Journal of Southeast University (English Edition) Vol. 35, No. 2, pp. 179 – 184 June 2019 ISSN 1003—7985

NO_x emission model for coal-fired boilers using partial least squares and extreme learning machine

Dong Ze[1,2] Ma Ning[1,2] Li Changqing[1]

([1]Hebei Engineering Research Center of Simulation and Optimized Control for Power Generation, North China Electric Power University, Baoding 071003, China)

([2]School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China)

Abstract: To implement a real-time reduction in NO_x, a rapid and accurate model is required. A PLS-ELM model based on the combination of partial least squares (PLS) and the extreme learning machine (ELM) for the establishment of the NO_x emission model of utility boilers is proposed. First, the initial input variables of the NO_x emission model are determined according to the mechanism analysis. Then, the initial input data is extracted by PLS. Finally, the extracted information is used as the input of the ELM model. A large amount of real data was obtained from the distributed control system (DCS) historical database of a 1 000 MW power plant boiler to train and validate the PLS-ELM model. The modeling performance of the PLS-ELM was compared with that of the back propagation (BP) neural network, support vector machine (SVM) and ELM models. The mean relative errors (MRE) of the PLS-ELM model were 1.58% for the training dataset and 1.69% for the testing dataset. The prediction precision of the PLS-ELM model is higher than those of the BP, SVM and ELM models. The consumption time of the PLS-ELM model is also shorter than that of the BP, SVM and ELM models.

Key words: NO_x emission; partial least squares; extreme learning machine; coal-fired boiler

DOI: 10.3969/j.issn.1003 – 7985.2019.02.006

With the increasing demand for environmental protection, reducing pollutant emissions has become an important and urgent problem to be solved for most coal-fired power plants in China. Nitrogen oxide (NO_x) is one of the main pollutants of coal combustion, and is also a contributor to global warming[1]. Combustion optimization technology has been proven to be an effective and economical method for the reduction of NO_x emissions from coal-fired boilers[2]. The reduction of NO_x emissions can be achieved by tuning the adjustable parameters of a boiler, such as coal feed quantity, excess air ratio, secondary air etc[3]. However, such optimization of operational parameters relies heavily on an accurate correlation between NO_x emissions and operational parameters. Therefore, a rapid and precise prediction model of NO_x emissions is required.

However, the generation mechanism of NO_x emissions is complex, and there is a serious coupling among each variable of a boiler, so it is difficult to set up an accurate mechanism model for real time control by mechanism modeling methods. Fortunately, the development of machine learning based on data drivers offers an alternative approach to constructing the NO_x emission model. Many researchers have paid attention to building the system model of NO_x emissions from coal-fired power plants. Wei et al.[4] used support vector regression (SVR), of which the parameters were optimized by the quantum genetic algorithm, to establish the NO_x emission model and obtained a good prediction result. Tan et al.[5] set up an efficient NO_x emission model based on the principle component analysis (PAC) and SVR. Although SVR has some disadvantages, such as difficulty in obtaining an optimum solution and costing much computational time[6], the artificial neural network (ANN), another computational intelligence-based method, was proposed to solve highly nonlinear problems. Ilamathi et al.[7] used the ANN to model NO_x emissions for tangentially fired boilers. However, the ANN suffers from some unsurmountable disadvantages, such as an abundance of controlling parameters, difficulty in obtaining a stable solution and risk of over-fitting[8]. The extreme learning machine (ELM) is a novel single hidden layer feed-forward network, which can overcome the drawbacks of the ANN to a certain degree. The ELM has an extremely fast learning algorithm and good generalization capability[9]. However, regarding the combustion process, there is a certain coupling relationship between some variables, such as coal feed, primary air flow, secondary air flow, and so on. If these variables are directly used as input factors of the model, it will inevitably lead to redundancy of the input information and affect the generalization ability of the model. In order to obtain an accurate NO_x emission model, it is necessary to eliminate the correlation between the variables and reduce the number of input variables before

Received 2018-11-30, **Revised** 2019-03-10.

Biography: Dong Ze (1970—), male, doctor, professor, dongze33@126.com.

