

Article

A Dynamic Model Incorporated with Delay Estimation and Variable Selection for Control of SCR De-NO_x Process

Ze Dong ¹, Ling Li ², Laiqing Yan ^{3,*}, Ming Sun ^{2,*} and Jinsong Li ⁴

¹ Hebei Technology Innovation Center of Simulation & Optimized Control for Power Generation, North China Electric Power University, Baoding 071003, China; dongze@ncepu.edu.cn

² School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; liling@shiep.edu.cn

³ School of Electric Power and Architecture, Shanxi University, Taiyuan 030006, China

⁴ Hebei Guohua Dingzhou Power Generation Co., Ltd., Dingzhou 073000, China; p0000463@chnenergy.com.cn

* Correspondence: yanlq@sxu.edu.cn (L.Y.); sunming@ncepu.edu.cn (M.S.); Tel.: +86-138-34685-455 (L.Y.)

Received: 29 October 2020; Accepted: 21 November 2020; Published: 23 November 2020



Abstract: In order to control NH₃ injection for the selective catalytic reduction of nitrogen oxide (NO_x) denitration (SCR de-NO_x) process, a model that can accurately and quickly predict outlet NO_x emissions is required. This paper presents a dynamic kernel partial least squares (KPLS) model incorporated with delay estimation and variable selection for outlet NO_x emission and investigated control strategy for NH₃ injection. First, k-nearest neighbor mutual information (KNN_MI) was used for delay estimation, and the effect of historical data lengths on KNN_MI was taken into account. Bidirectional search based on the change rate of KNN_MI (KNN_MI_CR) was used for variable selection. Delay-time difference update algorithm and feedback correction strategy were proposed. Second, the NH₃ injection compensator (NIC) and the outlet NO_x emission model constituted a correction controller. Then, its output and the output of the existing controller are added up to suitable NH₃ injection. Finally, the KNN_MI_CR method was compared with different algorithms by benchmark dataset. The field data results showed that the KNN_MI_CR method could improve model accuracy for reconstructed samples. The final model can predict outlet NO_x emissions in different operating states accurately. The control result not only meets the NO_x emissions standard (50 mg/m³) but also keeps high de-NO_x efficiency (80%). NH₃ injection and NH₃ escape are reduced by 11% and 39%.

Keywords: selective catalytic reduction nitrogen oxide denitration (SCR de-NO_x) process; dynamic kernel partial least squares (KPLS) model; variable selection; delay estimation; control strategy

1. Introduction

How to reduce nitrogen oxide (NO_x) emissions from coal-fired power plants has become an important task of environmental treatment. Selective catalytic reduction denitration (SCR de-NO_x) technology is one of the mature flue gas denitration technologies [1]. Many efforts have been made to optimize the SCR system, such as superior catalyst technology, enhanced de-NO_x control strategy, and byproduct limitation [2]. However, it is very difficult to control the amount of NH₃ injection accurately. First, new energy is integrated into the power grid on a large scale in recent years. The boiler load needs to be changed to stabilize the fluctuation from the output power of new energy frequently. Therefore, the wide range of operating states poses certain difficulties in NH₃ injection control. Second, the upper limitation of NO_x emission is decreased to 50 mg/m³ in China. In order to meet the national environmental standard, excessive NH₃ injection occurs easily. Hence, it is required

to control NO_x emissions within the upper limitation and to avoid NH₃ escape. Third, the SCR de-NO_x system has large inertia, and it also has a large delay in NO_x analysis. When the operating state changes, the SCR system cannot adjust NH₃ injection in time based on the sampling inlet NO_x concentration, which can make the outlet NO_x emission deviate from the set value. Furthermore, the SCR de-NO_x process, which has the feature of nonlinear and multivariable coupling, is complex. When boiler load is stable, outlet NO_x emission can fluctuate widely for the reasons of flue gas temperature, flue gas flow rate, NH₃ injection and oxygen content. If boiler load changes, outlet NO_x emission changes dramatically. The proportion integration differentiation (PID) cascade control cannot achieve optimal control, which ensures high de-NO_x efficiency and low NH₃ escape. Therefore, the establishment of an accurate SCR model and the research of optimal control strategies are significant.

