Engine Cylinder Head. Journal of Materials Engineering, 2012, 6, pp.16-20.
10. J. Yi, W. Xu, L. Wenchuan, D. Lei. Hybrid lot sizing production planning of manufacturing and remanufacturing to automobile engine. Computer Integrated Manufacturing Systems, 2013, 19, pp.774-781.

Research of Utility Boiler's NOx Combustion Optimization Based on Relevance Vector Machine

Yunfei Ma[1], Xiaofei Ma[2], Peifeng Niu[1]

1.Department of mechanical engineering, Yanshan University, Qinhuangdao, Hebei 066004, China
2.Department of Physical Education, Northeastern University, Qinhuangdao, Hebei 066004, China

Corresponding author is Xiaofei Ma

Abstract
In order to reduce NOx emissions from utility boilers, we draw a new machine learning method Relevance Vector Machine into the modeling of a 300MW pulverized coal boiler's NOx output and twenty-six inputs as drum secondary air, oxygen and so on, then we use Gravitational Search Algorithm to optimize the parameters of the model to obtain the optimal pattern, also we make comparisons of the outcome of Particle Swarm Optimization's and Genetic Algorithm's optimizing Relevance Vector Machine and Gravitational Search Algorithm's optimizing Support Vector Machine . Last we make the boiler adjustable variable input parameters as the optimization variables for the target of cutting down NOx emissions to achieve the appropriate input parameters of lower NOx emissions. The result shows: Gravitational Search Algorithm's optimizing Relevance Vector Machine gets better accuracy than the others, the model is well performed in the optimization of NOx emissions.
Keywords: NOX, OPTIMIZATION CONTROL, BOILER

1. Introduction

The nitrogen oxides are both a danger to human health and a destruction of toxic pollutants to the atmosphere. They are generated in the combustion process when the nitrogen compounds in the coal and oxygen in the air burn at high temperatures. Therefore, high NOx emissions of boilers is a problem worthy of attention in power plant, but it is difficult to establish accurate models of boiler because the boiler system is complex, coal and boiler operating conditions is also changing. So, it is difficult to achieve the best condition in the actual operation of the boiler system artificially adjusted [1].

The support vector machine model was introduced into the NOx emission projections of power plant in paper [1],in this paper ,a model was is established under the data of twelve different combustion conditions and analysis the influence on the prediction result when the support vector machine had different parameters .But the method didn't obtain the optimal parameters of the model which lack an optimization algorithm for parameter optimization. The gravitational algorithm is introduced into the model of SVM, the parameters were optimized to obtain the optimal values of the model parameters, but the support vector machine modeling method in paper [1] and [2] had some limitations, if the training data are more complicated, the effect of predicting is not good enough, but they didn't optimize NOx emissions of combustion, they just built a model not giving a reference for the adjustment operation.

In this paper, the idea of paper [1] and [2] is combined, a new machine learning method - Relevance Vector Machine is drew. The RVM which is a sparse Bayesian modeling method is better than SVM.RVM has few Optimization parameters, so it is simple and proved to have good performance in some literatures, the classification and prediction effect is also better. The RVM whose parameters are optimized by GSA is used to model the iputs and the NOx emissions of the boiler in this araticle. At last the optimal model is got. Compared with SVM optimizated by PSO and GSA, the accuracy of the modeling method based on RVM is higher. In the final, we use the low NOx emissions as the goal to adjust the input variables optimization on the basis of this model for low NOx combustion.

2. Relevance Vector Machine

RVM [3]which is a sparse probabilistic model similar to SVM was proposed by Michael E.Tipping. The training is conducted in a bayesian framework, under the structure of prior parameters based on active correlation decision theory (automatic Relevance determination, hereinafter referred to as ARI) to remove the relevant point, thereby gaining a sparse model. Compared with SVM, the biggest advantage of RVM is the kernel function of computation is greatly reduced, and also overcomes the fault of the selected kernel function must satisfy the Mercer condition.

