Intelligent Integrated Coking Flue Gas Indices Prediction

Yaning Li [1,2], Xuelei Wang [1,*], Jie Tan [1], Chengbao Liu [1,2], Xiwei Bai [1,2]

([1] *Institute of automation, Chinese Academy of Science. Beijing 100190, China*

[2] *University of the Chinese Academy of Sciences, Beijing 100049, China*
xuelei.wang@ia.ac.cn)

***Abstract*—Focus on the first China domestic coking flue gas desulfurization and denitriation integrated device, in order to solve the problem that the entrance parameters fluctuate and a detection lag exists due to the upstream coking workshop, which is extremely unfavorable to the optimal control of desulfurization and denitriation process. An intelligent integrated prediction model of flue gas SO2 concentration, O2 content and NOx concentration was proposed: the mechanism models of SO2, NOx concentration and O2 content were established according to the principle of material balance and reaction kinetics, respectively. For the prediction error, raw data was pretreated and the auxiliary variables were determined by principal component analysis, in order to improve the training speed and generalization ability of neural network, an improved RBFNN combining optimal stopping principle and dual momentum adaptive learning rate was proposed and used to compensate the error. Based on the practical data of two 55-hole and 6-meter top charging coke ovens in the coking group, the effectiveness and superiority of proposed model and method were verified by simulation via comparison of various models.**

Keywords—coking flue gas; mechanism model; neural networks; integrated modeling

I. INTRODUCTION

China is the world's largest coking production country with a huge annual emission of SO_2 and NO_x. With the formal implementation of "coking chemical pollutant emission standards" in January 1, 2015, the SO_2 and NO_x emission targets of coking industry are put forward strict and clear quantitative requirements[1]. Under this situation, Jiangxi Coking & Chemical Co. Ltd. of China has taken the lead in building and operating of coking gas desulfurization and denitriation integrated engineering(Fig.1), it has been proved good results from the actual operation, the emission of flue gas has reached the national standard[2]. However, the gas indices(SO_2 and NO_x concentration, O_2 content) are directly affected by the upstream coking process, but its working conditions are complex and changeable, at present, the control method is mainly based on manual observation of inlet flue gas indices, so as to adjust the operating parameters, as a result, the following problems are extremely unfavorable to the optimal control of the device:

When the coking production conditions change, there is bound to be a corresponding change for the generated flue gas indices, and reflecting in flue gas monitoring data of desulfurization tower inlet after a certain transmission delay. If only rely on the data of DCS system of the desulfurization and denitration unit, the operator may be aware of the change of the upstream condition after a long period of time because of the strong fluctuation of parameters. If adjust the relevant operating parameters at this time, it's bound to cause a serious lag because a large amount of flue gas has entered the follow-up process; besides, the formation mechanism of SO_2, NO_x and other indices are different, the traditional mechanism modeling method has a great error or even unable to be modeled, which increase the difficulty of solving such problems.

Therefore, it is urgent to establish a method of predicting indices of coke oven flue gas, in order to lay a foundation for optimal control of the desulfurization and denitration process.

Coking production using coking process management system (CPMS) for management nowadays, and the SIEMENS PCS7 process control system is used in the process of desulfurization and denitration, the interconnection of data provides the basis for data driven method of flue gas indices prediction. The neural network(NN) method doesn't need to establish the mathematical model of the object, and it has the advantages of parallel processing, strong learning ability and good robustness, as a result, it has been widely used in the modeling of complex systems. In recent years, many scholars at home and abroad have studied the neural network and intelligent integrated modeling of complex industrial processes[3-8], but there are few reports on the research of coking flue gas.

Prediction model is the basis of optimal control for desulfurization and denitration process, it is of great significance for the the timeliness and effectiveness of the subsequent control and optimization. In this paper, an intelligent integrated prediction model of SO_2, NO_x concentration and O_2 content is proposed: the mechanism models of SO_2, NO_x concentration and O_2 content are established according to the principle of material balance and reaction kinetics, a new RBF neural network training algorithm based on optimal stopping principle and double momentum adaptive learning rate is proposed and used to predict the dynamic error of the prediction after determining auxiliary variables via pretreatment and principal component analysis. The effectiveness of the proposed model and method is verified by the practical data of two 55-hole and 6-meter top charging coke ovens in the coking group.

978-1-5090-5504-3/17/$31.00 ©2017 IEEE
SNPD 2017, June 26-28, 2017, Kanazawa, Japan

Fig.1 Desulfurization and denitration integrated device for coking flue gas

II. INTELLIGENT INTEGRATED MODELING OF FLUE GAS PARAMETERS

We present an intelligent integrated prediction model strategy of coking gas parameters, as shown in Fig.2.

