**Preprints of the 21st IFAC World Congress (Virtual)**
**Berlin, Germany, July 12-17, 2020**

# Process operation optimization using system identification

**Chao Yang*. Yucai Zhu****

** College of Control Science and Engineering, Zhejiang University, Hangzhou, 310027 China (e-mail: ychao@zju.edu.cn)*
** * College of Control Science and Engineering, Zhejiang University, Hangzhou, 310027 China (e-mail: zhuyucai@zju.edu.cn).*

Abstract: Process optimization is an important topic in process industry, most process industry optimization works are based on mechanism models or performance test methods. However, it is very difficult to carry out optimization in actual operation because of the difficulty in obtaining the mechanism model, the difficulty in on-line measurement of objective function and the high test cost. In order to solve the problem, an online optimization method based on system identification is proposed. By replacing the unmeasurable variable with the measurable variable, the process model is identified on-line, and the gain of identified model is used as the optimization gradient to find the optimal variable value on-line. The method is verified using both simulation and real plant data.

*Keywords:* System identification; Process industry; On line optimization; Air volume optimization

## 1. INTRODUCTION

Process industry generally refers to the processing of raw materials such as petrochemicals, chemicals, and the energy industry. It is characterized by the continuous processing of raw materials for production, generally accompanied by chemical, physical, and phase changes. Process industry is dominant in the national economy, and its development status directly affects the country's economic foundation.

Process operation optimization has attracted many researchers in academia and industry to carry out research, and many academic and application results have been produced, which can be summarized into two types: one is the process optimization method based on mechanism model, the other is the process optimization method based on performance test.

The key point of optimization based on mechanism model lies in the modelling and solution problem of the mechanism model(Biegler, Grossmann, 2004).

Due to the complex nonlinear, time-varying, and coupling characteristics of the process industry, an important research direction of optimization based on mechanism is to obtain optimization propositions: in most researches, the optimization model is obtained through the material balance equation, energy conservation equation, working medium balance equation and so on; the constraints of the variables are given by simplifying the actual constraints; and a certain performance index(Jin, et al, 2015; Cong, et al, 2017) (such as benefit, minimum consumption) is optimized in combination with relevant methods of process characteristic evaluation (such as laws of thermodynamics). Modelling by process mechanism is costly (Wang, 2017), and in practical applications, the model needs to be greatly simplified. Because the real processes are mostly dynamic, the deviations between the model and the actual system will result in poor optimization results.

Another research direction of optimization based on mechanism model is how to solve the optimization proposition. In the past few decades, the solution of some typical optimization problems has been a research hotspot (Biegler, Grossmann, 1985; Padberg, 2013): such as linear programming (LP) and nonlinear programming (NLP) problems, and mixed integer linear programming (MILP) and mixed integer nonlinear programming (MINLP) problems. The solution difficulty of optimization propositions varies greatly according to the complexity of the problem. For a complex process industrial system, the computing time often increases exponentially with the problem size. Similar problem often restricts the application of optimization based on mechanism models in practice.

Another important method of process optimization is optimization based on performance test. Compared with the optimization based on the mechanism model, the optimization based on the performance test is more practical(Jiao, Guo, 2007) and avoids the difficulty in modelling and solving of the mechanism model. Most industrial systems are optimized by performance test before they are put into operation.

Although the optimization based on performance test method is more practical, performance test is time consuming because of the multiple operating conditions, the complex scenarios of the actual system and the fact that many variables related to the optimization objective cannot be measured online. In addition, due to the large interference of the test to the regular production, the test usually needs to be operated off-line, and regular production cannot be carried out, which makes the test costly. Taking the combustion optimization test of industrial boiler as an example: it is usually necessary to monitor dozens of variables at the same time to calculate the boiler efficiency, key variables such as carbon content of fly ash and composition of coal cannot even be measured directly, which makes the test time consuming and costly. Therefore, the performance test method is usually carried out only after the initial or retrofit of the system, on-line optimization cannot be realized.