RVM and SVM both have the same linear prediction model, such as formula (1):

$$y(x,w) = \sum_{i=1}^{N} w_i K(x,x_i) + w_0 \tag{1}$$

In the formula, w_i is the weights of the model, $K(x,x_i)$ is the kernel function.

The actual data contains noise, in order to establish a more accurate relationship between input and output, the algorithm joins the uncertainties in the model to describe the noise, such as formula (2):

$$t_n = y(x_n,w) + \varepsilon_n \tag{2}$$

In the formula, ε_n is noise, obey the normal distribution whose mean is 0 and the variance to δ^2.

To control the complexity of forecast model, a prior distribution is established for the model's each weight to show the credibility of a particular weight, assuming that weight prior distribution as shown in formula (3):

$$p(w\,|\,\alpha) = \prod_{i=0}^{N} N(w_i\,|\,0,\alpha_i^{-1}) \tag{3}$$

In the formula, vector $\alpha = (\alpha_0,\alpha_1,...\alpha_N,)^T$ contains N+1 super parameter, each parameter α_i corresponds to the weight ω_i, it is responsible for the size of the cumulative value.

RVM technology uses bayesian framework to manage uncertainty by learning from training data, which is the weight of the posterior distribution reasoning. The posterior distribution of the weight calculation as shown in formula (4).

$$p(w\,|\,t,\alpha,\delta^2) = \frac{p(t\,|\,w,\delta^2)p(w,\alpha)}{p(t\,|\,\alpha,\delta^2)} \tag{4}$$

Prediction which is based on the known training data set and a new input data forecast the new target data on the base of the technology of RVM, as shown in figure (5):

$$p(t_n\,|\,t) = \int p(t_n\,|\,w,\delta^2)p(w,\alpha,\delta^2\,|\,t)dwd\alpha d\delta^2 \tag{5}$$

In relevance vector machine, if different kernel functions is choosed, it will form different algorithms, the commonly used kernel function including polynomial, gaussian radial basis b-spline kernel function and so on, one of the most commonly used is gaussian radial basis kernel function, defined as follows:

$$K(x,x_i) = e^{\left[\frac{-\|x-x_i\|^2}{\delta^2}\right]} \tag{6}$$

In the formula, δ is the width parameter of kernel function to optimize, the gaussian radial basis function is used as the kernel function model in this paper.

3. Gravitational Search Algorithm

Gravitational Search Algorithm (Gravitational Search Algorithm, GSA) which is a intelligent optimization simulated the the law of universal gravitation Algorithm was put forward by E.R ashed, etc in 2009 , the principle of object movement to follow the law of dynamics, the effect of gravity made objects move toward the larger mass of object, and the biggest object in the quality has the optimal position, which can be the optimal solution of optimization problem, the Algorithm has simple operation, less parameter settings etc, the basic principle is as follows:

Suppose that a system have N individuals, the position of an individual is defined as follows:

$$X_i = (x_i^1, \ldots x_i^k \ldots x_i^n) \quad i = 1, 2, \ldots, N \tag{7}$$

Where x_i^k presents the position of the i agent in the k dimension.

At a specific time t, we define the force acting on mass i from mass j as following:

$$F_{ij}^k(t) = G(t)\frac{M_{pi}(t) \times M_{aj}(t)}{R_{ij}(t) + \Delta}(x_j^k(t) - x_i^k(t)) \tag{8}$$

Where $R_{ij}(t)$ can be calculated by $\| X_i(t), X_j(t) \|_2$, Δ is a small constant.

So the sum of kth components of the forces exerted from other agents:

$$F_i^k(t) = \sum_{j=1, j \neq i}^{N} rank_j F_{ij}^k(t) \tag{9}$$

Where $rank_j$ is a random number in [0, 1].