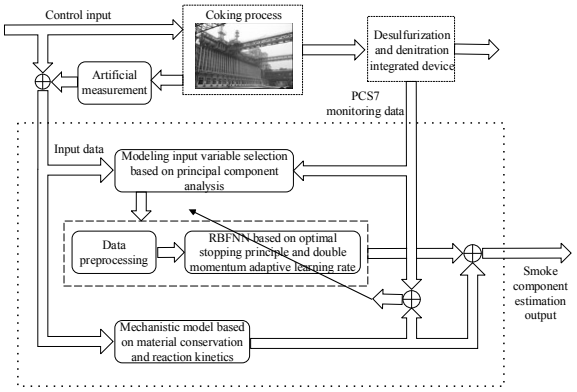


Fig.2 Modeling strategy of coking flue gas parameters prediction

The established prediction model can be expressed as:

$$C_{pi}(t+\tau_i) = C_{th} + C_{com} = f(\Theta(t))$$

where $C_{pi}(i = 1,2,3)$ is concentration and content of SO_2, O_2 and NO_x, respectively; $\tau_i(i = 1,2,3)$ is the lag time of the response of the inlet parameters of desulfurization tower corresponding to the change of working conditions caused by coke oven or the length of exhaust gas transmission pipeline; C_{th}、C_{com} represent mechanism model and neural network compensation model, respectively; $f(\cdot)$ denotes the complex nonlinear relation of integrated prediction model; $\Theta(t)$ is all the input vectors model calculation need at time t.

A. Mechanism Model Based on Material Balance

Flue gas in the inlet of the desulfurization and denitration integration device generated by two 55-hole and 6-meter top charging coke ovens in the coking plant, regard the two coke ovens as a generalized object, the schematic diagram of the gas flow in the combustion process as Fig. 3.

The dotted lines S1, S2 in Fig.3 indicate the leak in or out of raw gas, waste gas or air caused by improper sealing of the coke battery and conduit from main flue to the desulfurization tower.

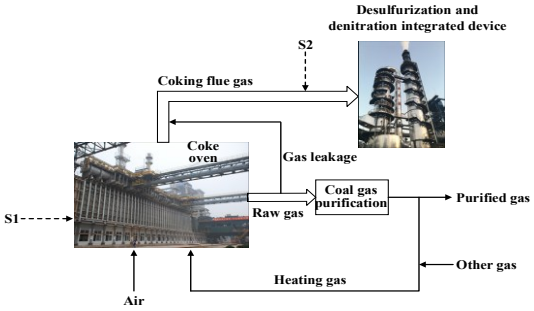


Fig.3 Gas flow in coke oven combustion

The amount of gas, air intake and combustion temperature change dramatically during the periodic reversing process of coke oven, it's difficult to use mathematical models to describe the rule, therefore, the prediction of flue gas parameters in this paper refers to the non-reversing process.

1) Gas flow setting model[9]

The heating gas consumption can be roughly calculated according to the design parameters if the coking time is determined, the actual situation may be adjusted according to the specific situation of coke oven.

$$V_0 = \frac{q_{sN}}{Q_{DW}t_g}B * 1000 \quad (1)$$

where V_0 is gas flow rate ($m^3 \cdot h^{-1}$) under standard condition (273.15K，101Kpa), q_s is heat consumption of wet coal ($kJ \cdot kg^{-1}$), N is the number of oven chambers, Q_{DW} is the low calorific value of coal gas ($kJ \cdot kg^{-1}$), t_g is the gross coking time (h), B is the average amount of charging wet coal per hole (t).

2) Air consumption model in combustion process

Under standard condition, the air consumption in combustion process can be obtained by the following models:

$$V_{O2n} = 10^{-2} * [0.5(R_{H_2} + R_{CO}) + 2R_{CH_4} + \frac{4n+1}{4}R_{C_nH_n} - R_{O_2}]$$
$$V_{an} = 1/P_{O_2} * V_{O2n}$$
$$V_{ard} = \alpha * V_{an}$$
$$V_{arw} = V_{ard} * (1 + \varphi * r_{H2O}) \quad (2)$$

where V_{O2n} is theoretical oxygen required for combustion, R_M is volume percentage of M in gas (%), V_{an} is theoretical air volume, P_{O_2} is the percentage of oxygen in the air, V_{ard} is actual dry air volume, V_{arw} is actual wet air volume, φ is relative air humidity, r_{H2O} is saturated water vapor content of dry air at present temperature ($m^3 \cdot m^{-3}$). The above calculation is under combustion of 1m^3 dry gas, all the volume unit is $Nm^3 \bullet Nm^{-3}$.

