Korean J. Chem. Eng., 35(9), 1779-1790 (2018)
DOI: 10.1007/s11814-018-0087-8

pISSN: 0256-1115
eISSN: 1975-7220

Comparative study of estimation methods of NO_x emission with selection of input parameters for a coal-fired boiler

Beom Seok Kim, Tae Young Kim, Tae Chang Park, and Yeong Koo Yeo†

Department of Chemical Engineering, Hanyang University, Seoul 04763, Korea
(*Received 13 February 2018 • accepted 27 May 2018*)

Abstract−This study focuses on estimation of NO_x emission and selection of input parameters for a coal-fired boiler in a 500 MW power generation plant. Careful selection of input parameters is required not only to improve accuracy of the estimation, but also to reduce the model dimensionality. The initial operating input parameters are determined based on operation heuristics and accumulated operation knowledge; the essential input parameters are selected by sensitivity analysis where the performance of the estimation model is assessed as one or some input parameters are successively eliminated from the computation while all other input parameters are retained. From the sequential input selection process, less than ten input parameters survived out of 36 initial input parameters. Auto-regressive moving average (ARMA) model, artificial neural networks (ANN), partial least-squares (PLS) model, and least-squares support vector machine (LSSVM) algorithm were proposed to express the relationship between the operating input parameters and the content of NO_x emission. Historical real-time data obtained from a 500 MW power plant coal-fired boiler were used to test the proposed models. It was found that principal components analysis (PCA) enhances the estimation performance of each model. Among the four proposed estimation models, the LSSVM model coupled with PCA scheme showed the minimum root-mean square error (RMSE) and the best R-square value.

Keywords: NO_x Emission, Parameter Reduction, ARMA, ANN, PLS, LSSVM, PCA

INTRODUCTION

Coal-fired power plants are dominant in industries of electrical power generation and will remain as important power sources to meet the demand for electricity in the future. Coal-fired power plants accounted for 48.3% of Korea's electricity in 2016. However, coal is not a clean energy resource because its combustion introduces many pollutants into the ambient surroundings. Nitrogen oxides (NO_x) are significant pollutants that affect the global atmosphere. NO_x emission is responsible for damage to the environment, formation of photochemical smog, the depletion of stratospheric ozone, and greenhouse effects. Therefore, reduction of NO_x emission has been a worldwide industrial concern, and various methods to minimize NO_x emission have been proposed. Coal-fired power plants are confronted by more rigorous challenges on NO_x emission due to strict environmental regulations. To reduce NO_x emission from coal-fired power plants, many approaches, including low-NO_x burner, selective catalytic reduction (SCR), selective non-catalytic reduction (SNCR), and high temperature air combustion technique, have been proposed. Most of these methods are effective in the reduction of NO_x emission, but there exist some constraints in the construction of relative facilities, and initial capital costs are very high.

As an alternative, the method of optimization of plant operating parameters has been proposed. In this approach, the operating parameters are carefully adjusted by modeling NO_x emission and employing a suitable optimization scheme. This method is attractive because it is more cost efficient and implemented more easily compared to aforementioned methods. This method requires an accurate NO_x emission model that represents the relationship between the operating parameters and NO_x emission. Many researchers have contributed to the development of the NO_x emission model. Choi [1] analyzed characteristics of flow pattern, combustion, temperature, and NO_x emission in the 500 MW tangentially fired pulverized coal boiler to develop a mathematical model for NO_x emission. Fiveland [2] used data obtained from the low NO_x burner of a utility boiler to develop a NO_x emission model. During the past few decades, the artificial neural network (ANN) has been widely used to model NO_x emission for coal-fired power plants due to its excellent nonlinear mapping capacities. In addition to ANN, computational fluid dynamics (CFD) has been widely used to represent coal combustion and behavior of NO_x emission. Zhou et al. [3] used ANN to develop a NO_x emission model for a pulverized coal-fired boiler and compared their results with the model by CFD. They found that the ANN model was more intuitive and easier to implement compared to CFD models. The support vector machine (SVM) has been experiencing a booming development as another computational intelligence method, and has been widely used to model NO_x emission. In particular, the support vector regression (SVR) was proposed to supplement ANN and to provide alternative solutions for highly nonlinear problems. Zhou et al. [4] employed the SVR to develop a NO_x emission model and optimized it by using the particle swarm optimization (PSO) method to find out that the NO_x emission reduced up to 32.67% when the

†To whom correspondence should be addressed.
E-mail: ykyeo@hanyang.ac.kr

load was 312.08 MW. Zheng [5] used SVR to construct a NO_x emission model and compared various optimization methods including genetic algorithm (GA), PSO, ant colony optimization (ACO) algorithm to find that the performance of the ACO algorithm is the best among these three schemes. The least-squares support vector machine (LSSVM), which is an improved variant of SVM and inherits the advantages of SVM including substantial generalization and a unique solution, significantly reduces the training time compared to CFD method. Ahmed et al. [6] employed the LSSVM to develop a NO_x emission model and compared it with the model obtained by using partial least square (PLS) algorithm. Lv et al. [7] combined the LSSVM-based model with fuzzy c-mean and PLS scheme to enhance the tracking performance of the model. The principal component analysis (PCA) method can be effectively used to improve the performance of prediction models. The PCA raises the training speed and improves the accuracy by reducing the dimensionality of data. Ding et al. [8] simplified their model by applying PCA to 23 sample indices and used a back propagation neural network to construct the prediction model for river water pollution. Nomikos [9] used multiway PCA to project past data of a batch process onto the lower dimensional space and found that large data reduction was achieved and the monitoring performance was significantly improved.

The performance of NO_x emission models is greatly affected by the type of input data from the coal-fired boiler. Therefore, much attention has to be given to the proper selection of input parameters. The input parameters that have only a slight effect on the model performance not only increase the model complexity, but also reduce the model accuracy. In this work, for the purpose of comparison, auto regressive moving average (ARMA), ANN, PLS, and LSSVM were used to estimate the NO_x emission based on the key operating variables for a 500 MW coal-fired power plant boiler. The set of input parameters in each method was determined by iterative elimination of the parameters that have a negative, or only a slight effect, on the model outcome. PCA was then applied to reduce the dimensionality of data and to improve the modeling outcome.

PROCESS DESCRIPTION

The coal-fired boiler in the present study uses pulverized coal combustion process, which burns coal pulverized up to 200 mesh by injecting the coal particles with air into the combustion chamber. The pulverized coal in the pulverizer is transported to the boiler burner via preheated primary air. The combustion chamber is divided into six layers and the pulverized coal is fed to each layer. The secondary air compressed by fans and preheated by the air preheater is fed into the combustion chamber and is burned with pulverized coal particles. The amount of oxygen supply required in combustion is adjusted by each damper, and additional secondary air (over-fire air) is introduced to the upper part of the burner to prevent incomplete combustion. The over-fire air flow is fed to the front wall and the rear wall of burner. This flow is controlled by four over-fire air dampers. The after-combustion heavy ash from the boiler is collected in the hopper, and the fly ash and combustion gas are transported to the SCR through ID fan. Fig. 1 represents the flow diagram for the underlying coal-fired boiler.