Foundation items: The National Natural Science Foundation of China (No. 71471060), Natural Science Foundation of Hebei Province (No. E2018502111).

Citation: Dong Ze, Ma Ning, Li Changqing. NO_x emission model for coal-fired boilers using partial least squares and extreme learning machine[J]. Journal of Southeast University (English Edition), 2019, 35 (2): 179 – 184. DOI: 10.3969/j.issn.1003 – 7985.2019.02.006.

modeling. Partial least squares (PLS) is a statistical method for dealing with the problem of correlated inputs and reducing the number of input variables. The PLS technique has been successful in many fields[10-12].

According to the above analysis, this paper proposes a NO_x emission model of the power plant boiler based on the combination of PLS and ELM, namely, the PLS-ELM model. The procedure of establishing the proposed PLS-ELM model consists of two stages: First, the PLS method is applied to extract the feature of the variables; secondly, the extracted feature information is used as input of the ELM model. To validate the performance, the proposed PLS-ELM model is developed as an effective analysis model for a 1 000 MW coal-fired power plant. Simulation results show that, compared with the BP neural network, SVM and ELM models, the PLS-ELM model can achieve not only a much greater accuracy, but also shorter modeling time-consumption.

1 Basis of PLS and ELM

1.1 PLS

PLS is a multivariate statistical technique that can obtain the orthogonal characteristic vectors of the independent variables and dependent variables, by mapping the high-dimensional data space between the independent variable and dependent variable to the corresponding low-dimensional feature space. Compared with the PCA, PLS can not only effectively overcome the common linear problem, but also strengthen the interpretation of the independent variable to the dependent variable when selecting the eigenvector.

The data set is assumed to consist of an input (predictor) variable matrix $\boldsymbol{X} \in \mathbf{R}^{N\times m}$ and an output (response) variable matrix $\boldsymbol{Y} \in \mathbf{R}^{N\times 1}$, and they are both mean-centered and scaled by the standard deviation, where the index N represents the number of samples, and m represents the number of dimensions of the predictor variables. The main component $\boldsymbol{t}$ is extracted in $\boldsymbol{X}$. When extracting principal components, $\boldsymbol{t}$ is required not only to carry the variation information of $\boldsymbol{X}$ as much as possible, but also to maximize the correlation between $\boldsymbol{t}$ and $\boldsymbol{Y}$. After the first component $\boldsymbol{t}$ is extracted, the regression of $\boldsymbol{X}$ and $\boldsymbol{Y}$ to $\boldsymbol{t}$ is carried out respectively, and then the next principal component is extracted from the residual information. The partial least squares algorithm used to extract principal component is detailed as follows[13]:

1) Obtain the output score $\boldsymbol{u}$: $\boldsymbol{u} = \boldsymbol{Y}$.

2) Calculate the input weight $\boldsymbol{w}$, and then normalize $\boldsymbol{w}$ to unit length: $\boldsymbol{w}^{\mathrm{T}} = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{X}/(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u})$, $\boldsymbol{w} = \boldsymbol{w}/\|\boldsymbol{w}\|$.

3) Calculate the input score $\boldsymbol{t}$: $\boldsymbol{t} = \boldsymbol{X}\boldsymbol{w}$.

4) Calculate the input load vector $\boldsymbol{p}$: $\boldsymbol{p}^{\mathrm{T}} = \boldsymbol{t}^{\mathrm{T}}\boldsymbol{X}/(\boldsymbol{t}^{\mathrm{T}}\boldsymbol{t})$.

5) Calculate the regression coefficient of internal model b: $b = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{t}/(\boldsymbol{t}^{\mathrm{T}}\boldsymbol{t})$.

6) Compute the residual matrix: $\boldsymbol{E} = \boldsymbol{X} - \boldsymbol{t}\boldsymbol{p}^{\mathrm{T}}$, $\boldsymbol{F} = \boldsymbol{Y} - b\boldsymbol{t}$. Replace $\boldsymbol{X}$ and $\boldsymbol{Y}$ by $\boldsymbol{E}$ and $\boldsymbol{F}$, respectively, and repeat steps 1) to 6) until the required components are extracted or the results satisfy the precision requirement.

