Modelling of thermal power unit target value based on hierarchical regression

Zhuang Shao, Fengqi Si * and Zhigao Xu

Key Laboratory of Energy Thermal Conversion and Control of Ministry of Education, School of Energy & Environment, Southeast University, Nanjing, China

*Corresponding author e-mail: fqsi@seu.edu.cn

Abstract. The operation optimization of thermal power units is of great significance for energy saving and consumption reduction. Under certain boundary constraints, determining the target value of adjustable parameters is the key to optimal operation of the unit. The basic method is to analyse the historical operation data, from where select the data with the best performance indicators in all range of working condition as the target sample, to establish the model between the boundary parameters and the adjustable parameters. The most difficult step is to solve the conditional probability distribution of the performance parameters with respect to the boundary parameters. In this paper, a hierarchical regression method is proposed to slice the raw data about the performance index, and the historical data is conditionally segmented. Based on this, a modelling framework is designed for the target value of the adjustable parameters of the thermal power unit. Finally, the operation data of an ultra-supercritical 1000MW thermal power unit is selected for verification. The results show that the method has the advantages of stable, and can provide the adjustment direction for the optimal operation of thermal power units.

1. Introduction

In recent years, in response to the pressure of pollution control, the requirements for energy-saving and emission reduction of thermal power units have been further improved. As for the energy-saving renovation is basically completed, refined operation has become the most urgent work [1]. At present, the adjustment of operating parameters is based on the operating personnel's experience or the set values of procedures. The set values make the unit stable and safe in operation, but is difficult in keeping state optimal. Due to the variation of the type of thermal power units, the condition of different power plants varies a lot. With many parameters to be optimized, the operation optimization needs to be customized according to the conditions of each factory. This is undoubtedly a complicated project.

The research of thermal power unit operation optimization is getting attention. Based on the mechanism of thermal system, many scholars use the design data or experiments data to carry out modelling research. In [2], the minimum coal consumption is taken as the objective function, and three sub-systems are optimized to determine the target values of the key parameters in each system. In [3], the target performance value and various losses are determined by simulation. The above methods have simplified assumptions about mathematical models, which have limitations in complex boundary conditions.

In recent years, with the accumulation of data, scholars do research on power plant operation optimization based on data mining [4,5]. In [4], multi-feature parameters are clustered by K-means algorithm and their reference values are determined. The key of above research method is how to determine the target value of efficiency index. It has certain reference for the operational personnel, but how to improve the current efficiency of the unit is still the problem.

To solve the above problems, this paper proposes a hierarchical regression method, sliced all data samples according to energy efficiency in full range of conditions for mining high level operational data. Based on the algorithm, a target value modelling framework for adjustable parameters of thermal power units is proposed. Finally, the framework is applied to establish the target value model of an ultra-supercritical 1000MW thermal power unit.

2. Methodology

Data driven target value model refers to the mapping from boundary parameters to target values. The training sample of boundary parameters - target value pairs comes from the historical data of better operational level. These historical data is taken from the data slice with better performance indicators obtained by the hierarchical regression. This section introduces hierarchical regression and target value modelling framework.

2.1. Hierarchical regression

The operating status of equipment are different even under the same boundary conditions. In the historical data, due to the different values or combinations of adjustable parameters, the performance have a certain range of variation, showing a specific data distribution pattern with statistical characteristics.

To analyze the variation of performance index with boundary parameters in the whole working condition, is to solve the conditional probability density function of performance parameters with respect to boundary parameters, as shown in Eq. 1.

$$E(P|B=b)=\int_{P} p f_{P|B}(p|b)dp \qquad (1)$$

Where B is the boundary parameters, P is the performance parameters. $f_{P|B}(p|b)=\frac{f_{P,B}(p,b)}{f_B(b)}$, where $f_B(b)$ gives the density of B.

There are two main methods for solving the probability density function [6]. One is assuming the probability density function, to solve the statistical parameters directly. The typical method is maximum likelihood estimation. The other is to solve the data distribution directly and establish a mapping relationship between the known distribution and the target distribution. The typical method is generating model.