MODELING OF NO_x EMISSION

1. Auto-regressive Moving Average (ARMA) Model

ARMA model can be used effectively to analyze historical time-series data and to predict future values. In an auto-regression model, a linear combination of past data is used in the estimation process. An auto-regressive process model of order p can be expressed as

$$X_t=\phi_1 X_{t-1}+\phi_2 X_{t-2}+\cdots+\phi_p X_{t-p}+e_t \quad (1)$$

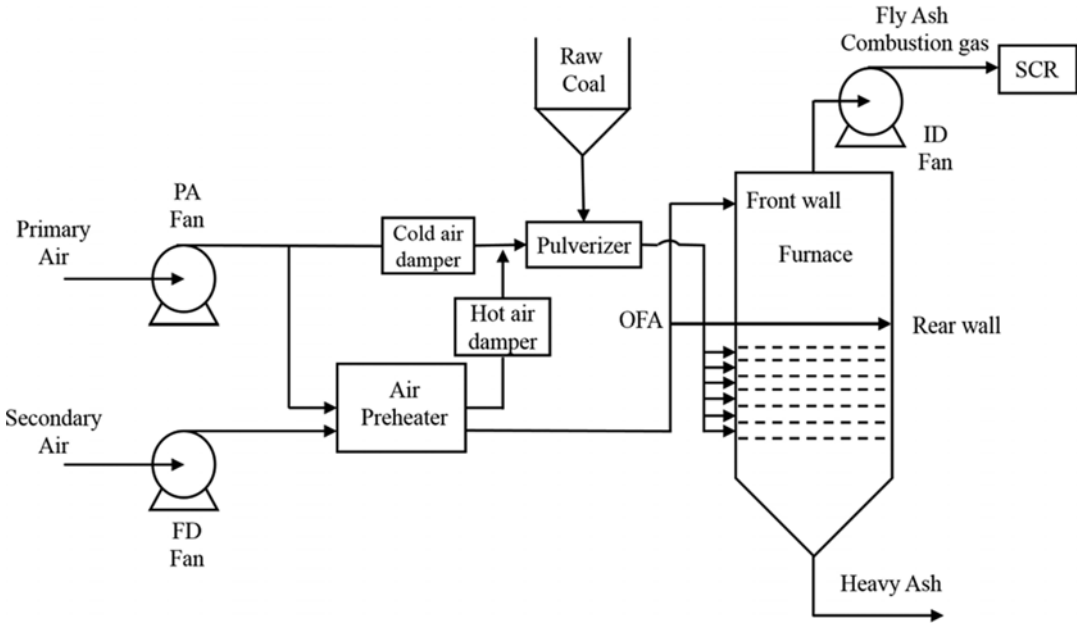

Fig. 1. Simplified process flow diagram of a coal-fired boiler.

where X_t, output of process model at time t, consists of past X values, ϕ is coefficient matrix, and e_t denotes modeling error. This model can be expressed as AR(p). Auto-regressive models are well known to be effective in treating a wide range of various time series patterns.

Rather than using past forecast variables, a moving average model employs past forecast errors.

$$e_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_q w_{t-q} \quad (2)$$

where w_t is white noise and θ is coefficient matrix. This model can be expressed as MA(q). X_t may be considered as a weighted moving average of errors of the past few forecasts. These two equations can be combined to give the model for X_t that can be expressed as ARMA (p, q) where p denotes the order of the autoregressive part and q denotes that of the moving average part. The ARMA model can be expressed as

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q} \quad (3)$$
$$w_t \sim WN(0, \sigma^2)$$

Eq. (3) can be represented using the backshift operator defined as $z^k X_t = X_{t-k}$. Using the z operator, Eq. (3) can be rewritten as

$$\phi(z) X_t = \theta(z) w_t \quad (4)$$

where

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$$
$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$$

This method is used to make linear equation applied to time series data.

2. Artificial Neural Network (ANN) Model

ANN represents the neural behavior of the human brain [10]. Compared to theoretical models, ANN establishes the relationships between inputs and outputs by training. A typical ANN is composed of neurons and transfer functions, and can be used to estimate the input-output relationship of a system. In general, ANN has single input layer, some hidden layers and single output layer. During the training process, the inputs are fed into the networks, multiplied by weight parameters, and summed to yield a node output which is transferred to the processing elements to produce an output. The input data are transported through the network and outputs are obtained from the output layers. Training process should be optimized to avoid over-training. Data used in the training and validation are divided into different sets: data set for training, data set for model validation, and data set to be used in the test. When input data is $x_i = (x_1, x_2 \ldots x_M)$ and output is $y_k = (y_1, y_2 \ldots y_N)$, y_k can be computed by Eq. (5).

$$y_k = F_O\left(\sum_{i=0}^{H} w_{ki} F_h\left(\sum_{j=0}^{M} w_{ji} x_i\right)\right) \text{ where } k=1, \ldots, N. \quad (5)$$

where H is the number of neurons in the single hidden layer and w is weight parameter of the each layer, and F_O and F_h are the output vectors from the output layer and the hidden layer, respectively. The model is trained with available data from a coal-fired boiler. When the training process is finished, the remaining data sets that were not used in the training are used to validate the model [11].

3. Partial Least-squares (PLS) Model

PLS methods involve an approximation of the input block, X, and output block, Y, and optimization of the correlation among data blocks [12,13]. The PLS model may be regarded as consisting of outer relations (individual X and Y blocks) and inner relations (linking both blocks). In the PLS model, the input and output blocks obtained from a coal-fired power plant are presented into the feature spaces with lower dimension. Projection of high dimensional blocks X and Y onto key factors (T, U) yields the most significant relationship among the feature vectors. The outer relations for the X and Y blocks can be represented as follows:

$$X = TP^T + E = \sum t_h p_h^T + E \quad (6)$$

$$Y = UQ^T + F = \sum u_h q_h^T + F \quad (7)$$

where P and Q are orthogonal matrices and E and F represent the errors. Each pair of latent variables (T and U) represents some of the variability contained in the input and output blocks. The first latent variables represent the variance in the data blocks, while the remaining latent variables represent noise contained in measurement and process. In the model building procedure, a threshold value for F is used [14]. The commonly used model to express the inner relation is a simple linear model given by

$$\hat{u}_h = b_h t_h \quad (8)$$

where $b_h = \hat{u}_h t_h / t_h^T t_h$ represents the coefficient in the linear regression. The quality variables are predicted by applying the PLS algorithm in the form of linear regression as follows [15]:

$$\hat{y}_h = x^T \hat{b} \quad (9)$$

4. Least-squares Support Vector Machine (LSSVM) Model

LSSVM has been widely applied in the modeling of nonlinear systems due to its ability of good generalization and low computational load [16]. If an individual learning subset L_k (k=1, …, T) consisting of data samples $\{x_i, y_i\}_{i=1}^n$ is given, the LSSVM model can be formulated as a typical optimization problem:

$$\min J(\omega, \xi) = \frac{1}{2}\omega^T \omega + \frac{1}{2}\gamma \sum_{i=1}^{n} \xi_i^2 \quad (10)$$

$$\text{s.t. } y_i = \omega^T \varphi(x_i) + b + \xi_i, \; i=1, \ldots, n$$

where φ denotes a nonlinear function that maps the input data into a space with higher dimension, b is the bias, ω is the weight vector, $\xi = [\xi_1, \ldots, \xi_n]^T$ is the error variable vector, and γ is the penalty factor. If data are noisy, smaller value of γ is preferred to avoid overfitting. Lagrange function and Karush-Kuhn-Tucker conditions are employed to solve the optimization problem. The final form of the LSSVM model to be used in the function estimation can be expressed as follows:

$$y(x) = \sum_{i=1}^{n} \alpha_i K(x, x_i) + b \quad (11)$$

where $\alpha = [\alpha_1, \ldots, \alpha_n]^T$ denotes the Lagrange multiplier, K is the kernel function used to represent the mapping procedure and to avoid computation of the function φ. In this study, the Gaussian radial basis function is used as K that is defined as follows:

$$K(x, x_i) = \exp(-\|x - x_i\|^2 / \sigma^2) \quad (12)$$

where σ is a tuning kernel parameter [7,17].