3) Model of waste gas and its composition

When gas is fully burnt, the waste gas contains only CO_2, SO_2, H_2O, N_2 and O_2 in excess air, normally, SO_2 and NO_x is ignored due to its relatively small volume fraction.

$$V_{RO_2} = V_{CO_2} + V_{SO_2} = 0.01(R_{CO_2} + R_{CO} + R_{H_2S} + \sum mR_{C_mH_n})$$
$$V_{H2O} = 10^{-2} * (R_{H_2} + R_{H_2S} + \sum \frac{n}{2}R_{C_mH_n} + R_{H_2S} + r'_{H2O} + R_{ard} * \varphi * r_{H2O})$$
$$V_{N2} = 10^{-2} * R_{N_2} + P_{N_2} * V_{ard}$$
$$V_{O2} = P_{O_2} * V_{ard} - V_{O2n}$$

$$V_{m3}=V_{RO_2}+V_{H2O}+V_{N2}+V_{O2} \quad (3)$$

where V_{RO_2} means the sum volume of CO_2 and SO_2, $V_{H2O}, V_{N2}, V_{O2}, V_{m3}$ represent the volume of H_2O, N_2, O_2 and waste gas generated by $1m^3$ dry gas combustion, r'_{H2O} is saturated water vapor content in coal gas at current temperature ($m^3 \cdot m^{-3}$), P_{N_2} is volume ratio of N_2 in air.

4) Oxygen content in waste gas

The model of volume ratio of oxygen content in waste gas can be expressed from the first two sections:

$$C_{Oth}=\frac{V_{O2F}}{V_{m3}}=\frac{P_{O_2}*V_{ard}-V_{O2n}}{V_{m3}}*100\% \quad (4)$$

where C_{Oth} is the percentage of oxygen content in waste gas ($\%$).

5) SO_2 concentration in waste gas

According to Ref.[10,11], under coke oven gas heating, SO_2 from H_2S in heating gas accounts for about 30% of that from waste gas; SO_2 from organic sulphur in heating gas accounts for about 10%; SO_2 from sulphur in raw gas channeling to the combustion system accounts for about 60%. The organic sulphur is difficult to measure, so the following will carry on calculation from two sections of recycled coal gas and gas leakage.

Assuming oxygen is sufficient, H_2S burns into SO_2:

$$2H_2S+3O_2=2SO_2+2H_2O$$

The SO_2 concentration generated by recycled coal gas combustion can be calculated as:

$$C_{Sth1}=\frac{M_{O2F1}}{V_0*V_{m3}}=\frac{(V_0*a_1*1.882)}{V_0*V_{m3}}=\frac{a_1*1.882}{V_{m3}} \quad (5)$$

where C_{Sth1} is the SO_2 concentration generated by recycled coal gas combustion ($mg \cdot m^{-3}$), M_{O2F1} is the quality of SO_2 generated by recycled coal gas combustion per hour (mg), a_1 means the content of H_2S in recycled coal gas ($mg \cdot m^{-3}$), 1.882 is the molar mass ratio of H_2S to SO_2.

The SO_2 concentration produced by leakage of raw gas can be calculated by the following formula:

$$C_{Sth2}=\frac{M_{O2F2}}{V_0*V_{m3}}=\frac{(V_{raw}*P_{lea}*a_2*2*10^3)}{V_0*V_{m3}} \quad (6)$$

where, C_{Sth2} is SO_2 concentration produced by raw gas leakage ($mg \cdot m^{-3}$), M_{O2F2} is the quality of SO_2 produced by raw gas leakage per hour (mg), V_{raw} is the generating amount of raw gas per hour (m^3), P_{lea} is the average leakage rate of raw gas, a_2 is the sulfur content in raw gas, generally 6-10$g \cdot m^{-3}$, number 2 is the molar mass ratio of S to SO_2.

In general, calculation model of coke oven gas SO_2 concentration is:

$$C_{Sth}=C_{Sth1}+C_{Sth2}=\frac{V_0*a_1*1.882+V_{raw}*P_{lea}*a_2*2*10^3}{V_0*V_{m3}} \quad (7)$$

Although the theoretical value of oxygen content and SO_2 concentration can be calculated based on material constant calculation, in fact, the amount of the leakage of waste gas, wind, organic sulfur content and other parameters, and leakage rate of raw gas in carbonization chambers are difficult to estimate, therefore, there will be a large error compared with the flue gas detection data of desulfurization tower inlet if the prediction only relying on the mechanistic model.