Hence, by the law of motion, the acceleration of the agent i at time t, and in direction kth, $a_i^k(t)$ is given as follows:

$$a_i^k(t) = \frac{F_i^k(t)}{M_{ii}(t)} \tag{10}$$

Therefore, its position and its velocity could be calculated by the laws of motion as follows:

$$\begin{cases} v_i^k(t+1) = rank_i \times v_i^k(t) + a_i^k(t) \\ x_i^k(t+1) = x_i^k(t) + v_i^k(t+1) \end{cases} \tag{11}$$

GSA impacts on objects through the interaction between different objectss, it make the small quality objects move towards the big quality according to the fomular (11), the optimal position of the object is the solution of the optimization problem.

4. Modeling the NOx Emission

4.1 The experimental data

There are three ways to generate NOx, fuel nitrogen, heat nitrogen and quick nitrogen. The generation of rapid nitrogen, is very little, the main influence is by the first two. The generation of fuel nitrogen mainly depends on the nitrogen content and the way of supply air volume, which is corresponding to theprimary air and secondary air , coal characteristics and coal feeder's speed of modeling data ,as one of the secondary air, coal characteristics and coal feeder speed; the generation of thermal nitrogen has much to do with the temperature of furnace, which is corresponding to the blast's temperature, exhaust gas temperature and the excess oxygen in flue gas. Therefore boiler load, into the air and the wind temperature, exhaust temperature, furnace exit oxygen, coal quality characteristic parameters are chose as the input data of NOx modeling (this paper NOx is converted to 6% of the value of flue gas oxygen content of) [4-6].

For the study of the adjustable variables's influence on NOx emissions in power plant boiler combustion system, 20 groups experimental data of working condition were collected, he data table is shown in table 1, each working condition includes coal, coal feeder speed, secondary air, oxygen, exhaust gas temperature and so on ,the total parameters is 26. First 17 sets of data are the training data, the remaining three sets of data are prediction data. the first 3 groups are under the condition is of 220MW of boiler load, 4 to 17 groups are under the condition is of 230MW of boiler load conditions, 1 to 5 sets of data is under the condition that the opening of burning wind is 0 ,6 to 17 is under the condition of different opening of burning wind, 20 groups under conditions of different parameters are on behalf of the different influence of various parameters on the NOx emission, it is a symbol of the complex operation of boiler[7].

4.2. The determination of kernel function parameters

this paper uses discrete value δ to train the model, due to the selection of range of δ is not exactly criterion; δ are from 0 to 100 and the nterval is 5, observing the changes of the predicted model's error ,it is found that error reduced gradually when t δ is 0 to 30, if δ is from 30 to 100,the growth of ,error is repetitive ,then if δ a take larger value, the error becomes even larger, so the optimal value should be around 30, this paper selects the optimum range of δ is [0 100].

4.3. Establishing the model

GSA is used to optimize the width of the SVM parameters, then NOx emission characteristics model is established ,after three groups which are used to predict the NOx emissions of the resulting model is the test data .the steps of modeling are as follows:

1. Initializing the population of GSA, the number of iterations, δ to be optimized.

2. Putting the initial value of δ into RVM, then RVM is trained by training data, the parameters of model are obtained, finally the training data is used to predict the model, mean square error which is cacu-

Table 1. Operating data of boiler (1)