5. Principal Component Analysis (PCA)

PCA converts a data set containing p correlated variables to a new data set consisting of p uncorrelated orthogonal variables. These variables, referred to as principal components (PCs), are linear functions of the original variables. The sums of variances of principal components and original variables are identical with each other. Algebraically, for p original variables, $x_1, x_2, \ldots, x_p$, these can be represented as

$$PC1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p = \sum_{j=1}^{p} a_{1j}x_j \quad (13)$$

$$PC2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p = \sum_{j=1}^{p} a_{2j}x_j \quad (14)$$

$$\vdots$$

and so on for all p PCs. The variances in the PCs are represented by the eigenvalues, and the weights (a_{ij}) are represented by the eigenvectors obtained from the covariance matrix. The goal is that the first k principal components (k<<p) should retain most of the information contained in p original variables to reduce the dimensionality of the data. This means that if the correlations among the original variables are high, the first k PCs will represent most of the total variance and may be used to express multivariate patterns [18-21]. In this study, PCA is combined with other model type to enhance the tracking performance.

RESULTS AND DISCUSSION

Data points used in the simulations were obtained from the operation of 500 coal-fired boiler with the sampling time of 1 min. The outliers exhibiting severe fluctuation or erroneous behavior caused by either faulty sensors or human error were eliminated from the data set used in the simulations. As a result, 3300 data points were acquired over three consecutive days of operation. Among these operational data points, a total of 3000 data points were selected for model training, and the remaining 300 data points were used in the model validation process. As a basic case we selected 41 input parameters and chose the concentration of NO_x emission as the output variable (see Table 1).

The input parameters that did not change at all during the sampling period could be excluded from the simulations. From careful check, effects of five input parameters (fg_{O2}, $od_{A,B,C,D}$) on NO_x emission were found to be negative, and the remaining 36 input parameters were considered in the estimation of NO_x emission hereafter. After elimination of outliers, the Kalman filter was employed to pretreat data points followed by normalization process to constrict the scale of data points within the range of [0, 1]. In this study, the root-mean square error (RMSE), defined as follows, was used to assess the model performance:

$$RMSE = \sqrt{\frac{\Sigma(y_k - y_{ref})^2}{N}} \quad (15)$$

where y_k are the estimated NO_x emission, y_{ref} are the measured NO_x emission, and N is the number of data samples. Fig. 2 shows the results of NO_x emission for each model. The order of the ARMA model is 2, and the number of layers and nodes of the ANN model is 1 and 5, respectively. The number of principal components in the PLS model is determined by cumulative percentage. Only the principal components that make the cumulative percentage over 90% were selected. The number of principal components used to generate Fig. 2 was 26. In the LSSVM model, the penalty factor, γ, was set to 2 and the tuning kernel parameter, σ, was set to 2.7. Values of RMSE were 0.2979 for ARMA model, 0.4446 for ANN model, 0.1378 for PLS model, and 0.1360 for LSSVM model. Even though the LSSVM model displays the smallest RMSE, we can see that the tracking performance of this model is not satisfactory.

Input parameters with little or negative contribution to the output parameter (concentration of NO_x emission) can be excluded to avoid model complexity. The variation in the model accuracy for a specific set of operational data was analyzed as one or several input parameters are omitted from the simulations while retaining all other input parameters. For accurate analysis, only the input parameters are varied, and the data set related to model training and validation should remain unchanged for each set of input parameters. In this study, the value of RMSE obtained when all initial input parameters were employed was regarded as the base case. One input parameter was omitted and RMSE was calculated using the remaining input parameters. If the resultant RMSE was smaller than the base case, the omitted input parameter was excluded from further computations because the contribution of this parameter to the output parameter can be neglected. This process was repeated for

Table 1. List of input and output parameters used in the present study

Input parameters			
P: Power	(MW)	fg_{O2}: Flue gas O2 bias	(%)
Fr_{A-F}: Coal feeding rate (A-F)	(ton/hr)	TF: Total air flowrate	(Nm^3/sec)
$FD_{A,B}$: FD Fan blade control bias (A, B)	(%)	ph_{A-F}: Primary air hot damper position (A-F)	(%)
$Fz1_{A,B}$: FD Fan blade z-1 (A, B)	(%)	pc_{A-F}: Primary air cold damper position (A-F)	(%)
$Fz2_{A,B}$: FD Fan blade z-2 (A, B)	(%)	T_{A-F}: Pulverizer outlet temp (A-F)	(°C)
$Iz1_{A,B}$: ID Fan blade z-1 (A, B)	(%)	$od_{A,D}$: Over-fire air damper (front wall A, D)	(%)
$Iz2_{A,B}$: ID Fan blade z-2 (A, B)	(%)	$od_{B,C}$: Over-fire air damper (rear wall B, C)	(%)
Output parameters			
NO_x emission	(ppm)		

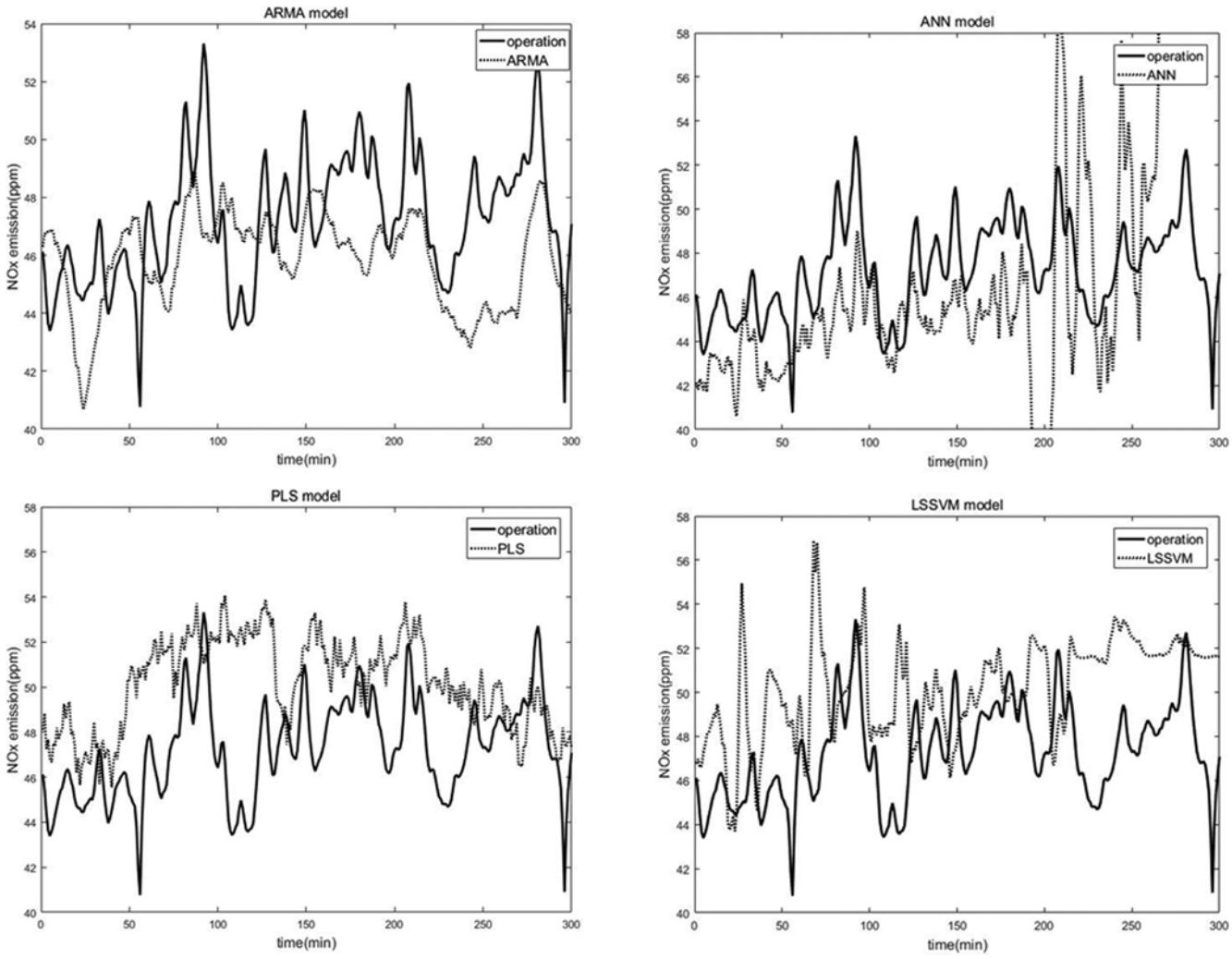

Fig. 2. Estimation of NO_x emission: base case.

all the input parameters using each model type until the RMSE no longer decreased.