The current control strategy mainly uses a transfer function to control NH₃ injection. Considering the strong coupling between boiler parameter and SCR process parameter, the transfer function is difficult to reflect SCR de-NO_x process accurately. With the application of a plant-level supervisor information system, operating data can be viewed and recalled easily. Therefore, data-driven modeling and control methods are widely used. Data-driven modeling methods, such as Hammerstein–Wiener [3], radial basis function autoregressive exogenous [4], online support vector regression [5], neural network, nonlinear autoregressive exogenous and polynomial fitting [6], have been used to predict NO_x emission. Although most models are applied for diesel engines, the reaction principle is the same as the SCR de-NO_x system of a coal-fired power plant. The differences are that the boiler parameter affects outlet NO_x emissions, and the NO_x analyzer has a large lag in the coal-fired power plant. In addition, the advanced thermal management control strategies of catalysts have been developed to reduce emission concentrations [7].

Variable selection and delay estimation are investigated to simple model structure and improve model accuracy. Mutual information (MI), which denotes statistical dependence between any two variables, is a measure of variable correlation [8]. Therefore, it can express a linear and nonlinear correlation [9]. In order to improve the model accuracy, Stanišić et al. [10] used the MI method to select input variables. However, how to select an evaluation strategy and how to estimate the MI value accurately are two main difficulties in variable selection by the MI method. In their evaluation strategy, Battiti et al. proposed the mutual information feature selection (MIFS) method [11], which got a good result. After this, Peng et al. [12] and Fleuret [13] proposed a different evaluation strategy. The above methods mainly use one-dimensional MI to evaluate the variable set. When there are redundant variables, the subset of the optimal variables cannot be obtained. The literature [10] used a single evaluation function as an indicator, which often leads to an imbalance between the relevant variable and redundancy variables. Considering the large time-delay process, the effect of time-delay needs to be considered in variable selection. Ludwig et al. [14] used the MI method to estimate the time-delay and got better results in a case. However, the method ignores the effect of variable dimension on the MI value. In general, output variable $y(t)$ not only related to input variable $x_i(t)$ at time t , but also related to historical input $x_i(t-1)$, $x_i(t-2)$, ..., $x_i(t-l_i)$ before time t . Here, l_i is defined as the historical data length of the input variable x_i . Therefore, it is necessary to determine the historical data length of each input variable when the time-delay is estimated.

Recently, adaptive nonlinear model predictive controllers, such as model-based predictive controller [15], the generalized predictive controller [16], dynamic matrix controller [17], nonlinear autoregressive moving average with exogenous inputs (NARMAX) based predictive controller [18], the neural network-based predictive controller [19] and subspace predictive controller [20], have been used to control the NH₃ injection and got better effect compared with the conventional PID controller. However, some models (e.g., [15,16]) involved high order nonlinear optimization, which is time-consuming and gets a local minimum rather than a global minimum easily. Not all of the operating states are covered in the NARMAX prediction model. The neural network prediction model is difficult to be predicted online, quickly and accurately. The input variable of the subspace predictive model only relates to NH₃ injection. Hence, it cannot reflect other variables that affect outlet NO_x emission correctly. In addition, the control effect is also affected by the constraints,

such as the NH₃ valve opening and the increment of NH₃ valve opening, which are neglected in the above literature.

The objective of this paper is to model outlet NO_x emissions under the different operating state, thereby to facilitate controller design. The kernel partial least squares (KPLS) model shows a good effect when it deals with the high dimensionality and colinearity data. In this paper, a dynamic KPLS model incorporated with variable selection and delay estimation is proposed. The k-nearest neighbor mutual information (KNN_MI) was used for delay estimation, and the effect of historical data lengths on KNN_MI was taken into account. Based on the results of delay estimation, a bidirectional search based on the change rate of KNN_MI (KNN_MI_CR) was used for variable selection. In order to improve model update accuracy, delay–time difference (DTD) update algorithm and feedback correction strategy are proposed. In order to take advantage of the existing controller, which has rapid response-ability to deal with disturbance signals, a parallel control structure that can identify the dynamic model online by using operation data is proposed. The existing conventional controller is used as the main controller. The NH₃ injection compensator (NIC) combined with the outlet NO_x emission model are used as correction controller, which can ensure the outlet NO_x emission is stable near the set value. The principle of NIC is to eliminate the adverse effects of the large inertia of the SCR system and the lag of NO_x analysis through feed-forward compensation to a certain extent.

The paper is organized as follows: In Section 2, the SCR de-NO_x process and its control strategy are described. In Section 3, the delay estimation and variable selection method are presented. The experimental results of the benchmark dataset for variable selection are presented in Section 4. The field data experiment for the SCR de-NO_x process is investigated in Section 5, and the predictive accuracy of the outlet NO_x emission model and the control effect of the proposed control strategy are analyzed. Finally, concluding remarks are given in Section 6.