Operating number	load (MW)	the speed of primary air (m·s-1)				the speed of secondary air (m·s-1)					the degree of valve plate for the combustion air (%)			Exhaust temperature (/ °C)	Oxygen volume fraction (%)
		A	B	C	D	E	F	G	H	I	OFA up	OFA down	SOFA		
1	320	29.7	29.3	28.8	28.2	36.1	36.9	35.4	36.2	33.3	0	0	0	142.2	3.65
2	321	31.9	29.9	28.8	28.6	40	42.1	39.1	40.8	35.6	0	0	0	145.8	5.03
3	320.7	29.8	29.8	29.4	28.3	47.7	49	45.8	48.3	40.1	0	0	0	143.2	6.05
4	330.8	29.4	29.3	29.6	28.5	42	43.5	41.3	41.2	35.4	0	0	0	144.3	4.5
5	330.5	29.3	29.3	29.1	28.7	41.7	43	39.9	39.5	35.5	0	0	0	147	4.63
6	329.7	28	27.1	27.5	29.2	39.1	39.3	37.1	37.3	36.6	100	0	0	141.5	5.03
7	329.7	31.1	27.6	27.5	27.3	40.2	40.5	38.1	39.2	37.5	100	0	100	144.3	5.53
8	329.7	29.7	30.3	27.6	27.2	44.1	43.7	42.5	43.3	40.3	0	0	100	145	5.21
9	329.9	29.6	29.2	28.1	26.9	45.2	47.1	44.2	45.9	43.4	0	0	0	144.4	5.25
10	329.4	29.5	29.4	28.1	27.3	40.5	41	39	40.1	39.2	0	0	100	143.9	5.22
11	329.8	29.6	29.2	27.5	27.3	38.6	39	35.6	36.9	36.7	100	0	100	144.2	5.37
12	331	28.9	28.4	28.7	29.4	33.1	36.2	35.1	37.1	35.4	100	0	0	154.8	4.85
13	330	27.6	29	29	29	31.1	36	31.9	33.4	33.9	100	100	0	155.6	4.75
14	330	28.8	28.4	28.6	29.4	30.3	33.7	32.2	32.5	32	100	100	50	155.8	4.65
15	329.6	28.9	28.2	28.7	29.1	30.3	32.9	31.2	32.4	31.7	100	100	100	154.2	4.6
16	330	28.6	28.4	28.9	28.8	34.5	38.1	36.6	38	35.6	0	0	100	153.6	4.95
17	329.5	28.5	28.3	28.7	29.3	32.6	37.4	34.5	37.2	35.3	0	100	100	153.8	4.8
18	319.8	29.4	29.2	30.2	29.3	38.6	40.6	37.9	39.1	36.9	0	0	0	141.5	4.81
19	320.2	29.5	28.8	28.9	28.6	40.2	40.9	38	39.2	35.7	0	0	0	142.8	4.33
20	330	29.3	27.6	28.3	28	38.7	39.7	36.6	37.2	37.5	50	50	100	150.5	5.69

lated by the predictive value and the actual output determines the fitness of particle is good or bad.

3. Bring the fitness of particle back to GSA.GSA is made use of iterating the particle's position, the optimal particle the new set of δ

4. Judging whether t iterations is the maximum number, if it is ,directly skipping to step 5, otherwise returning to step 2.

5. The obtained optimum δ is used to train RVM, then the parameters of the model is got, finally the model is used to predict the 20 groups data.

4.4. Analyzing the result

Under the condition that the initial population of GSA is 50 and the number of iterations is 50, GSARVM is trained by the 17 training data, the optimal δ is 30.38, then 20 groups data are predicted. The contrast curve between the output of the prediction model and the output of actual is shown in figure 1, Error of the predicted values and the actual value is shown in figure 2.

Based on Figure 1 and Figure 2, the data from situation 1 to 17 are trained and the forecast error is almost 0. This indicates that relevance vector machine method based on the gravitational search algorithm has excellent differentiated ability. For the data from situation 18 to 20, cause the models are not trained, the errors are -7.2, 7.2 and 13 for situation 18, 19, 20. Comparing with the NOx value from 550 to 950, data