For the ARMA model, RMSE for the estimation using 36 input parameters (base case) was 0.2979. One input parameter was omitted sequentially from 36 input parameters and RMSE was calculated for the estimation based on the remaining input parameters. Fig. 3(a) shows the value of RMSE for the base case (0.2979) and those for the 36 estimations with one input parameter omitted sequentially. We can see that the calculated RMSEs for 18 cases are less than the base case RMSE. This indicates that these 18 input parameters have negative effects on the content of NO_x emission. Next, the ARMA model was constructed again using the remaining 18 parameters except for the 18 parameters that reduced the RMSE. RMSE for the estimation using 18 input parameters (new base case) was 0.1293. Again, one input parameter was omitted sequentially from 18 input parameters and RMSE was calculated for the estimation based on the remaining 18 input parameters. Fig. 3(b) shows the RMSE for the new base case and 22 RMSEs for sequential estimations. Three parameters reduced the RMSE, which indicated that these input parameters could be excluded. When the ARMA model was reconstructed using the remaining 15 parameters, the calculated RMSE value was 0.1089, indicating improved estimation over previous cases. Further iteration showed no improvement over previous estimation, which means reduction of input parameters was no longer required. Thus the remaining 15 input parameters ($Fr_{C,F}$, $FD_{A,B}$, $Fz1_{A,B}$, $Fz2_{A,B}$, $Iz1_B$, ph_B, $pc_{A,B,D,E}$, T_D) were chosen as the input parameters for the ARMA model.

The RMSE for the initial base case ARMA model was reduced significantly when input parameters P, Fr_E, ph_C, ph_D and T_F were excluded sequentially. In the second iteration procedure, the input parameters reduced RMSE by 0.0185 and 0.0082, respectively. When the parameter pc_B was excluded in both the first and second iteration procedures, the increase in the RMSE was 0.0526 and 0.0542, respectively, which indicates that the parameter pc_B greatly influences NO_x emission.

For the ANN model, RMSE for the estimation using 36 input parameters (base case) was 0.4446. One input parameter was omitted sequentially from 36 input parameters and RMSE was calculated for the estimation based on the remaining input parameters. Fig. 4(a) shows the value of RMSE for the base case (0.4446) and

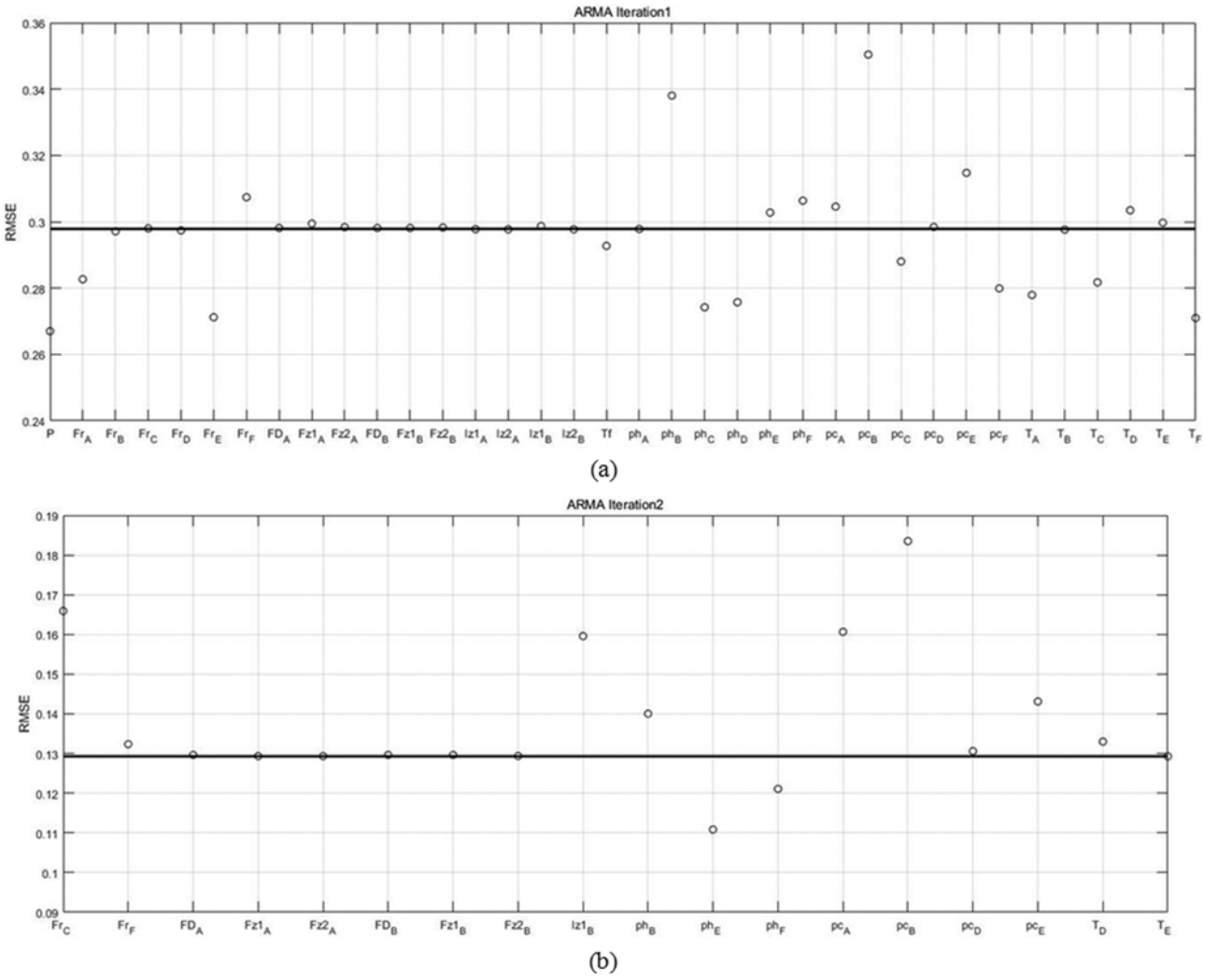

Fig. 3. Effects of the omission of each input parameter on the RMSE for ARMA model: (a) The first and (b) the second iteration.

those for the 36 estimations with one input parameter omitted sequentially. We can see that the calculated RMSEs for 14 cases are less than the base case RMSE. This indicates that these 14 input parameters have negative effects on the content of NO_x emission. Next, the ANN model was constructed again using the remaining 22 parameters except for the 14 parameters that reduce the RMSE. RMSE for the estimation using 22 input parameters (new base case) is 0.3399. Again, one input parameter was omitted sequentially from 22 input parameters and RMSE was calculated for the estimation based on the remaining 21 input parameters. Fig. 4(b) shows the RMSE for the new base case and 22 RMSEs for sequential estimations. It is observed that 14 parameters reduce the RMSE, which indicates that these input parameters can be excluded. When the ANN model was reconstructed using the remaining eight parameters, the calculated RMSE value was 0.1988, indicating improved estimation over previous cases. The same procedure was repeated on the new ANN model based on eight input parameters, and the results are shown in Fig. 4(c). Four input parameters turned out to be omissible from eight parameters. Further iteration showed no improvement over previous estimation, which meant reduction of input parameters was no longer required. Thus, the remaining four input parameters ($Iz2_A$, ph_C, pc_A, T_D) were chosen as the input parameters for the ANN model. The RMSE for the initial base case ANN model was the largest among the four models considered in this work, but the RMSE value was greatly reduced through iterative reduction procedure of input parameters. The prediction results based on the final four input parameters were the second best among the four estimation models. It is confirmed that minimal use of input parameters is one of the most significant advantages of the ANN model.