$\boldsymbol{T} = \{\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_A\}$ is the score matrix, which is the characteristic matrix of the samples, and is used as the input of the ELM model in this paper. $\boldsymbol{P} = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_A\}$ is the load matrix. $\boldsymbol{W} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_A\}$ is the coefficient matrix. A represents the number of extracted principal elements.

1.2 ELM

ELM is a novel single hidden layer feed-forward neural network[14]. The basic idea of ELM is to randomly assign the input weight values and hidden layer biases. Then, the ELM will become a linear system, of which the output weights can be analytically calculated by the least square method. Here, we simply review the learning algorithm of ELM.

Suppose that there are N training samples $(\boldsymbol{x}_i, \boldsymbol{t}_i)$, in which $\boldsymbol{x}_i = \{x_{i1}, x_{i2}, \ldots x_{in}\}^{\mathrm{T}} \in \mathbf{R}^n (i = 1, 2, \ldots, N)$, the n-dimensional feather vectors, are the input parameters of ELM and $\boldsymbol{t}_i = \{t_{i1}, t_{i2}, \ldots, t_{im}\}^{\mathrm{T}} \in \mathbf{R}^m$, and the m-dimensional target vectors, are the output parameters of ELM. The n and m are equal to the number of input layer nodes and the number of output nodes of the ELM model, respectively. Here, the number of hidden layer nodes of the model is l and the activation function is $g(\cdot)$, then the output of the ELM model can be calculated by the following form:

$$\sum_{i=1}^{l} \boldsymbol{\beta}_i g_i(\boldsymbol{x}_j) = \sum_{i=1}^{l} \boldsymbol{\beta}_i g(\mathbf{w}_i \cdot \boldsymbol{x}_j + b_i) = \boldsymbol{o}_j \qquad j = 1, 2, \ldots, N \tag{1}$$

where $\boldsymbol{\beta}_i = \{\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}\}^{\mathrm{T}}$ denotes the output weights vector, which connects the i-th hidden node with the output nodes. Simultaneously, $\boldsymbol{w}_i = \{w_{i1}, w_{i2}, \ldots, w_{in}\}^{\mathrm{T}}$ is the input weight vector, which connects the i-th hidden node with the input nodes. b_i represents the bias of the i-th hidden node. Previous studies have shown that the output value of the ELM model can be fitted to samples with zero error. So, a derivation equation can be obtained as follows:

$$\sum_{j=1}^{l} \| \boldsymbol{o}_j - \boldsymbol{t}_j \| = 0 \qquad j = 1, 2, \ldots, N \tag{2}$$

Eq. (1) can be transformed as

$$\sum_{i=1}^{l} \boldsymbol{\beta}_i g(\boldsymbol{w}_i \cdot \boldsymbol{x}_j + b_i) = \boldsymbol{t}_j \qquad j = 1, 2, \ldots, N \tag{3}$$

Eq. (3) can be simply written as

$$\boldsymbol{H}\boldsymbol{\beta} = \boldsymbol{T} \tag{4}$$

where

$$\boldsymbol{H}=\begin{bmatrix} g(w_1\cdot\boldsymbol{x}_1+b_1) & \dots & g(w_l\cdot\boldsymbol{x}_1+b_l) \\ \vdots & & \vdots \\ g(w_1\cdot\boldsymbol{x}_N+b_1) & \dots & g(w_l\cdot\boldsymbol{x}_N+b_l) \end{bmatrix}_{N\times l} \tag{5}$$

$$\boldsymbol{\beta}=\begin{bmatrix} \boldsymbol{\beta}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{\beta}_l^{\mathrm{T}} \end{bmatrix}_{l\times m}, \quad \boldsymbol{T}=\begin{bmatrix} \boldsymbol{t}_1^{\mathrm{T}} \\ \vdots \\ \boldsymbol{t}_N^{\mathrm{T}} \end{bmatrix}_{N\times m} \tag{6}$$

In the form, $\boldsymbol{H}$ is the hidden layer output matrix. In the training process of ELM, the input weight and bias values of ELM are generated randomly. Then, the output matrix $\boldsymbol{H}$ can be obtained, so that the ELM learning training problem is transformed into a least squares norm problem for solving the output weight, and that is

$$\hat{\boldsymbol{\beta}}=\boldsymbol{H}^{+}\boldsymbol{T} \tag{7}$$

where $\boldsymbol{H}^{+}$ is the Moore-Penrose generalized inverse of the hidden layer output matrix $\boldsymbol{H}$.

