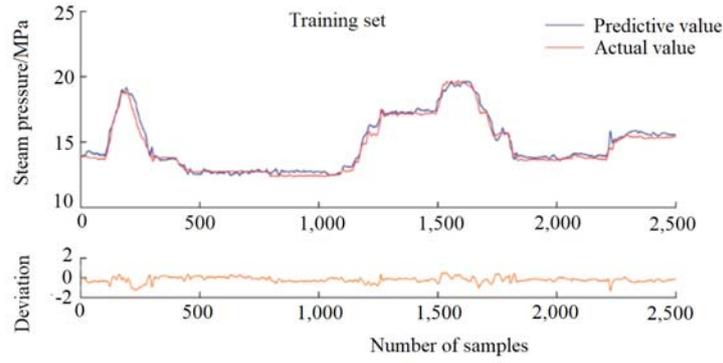
**Figure 6.** The identification results and errors of NO<sub>x</sub> emission training set.



**Figure 7.** The identification results and errors of NO<sub>x</sub> emission testing set.



**Figure 8.** The identification results and errors of main steam pressure training set.



**Figure 9.** The identification results and errors of main steam pressure testing set.

**Table 3** shows the comparison of root mean square error, mean absolute error and calculation time between GBDT-BiGRU and RNN.

**Table 3.** Comparison of  $\delta_{RMSE}$ ,  $\delta_{MAE}$  and computation time between GBDT-BiGRU and RNN

Project	Parameter		$\delta_{RMSE}$	$\delta_{MAE}$	Time/s
Bi-GRU	Main steam flow	Training data	10.757	7.051	697
		Test data	16.927	11.974	243
	Main steam pressure	Training data	0.267	0.191	697
		Test data	0.327	0.254	243
RNN	NO <sub>x</sub> emission	Training data	12.439	9.169	697
		Test data	14.518	12.359	243
	Main steam flow	Training data	12.233	9.836	1161
		Test data	18.431	12.994	431
	Main steam pressure	Training data	0.292	0.208	1161
		Test data	0.271	0.187	431
	NO <sub>x</sub> emission	Training data	13.063	10.054	1161
		Test data	15.011	13.026	431

It can be seen from **Table 3** that the results of Bi-GRU model are far better than those of traditional RNN model. The reason why Bi-GRU algorithm has better performance and model accuracy than other methods is that it can process a large number of time series data and carry important information of initial learning for a long time. Compared with RNN model, Bi-GRU model has a simpler structure, which can maximize the accuracy of the model and greatly reduce the calculation time.

## 5. Conclusion

With the deepening of intelligent power plant, the amount of DCS data on industrial site is increasing geometrically, and the traditional rough boiler combustion is gradually changing to fine combustion. Based on the operation data of a 660 MW power plant, the boiler combustion system is modeled in this paper. In order to obtain a high-precision dynamic model, the boiler combustion system model is established by using Bi-GRU. Considering the sensitivity of combustion system to load changes, three models are established under

low, medium and high loads. The actual data simulation calculation shows that the model in this paper can better reflect the change trend of the output of the boiler combustion system. Compared with other methods, Bi-GRU has more outstanding performance and model accuracy. At the same time, the model has simple structure and shorter calculation time, which provides a basis for further study of dynamic control and optimization of boiler combustion system.

## Acknowledgement

This work was supported by Fundamental Research Funds for the Central Universities (JB2017169).

## Conflict of interest

The authors declared that they have no conflict of interest.

