

Study on boiler's comprehensive benefits optimization based on PSO optimized XGBoost algorithm

Wang Hui¹, Zhang Guobao^{1,*}, Huang Yongming¹, and Zhang Yongchun²

¹College of Automation, Southeast University, Nanjing, China, 210000

²Jiangsu special equipment safety supervision and Research Institute Nanjing China, 210036

Abstract. To predict the boiler's combustion efficiency and NO_x emissions, this paper introduced a particle swarm optimization optimized XGBoost algorithm. The results show that the MAPE can reach 0.107% and 3.732% respectively on the verification set, which is better SVM, LR and ANN. At the same time, this paper presents a comprehensive benefits evaluation function considering economic and environmental benefits to optimize the multi-objective optimization problem of boiler's combustion efficiency and NO_x emission. Based on the operation data of a 300 MW Circulating Fluidized Bed, the experimental results show that: the comprehensive benefits evaluation function can reasonably balance boiler's combustion efficiency and NO_x emissions to achieve the optimal comprehensive benefit.

1 Introduction

Coal is the main fuel of power plant in China^[1]. To achieve the purpose of economic performance and environmental protection, boiler's optimization should not only improve the combustion efficiency, but also reduce NO_x emissions. At this time, it is very important to establish high precision predictive model.

The combustion process of boiler is a complex system with large time delay and nonlinearity. There are many factors that influence the combustion of boiler and they are coupled with each other^[2]. With the development of artificial intelligence technology, many scholars began to apply statistical methods to boiler's modeling, such as Linear Regression (LR)^[3] and Support Vector Regression (SVR)^[4, 5] have achieved good results. Because of its strong nonlinear fitting ability, ANN is used to predict boiler's combustion efficiency and NO_x emissions in papers^[6-8]. However, through the research of papers^[9, 10], the advantage of ANN is that it can extract information from some weak feature data such as text and audio. In the field of traditional data mining, because there are rich expert experience to guide feature selection, Decision Tree algorithms such as XGBoost^[11] with better interpretation are more suitable.

Aiming at the multi-objective optimization problem of boiler combustion efficiency and NO_x emissions, this paper^[12] takes the coal cost and NO_x emissions cost of power generation per unit of power plant boiler as the optimization objective, and adds them together. This method does not take into account the difference in the magnitude of coal cost and emission cost. In paper^[13], NO_x emissions and combustion efficiency is transformed into a single objective by using linear weighting method. The weight setting of this method is subjective, and can

not reflect the marginal effect of reducing NO_x emission and improving boiler combustion efficiency.

In this paper, XGBoost algorithm based on particle swarm optimization (PSO) is proposed to build the prediction model of boiler combustion efficiency and NO_x emissions, which is superior to the traditional LR, SVR and ANN. Based on the prediction model established by this algorithm, a nonlinear weighting evaluation function is proposed which unifies the boiler combustion efficiency and NO_x emissions into comprehensive benefits, and then PSO was used to optimize the boiler's parameters.

2 Prediction model of boiler thermal efficiency and NO_x emission

The NO_x emissions and boiler combustion efficiency could be described by nonlinear prediction models. A PSO optimized XGBoost algorithm is introduced in this section. Then The prediction models are constructed and evaluated in Section 2.3.

2.1 Experimental data

The experimental data is operation data of a 300 MW circulating flow bed data as shown in Table 1. Boiler combustion efficiency and NO_x emissions are the targets of the prediction models, and the remaining variables are used as inputs. Coal feed, air flow, and air temperature in the input variables are the most presentative controllable variables. The oxygen content of flue gas, the temperature of flue gas and the carbon content of flue gas are state variables

* Corresponding author: guobaozh@seu.edu.cn

Tab.1. Boiler operation variables

Category	Name	Unit	Count
Controllable variables	Coal feed	t/h	4
	Primary air flow	kNm ³ /h	2
	Primary air temperature	°C	2
	Secondary air flow	kNm ³ /h	2
	Secondary air temperature	°C	2
State variables	Oxygen content of the flue gas	%	1
	Temperature of the flue gas	°C	1
	Carbon content of the flue gas	%	2
Output variables	NO _x Emissions	mg/m ³	1
	Boiler combustion efficiency	%	1

2.2 PSO optimized XGBoost algorithm

2.2.1 XGBoost(e Xtreme Gradient Boosting)

XGBoost is an integrated decision tree algorithm, which can improve the prediction accuracy by continuously generating new decision trees to fit residuals. Each decision tree generated by XGBoost in the training process is a function f . Finally, the results of K decision trees are accumulated as the prediction results, as follows.