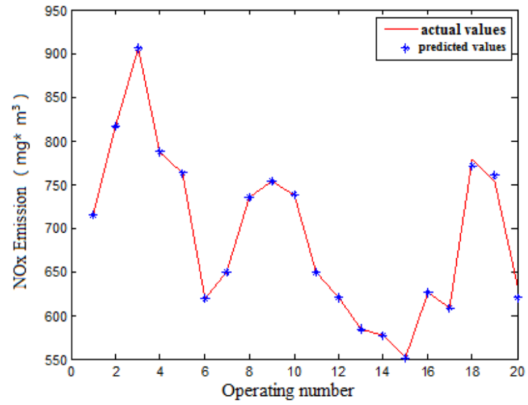

Figure 1. Comparison of predicted values and actual values based on GSARVM

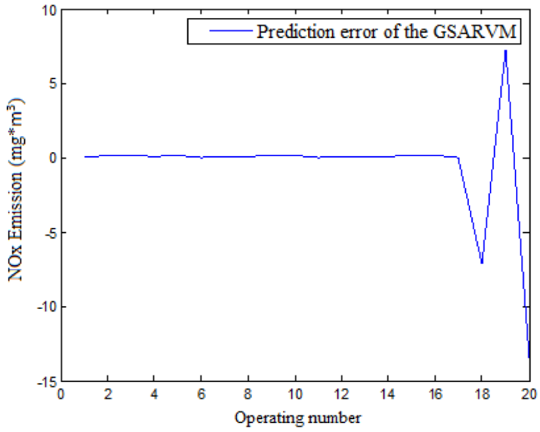

Figure 2. The forecast error of GSARVM

Table 2. Operating data of boiler (2)

Operating number	properties of the coal of power plant(%)						Qar(kg·kg-1)	speed of the coal feeder(r·min-1)				NOx mass concentration (mg·m-3)
	ω(Car)	ω(Har)	ω(Oar)	ω(Nar)	ω(War)	ω(Aar)		I	II	III	IV	
1	56.88	3.5	5.38	1.16	10.8	21.6	18.796	334	355	389	377	716.05
2	57.56	3.62	5.38	1.15	12.2	19.15	17.914	378	345	388	399	817.65
3	62.84	3.96	6.48	1.25	10.5	14.52	21.336	327	351	393	386	906.05
4	62.84	3.96	6.48	1.25	10.5	14.52	21.336	341	359	343	401	787.4
5	60.49	3.7	5.96	1.24	8.4	19.83	21.023	327	389	417	336	763.3
6	60.49	3.7	5.96	1.24	8.4	19.83	21.023	360	430	463	322	619.05
7	60.49	3.7	5.96	1.24	8.4	19.83	21.023	343	432	465	332	649.75
8	60.49	3.7	5.96	1.24	8.4	19.83	21.023	347	459	481	310	735.1
9	60.49	3.7	5.96	1.24	8.4	19.83	21.023	345	451	492	321	754.35
10	60.49	3.7	5.96	1.24	8.4	19.83	21.023	340	453	488	311	738.2
11	60.49	3.7	5.96	1.24	8.4	19.83	21.023	344	456	479	319	650.35
12	58.37	3.81	5.1	1.21	6.4	25	20.29	341	344	411	415	621
13	58.37	3.81	5.1	1.21	6.4	25	20.29	339	340	415	419	584.5
14	58.37	3.81	5.1	1.21	6.4	25	20.29	342	345	420	414	577
15	58.37	3.81	5.1	1.21	6.4	25	20.29	333	344	419	422	552
16	58.37	3.81	5.1	1.21	6.4	25	20.29	336	339	420	418	626.5
17	58.37	3.81	5.1	1.21	6.4	25	20.29	340	340	417	417	609
18	56.88	3.5	5.38	1.16	10.8	21.6	18.796	337	342	391	396	778.45
19	56.88	3.5	5.38	1.16	10.8	21.6	18.796	320	343	397	396	753.5
20	60.49	3.7	5.96	1.24	8.4	19.83	21.023	330	341	422	455	634.65

from situation 18, 19, 20 are more accuracy. So the result is acceptable.

Set the quantity of particle to 50, set iterative time to 50. The forecast values obtained from different algorithms (PSO, GA, GSA) are showed as Table.