2 Modeling NO_x Using PLS-ELM

2.1 Boiler introduction and data preparation

A 1 000 MW ultra-supercritical variable pressure once-through boiler was chosen as the research object in this work. This boiler is 65.5 m in height and has a cross section of 32.08 m × 15.67 m. The boiler adopts the Π type arrangement, the single furnace, the improved low NO_x PM (pollution minimum) main burner, the MACT (Mitsuibishi advanced combustion technology) type low NO_x grading air supply combustion system and the reverse double tangent combustion mode. The water cold wall adopts the vertical rising mode of the inner threaded pipe.

In this work, 3 651 measurements encompassing approximately 61 h of boiler operation were obtained from the DCS (distributed control system) database of the power plant. The sampling interval was 1 min and it was proved by Smrekar et al. [15] to be the most suitable data-sampling interval for modeling NO_x emissions. Based on the basic knowledge of boilers and the engineers' experience, twenty-three variables, including unit load (one), pulverized coal flow rate (six), primary air flow rate of coal pulverized (six), OFA air damper opening percentage (two), secondary air damper opening percentage (six), total secondary air flow rate (one) and total air flow rate (one) were employed as inputs of the NO_x emission model, and the only output was the NO_x emissions. The variables employed and their ranges are listed in Tab. 1. The keynote of this paper is to establish the relationship between operational parameters and NO_x emissions, as well as demonstrating the validity of the hybrid PLS-ELM NO_x emission model, therefore, the influence factor of coal properties was ignored. Through investigation, the unit fluctuates within the range of 700 to 1 000 MW under normal operating conditions, so the sampled data can meet the requirement of NO_x emission modeling for the unit. In addition, the NO_x emission data obtained in this study is represented on a dry gas basis at 6% O_2.

Tab. 1 The range of each sample variable

Parameter	Range
Unit load/ MW	704.5 to 991.4
Pulverized coal flow rate (A-F)/($t\cdot h^{-1}$)	0 to 76.8
Primary air flow rate of coal pulverized/($t\cdot h^{-1}$)	0 to 166.2
OFA air damper opening percentage/%	12 to 100
Secondary air damper opening percentage/%	15 to 100
Total secondary air flow rate/($t\cdot h^{-1}$)	2 054.3 to 2 997.1
Total air flow rate/($t\cdot h^{-1}$)	2 779.7 to 3 852.5
NO_x emissions/ ($mg\cdot Nm^{-3}$)	135.5 to 275.1

2.2 Modeling process

In this study, PLS was used to extract the features of the input variables to reduce the number of the variable dimensions and the coupling of the input variables, since some variables, such as unit load, pulverized coal flow rate, total secondary air flow rate and total air flow rate, have strong coupling correlation. The feature matrix extracted by the PLS takes the place of the original input samples as input for the ELM model.

The modeling process can be made up of three steps, which are represented in Fig. 1.