B. NOx Generation Model Based on Reaction Kinetics

NO_x, mainly refers to the mixture of NO and NO_2, is produced when gas is burning in the coke oven. Due to NO will eventually be converted to NO_2 in the air, NO_x is all calculated by NO_2 in the estimation of NO_x emission. Some research shows that, NO accounts for more than 95% of NO_x produced by combustion, the rest is mainly NO_2, so we mainly study the formation mechanism of NO instead of NO_x. The nitrogen oxides formed during the combustion of coke oven can be divided into three types[12]: (1) thermal NO; (2) prompt NO; (3) fuel type NO. According to Ref.[13], the NO generated in the process of coke oven combustion is mainly thermal NO. The proportion of fuel type NO is not more than 5% with coke oven gas heating, and it's almost all thermal NO with lean gas heating.

According to the Zeldovich mechanism[14], the oxidation of N_2 in air at high temperature is carried out by a series of unbranched chain reactions, that is:

$$N_2+O_2 \leftrightharpoons N+NO$$
$$O_2+N \leftrightharpoons NO+O$$

If there is water vapour in the combustion process, the following reaction will occur:

$$N+OH \leftrightharpoons NO+H$$

The thermodynamic NO generation rate model based on the Arrhenius formula is:

$$\frac{dC_{NO}}{dt}=3*10^{14}\exp\left(\frac{-542000}{RT}\right)*C_{N_2}*C_{O_2}^{0.5} \quad (8)$$

where t is time (s), C_M is concentration of M ($mg \cdot m^{-3}$), T is maximum combustion temperature (K), R is general gas constant which usually take 8.314 ($J \cdot mol^{-1} \cdot k^{-1}$).

The concentration calculation model is:

$$C_{Nth}^{i}=\frac{dC_{NO}}{dt}\Delta t+C_{Nth}^{i-1} \quad (9)$$

where C_{Nth}^{i} is theoretical value of NO_x concentration at present time ($mg \cdot m^{-3}$), Δt is sampling interval (s), i is computing times.

In calculation process, take theoretical air volume as N_2 concentration, take the results in section 1.1.4 as O_2 concentration, and the measurement data of single vertical flue in order to reduce the error. Since the actual combustion temperature is between theoretical combustion temperature and measured temperature of fire brick masonry[13], because the actual temperature of vertical flue reflects the high temperate zone of combustion, so we take longitudinal temperature as the approximate value of combustion temperature.

Because 1) Dynamic model (8) is a general model for generating rate of thermal NO, the NO_x concentration of coking flue gas is also related to the structure of coke oven and oxygen content of the flue gas; 2) In the calculation process of the model, the temperature is roughly approximated, the error between the

measured value and the actual value of each parameter exists; 3) The accidental error caused by the leak in or out of gas in coke oven and conduit from main flue to the desulfurization tower cannot be estimated. Therefore, the NO_x concentration model based on the reaction kinetics has a poor effect.

C. Error Compensation Model Based on NN

An improved RBF neural network is proposed, using gradient descent method with double momentum adaptive learning rate to optimize weights of the network, and the optimal stopping strategy is used for training. The specific algorithms are as follows.

1) Dual momentum adaptive gradient descent algorithm

In view of the advantages of RBF neural network in approximation ability, classification ability and learning speed, [15-16], RBF neural network is used to compensate the modeling error in this paper. A typical three layer RBFNN nonlinear mapping is:

$$C_{com}(k) = \sum_{i=1}^{N} \omega_i G_i(||\upsilon - c_i||)$$
$$G_i(||\upsilon - c_i||) = \exp\left(-||\upsilon - c_i||^2 / 2\sigma_i^2\right), i = 1, \ldots \ldots \mathrm{N} \quad (10)$$

where υ is input vector of the network, $C_{com}(k) = [C_{scom}(t), C_{ocom}(t), C_{ncom}(t)]$ denote the error compensation value of SO_2 concentration, O_2 content and NO_x concentration, respectively, $G_i(\cdot)$ is the Gauss function of hidden layer, c_i and σ_i is the center point and width vector of $G_i(\cdot)$, ω_i is the weights of network, N is the number of hidden layer node.

c, σ and ω are trained via a supervised learning algorithm, gradient descent method, let $C_p(k)$ and $C(k)$ represent the actual output and expected output of gas indices, respectively. The network approximation error is:

$$e(k) = C(k) - C_p(k) \quad (11)$$

Define the objective function:

$$E = {}^1/_2\, e(k)^2 \quad (12)$$

The modification algorithm of the weights of output layer and the parameters of membership function is:

$$\Delta\omega_i(\mathrm{k}) = -\eta\frac{\partial E}{\partial \omega_i} = \eta(C(k) - C_p(k))G_i$$
$$\omega_i(\mathrm{k}) = \omega_i(\mathrm{k}-1) + \Delta\omega_i(\mathrm{k}) + \alpha(\omega_i(\mathrm{k}-1) - \omega_i(\mathrm{k}-2))$$
$$\Delta b_i(\mathrm{k}) = -\eta\frac{\partial E}{\partial b_i} = \eta(C(k) - C_p(k))\omega_i G_i \frac{||\upsilon - C_i||^2}{\sigma_i^3}$$
$$b_i(\mathrm{k}) = b_i(\mathrm{k}-1) + \Delta b_i(\mathrm{k}) + \alpha(b_i(\mathrm{k}-1) - b_i(\mathrm{k}-2))$$
$$\Delta c_{ij}(\mathrm{k}) = -\eta\frac{\partial E}{\partial c_{ij}} = \eta(C(k) - C_p(k))\omega_i G_i \frac{\upsilon - c_{ij}}{\sigma_i^2}$$
$$c_i(\mathrm{k}) = c_i(\mathrm{k}-1) + \Delta c(\mathrm{k}) + \alpha(c_i(\mathrm{k}-1) - c_i(\mathrm{k}-2)) \quad (13)$$

where η denotes the learning rate, α is momentum factor, $\eta \in (0,1)$, $\alpha \in (0,1)$.

The choice of learning rate η is important in the process of weight optimization, unsuitable η may cause slow convergence, oscillation and even divergence and other problems of the network, there is no rule to choose it at present, it's often determined by experience and repeated experiments[17].

An adaptive gradient descent algorithm based on double momentum is proposed, the learning rate η and its change rate β are adjusted at each step in learning process, which are related to the error function and the error variation, respectively. The specific method is as follows:

Construct the nonlinear function:

$$\eta(\mathrm{e(k)}) = \mathrm{q} + \mathrm{dsech}(\beta \mathrm{e(k)}) \quad (14)$$

where $\mathrm{q}, \mathrm{d} \in \mathrm{N}^+$, d is the change interval of η, β is the change rate of η, sech is a hyperbolic secant function:

$$\mathrm{sech(x)} = \frac{2}{e^x + e^{-x}}, -\infty < \mathrm{x} < +\infty \quad (15)$$

β is adjusted as follows:

$$\beta = \mathrm{e}^{b||\Delta e(\mathrm{k})||}, 0 \le ||\Delta e(\mathrm{x})|| \le 1 \quad (16)$$

where b is a normal number gain, $\Delta e(\mathrm{k}) = \mathrm{e(k)} - \mathrm{e(k-1)}$.

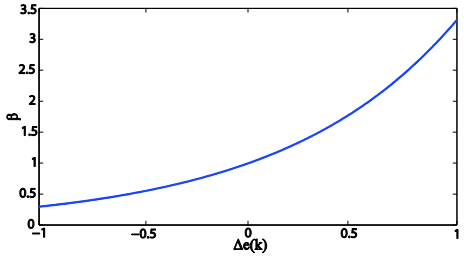


Fig.4 β curve

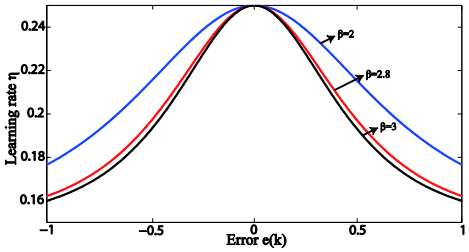


Fig.5 Learning rate curve

The above figures are function curve of β and error, and function curve of learning rate η under different β and error, respectively. It can be seen that the learning rate becomes smaller along with the error absolute value becomes larger, so as to achieve self-adjustment. Besides, β is adjusted dynamically according to $\Delta e(\mathrm{k})$, that is, when $\Delta e(\mathrm{k})$ becomes larger, the reduction rate of η will be reduced, or the increment rate of η will be increased. After the improvement of dual momentum adaptive adjustment, the number of network learning is greatly reduced, and the convergence rate is also greatly accelerated.

2) Optimal stopping strategy

Too much training of neural network will increase the training time, more seriously, it may cause over-fitting results. Some studies showed that, the generalization error of neural network will decrease with the decrease of training error, but after reaching the minimum point, the generalization error will increase gradually although the training error decreases, that is

Identify applicable sponsor/s here. If no sponsors, delete this text box (*sponsors*).

so-called over-fitting. Optimal stopping strategy[18] can be expressed as follows: the sample data are randomly divided into training set, validation set and test set before training, when start training, the validation set is used to monitor the training process. When it starts to over-fitting, the verification error will increase, and the network training should be stopped in advance, meanwhile, the network parameters with the minimum error are returned.