## References

1. Wang Z, Liu H, Chen Y. Process control systems and instruments. Beijing: Machinery Industry Press;

2006. p. 1.
2. Pan H, Zhong W, Wang Z, *et al.* Optimization of industrial boiler combustion control system based on genetic algorithm. *Computers and Electrical Engineering* 2018; 70: 987–997.
  3. Tang Z, Zhang Z. The multi-objective optimization of combustion system operations based on deep data-driven models. *Energy* 2019; 182: 37–47.
  4. Zhao M, Wang P, Liang J, *et al.* Modeling of reference value of boiler operating parameters based on weighted fuzzy C-means clustering algorithm. *Thermal Power Generation* 2019; 48(1): 12–17.
  5. Li C, Bai F, Hu S. Establishment of system model and parameters optimization of HRSG combined cycle units. *Thermal Power Generation* 2021; 50(2): 74–83.
  6. Zhen C, Liu H. Prediction model of NO<sub>x</sub> emission from coal-fired boiler based on multi-model clustering ensemble. *Thermal Power Generation* 2019; 48(4): 33–40.
  7. Niu Y, Du M, Ge W, *et al.* A dynamic nonlinear model for a once-through boiler-turbine unit in low load. *Applied Thermal Engineering* 2019; 161: 113880.
  8. Xu L, Yuan J. Online identification of the lower heating value of the coal entering the furnace based on the boiler-side whole process models. *Fuel* 2015; 161: 68–77.
  9. Henryk R, Wojciech S. Hybrid model of steam boiler. *Energy* 2009; 35(2): 1107–1113.
  10. Eddy HC, Gao H. Estimation of NO<sub>x</sub> emissions from coal-fired utility boilers. *Fuel* 2010; 89(10): 2977–2984.
  11. Sandro DS, Olivier J, Myriam LL, *et al.* Using a genetic algorithm and CFD to identify low NO<sub>x</sub> configurations in an industrial boiler. *Fuel* 2015; 158: 672–683.
  12. Choi CR, Kim CN. Numerical investigation on the flow, combustion and NO<sub>x</sub> emission characteristics in a 500 MWe tangentially fired pulverized-coal boiler. *Fuel* 2009; 88(9): 1720–1731.
  13. Dong X, Wang H, Zhao X. Model reference neural network control for boiler combustion system. *Proceedings of 2005 International Conference on Machine Learning and Cybernetics*; 2005 Aug 1821; Guangzhou, China: IEEE; 2005. p. 12.
  14. Gu Y, Zhao W, Wu Z. Online adaptive least squares support vector machine and its application in utility boiler combustion optimization systems. *Journal of Process Control* 2011; 21(7): 1040–1048.
  15. Chen J, Chan L, Cheng Y. Gaussian process regression based optimal design of combustion systems using flame images. *Applied Energy* 2013; 111: 153–160.
  16. Xu X, Chen Q, Ren M, *et al.* Combustion optimization for coal fired power plant boilers based on improved distributed ELM and distributed PSO. *Energies* 2019; 12(6): 1036.
  17. Rao H, Shi X, Rodrigue AK, *et al.* Feature selection based on artificial bee colony and ent boosting decision tree. *Applied Soft Computing* 2018; 74: 634–642.
  18. Tian D, He G, Wu J, *et al.* An accurate eye pupil localization approach based on adaptive gradient boosting decision tree. *2016 Visual Communications and Image Processing*; 2016 Nov 27–30; Chengdu, China: IEEE; 2016.
  19. Xu Z X, Huang G, Weinberger KQ, *et al.* Gradient boosted feature selection. *Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining*; 2014 Aug 24; 2014. p. 522–531.
  20. Gtanes A, Mohamed AR, Hinton G. Speech recognition with deep recurrent neural networks. *IEEE International Conference on Acoustic*; 2013 May 26–31; Vancouver, BC, Canada: IEEE; 2013.
  21. Cho K, Merriënboer BV, Gulcehre C, *et al.* Learning phrase representations using RNN encoder-decoder for statistical machine translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*; 2014 Oct 25–29; Doha, Qatar; 2014. p. 1724–1734.
  22. Jaihuni M, Basak JK, Khan F, *et al.* A partially amended hybrid Bi-GRU-ARIMA model (PAHM) for predicting solar irradiance in short and very-short terms. *Energies* 2020; 13: 435.