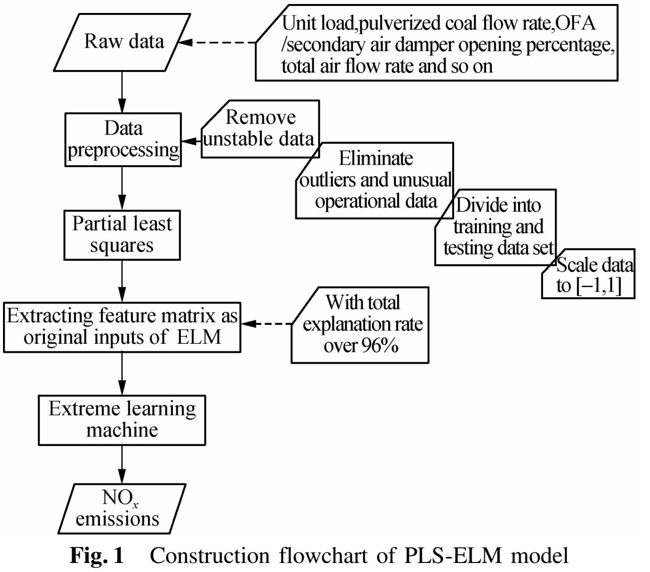

Fig. 1 Construction flowchart of PLS-ELM model

Step 1 Data preprocessing

Data preprocessing was further divided into four procedures before using sampled data to set up the hybrid PLS-ELM model. In the first step, in order to obtain data under steady state conditions, the unreliable data was removed. In the second step, all unusual and outliers operations were eliminated from the rest of the data. In the third step, the remaining data were divided into two groups, as illustrated in Fig. 2. The first group, which is referred to as the training set, was used for training the model. The second group, which is referred to as the testing set, was used for verifying the model. What should be emphasized is that the testing set has never been used in the model training phase. Therefore, the testing set was guaranteed as new to the model and contributed well to the reliability for proving the generalization capability of the model. In the fourth step, to eliminate the errors caused by the difference of dimension and range of variables, all data were scaled to [−1, 1].

Step 2 PLS execution

After data preprocessing, PLS was executed to obtain the latent correlation among the operational variables and lay down the foundation for reducing input variables.

Step 3 Extracting feature matrix as the original inputs of the ELM

Based on the result of the PLS, feature matrix $\boldsymbol{T}=\{\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_A\}$ was selected as the input variable for the ELM. The number A was determined based on the total explanation rate of the principal elements extracted. In this paper, the value of A is 5 with a total explanation rate greater than 96%. Moreover, the number of hidden layer nodes of the ELM model is 30.

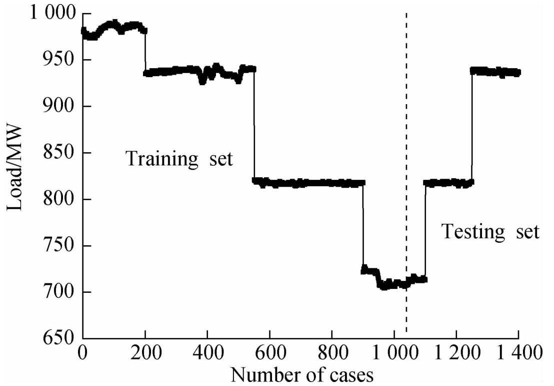

Fig. 2 Load dynamics of the sampled data

3 Results and Discussion

3.1 NO_x prediction using PLS-ELM model

The learning and generalization ability of the PLS-ELM model depends on the number of principal components of PLS since PLS is used to extract some principal components from candidate input variables. Fig. 3 shows the explanation rate of the first ten principal components, however, the remaining principal components' explanation rates are too low to be shown in the figure. In addition, it can be seen that the explanation rates of the first principal component is 74.3%, which indicates that the original inputs have a significant linear relationship with one another. Therefore, it is meaningful to employ PLS for data compression prior to the ELM modeling process. The explanation rates of the first five principal components is greater than 96%, which means that the variance interpretation information obtained by the extraction of the new principal component is small, so the residuals can be considered to be noise interference. If the number of principal components is added, the noise will be introduced and the complexity of the model is increased. Therefore, the first five principal components with a total explanation rate of over 96% were applied as inputs to the following ELM model.

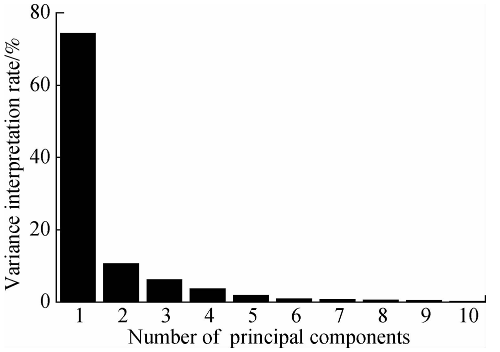

Fig. 3 Explanation rate of the principal components

The prediction result of the PLS-ELM NO_x prediction model and the measured values from the DCS database for the training dataset and testing dataset are shown in Fig. 4. It can be seen from Fig. 4 that the predicted value curve of the PLS-ELM model fits well with the training and testing dataset curves, indicating that the model has a good fitting and prediction ability. Furthermore, other three criteria are calculated. The root mean squared error (RMSE) is 3.733 for the training dataset and 3.817 for the testing dataset. In addition, the correlation coefficients are 0.994 and 0.992, and the mean relative errors (MRE) are 1.58% and 1.69%, respectively. The results demonstrate that the prediction accuracy of PLS-ELM model is