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) \quad f_k \in F \quad (1)$$

Where: x_i – the i th sample; \hat{y}_i – target value; K – total number of decision trees; f_k – k th decision tree model; F – set of decision tree models.

When XGBoost generates a decision tree, the objective function in (2) is minimized to determine the structure of the decision tree. The objective function consists of two parts: the loss function L is used to ensure the accuracy of the decision tree, and the regular term Ω is used to suppress the complexity of the model and prevent overfitting.

$$obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2)$$

Because XGBoost generates a new decision tree model f_t by fitting prediction residuals in each round, when XGBoost iterates to the t round, the objective function can be rewritten as equation (3)

$$obj^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (3)$$

The regular term formula is as follows:

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (4)$$

Where: γ and λ – penalty coefficients; t – number of leaf nodes; ω_j – weight of leaf nodes.

XGBoost uses the second-order Taylor expansion in the loss function approximation, and uses the information of the first-order and second-order derivatives to speed up the convergence rate. At the same time, it avoids over fitting effectively because of the addition of the regular term.

2.2.2 PSO optimized XGBoost algorithm

When XGBoost is used to build prediction models, different parameter settings will affect the final results. For the parameter selection of XGBoost, the traditional methods include empirical method, random walk and grid search. The empirical method is lack of theoretical guidance; the speed of grid search is slow; random walk is easy to fall into local optimum.

PSO is a global random search algorithm based on swarm intelligence. It is mostly used in some application scenarios of parameter optimization. For example, PSO is applied to find the optimal parameter combination to improve the prediction accuracy of SVM in paper [14]. In paper [15], it's applied to search a group of parameters which make the NO_x emissions lowest.

This paper used PSO (particle swarm optimization) to set XGBoost parameters. Firstly the initial population and the number of iterations were set, and then the parameters which need to be optimized were chose. The prediction accuracy of the model was taken as the fitness function, and training was carried out until the iteration rounds were reached. The pseudo code of PSO-XGBoost algorithm is as follows.

Algorithm 1 PSO-XGBoost

Input: N : Iteration round, P : Population size

Output: Optimal parameters

```

1: function Fit(X)
2:   result ← 0
3:   trainXGB(w)
4:   result ← evaluateXGB(X)
5:   return result
6: end function
7: function PSO-XGBOOST(N, P)
8:   for i ← 1 to P do
9:     initialize Xi
10:    initialize Vi
11:    pBesti ← Fit(Xi)
12:  end for
13:  gBest ← min{pBesti}
14:  for p ← 1 to N do
15:    for i ← 1 to P do
16:      update Xi
17:      update Vi
18:    end for
19:    if Fit(Xi) < Fit(pBesti) then
20:      pBesti ← Xi
21:    end if
22:    if Fit(pBesti) < Fit(gBesti) then
23:      gBesti ← pBesti
24:    end if
25:  end for
26: end function

```

Fig.1. PSO-XGBoost

2.3 Build the model

This section used PSO-XGBoost to build predictive models and chose three of XGBoost' parameters which have great influence on the accuracy for optimization, including maximum tree depth, learning rate and number of decision trees. The initial iteration is set to 15, and population size is 80. The results are shown in figures 3 and 4.

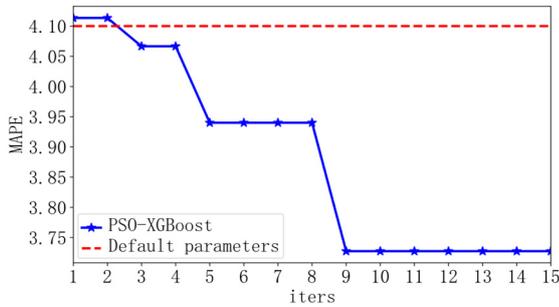


Fig.2. NO_x prediction model

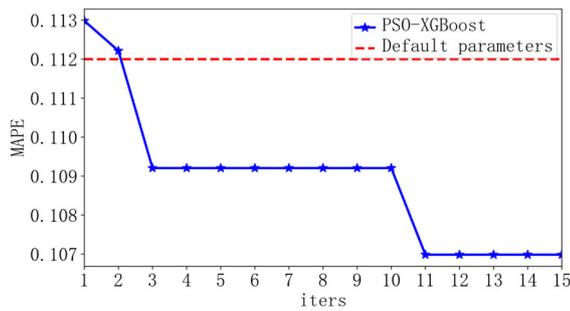


Fig.3. Combustion efficiency prediction model

According to figure 2, under the default parameters setting, the MAPE of NO_x emission prediction model is 4.100, and the final MAPE after optimization is 3.732; according to figure 4, under the default parameters setting, the MAPE of combustion efficiency prediction model is 0.112, and the MAPE after particle swarm optimization is 0.107. The experimental results show that PSO-XGBoost is better than the XGBoost with default parameters

SVR, ANN and LR, which are often used in boiler combustion efficiency and NO_x emission prediction, are compared with the proposed method. The experimental results are shown in table 2 and table 3.