For the PLS model, RMSE for the estimation using 36 input parameters (base case) was 0.1906. One input parameter was omitted sequentially from 36 input parameters and RMSE was calculated for the estimation based on the remaining input parameters. Fig. 5(a) shows the value of RMSE for the base case (0.1906) and those for the 36 estimations with one input parameter omitted sequentially. We can see that the calculated RMSEs for 22 cases are less than the base case RMSE. This indicates that these 22 input parameters have negative effects on the content of NO_x emission. Next, the PLS model was constructed again using the remaining

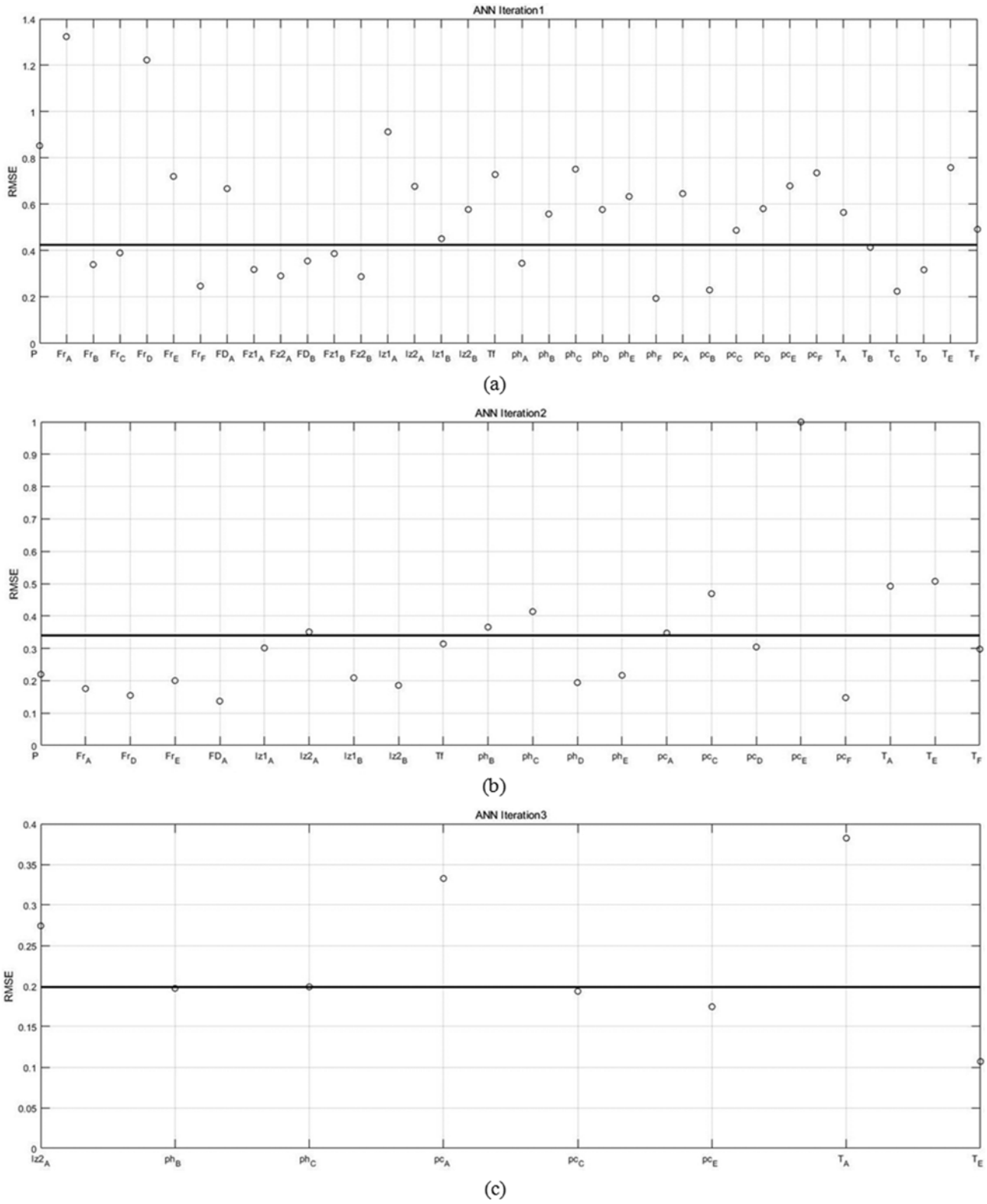

Fig. 4. Effects of the omission of each input parameter on the RMSE for ANN model: (a) The first, (b) the second, and (c) the third iteration.

14 parameters, except for the 22 parameters that reduce the RMSE. RMSE for the estimation using 14 input parameters (new base case) was 0.1293. Again, one input parameter was omitted sequentially from 14 input parameters and RMSE was calculated for the estimation based on the remaining input parameters. Fig. 5(b) shows the RMSE for the new base case and 14 RMSEs for sequential estimations. It was observed that four parameters reduced the RMSE, which indicated that these input parameters could be