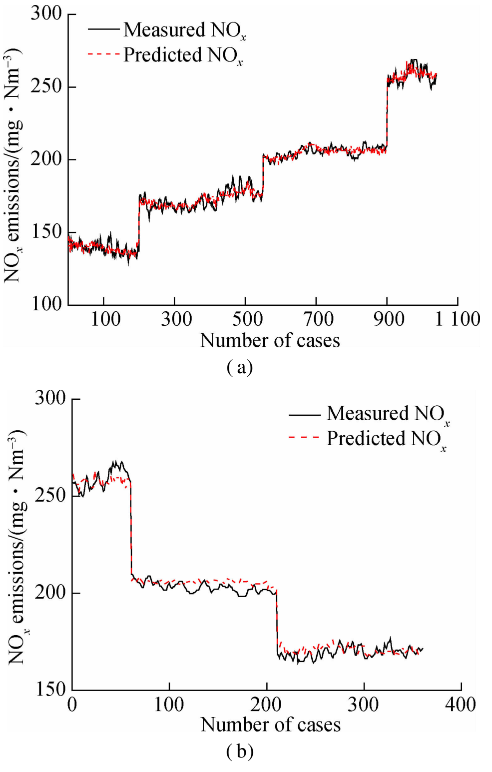

Fig. 4 Prediction results of NO_x emission. (a) Training dataset; (b) Testing dataset

high for both the training and testing datasets. Moreover, the model took 0.5 s for the entire modeling process on a personal computer, which indicates that the proposed PLS-ELM NO_x prediction model is relatively fast in setting up the model of NO_x emission for the coal-fired boiler.

3.2 Comparison of prediction performance between the BP neural network, SVM, ELM and PLS-ELM models

In order to further evaluate the performance of the PLS-ELM model, the BP neural network, SVM and ELM are employed to predict NO_x emissions in this paper. It should be noted that all the models used the same training and testing datasets.

The BP neural network uses gradient descent with momentum and adaptive learning rate back propagation to train the model. The number of hidden layer nodes of the BP neural network is the same as that of the PLS-ELM model. The RMSE of the BP model is 3.524 for the training dataset and 5.688 for the testing dataset. The consumption time of the total simulation process is 109 s.

In this work, the optimal parameters of the SVM model are determined by a grid search based on cross validation criterion. Specifically, two optimized parameters are the penalty coefficient c and the kernel parameter γ. The range of parameter c is $[2^{-2}, 2^{-1}, \ldots, 2^{11}, 2^{12}]$, and the range of kernel parameter γ is $[2^{-10}, 2^{-9}, \ldots, 2^{3}, 2^{4}]$. The RMSE value of the SVM model is 4.589 for the training dataset and 5.548 for the testing dataset, the corresponding MRE values are 1.86% and 2.24%. 362 s is required for the SVM to finish the complete simulation.

The traditional ELM is also applied for modeling the NO_x emissions. Unlike the PLS-ELM model, the data is directly applied to the ELM model without the process of feature extraction. The number of hidden layer nodes of the traditional ELM is the same as that of the PLS-ELM model. The traditional ELM model has a RMSE of 5.086 for the training dataset and 5.581 for the testing dataset. In addition, the corresponding MRE values are 2.13% and 2.29%, respectively. The time consumed by the simulation process is 0.8 s. For the training and testing datasets, the PLS-ELM model performs better than the traditional ELM model. This consequence may be a result of the majority of input variables and coupling between variables, which adds complexity to the model. So, the fitting and generalization capacity of the traditional ELM decrease. Therefore, it is necessary to use PLS to reduce the dimensions of the input before the ELM modeling process. Compared with the traditional ELM model, the PLS-ELM model is more suitable for establishing the NO_x emission model of the fired-boiler.