Tab.2. Comparison of NO_x prediction models

	MAE	MSE	MAPE
SVR	26.701	1101.482	20.651
LR	18.315	516.63	16.691
ANN	6.265	14.384	5.238
PSO-XGBoost	3.732	53.713	4.319

Tab.3. Comparison of efficiency prediction models

	MAE	MSE	MAPE
SVR	0.216	0.085	0.233

LR	0.131	0.038	0.154
ANN	0.124	0.027	0.137
PSO-XGBoost	0.107	0.011	0.116

It can be seen that the performance of PSO-XGBoost algorithm is better than that of SVM, LR and ANN.

3 Boiler's combustion optimization based on comprehensive benefits

3.1 Boiler's comprehensive benefits

The essence of combustion optimization that integrates economic and environmental benefits of boiler is to improve combustion efficiency of boiler and reduce NO_x emission at the same time. NO_x emission and thermal efficiency have positive correlation, as shown in the figure below.

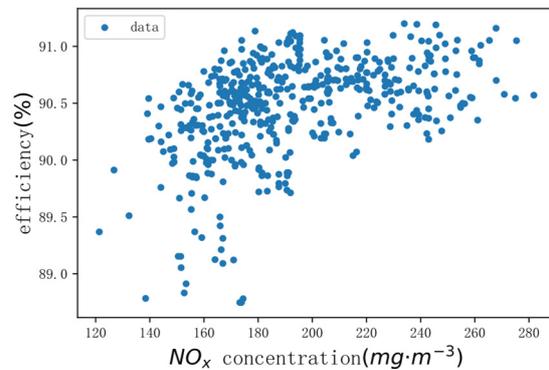


Fig.4. Relation between NO_x and efficiency

As can be seen from Figure 4, boiler's NO_x emission and combustion efficiency are positively correlated, so it is difficult to find an optimal scheme for both. When the NO_x emission decreases to a certain value, the combustion efficiency will decrease rapidly; when the thermal efficiency increases to a certain value, the NO_x emission will increase rapidly.

In paper [15], a linear weighting method is adopted to unify the boiler thermal efficiency and NO_x emission as the following optimization objectives.

$$f = \alpha \times \eta + (\alpha - 1) \times C_{NO_x} \quad (5)$$

Where: η - the normalized combustion efficiency; C_{NO_x} - the normalized NO_x emission concentration; α - the weight. By changing the value of α , the weight of combustion efficiency and NO_x emission can be adjusted. However, in this method, the weight α can only be set by experience, and it is difficult to adjust the parameters.

Therefore, this paper defines a nonlinear weighted comprehensive benefit evaluation function as follows.

$$f = \eta - (e^{C_{NO_x}} - 1) \times C_{NO_x} \quad (6)$$

Compared with the linear weighting method, the objective function sets exponential weight for NO_x emissions. When the NO_x emissions concentration is low, the objective function pays more attention to the improvement of combustion efficiency. When NO_x emissions increases, the weight of NO_x emissions

increases, and the objective function will pay more attention to NO_x emissions.

3.2 Optimization experiment of boiler's comprehensive benefits

PSO is used to optimize the boiler's operation parameters, and the comprehensive benefits evaluation function is set as the fitness function. The linear weighted evaluation function was added to the control experiment, and the values of α were set to 0.1, 0.9 and 0.5, respectively. The experimental results are shown in Table 4.

Tab.4. Boiler optimization results

	NO _x emission / mg·m ⁻³	Combustion efficiency /%
$\alpha = 0.1$	129.73	90.72
$\alpha = 0.9$	282.007	91.373
$\alpha = 0.5$	135.168	90.856
Nonlinear weighting	197.443	91.299

It can be seen from table 4 that NO_x emission and boiler thermal efficiency are positively correlated in the process of parameter adjustment, and they can not reach the optimum at the same time. It can also be seen that when the comprehensive benefit function is used for optimization, the boiler's combustion efficiency can be optimized and the NO_x emission can also be kept at a low value.

Add the above experimental results to figure 2 to get the following figure.