To sum up, the main problems of the existing process optimization methods are that the optimization cost is high, the methods are complex, and the optimization cannot be carried out online.

Aiming at the above problems, a process optimization method based on system identification is proposed, which can guarantee the online application in the system operation and effectively reduce the optimization cost.

## 2. OPTIMIZATION USING SYSTEM IDENTIFICATION METHOD

The process optimization can be considered as a mathematical optimization problem. In a mathematical optimization problem, an objective (loss) function is maximized (minimized) by adjusting the related optimizing variables (Forst, Hoffmann, 2010). The objective function for the process optimization is normally the efficiency or some other benefit indexes, which should be maximized; equivalently, the loss function is the total energy loss or some other punishment indexes, which should be minimized.

$$\begin{aligned} &Min\ \Phi(x) \\ &s.t.\ g(x) \leq 0 \quad (1) \\ &c(x) = 0 \end{aligned}$$

In order to measure the objective function, very dedicated test needs to be conducted with strict test conditions and very long test time, for example, the thermal performance test for power plant always last for several weeks. It is difficult to use this method during normal operation, because it is not possible to calculate the objective function accurately as some related variables cannot be measured on-line.

One may ask if it is possible to perform process optimization without using objective functions that cannot be measured online and without using time consuming test.

Take a step back now, the ultimate goal of process optimization is to maximize the efficiency, if the unmeasurable variable can be replaced by some on-line measurable variables, the optimization will be available online. For example, if the goal is to optimize boiler efficiency in a power plant, although boiler efficiency cannot be measured online, but when the turbine is in closed loop operation, the absolute internal efficiency of the turbine is stable, then generated electrical power for a given amount of coal flow is directly related to boiler efficiency, therefore, one can use the real power output of the unit as the objective function to perform online optimization.

### *2.1 The optimization procedure*

As the objective function can be measure on line, an optimization procedure based on multivariable system identification is developed.

Denote $y(t)$ as the objective function at time $t$, $u(t)$ is the input vector of independent variables that affect the objective function $y(t)$. The input vector $u(t)$ contains variables can be determined according to the field process experience.

The relationship between input vector $u(t)$ and the output $y(t)$ can be described by a set of nonlinear difference equations:

$$y(t+1) = f(y(t), u(t) \quad (2)$$

Where function $f$ is unknown and assumed differentiable in its arguments. Denote $u_i(t)$ as an optimization variable in vector $u(t)$. Then, optimization using the well-known hill-climbing is as follows.

### *2.2 Hill-climbing optimization*

(1) If gradient $\partial y/\partial u_i > 0$, then increase $u_i(t)$ by a certain amount;

(2) If gradient $\partial y/\partial u_i < 0$, then decrease $u_i(t)$ by a certain amount;

(3) If gradient $\partial y/\partial u_i = 0$, then the objective function $y(t)$ is at its (local) maximal, keep $u_i(t)$ unchanged.

At a stationary point of $y^*$, the gradients of the unknown function $f$ to all the input variables, including the optimization variables, are the steady state gains of the linear model at $y^*$. These gains can be obtained by identifying a linear dynamic model with vector $u(t)$ as the inputs and vector $y(t)$ as the output.

Applying the hill-climbing optimization to the process and doing it for all optimization variables, one obtains the following procedure.

Step 0. Define the optimization model.

The model output is the objective function, the model inputs are the operation variables and the optimization variables.

Step 1. Model identification.

Collect normal operation data at different working conditions, and perform model identification (Zhu, 2001) for the optimization model at different working conditions. If some of the input variables are not sufficiently exciting, add test signals on them. Compared with the traditional performance test method, the proposed method does not require strict test conditions, such as maintaining a certain working condition for a certain time, which enables the optimization to be carried out online.

Step 2. Adjusting optimization variables by hill-climbing.