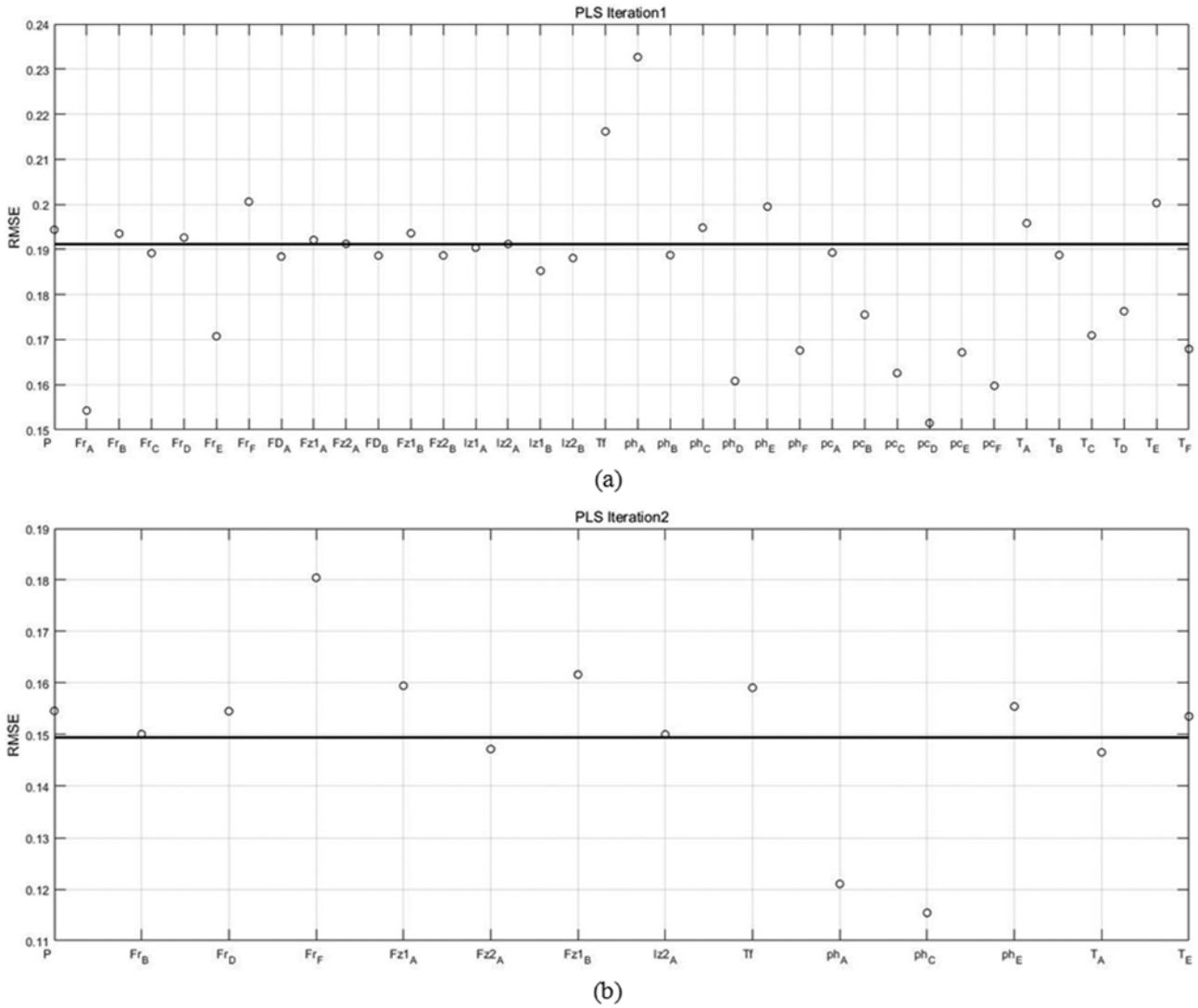

Fig. 5. Effects of the omission of each input parameter on the RMSE for PLS model: (a) The first and (b) the second iteration.

excluded. When the PLS model was reconstructed using the remaining ten parameters, the calculated RMSE value was 0.1276, indicating slightly improved estimation over previous cases. Further iteration showed no improvement over previous estimations, which meant reduction of input parameters was no longer required. Thus the remaining ten input parameters (P, $Fr_{B,D,F}$, $Fz1_{A,B}$, $Iz2_A$, TF, ph_E, T_E) were chosen as the input parameters for the PLS model.

The variation of the RMSE by excluding an input parameter was larger in the PLS model compared with the other models. The RMSE for the initial base case PLS model was reduced up to more than 0.03 when input parameters Fr_A, ph_D, pc_D and pc_F were excluded sequentially. In the second iteration procedure, the effects of the input parameters ph_A and ph_C on the NO_x emission were larger than other input parameters.

For the LSSVM model, RMSE for the estimation using 36 input parameters (base case) was 0.1360. One input parameter was omitted sequentially from 36 input parameters and RMSE was calculated for the estimation based on the remaining input parameters. Fig. 6(a) shows the value of RMSE for the base case (0.1360) and those for the 36 estimations with one input parameter omitted sequentially. The calculated RMSEs for 16 cases are less than the base case RMSE. This indicates that these 16 input parameters have negative effects on the content of NO_x emission. Next, the LSSVM model was constructed again using the remaining 20 parameters, except for the 16 parameters that reduce the RMSE. RMSE for the estimation using 20 input parameters (new base case) was 0.1152. Again, one input parameter was omitted sequentially from 20 input parameters, and RMSE was calculated for the estimation based on the remaining input parameters. Fig. 6(b) shows the RMSE for the new base case and 20 RMSEs for sequential estimations. It is observed that 13 parameters reduce the RMSE, which indicates that these input parameters can be excluded. When the LSSVM model was reconstructed using the remaining seven parameters, the calculated RMSE value was 0.1087, indicating slightly improved estimation over previous cases. Further iteration showed no improvement over previous estimations, which means reduction of input parameters was no longer required. Thus the remaining seven input parameters (P, Fr_E, TF, $ph_{B,D}$, pc_D, T_A) were chosen as the

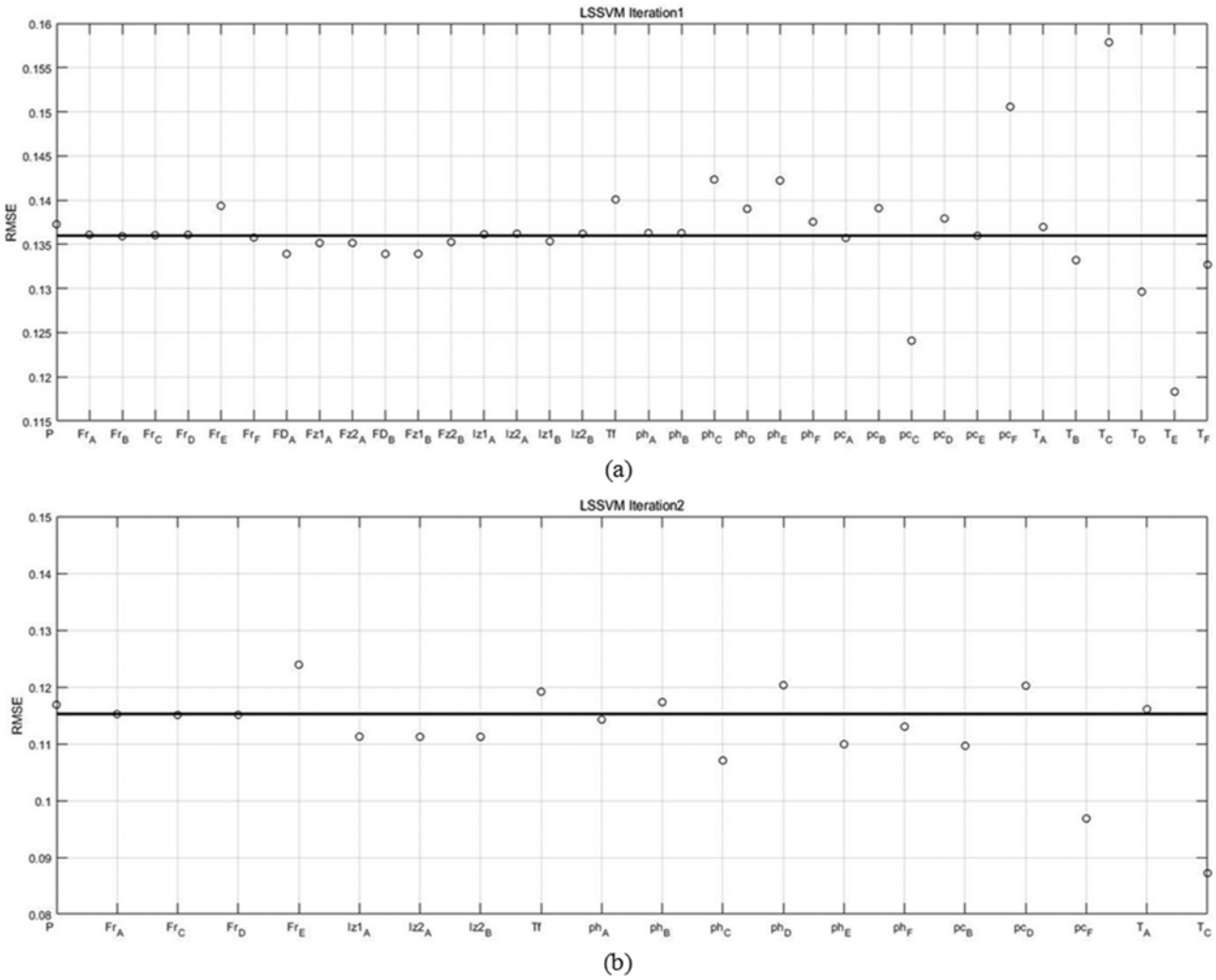

Fig. 6. Effects of the omission of each input parameter on the RMSE for LSSVM model: (a) The first and (b) the second iteration.

input parameters for the LSSVM model.

The LSSVM model exhibits the least base case RMSE among the four estimation models considered in the present study. The maximum decrease in RMSE was observed when the input parameter T_E was excluded. In the second iteration procedure, effects of the input parameters pc_F and T_E on the NO_x emission are larger than other input parameters.