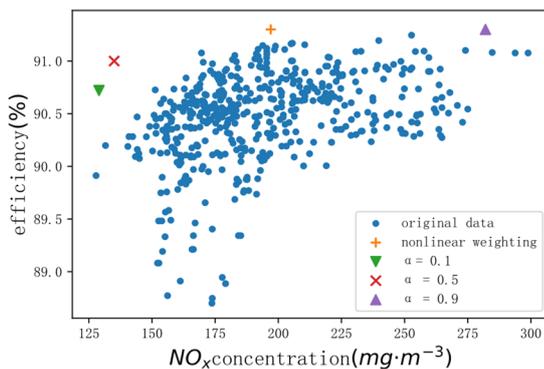


Fig. 5. Optimization results

It can be seen from Figure 5 that the four groups of optimization results in the experiment fall above the original data, which indicates that the combustion of boiler can be improved by PSO; while the optimization results using the comprehensive benefits function as the evaluation function fall into a more reasonable optimization range where efficiency and NO_x emissions are linearly correlated.

4 Conclusion

Based on the operation data of a 300 MW Circulating fluidized bed, PSO-XGBoost algorithm is proposed to predict the combustion efficiency and NO_x emissions of the boiler. Then a nonlinear weighting comprehensive benefit evaluation function is proposed for the multi-objective optimization of boiler's combustion efficiency and NO_x emissions. Based on the established prediction model, a boiler parameters optimization experiment was carried out. The results show that:

(1) The performance of PSO-XGBoost algorithm is better than that of SVM, LR and ANN.

(2) Compared with the linear weighting method, the nonlinear comprehensive benefit evaluation function in this paper can improve the combustion efficiency and avoid too high NO_x emissions at the same time without parameters adjustment

Acknowledgement

The authors wish to acknowledge the Science and technology project of Jiangsu market supervision and Administration Bureau (NO. KJ196045)

References

1. S. Chang, J. Zhuo, S. Meng, et al. Clean Coal Technologies in China: Current Status and Future Perspectives[J]. Engineering, 2(4), 447-459 (2016)
2. L. Zheng, H. Zhou, K. Cen, et al. A comparative study of optimization algorithms for low NO_x combustion modification at a coal-fired utility boiler[J]. Expert Systems with Applications, 36(2), 2780-2793 (2008)
3. Y. Shen, X. Gao, M. Zhang, Prediction of NO_x emission in boiler based on on-line operation data [J]. Electric Power, 44 (12), 64-67 (2011)
4. W. Xu, C. Taihua, The balanced model and optimization of NO_x emission and boiler efficiency at a coal-fired utility boiler[C]// IEEE Conference Anthology. IEEE (2013)
5. W. Zheng, C. Wang, Y. Yang, et al. Multi-objective combustion optimization based on data-driven hybrid strategy[J]. Energy, 191, 116478 (2019)
6. Y. Shi, W. Zhong, X. Chen, et al. Combustion optimization of ultra supercritical boiler based on artificial intelligence[J]. Energy, 170(MAR. 1), 804-817 (2019)
7. Y. Lu, X. Peng, K. Zhao. Hybrid Modeling Optimization of Thermal Efficiency and NO_x Emission of Utility Boiler[J]. Proceedings of the CSEE, 31 (26), 16-22 (2011)
8. X. Huang, H. Wang, Modelling of Combustion Process Characteristics of a 600 MW Boiler by T-S Fuzzy Neural Networks[C]// International Conference on Energy & Environment Technology. IEEE, 2009.
9. E. A. Bayrak, P. Kirci and T. Ensari, Comparison of Machine Learning Methods for Breast Cancer

- Diagnosis[C]// Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), 1-3 (2019)
10. P. R. Hegde, M. M. Shenoy and B. H. Shekar, Comparison of Machine Learning Algorithms for Skin Disease Classification Using Color and Texture Features [C]// International Conference on Advances in Computing, Communications and Informatics (ICACCI), 1825-1828 (2018)
 11. T. Chen, C. Guestrin, XGBoost: A Scalable Tree Boosting System[J]. 2016.
 12. Y. Zhang, Y. Ding, H. Zhang, Z. Wu, Study on Optimized Boiler Operation Control for Harmonized Coordination Between Economic Efficiency and Environmental Protection[J]. Journal of Chinese Society of Power Engineering, 5, 676-679 + 706 (2005)
 13. C. Xu, J. Lu, Y. Zheng, X. Feng, A Boiler Combustion Global Optimization on Efficiency and Low NOx Emissions Object[J]. Proceedings of the CSEE, 4, 46-50 (2006)
 14. P. Duan, K. Xie, T. Guo, et al. Short-Term Load Forecasting for Electric Power Systems Using the PSO-SVR and FCM Clustering Techniques[J]. Energies, 4(1), 173-184 (2011)
 15. X. Xu, Q. Chen, M. Ren, et al. Combustion Optimization for Coal Fired Power Plant Boilers Based on Improved Distributed ELM and Distributed PSO[J]. Energies, 12, 6(2019)