Fig. 5 shows the modeling errors of the testing dataset for each of the four models. It can be clearly seen that the modeling errors of the proposed PLS-ELM model are closer to zero compared to those of other models. The performance criteria of the models, such as RMSE, MRE and R-value on the testing dataset, as well as computation time, are listed in Tab. 2. As the four models use the same training and testing datasets, the prediction of the testing dataset is reliable for proving the models' generalization and prediction accuracy capability. The BP model is observed to have the worst performance among the four models. The prediction accuracy of the SVM model is slightly better than that of the ELM model, but not as good as the PLS-ELM model. Compared with the ELM model, the performance of the PLS-ELM model has been significantly improved, which again demonstrates the validity of PLS feature extraction for enhancing the generalization ability of the model. Moreover, the computation time of modeling process is also important for online combustion optimization. Significantly, the establishment of the PLS-ELM model is completed in 0.5 s, which is much less time compared with BP and SVM models. This is also an outstanding advantage of the ELM method. It can be concluded that the PLS-ELM model performs better not only in terms of its prediction accuracy but also in its computation time. The results in this work indicate that the PLS-ELM NO_x model is suitable for online reduction of NO_x emissions from the coal-fired boilers of power plants.

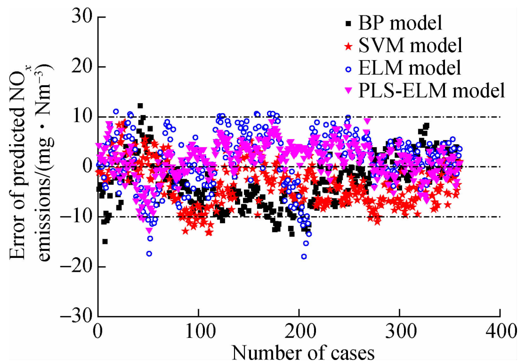

Fig. 5 The errors of testing dataset for various models

Tab. 2 The performance criteria of various models

Models	RMSE	MRE/%	R	Computation time/s
BP	5.688	2.33	0.978	109
SVM	5.548	2.24	0.985	362
ELM	5.581	2.29	0.982	0.8
PLS-ELM	3.817	1.69	0.992	0.5

4 Conclusions

1) The proposed PLS-ELM model applied the PLS to extract feature information with input variables, which can reduce the dimensions and coupling of the input vectors.

2) In order to verify this model's performance, a large amount of real-time data was obtained from the DCS database of a 1 000-MW coal-fired power plant. The simula-

tion results show that the proposed hybrid PLS-ELM model has a high predictive accuracy and good generation capacity.

3) The modeling performance of PLS-ELM was also compared to the BP neural network, SVM and ELM models, and the results demonstrate that the PLS-ELM is better than the other three models in the aspects of its predictive accuracy and computational cost.

References

[1] Molina A, Eddings E G, Pershing D W, et al. Char nitrogen conversion: Implications to emissions from coal-fired utility boilers[J]. *Progress in Energy and Combustion Science*, 2000, **26**(4/5/6): 507-531. DOI: 10.1016/s0360-1285(00)00010-1.

[2] Belošević S V, Tomanović I, Crnomarkovic N, et al. Modling and optimization of processes for clean and efficient pulverized coal combustion in utility boilers [J]. *Thermal Science*, 2016, **20**(suppl. 1): 183 − 196. DOI: 10.2298/tsci150604223b.

[3] Luo Z X, Wang F, Zhou H C, et al. Principles of optimization of combustion by radiant energy signal and its application in a 660 MWe down- and coal-fired boiler[J]. *Korean Journal of Chemical Engineering*, 2011, **28**(12): 2336 − 2343. DOI: 10.1007/s11814-011-0098-1.

[4] Wei Z B, Li X L, Xu L J, et al. Comparative study of computational intelligence approaches for NO_x reduction of coal-fired boiler[J]. *Energy*, 2013, **55**: 683 − 692. DOI: 10.1016/j.energy.2013.04.007.