Table 3. The comparision of predicted values based on diffierent optimization algorithm

Operating number	Actual Values	GSARVM predicted values	PSORVM predicted values	GARVM predicted values	GSASVM predicted values
18	778.45	771.24	793.37	722.08	715.45
19	753.5	760.7	776.06	722.5	712.5
20	634.65	621.09	621.59	621	667.45

Diagram based on the forecast error of GSARVM model is showed in Figure 2. After optimized, the value of δ is 30.38. The value of δ is 23.78 after optimized from model GSARVM, showed as Table 3. And the value of δ is 35.46 obtained from model PSORVM. Figure 2 and Table 3 show that GSARVM algorithm is more accuracy than others and the performance of GSARVM is also more excellent than PSO and GA. The identity ability of relevance vector machine is more excellent than supporting vector machine.

5. The Optimization of NOx Combustion

There are 26 groups of input parameters for the coal boiler. 17 groups of input parameter are adjustable variables, including primary air, secondary air, combustion air, speed of coal service machine, oxygen and so on. 9 groups are unadjustable variables, including the quality of coal, load and so on. NOx output values are affected by these 26 groups of variables. Now these 17 groups of adjustable variable are used as optimized objects, 9 groups of unadjustable variable as conditions, to optimize the model for low level NOx combustion.[8-10] The range of been optimized variables are listed as below:

$$\begin{cases} 25 < x_1,\dots x_4 < 35 \\ 30 < x_5,\dots x_9 < 50 \\ 0 < x_{10}, x_{11}, x_{12} < 100 \\ 320 < x_{13} < 380 \\ 340 < x_{14} < 460 \\ 300 < x_{15} < 490 \\ 390 < x_{16} < 460 \\ 3 < x_{17} < 6 \end{cases} \tag{12}$$

$x_1,\dots,x_4$ are the speed of primary air for the 4th floor

$x_5,\dots,x_9$ are the speed of secondary air for the 5th floor

x_{10}, x_{11}, x_{12} are the degree of valve plate for the combustion air

$x_{13},\dots x_{16}$ are the speed of coal service machine.

x_{17} is the oxygen content.

Choose the data of situation 3 from 20 groups of data to optimize. Set the quantity of particle to 100, iterative time to 500. The process of optimizing is introduced as below procedure:

Initialize the quantity of particle, iterative time, dimensions of variable and the value of δ

The variables to be optimized are input into model as trained data. So the NOx output is obtained to estimate the particles.

Return the dimensions of variable into gravitational search algorithm. Optimize the iterative times to obtain new optimized values of variable.

Estimate whether the iterative time has reached the maximum. If yes, jump to procedure 5. If no, return to procedure 2.

Output the values of optimized variable. Calculate the output value of NOx under the situation that the optimized variables are used as input parameters. The transformation of optimized process is showed in Figure 3. And the comparing between adjustable variables unoptimized and variables optimized is showed in Table 4 and Table 5.

Table 4. The comparision of parameters before and after NOx optimization(1)

Operating number 3	the speed of primary air ($m \cdot s^{-1}$)				the speed of secondary air ($m \cdot s^{-1}$)				
	A	B	C	D	E	F	G	H	I
Unoptimization	29.8	29.8	29.4	28.3	47.7	49	45.8	48.3	40.1
optimization	27.6372	28.5576	31.2814	30.0758	35.192	47.328	32.362	34.698	45.688

Table 5. The comparision of parameters before and after NOx optimization(2)

Operating number 3	OFA (up)	OFA (down)	SOFA	oxygen content	speed of the coal feeder ($r \cdot min^{-1}$)				NOX mass concentration
	%	%	%	%	I	II	III	IV	mg·m-3
Unoptimization	0	0	0	6.05	327	351	393	386	906.05
optimization	88.1795	22.3698	40.2446	4.5842	350.344	396.577	408.661	312.036	550.600