Particularity of modelling problems with target values, the data hierarchical regression method is designed, which only solve the conditional expectation of data distribution [8], as shown in Equ. 2, avoiding the complexity of solving the conditional probability density function.

$$\hat{\beta}=\arg\min_{\beta}(|p-\hat{p}|^2) \qquad (2)$$

Where $\hat{p}=f(b,\beta)$ is regression model, b is boundary parameter, p is performance parameter, $\hat{p}$ is an estimate of the expected performance parameter, β is the parameter of regression model, $\hat{\beta}$ is an estimate of the model parameters.

After first regression of all samples, the raw data is divided into upper and lower slices along the performance parameter direction. Then use the upper and lower slices of the data to establish two other regression models respectively, to divide the data into more slices [7]. After n rounds of hierarchical regression, the data is divided into 2^n slices, as shown in Fig. 1, where $n=2$.

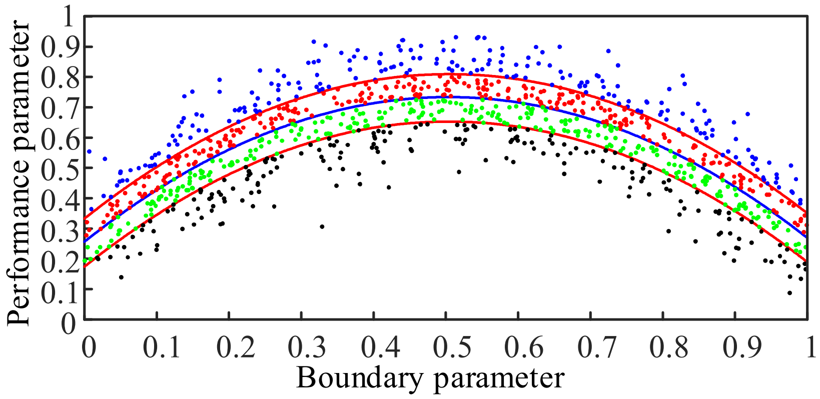

Figure 1. Divided raw data by hierarchical regression for 2 rounds.

2.2. *Target value modelling framework*

Based on the hierarchical regression, the samples can be divided into multiple groups along the direction of performance parameters, which is convenient for obtaining sample slices with higher operation level beyond other samples.

Extracting the adjustable parameters and boundary parameters from these data as training samples, a target value regression model can be established. The specific modelling framework process is shown in Fig. 2.

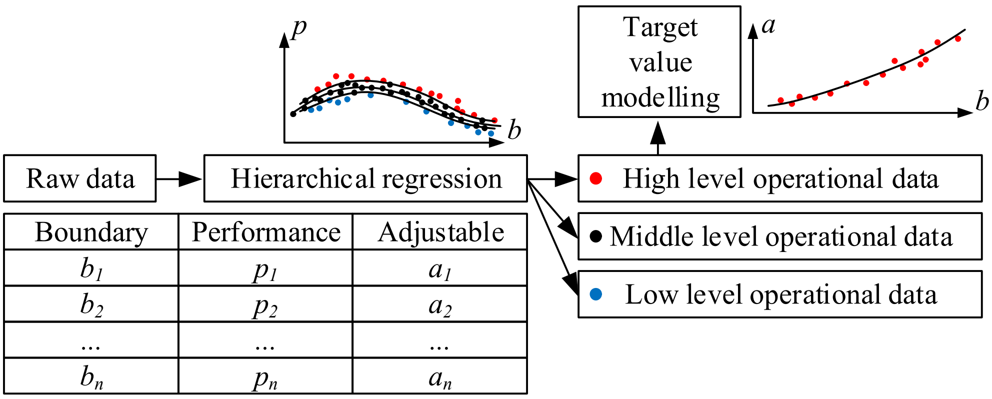

Figure 2. The flow chart of target value modelling framework.

Noted that in the hierarchical regression step, all of the raw data is used, where taking the boundary parameters as features and the performance parameters as labels. In the target value modelling step, only the high level operational data is used, where taking the boundary parameters as features and the adjustable parameters as labels.

3. Case Study

In order to demonstrate the effectiveness of the hierarchical regression and target value modelling framework proposed in this paper, two cases are designed in this section.

Case 1 is the hierarchical regression for different standard distribution of boundary parameters and performance parameters, respectively. Case 2 is a modelling example of the oxygen content target value of an ultra-supercritical 1000MW thermal power unit.