Increase the optimization variable if its gain is positive; decrease it when negative. The adjustment size can be determined from process knowledge and operation experience.

Step 3. Optimization completed

If the gain of the optimization variable is not zero, go to Step 1; if the gain is zero, stop adjusting the optimization variable. Fig.1. shows a flow diagram of the optimization procedure.

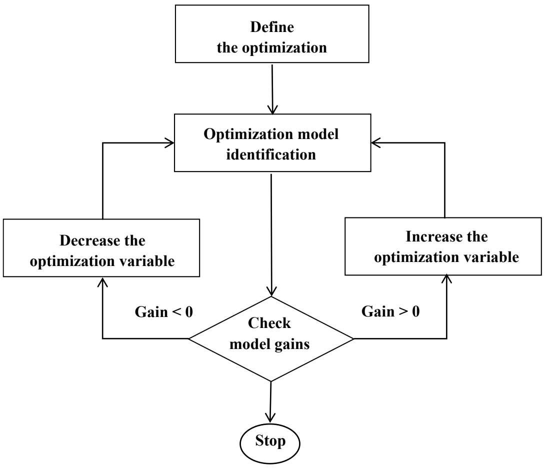


Fig.1. Flow diagram of identification-based optimization

Comments:

The optimization procedure is multi-variable meaning that it can be used to adjust several optimization variables simultaneously.

The advantage of the optimization approach is its low cost. The traditional process optimization is based on the performance test lasts several weeks or even months; it disturbs normal operation and costs much manpower. Optimization based on mechanism model is costly and less accurate. The new optimization approach only needs several days of operation data per iteration and its cost is only a fraction of that of the other method. It can be used when necessary.

However, the identification-based optimization is not as rigorous as the performance test. Therefore, one need to use the information provided from field test in identification-based approach, for example, in determining the feasible ranges of the optimization variables.

There is no need to perform process optimization daily. It is only necessary after long time of operation, or when there is raw material type change or equipment maintenance. Although the parameters of the raw material may change randomly, such as the calorific value of coal, as the parameters do not change widely, the disturbance can be regarded as stationary random disturbance, it does not have much effect on the proposed method. However, if the parameters of the raw material change a lot, such as coal type changes, then the proposed optimization method should be performed again.

What needs to be emphasized is that when the operation conditions or raw material parameters change a lot, the change can be found by system identification method which can be regarded as a system identification based diagnostic method, for example, add a little excitation signal to the optimization variable, if the gain deviation is found to be significant, then the optimization should be performed. The diagnostic method is not the key point of this article, but it is feasible.

In fact, the identification-based optimization can be used for all process units in all industries where the objective function can be measured online or can be calculated from online measured signals.

## 3. CASE STUDY

### *3.1 Case description*

Coal-fired power generation is an important part of power production, in which the quality of boiler operation is directly related to the unit economy.

The heat loss is directly related to air volume, the qualitative analysis is shown in Fig.2, where: $q_2$ is flue gas heat loss, $q_3$ is non-full burning loss, $q_4$ is solid incomplete combustion heat loss, they are the main components of boiler heat loss. Generally speaking, the coal rate of boiler operation is based on the power demand of the unit; the air volume required for fuel combustion is determined by multiplying the amount of fuel by the pre-designed air-fuel ratio (AFR). Whether the AFR is reasonable or not should be judged by the oxygen content in the flue gas.