[5] Tan P, Zhang C, Xia J, et al. NO_x emission model for coal-fired boilers using principle component analysis and support vector regression [J]. *Journal of Chemical Engineering of Japan*, 2016, **49**(2): 211 − 216. DOI: 10.1252/jcej.15we066.

[6] Wang L P. Support vector machines: Theory and applications [M] //*Studies in Fuzziness and Soft Computing*. New York: Springer-Verlag, 2005.

[7] Ilamathi P, Selladurai V, Balamurugan K, et al. ANN-GA approach for predictive modeling and optimization of NO_x emission in a tangentially fired boiler [J]. *Clean Technologies and Environmental Policy*, 2013, **15**(1): 125 − 131. DOI: 10.1007/s10098-012-0490-5.

[8] Lopes C, Perdigão F. Event detection by HMM, SVM and ANN: A comparative study[C]// *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer-Verlag, 2008. DOI: 10.1007/978-3-540-85980-2_1.

[9] Li G Q, Niu P F, Duan X L, et al. Fast learning network: A novel artificial neural network with a fast learning speed [J]. *Neural Computing and Applications*, 2014, **24**(7/8): 1683 − 1695. DOI: 10.1007/s00521-013-1398-7.

[10] Singh K P, Ojha P, Malik A, et al. Partial least squares and artificial neural networks modeling for predicting chlorophenol removal from aqueous solution[J]. *Chemometrics and Intelligent Laboratory Systems*, 2009, **99**(2): 150 − 160. DOI: 10.1016/j.chemolab.2009.09.004.

[11] Ronen D, Sanders C F W, Tan H S, et al. Predictive dynamic modeling of key process variables in granulation processes using partial least squares approach[J]. *Industrial & Engineering Chemistry Research*, 2011, **50**(3): 1419 − 1426. DOI: 10.1021/ie100836w.

[12] Huang Z Y, Yu Y L, Gu J, et al. An efficient method for traffic sign recognition based on extreme learning machine[J]. *IEEE Transactions on Cybernetics*, 2017, **47**(4): 920 − 933. DOI: 10.1109/tcyb.2016.2533424.

[13] Baffi G, Martin E B, Morris A J. Non-linear projection to latent structures revisited (the neural network PLS algorithm)[J]. *Computers & Chemical Engineering*, 1999, **23**(9): 1293 − 1307. DOI: 10.1016/s0098-1354(99)00291-4.

[14] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: Theory and applications [J]. *Neurocomputing*, 2006, **70**(1/2/3): 489 − 501. DOI: 10.1016/j.neucom.2005.12.126.

[15] Smrekar J, Potočnik P, Senegačnik A. Multi-step-ahead prediction of NOx emissions for a coal-based boiler[J]. Applied Energy, 2013, **106**: 89 − 99. DOI: 10.1016/j.apenergy.2012.10.056.

基于偏最小二乘和超限学习机结合的电站锅炉 NO_x 排放建模

董　泽[1,2]　马　宁[1,2]　李长青[1]

([1]华北电力大学河北省发电过程仿真与优化控制工程技术研究中心, 保定 071003)

([2]华北电力大学控制与计算机工程学院,北京 102206)

摘要:为了降低 NO_x 排放量,需要建立一个实时准确的 NO_x 排放模型. 提出了基于偏最小二乘(PLS)和超限学习机(ELM)相结合的 PLS-ELM 模型用于建立电站锅炉 NO_x 排放模型. 首先,根据机理分析确定 NO_x 排放模型的初始输入变量,然后,利用 PLS 对初始输入数据进行特征提取,最后,将提取后的信息作为 ELM 模型的输入. 利用某 1 000 MW 电站锅炉分散控制系统(DCS)历史数据库中的现场运行数据对 PLS-ELM 模型进行训练和验证,并将模型的性能与 BP 神经网络、SVM 和 ELM 模型进行了对比. PLS-ELM 模型对训练数据集和测试数据集的平均相对误差(MRE)分别为 1.58% 和 1.69%. 仿真结果表明:PLS-ELM 模型的预测精度和模型的耗时均优于 BP、SVM 和 ELM 模型.

关键词:NO_x 排放;偏最小二乘;超限学习机;电站锅炉

中图分类号:TK22