The learning samples are divided into training sample D_{tr} and validation sample D_{im}, and the error expressions are defined as following:

$$E_{tr} = \frac{1}{n_{tr}} \sum_{i=1}^{n_{tr}} (y_i - \hat{y}_i)^2 \quad (17)$$

$$E_{im} = \frac{1}{n_{im}} \sum_{i=1}^{n_{im}} (y_i - \hat{y}_i)^2 \quad (18)$$

where i represents the sample number, n_{tr} and n_{im} are the batch for training samples and validation samples, y and $\hat{y}$ are the actual value and prediction value of flue gas parameters, respectively.

The training steps are as follows:

1) Parameter initialization: learning time k=1, the maximum number of learning time k_{max}, precision target E_{trmin}, threshold $s, (s > 1)$.

2) Training the network for one round using the method described in section 1.1.3, store the corresponding connection weights and the corresponding $E_{tr}(k)$.

3) Calculate the relative error square and mean value of the predictive value of the model for validation sample D_{im}. When k=1, let $E_{immin} = E_{im}$ and save the current connection weights and other information as the best connection parameters of neural network.

4) Define the termination conditions:

a) $k = k_{max}$

b) $E_{tr} \leq E_{trmin}$

c) $\begin{cases} E_{immin} < E_{im}(k+i), i = 0,1, \dots \dots, s-1 \\ \sum_{i=0}^{s-1} E_{tr}(k+i) < \sum_{i=1}^{s} E_{tr}(k-i) \end{cases}$

In which, the two expressions in condition c) need to be satisfied simultaneously, the reason for this condition is that the verification error may temporarily increase due to the turbulence in training process, which will lead to premature end of the training. To avoid this situation, a threshold is set, the training process wouldn't stop when the verification error getting larger but to continue. If the verification error becomes larger for s times, it's considered the network is over-fitting, stop the training.

If one or more of conditions a)~c) are satisfied, stop the training and the optimal connection weights at time k are used as the final connection weights of network; otherwise, let $k = k + 1$ and go to step 2).

The algorithm flow is shown in Fig 6, where TC is the stop condition.

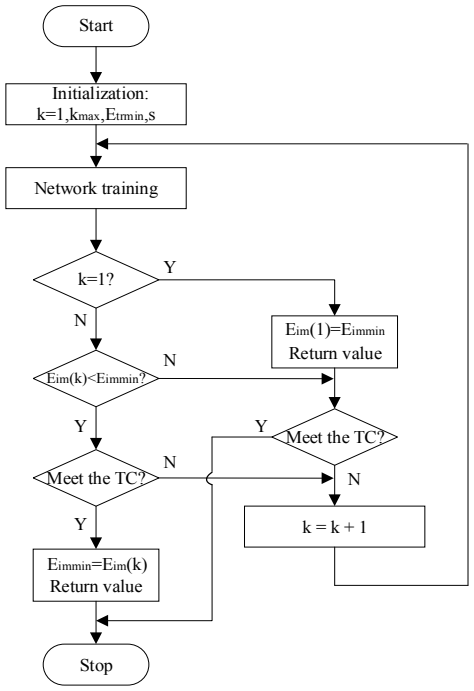


Fig.6 The program flow chart of optimal stopping method

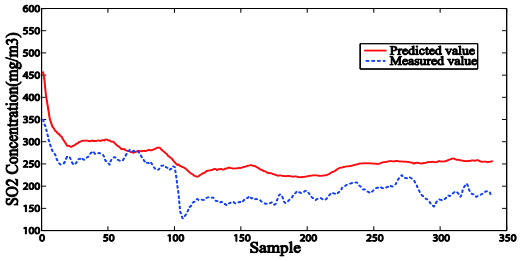
III. INDUSTRY EXPERIMENT

A. Mechanism Model Validation

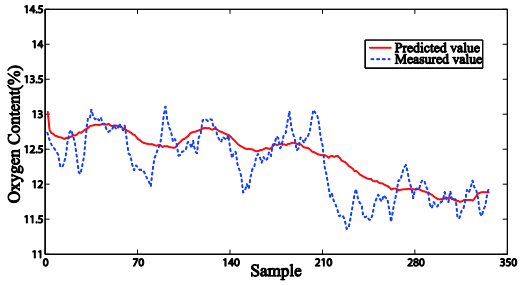
The mechanism model is established first, the data of coking plant in October to December, 2015 is collected and used for modeling, the prediction results of three parameters are shown in Figure 8 (a), (b), (c), respectively.

It can be seen that the change trend of the three parameters is basically in line with the actual situation, but the prediction accuracy is significantly worse. According to statistical analysis, the average relative error of SO_2 concentration is 19.2%, the maximum relative error is nearly 60%; the average relative error of O_2 content is 11.2%, the maximum relative error is about 17.5%; the average error of NO_x concentration is very large, only the trend is consistent with the fact.