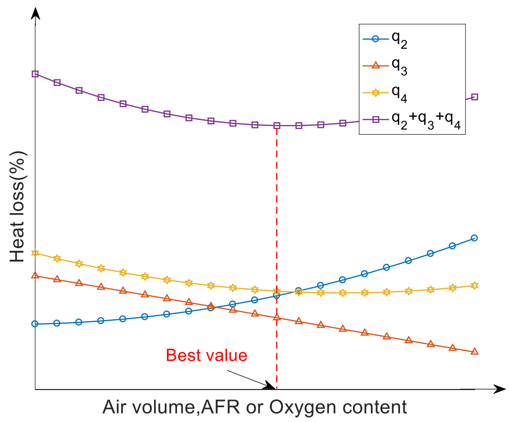


Fig. 2. Boiler combustion characteristic curve

In order to maintain the efficient operation of the boiler, the coal-fired units need to determine an optimal air volume, some literatures also call it optimal oxygen content, optimal AFR, because adjusting oxygen or AFR is the main way to adjust air volume. Usually, engineers use a mechanism-based model (Zhang, 2015) or an on-line test to determine the optimal air volume. However, it is very difficult to obtain the mechanism model of boiler (Williams, 2018), which includes complicated mass transfer and heat transfer process. On-line test optimization is difficult to be used in practical operation because of its high cost and interference to normal operation of boiler (China National Standardization Management Committee, 2015).

Air volume control logic is as shown in Fig.3. Generally speaking, the coal feeding rate is changed according to the boiler power output demand, during system operation, air volume setting value is determined by coal feed rate and AFR. However, when the working condition changes (which happens frequently), the excess air in the furnace cannot be kept stable and the economic combustion cannot be maintained. In order to keep economic combustion at any time, it is necessary to check the oxygen content in the boiler frequently, and adjust (correct) the air volume according to the oxygen content, therefore, the function of oxygen control loop is to correct the AFR value.

Therefore, to obtain the optimal air volume is to obtain the optimal setpoint of oxygen content and AFR. This case will be optimized by the proposed method

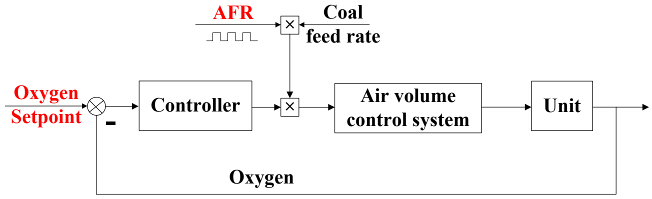


Fig.3. Boiler air volume Control Logic Diagram

.