Table 4 and Table 5 show that the output value of NOx have changed from 906.65 mg/m^3 to 550.600 mg/m^3 after optimized. The drop scope is about 38.9%. After optimizing, other variables also have some changes. For primary air and secondary air, there is a little change. But there is a dramatic drop for the oxygen. It indicates that NOx emission can be reduced when the oxygen reduced. For the speed of coal service machine, number 1, 2,3 have a little extent grow. But number 4 has a big extent drop. This also indicates that NOx emission can be reduced when the speed of coal service machine is down. Above evidence can proof that utility boilers' NOx combustion optimization based on relevance vector machine has excellent performance and an acceptable technology of low level NOx combustion.

Conclusions

This paper introduces a new machine learning method - Relevance Vector Machine into the modeling of a 300MW pulverized coal boiler. The model is built based on 20 groups of actual data obtained from site, and Gravitational Search Algorithm is used to optimize the parameters of the model. The model get from GSARVM is more accuracy comparing with other algorithm. By using the boiler adjustable variable input parameters as the optimization variables, the NOx emission has been reduced to a very low level after adjustable variable input parameters optimized and meets the requirement which the Nation has ordered for the NOx emission. The lower NOx emission has been realized based on the new model.

Acknowledgements

This paper is supported by the Science and Technology Planning Project of Hebei Province (13211610) and Doctoral Program of Yan Shan university.

References

1. Wang Chunlin, Zhou Hao, Li Guoneng, Cen Kefa. Support vector machine modeling on NOx emission property of high capacity power station boiler. Journal o f Zhejiang University, 2006, 40(10), p.p.1787-1791
2. Niu Peifeng, XIAO Xingjun, LI Guoqiang, MA Yunfei, CHEN Guilin, ZHANG Xianchen. Parameter Optimization for NOx Emission Model of Power Plant Boilers Based on Gravitational Search Algorithm . Journal of Chinese of Society of Power Engineering, 2013, 33(2), p.p. 101-106.
3. M E. Tipping. Sparse Bayesian learning and the relevance vector machine. Journal of Machine Learning Research, 2012, 20, p.p.211-244.
4. Zhang Yi, Chen Biao, Ding Yanhui, Zhang Qingfeng, Wu Zhansong. Experimental investigation of operating strategy of low NOx high efficiency coal-fired utility boilers. Tsinghua

University(Sci & Tech) , 2006, 46, p.p.666-669 .

5. Wang Xuedong, Xin Hongchang, Luan Tao, Cheng Lin. Research and Test on Influence of Boiler Combustion Adjusting on NOx Emission of 330MW Unit. Power System Engineering, 2014, 23(3) p.p.7-10.
6. Wang Peihong, Li Leilei, Chen Qiang, Dong Yihua. Response Characteristics Model of NOx Emission and Efficiency for Power Station Boiler.Power Engineering, 2013, 24(2), p.p.254-259.
7. Fei Jun, Sun Rui, Zhang Xiaohui, Zhang Yong, Sun Shaozeng, Qin Yukun. Characteristics of NOx Emission in Pulverized Coal Fired Boiler Under Different Combustion Conditions.Journal of Power Engineering,2009,29 p.p.813-817.
8. Ma Yunfei, Niu Peifeng, Zhao Yantao, Ma Xiaofei. Adaptive Particle Swarm-Based Fuzzy Clustering Algorithm in the Application of Steam Drum Pulverized Coal Fired Boiler. International Journal of Advancements in Computing Technology, 2012, 11, p.p.444-452.
9. Peifeng Niu. Yunfei Ma.Pengfei Li.Yang Zhang. Hybrid neural network in circulating fluidized bed boiler basedon information fusion clustering control. Neural Comput & Applic .2013, 23, p.p.1949-1962.
10. Kim W, Song H. Filtering of erroneous positioning data with Iterative application of one class support vector machine. Database Theory and Application, 2012, 5, p.p.67-88.