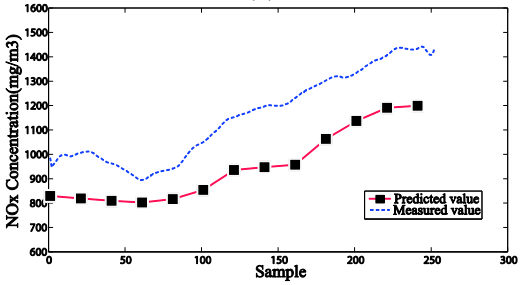
The main reason may be: after calculation, it is found that the total amount of waste gas is lower than the actual amount, which indicates there is a lot wind leaking in the system, result in higher SO_2 concentration and O_2 content, while the actual combustion temperature of gas is higher than that of the longitudinal temperature, so the NO_x concentration is lower.



(a)



(b)



(c)

Fig.8 Flue gas indices prediction based on mechanism model

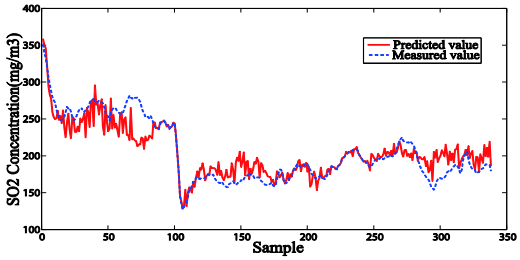
B. Integrated Model Validation

The industrial field and artificial measurement data is mainly used for experimental process. The data obtained from the DCS system cannot be directly used as the input of the compensation model due to some other interference factors, data pretreatment is needed, including error processing and data transformation. For error and noise, the 3σ criterion is used to eliminate the abnormal value firstly, then the wavelet denoising[19] and the moving average digital filtering method are used together to smooth the noise; data transformation is to normalize the input before modeling, in order to eliminate the influence of different dimension, the paper uses Max-min standard method[19].

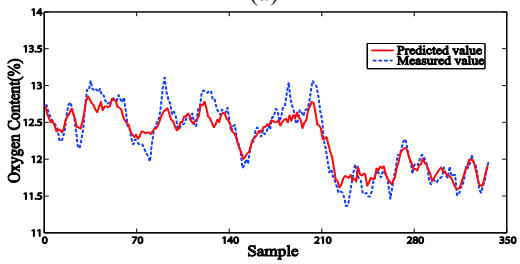
The main variables that affect the gas indices are a lot, including composition of coal (such as moisture, degree, ash, sulfur, etc.), charging quantity, gas flow(pusher and coke side) and its composition, heat value of gas, excess air factor, air volume, pressure of pusher and coke side, longitudinal temperature, rate of flow of raw gas and its composition, coking time etc. There are correlations between variables, if all are used to establish the prediction model, the computational complexity of modeling process will be increased, which would affect the accuracy and validity of model, so the principal component analysis (PCA) is used to process the sample data[20] to determine the most critical parameters affect SO_2 concentration, O_2 content and NO_x concentration of flue gas, in order to construct the auxiliary variables. In this paper, the cumulative variance contribution rate is 80%, taking NO_x concentration of flue gas as an example, the first 7 items of principal components are sulfur content of coal(%), rate of flow of gas($m^3 \cdot h^{-1}$), sub flue pressure of pusher side(K*Pa*), excess air factor, air volume($m^3 \cdot h^{-1}$), longitudinal temperature(℃) and rate of flow of raw gas($m^3 \cdot h^{-1}$), the cumulative variance contribution rate is 82.3%, therefore, the 7 variables are selected to form a new sample set as the auxiliary variable of NO_x concentration modeling.

The final 550 samples are randomly divided into 350 groups of training sample D_{tr}, 150 groups of validation sample D_{im}, and 50 groups of test sample D_{te}, the algorithm proposed in section 1.3 is used for modeling, the values of the parameters are: s=50, k_{max} =10000, E_{trmin} =0.5%, momentum factor α=0.05, the parameters related to learning rate are: b=1.2, q=0.15, d=0.15.

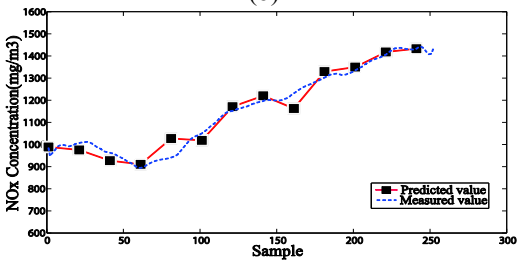
After the training of neural network, composited with mechanism model as the integrated model(M1) to predict the three indices. The results are shown in Figure 9, and the performance of the model is shown in table 1.