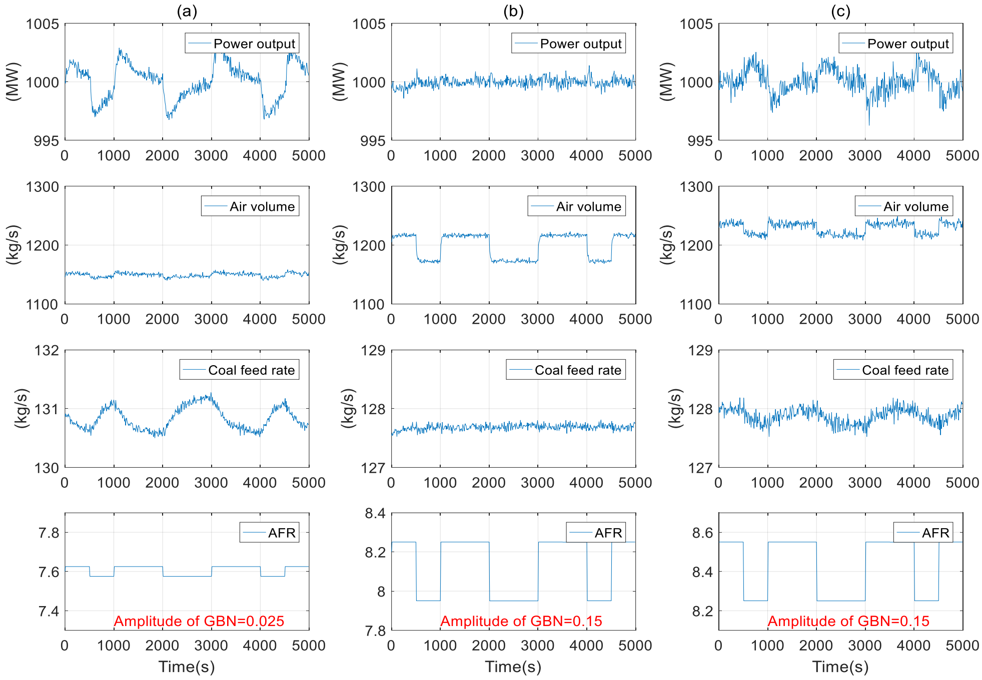


Fig.4 Cross validation of power output, air volume, coal feed rate at different AFR. (a) AFR=7.6; (b) AFR=8.1; (c) AFR=8.4.

*3.2 Simulation result*

Apros Process Simulation Software is used in carrying out the simulations for the study. Predictive error method (ARMAX model) is used to ensure the consistency of identified model in close loop.

The research object is the boiler in a 1000MW unit, and the main task is to optimize the air volume to maximize the boiler efficiency. Online Air volume optimization is performed with the following procedure.

Step 0. Define the optimization model.

Although boiler efficiency cannot be measured online, when the turbine is in closed loop operation, the absolute internal efficiency of the turbine is stable, then power output for a given amount of coal flow is directly related to boiler efficiency. So, power output is used as the objective function instead of boiler efficiency. Coal feed rate and air volume are used as operation variables.

Step 1. Model identification.

The pre-set AFR value is 7.6, by adding excitation signals to AFR setpoint, air volume changes, and identification experiments are carried out to identify the air volume-power output model, and the coal feed rate-power output model. The gain of air volume-power output model is used as gradient.

Step 2. Adjusting optimization variables by hill-climbing.

When the value of AFR is set to 7.6, the gain of air volume-power output model is 0.5, this means that each unit increase in air volume is associated with 0.5MW increase in power output. Obviously, increasing air volume can increase profits, so AFR is set to 7.7 (add 0.1 to the original value) then to repeat Step 1.

Step 3. Optimization completed

When ARF gradually increased to 8.1, the gain of air volume-power output model is about 0, and it is very difficult for the test signal to generate excitation to the system. If AFR continues to increase to 8.2, the gain of air volume-power output model is -0.1, this means that each unit increase in air volume is associated with -0.1MW decrease in power output.

The objective function reaches the inflection point when AFR value is 8.1 as the gain is 0. And the optimization is ended. The optimal set value of AFR is 8.1, and the corresponding oxygen content value (2.75%) is the optimal oxygen content value.

The proposed method only requires eight hours of operating data (AFR changes from 7.6 to 8.1), it is very time saving and less costly compared to traditional performance test method.

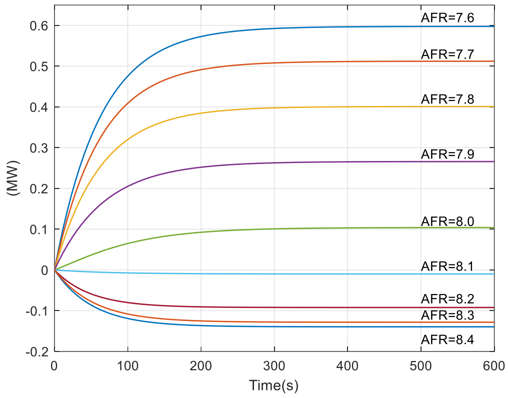


Fig.5. Air volume-Power output model step response

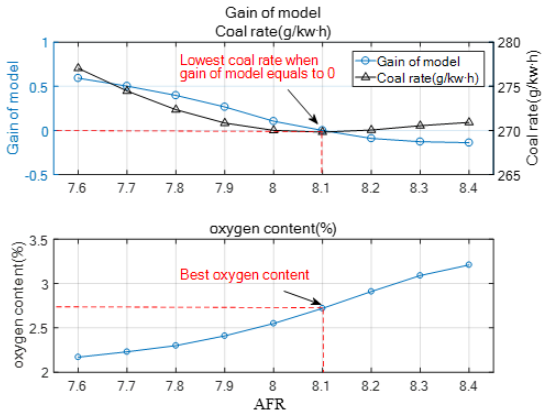


Fig.6. Comparison of experimental results

To better illustrate the changing trend of optimization function, the experimental range of AFR was expanded to 7.6-8.4. The step responses of the identified models obtained by identification experiments under different AFR values are shown in Fig.5. The variation trend of coal consumption with gain is shown in Fig.6.