After the iterative input reduction procedures to achieve the minimum RMSE, the number of final remaining input parameters was 15, 4, 10 and 7 for ARMA, ANN, PLS and LSSVM model, respectively. Table 2 shows the resultant input parameters surviving from the sequential input selection processes. The parameters that provide minimum RMSE are different for each model except some common input variables such as pc_A and T_D for ARMA and ANN models, Fr_F and $Fz1_{A,B}$ for ARMA and PLS models, ph_B and pc_D for ARMA and LSSVM models, and $Iz2_A$ for ANN and PLS models. It is confirmed that these are key input parameters that have an important effect on the NO_x emission.

Table 2. The resultant input parameters survived from the sequential selection processes

Model	Input parameters
ARMA	$Fr_{C,F}$, $FD_{A,B}$, $Fz1_{A,B}$, $Fz2_{A,B}$, $Iz1_B$, ph_B, $pc_{A,B,D,E}$, T_D
ANN	$Iz2_A$, ph_C, pc_A, T_D
PLS	P, $Fr_{B,D,F}$, $Fz1_{A,B}$, $Iz2_A$, TF, ph_E, T_E
LSSVM	P, Fr_E, TF, $ph_{B,D}$, pc_D, T_A

Using these key input parameters shown in Table 2, the NO_x emission was estimated using four models. Results of estimations are shown in Fig. 7. Values of RMSE are 0.1089 for ARMA model, 0.1279 for ANN model, 0.1276 for PLS model, and 0.1087 for LSSVM model. Values of RMSE are reduced significantly compared to the base cases shown in Fig. 2. Even with the reduced RMSE values, the tracking performance of each model is still not satisfactory.

As mentioned, PCA reduces dimensionality of input parameters and transforms the original feature data vector into the principal component space. Therefore, we can expect that combination of the estimation models with PCA may enhance the tracking performance of the model. Fig. 8 shows results of NO_x emission obtained from each model combined with PCA, and Table 3 shows the resul-

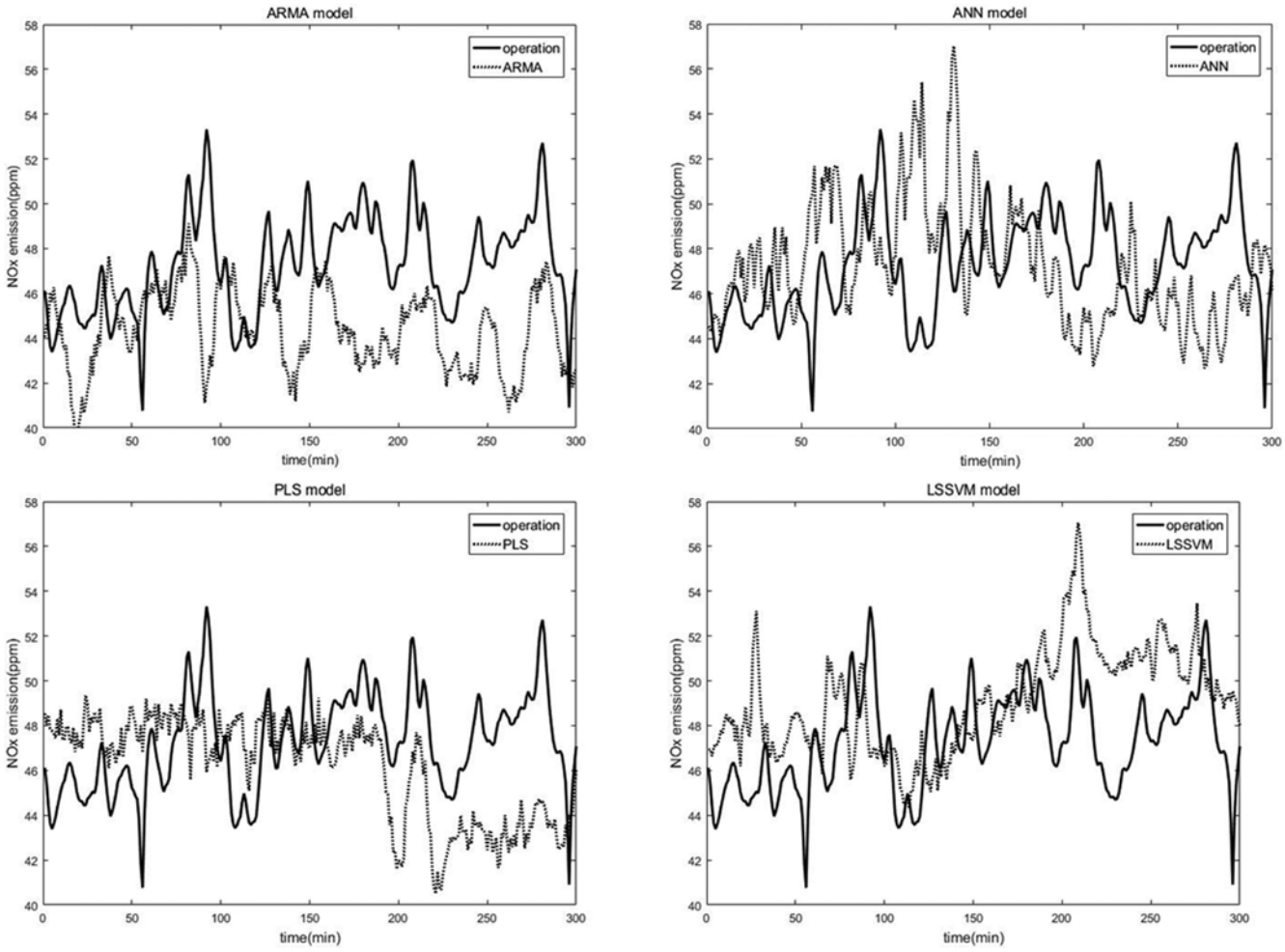

Fig. 7. Results of estimation of NO_x emission: Selected input parameters.

Table 3. The RMSE values for each model

Model	Base case	Selected parameters	Selected parameters combined with PCA
ARMA	0.2979	0.1089	0.1044
ANN	0.4446	0.1279	0.1184
PLS	0. 1378	0.1276	0.1205
LSSVM	0.1360	0.1087	0.1016

Table 4. R-squared values for each model

Model	Base case	Selected parameters	Selected parameters combined with PCA
ARMA	0.9646	0.9953	0.9954
ANN	0.9815	0.9939	0.9959
PLS	0.9929	0.9936	0.9942
LSSVM	0.9933	0.9958	0.9972

tant values of RMSE for each model.

To confirm the model fitting performance shown in Table 3, R-squared is employed as an additional criterion. The R-squared is defined as follows:

$$R^2 = 1 - \frac{SS_r}{SS_t} \left(SS_r = \Sigma_i (y_k - y_{ref})^2,\ SS_r = \Sigma_i (y_k - \bar{y})^2\right) \quad (16)$$

where $\bar{y}$ is the mean of the measured data. The closer the R-squared is to 1, the better is the fitting of the model. Values of R-squared for each case are shown in Table 4.

The value of RMSE for each model is reduced when combined with PCA and R-squared is close to 1. Though all models exhibit improved estimation results, there are differences in performance among the four models. The ANN model experienced the best improvement by reducing the RMSE by 73%. Among other models, the LSSVM model yields the least RMSE for base case, estimation using selected input parameters, and estimation combined with PCA. For all cases, the LSSVM model has the best RMSE and R-square values and shows enhanced fitting performance by using selected parameters and applying PCA. It is obvious that the LSSVM model combined with PCA shows the best prediction performance among the four models considered in the present study.