3.1. Standard distribution hierarchical regression

For steady thermal system, performance and boundary have a fixed relationship. However, since the distribution of boundary parameters and adjustable parameters are determined externally, the distribution of samples varies in different periods. The hierarchical regression proposed has good effect on different distributions, which has a good stability in applications.

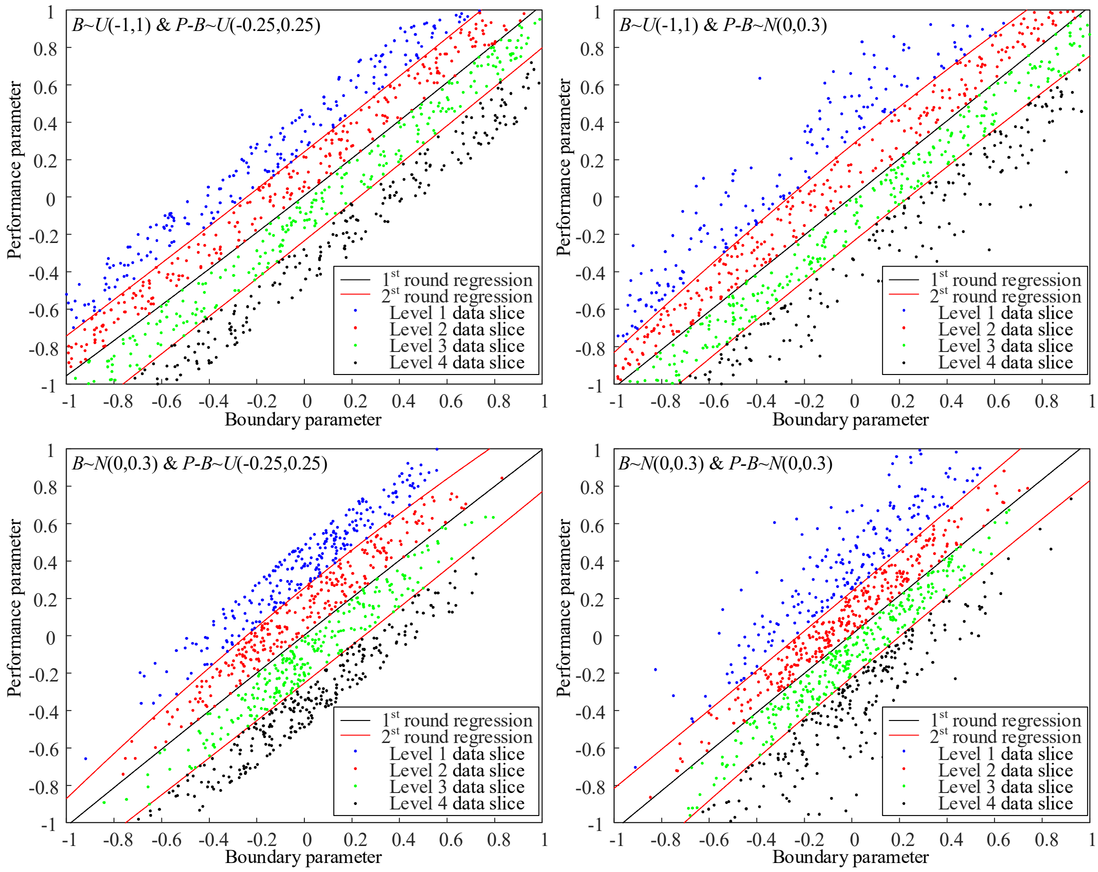

Figure 3. The results of the hierarchical regression for different distribution combinations of boundary parameter and performance parameter.

Assuming the performance and boundary satisfy the uniform distribution and the normal distribution respectively, the results of the hierarchical regression are shown in Fig. 3. It can be seen that the sample distribution has little effect on hierarchical regression.

3.2. Oxygen content target value modelling example

Collect historical data (dynamic process data is rejected) from an ultra-supercritical 1000MW thermal power unit, taking boiler efficiency (positive balance) as performance parameter, load as boundary parameter, and oxygen content (economizer outlet) as adjustable parameter [9,10].