(a)



(b)



(c)

Fig.9 Flue gas indices prediction based on integrated model

Table 1 Modeling performance index

Index	Value		
	SO_2	O_2	NO_x
MRE	6.25%	1%	2.64%
Max-MRE	23%	4.03%	9.07%

The test samples are verified by the model, the average relative error of three indices are 8.7%, 3.2%, 7.66%, respectively.

In order to verify the performance of the proposed method, the following models are established: 1) The traditional RBF neural network is used as compensate model; 2) Only the optimal stopping method is adopted, and the RBFNN is established without the double momentum learning rate algorithm. The above models are recorded as M2 and M3, respectively, the training data is the union of training samples and testing samples, and the training termination condition of

M1 is k=10000 or the training error is less than 1%, M2 is the same as M1, other parameters are determined by repeated experiments. The prediction accuracy (MRE) of the two models to test samples as shown in table 2:

Table2 Comparison of model performance index

Gas Indice	Model	
	M2	M3
SO_2	11.8%	9.26%
O_2	6.25%	3.54%
NO_x	13.2%	7.48%

C. Results and analysis

It can be seen from table 1 and the results in training samples that the prediction accuracy of the three indices can meet the technological requirements by using the intelligent integrated model. Compared the training process time of M1, M2 and M3, it is shown that the proposed double momentum algorithm and the optimal stopping method both can reduce the training time, and as can be seen from table 2, the prediction accuracy of the model M2 for flue gas indices is obviously lower than that of M1 and M3, which shows that the optimal stopping method can avoid over-fitting and improve the generalization performance of the network.

IV. CONCLUSION

This paper takes the coking production process as the background, an intelligent integrated prediction model of SO_2 concentration, oxygen content and NO_x concentration is proposed, respectively. It can be seen from the model training and validation results based on data from two 55-hole and 6-meter top charging coke ovens in the coking plant, the prediction model established in this paper can accurately reflect the change trend of flue gas with a higher prediction accuracy, solving the problems of detection delay due to the change of working conditions of the upstream coking section in the process of desulfurization and denitration. Lays a foundation for the optimal operation of the desulfurization and denitration integrated device.

ACKNOWLEDGMENT

This work was financially supported by the National Natural Science Foundation of China (U1201251).

REFERENCES

[1] People's Republic of China environmental protection department. Coking chemical industry pollutant emission standards: GB 16171-2012. Beijing: China Environmental Science Press, 2012.

[2] WU X P, WANG X L, SONG Y H. Fuel & Chemical Processes, 47(2016) 31-34.

[3] WANG Y L. Intelligent integrated modeling theory and its application to optimal control for non-ferrous metallurgical process. Hunan: Central South University, 2001.

[4] GUI W H, WANG Y L, YANG C H, Journal of Measurement and Control, 40(2007) 176-181.

[5] YANG C H, GUI W H, KONG L S, et al, Computers and Chemical Engineering, 33(2009) 1289-1297.

[6] WANG Y L, GUI W H, YANG C H, et al. Control Theory & Applications, 26(2009) 861-866.

[7] WU Z W, CHAI T Y, WU Y J. ACTA AUTOMATICA SINICA, 39(2013) 2002-2011.

[8] SONG H D, ZHOU P, WANG H, et.al. ACTA AUTOMATICA SINICA, 42(2016) 1664-1679.

[9] YAO Z Z, ZHENG M D. Coking. Metallurgical industry press, Beijing, 2012.

[10] JI G X. Coal Chemical Industry, 42(2014) 35−38.

[11] QI W. Clean Coal Technology, 20(2014) 34-36,40.

[12] HAO J M, MA G D, WANG S X. Air Pollution Control Engineering. Higher Education Press, Beijing, 2010.

[13] ZHONG Y F, Fuel and chemical engineering, 40 (2009) 5-8,12.

[14] Zeldovich J, Acta Physiochimica USSR.Moscow, 21(1946) 577-628

[15] HARTMAN E J, KEELER J D, KOWALSKI J M, Neural Comput, 2 (1990) 210-215.

[16] PARK J, SANDBERG L W, Neural Comput, 3 (1991) 246-257

[17] ZHANG J Y, MA Z H, QIAN X B. ACTA AUTOMATICA SINICA, 38(2012) 1017-1024.

[18] YANG Q D, WANG F L, CHANG Y Q, Control and Decision, 23(2008) 869-874.

[19] Yaning Li, Xuelei Wang, Jie Tan. IEEE/ACIS International Conference on Computer and Information Science, Japan, 2016.

[20] GOOD R P, KOST D, CHERRY G A, IEEE Transactions on Semiconductor Manufacturing, 23(2010) 201-209.