CONCLUSIONS

A comparative study on the estimation of NO_x emission and

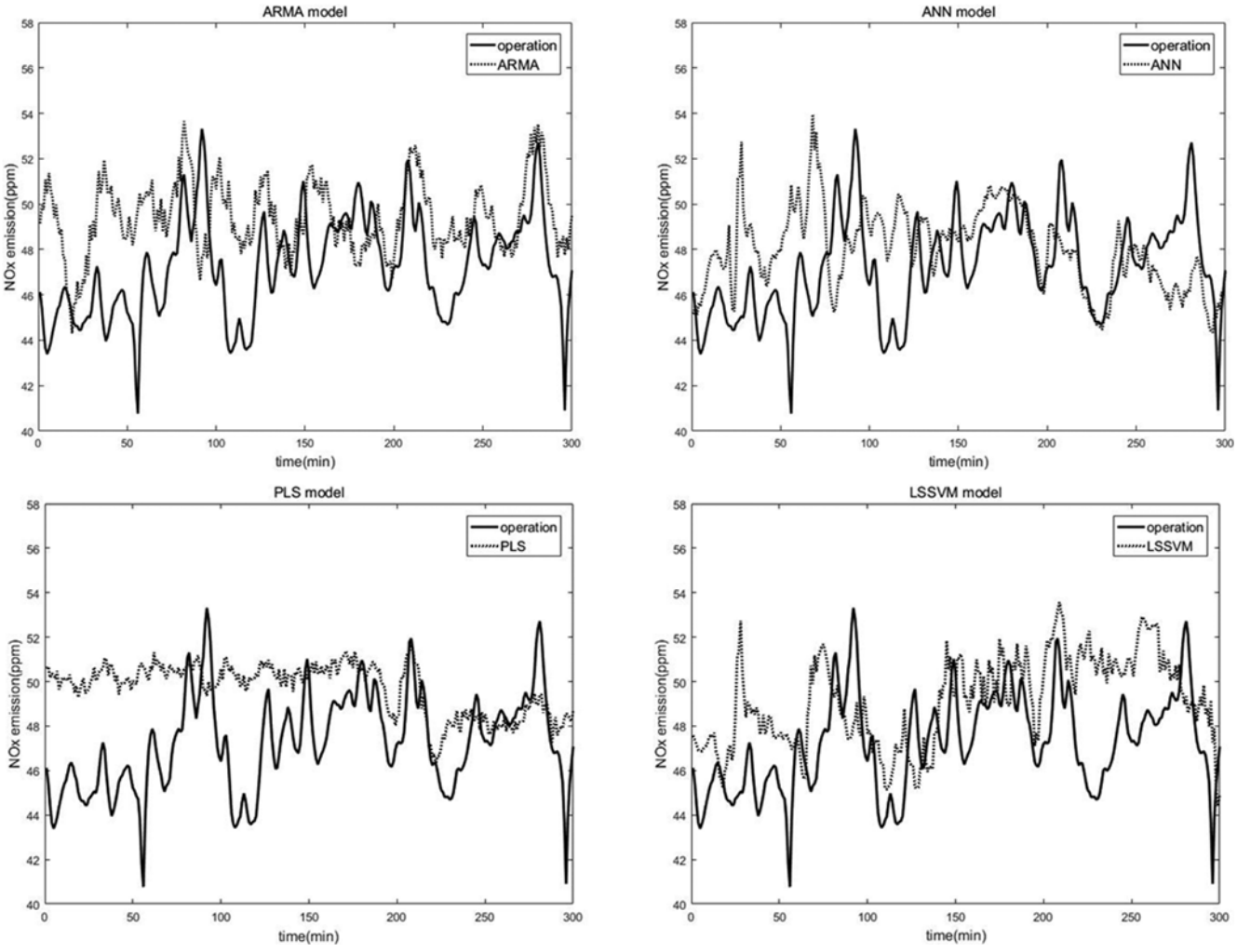

Fig. 8. Results of estimation of NO_x emission: Selected input parameters combined with PCA.

selection of input parameters was performed for a coal-fired boiler in 500 MW power generation plant. In the sequential selection procedure of input parameters, the input parameters that exert negative effects on the RMSE are excluded from further consideration and estimations are performed with remaining parameters. In the estimation of NO_x emission, ARMA, ANN, PLS and LSSVM models were used. Without input selection procedure, all models showed poor estimation performance due to the inclusion of input parameters with little or negative effects on NO_x emission while increasing complexity in the estimation. After the iterative input reduction procedures to achieve the minimum RMSE, the numbers of final remaining input parameters out of 36 initial parameters were 15 for ARMA model, 4 for ANN model, 10 for PLS model, and 7 for LSSVM model. The use of PCA enhanced the estimation performance of each model. Values of RMSE obtained from estimations with selected input parameters and coupled with PCA were 0.1044 for ARMA model, 0.1184 for ANN model, 0.1205 for PLS model, and 0.1016 for LSSVM model. The R-squared values were 0.0.9954 for ARMA model, 0.9959 for ANN model, 0.9942 for PLS model, and 0.9972 for LSSVM model. Among the four proposed estimation models, the LSSVM model coupled with PCA scheme showed the minimum RMSE and the best R-squared values.

ACKNOWLEDGEMENTS

This work was supported by Korea Research Foundation Grant funded by the Korean Government (NRF-2017R1A2B1005649).

REFERENCES

1. C. R. Choi and C. N. Kim, *Fuel*, **88**, 1720 (2009).
2. W. A. Fiveland and C. E. Latham, *Combust. Sci. Technol.*, **93**(1), 53 (1993).
3. H. Zhou, K. Cen and J. Fan, *Energy*, **29**, 167 (2004).
4. H. Zhou, L. Zheng and K. Cen, *Energy Convers. Manage.*, **51**, 580 (2010).
5. L. G. Zheng, H. Zhou, K. F. Cen and C. L. Wang, *Expert. Syst. Appl.*, **36**(2), 2780 (2009).
6. F. Ahmed, H. J. Cho, J. K. Kim, N. U. Seong and Y. K. Yeo, *Korean J. Chem. Eng.*, **32**(6), 1029 (2015).
7. Y. Lv, J. Liu, T. Yang and D. Zeng, *Energy*, **55**, 319 (2013).
8. Y. R. Ding, Y. J. Cai, P. D. Sun and B. Caen, *J. Appl. Res. Technol.*,

12, 493 (2014).
9. P. Nomikos and J. F. MacGregor, *AIChE J.*, **40**(8), 1361 (1994).
10. S. Haykin, *Neural networks: A comprehensive foundation (2nd Ed.)*, Prentice Hall, New Jersey (1999).
11. J. Smrekar, M. Assadi, M. Fast, I. Kustrin and S. De, *Energy*, **34**, 144 (2009).
12. P. Geladi and B. R. Kowalski, *Anal. Chim. Acta*, **185**, 1 (1986).
13. M. A. Sharaf, D. L. Illman and B. R. Kowalski, *Chemometrics*, Wiley, New York (1986).
14. G. Baffi, E. B. Martin and A. J. Morris, *Comput. Chem. Eng.*, **23**(3), 395 (1999).
15. T. Y. Kim, B. S. Kim, T. C. Park and Y. K. Yeo, *Korean J. Chem. Eng.*, **34**(7), 1952 (2017).
16. J. A. K. Suykens and J. Vandewalle, *Neural Process Letters*, **9**(3), 293 (1999).
17. T. van Gestel, J. A. K. Suykens, B. Baesens, S. Viaene, J. Vanthienen, G. Dedene, B de Moor and J. Vandewalle, *Mach Learning*, **54**(1), 5 (2004).
18. R. L. Olsen, R. W. Chappell and J. C. Loftis, *Water Res.*, **46**, 3110 (2012).
19. G. H. Dunteman, *Principal components analysis*, Sage University Paper Series on Quantitative Applications in the Social Sciences, California (1999).
20. H. Jeong, S. Cho, D. Kim, H. Pyun, D. Ha, C. Han, M. Kang, M. Jeong and S. Lee, *Int. J. Hydrogen Energy*, **37**, 11394 (2012).
21. J. Zeng, K. Liu, W. Huang and J. Liang, *Korean J. Chem. Eng.*, **34**(8), 2135 (2017).