When AFR varies from 7.6 to 8.4, the model gain changes from positive to negative. When the model gain is 0, the coal consumption of the unit is the lowest (270g/kw·h), and the corresponding oxygen content and AFR are the optimal values, and the corresponding boiler efficiency is the highest. It can be seen that when the system identification method is used for optimization, when the model gain reaches 0, it should be the optimal point, and it is also the goal of our optimization work.

The coal rate decreased by 8g/kw·h (from 278g/kw·h to 270g/kw·h) by optimization, suppose that the unit has an effective operating time of 5,000 hours per year, then the annual power generation is 5 billion kw·h, and 40,000 tons of coal will be saved.

### *3.3 Field test results*

The field test is done in the #1 Power Generation Unit at a power plant in Foshan, Guangdong Province, China.

Test signals with small amplitudes have been added at the setpoints of the PID control loops, which have not caused any problem for normal operation of the unit.

Here oxygen content and the coal feeder separator velocity will be optimized in order to improve unit efficiency. To this end a model is identified.

As discussed in Section 2, the model output is the objective function which is also the power output. In total 5 inputs are used in the model: (1) boiler main control which is a combination of coal feed, air flow and water flow; (2) turbine valve opening; (3) water flow for middle temperature; (4) oxygen content; (5) burner angle.

The optimization model has been identified using three days of normal operation data and two set of models are obtained: high power output model and low power output model; see Fig.9 and Fig.10 for the result. The model gains from oxygen content to real power output are negative at both high-power output and low-power output. This implies that oxygen content is set too high for the whole power output range. Based on the model and the experience of the authors, it is recommended an oxygen content reduction of 0.3~0.4%.

Based on the modelling result, the oxygen content setpoint has been reduced by 0.3% for the whole power output range and tested. Testing result has shown that oxygen content reduction resulted in 10% NOx reduction and 0.3% saving of coal. No more reduction is permitted from the operation staff and, therefore, second iteration of optimization was not carried out.

The coal feeder separator velocity influences the fineness of the coal powder which can be optimized. Currently this velocity is set constant. Some small test signal is used to identify models that includes coal feeder separator velocity as the 6th input of the above model. The model from the separator velocity to the real power output is very poor according to model validation, which is due to too small amplitude of the test signal. No conclusion can be draw from the poor model and no adjustment of the separator velocity was made.

Other variables, such as primary air pressure and the distribution of damper opening, can also be optimized, but are not included due to limited project scope.

## 4. CONCLUSIONS

This paper realizes on-line optimization of process industry by system identification method. This method does not need to use mechanism model, does not need to calculate unmeasured objective function, and has little interference to the normal operation. The experimental results show that the model obtained by identification method can accurately determine the optimal direction of operation variables and finally obtain the optimal point. The same method can also be used for more complex multi-variable and multi-objective optimization problems, which will be validated by selecting CCHP unit operation optimization as an example in the following work.