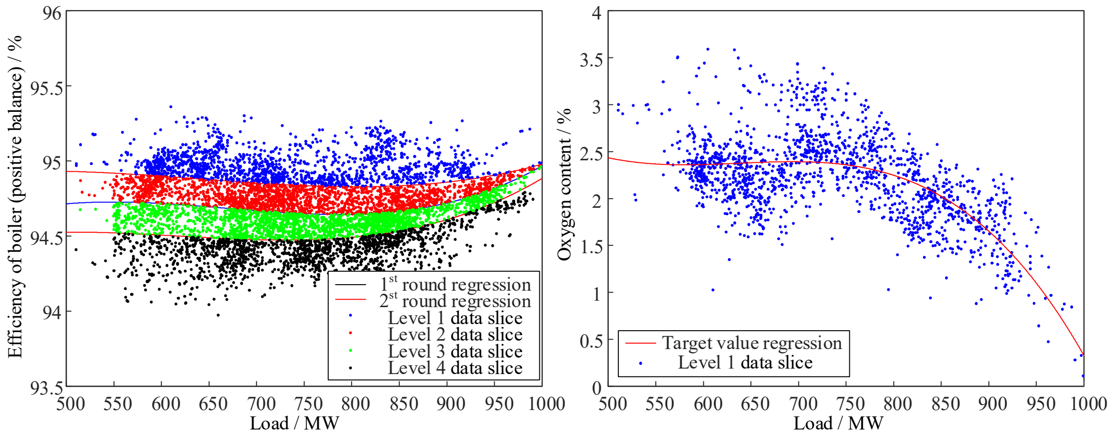

Figure 4. Oxygen content target value modelling.

Divide the raw data into 4 data slices along the boiler efficiency with hierarchical regression. Using the level 1 data slice (maximum boiler efficiency), a mapping from load to oxygen content is established, as shown in Fig. 4, which is the target value model of oxygen content with respect to load.

4. Conclusion

Considering that the distribution of performance parameters in the steady state historical data of thermal power equipment under the same boundary conditions depends on the changes of adjustable parameter, hierarchical regression method is proposed to slice data along the direction of thermal power unit performance indicators into data sets with different operating levels. Based on the hierarchical regression, a thermal power unit target value modelling framework is proposed.

The hierarchical regression method is stable to the variance of the distribution of samples. The target value modelling framework is efficient and easy to implement. The cases prove that the method proposed in this paper has high practical value.

References

[1] Xiong J, Zhao H, Zhang C, et al. Thermoeconomic operation optimization of a coal-fired power plant[J]. Energy, 2012, 42(1):486-496.
[2] Hong-Hua H U, Huang T H, Wei-Guo A I. Study and determination of operation optimization target values of large size fossil-fired power unit[J]. Electric Power, 2004.
[3] Yang Y. Theory of Energy Conservation Potential Diagnosis for a Coal fired Unit and its Application[J]. Proceedings of the Csee, 1998.
[4] Xu J, Gu Y, Wang Z, et al. Research on Indexes of Energy Efficiency and Its Reference-value for Coal-fired Power Units Based on Data-mining[J]. Proceedings of the Csee, 2017, 37(7):2009-2015.
[5] Jianqiang L I, Chen X, Zhao K. Target Value Prediction of Condenser Vacuum Degree Based on PSO-SVR Model[J]. Electric Power Science & Engineering, 2017.
[6] Klenke A. Conditional Expectations[M]// Probability Theory. Springer London, 2014:47-78.
[7] He H, Starzyk J A. Online Dynamic Value System for Machine Learning[C]// International Symposium on Neural Networks: Advances in Neural Networks. Springer-Verlag, 2007:441-448.
[8] Edition S. Applied Linear Regression Models[J]. Technometrics, 2016, 39(3):342-342.
[9] Shen F P, Zhou K Y, Jian-Qun X U, et al. Analysis and Comparison on Calculation Pattern of Boiler Efficiency[J]. Boiler Technology, 2004.
[10] Jun-Jie G U, Kong D Q, Gao D M, et al. Calculation of the optimal set value of flue gas oxygen content for the optimization of combustion in power plant boiler[J]. Journal of North China Electric Power University, 2007.