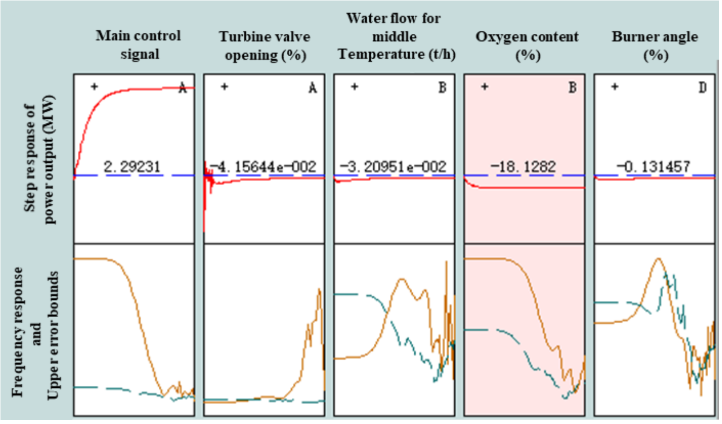


Fig.9. High power output work condition optimization model. Step responses (upper row) and frequency responses and upper error bounds (lower row) of the model where input 4 is oxygen content. The oxygen content model has a gain of -18.13

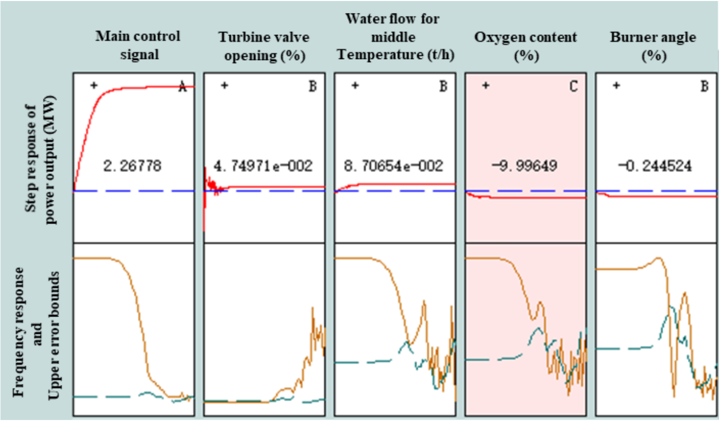


Fig.10. High power output work condition optimization model. Step responses (upper row) and frequency responses and upper error bounds (lower row) of the model where input 4 is oxygen content. The oxygen content model has a gain of -10.00

## REFERENCE

Biegler, LT, Grossmann, IE, 1985. Strategies for the optimization of chemical processes, *Reviews in Chemical Engineering*, 3(1), pp.1-48.

Biegler, LT, Grossmann, IE, 2004. Challenges and research issues for product and process design optimization. In *Proc. 6th Intl. Conf. on Foundations of Computer-Aided Process Design (FOCAPD), Austin, TX, CACHE Corp*, pp.99-117.

China National Standardization Management Committee, 2015. *People's Republic of China national standard of Test procedures of power plant boiler performance*. GB10184-2015.

Cong, W, Kilkis, S, Johan, T, et al, 2017. Multi-objective Optimization and Parametric Analysis of Energy System Designs for the Albano University Campus in Stockholm, *Procedia Engineering*, 180, pp.621 - 630.

Forst, W, Hoffmann, D, 2010. *Optimization-theory and practice*. Springer, New York.

Jiao, M, Guo, B, 2007. Comparison and analysis on coal consumption rate of fossil-fired power unit in thermal test and actual operation [J], *Electric Power*, 40(11), pp.82-84.

Jin, KM, Shik, KY, Hee, CM, et al, 2015. Multi-objective optimization design for a hybrid energy system using the genetic algorithm, *Energies*, 8(4), pp.2924-2949.

Padberg, M, 2013. *Linear optimization and extensions*. Springer, New York.

Wang, Z, 2017. Modeling of the Heating Network for Multi-district Integrated Energy System and Its Operation Optimization, *Proceedings of the CSEE*, 37(5), pp.1305-1315.

Williams, FA, 2018. *Combustion theory*. CRC Press, Florida.

Zhang, GQ, 2015. *Study on Optimum Excess Air Coefficient for Power Plant Boilers*, pp.1374-1378.

Zhu, Y, 2001. *Multivariable system identification for process control*. Elsevier Science Ltd, Oxford.