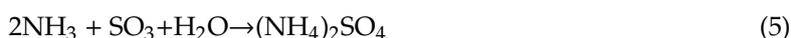
2. The SCR de-NO_x Process and Its Control Strategy

2.1. The SCR de-NO_x Process

For a 1000 MW ultra-supercritical coal-fired power plant, the SCR de-NO_x system is shown in Figure 1. The temperature of the SCR de-NO_x system is controlled by the bypass gas damper of the economizer in the upstream flue. Mixed with dilution air, the ammonia gas is sprayed by the NH₃ flow rate valve through the ammonia injection grid (AIG). In the reaction zone, the selective catalytic reaction occurs on the stainless steel plates, which are supported by the V₂O₅-WO₃-TiO₂ catalyst. Finally, the NO_x in the flue gas is converted into harmless N₂ and H₂O, and thereby the flue gas is purified. The main reaction of the SCR de-NO_x process is described as follows [1]:



Actually, adverse side reaction will occur as follows [1]:



It can be seen that the excessive NH₃ injection leads to side reaction, and de-NO_x efficiency is decreased because of the regeneration of NO_x. Excess NH₃ reacts with SO₃ in the flue gas to form NH₄HSO₄ and (NH₄)₂SO₄, which not only reduce the catalyst activity but also cause plugging and corrosion in the heated surface of the air preheater. In addition, increased NH₃ escape could lead to secondary pollution. Secondary pollution refers to new pollutants, that is, unreacted ammonia gas.

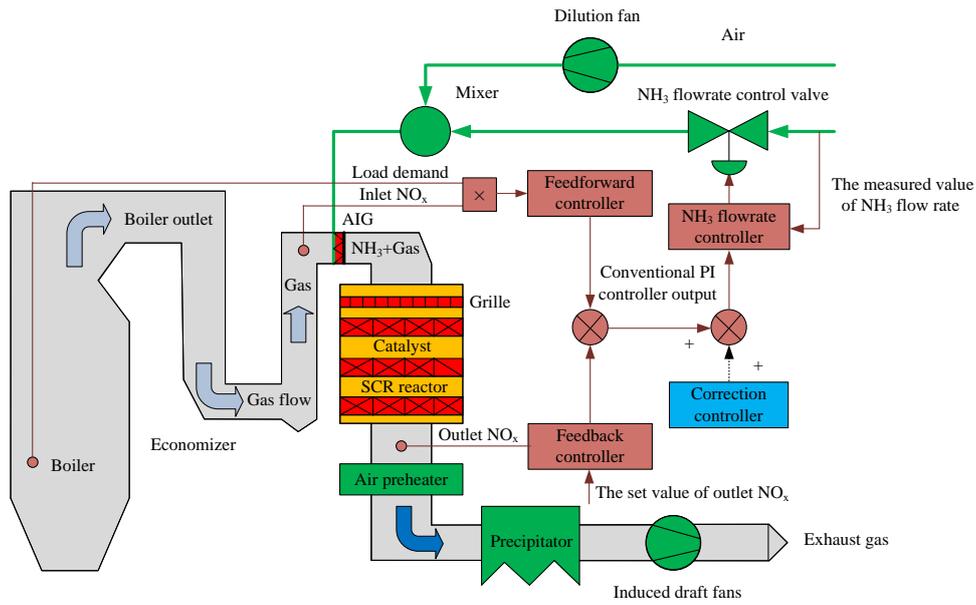


Figure 1. The selective catalytic reduction nitrogen oxide (NO_x) denitration (SCR de-NO_x) system.

2.2. The SCR de-NO_x Process Control Strategy

Under the variable state of the power plant, the SCR de-NO_x process shows nonlinear, strong disturbance and large time-delay in NO_x analysis, so it is hard to achieve a good control effect by the existing PID control. Actually, except for controlling the temperature of the reaction zone by the bypass flue gas, it is necessary to design a better automatic control system which can control the NH₃ injection strictly. In order to ensure the most suitable NH₃/NO_x ratio, which is related to de-NO_x efficiency and ensure the NH₃ escape in a reasonable range, the proposed control system includes the main controller and the correction controller, which are in parallel operation. The main controller is the existing PID controller. The correction controller, which is shown in Figure 1, is investigated in this paper. The parallel control system is shown in Figure 2. The dynamic KPLS model is used as an outlet NO_x emission predictive model, which can be identified offline by the field data.

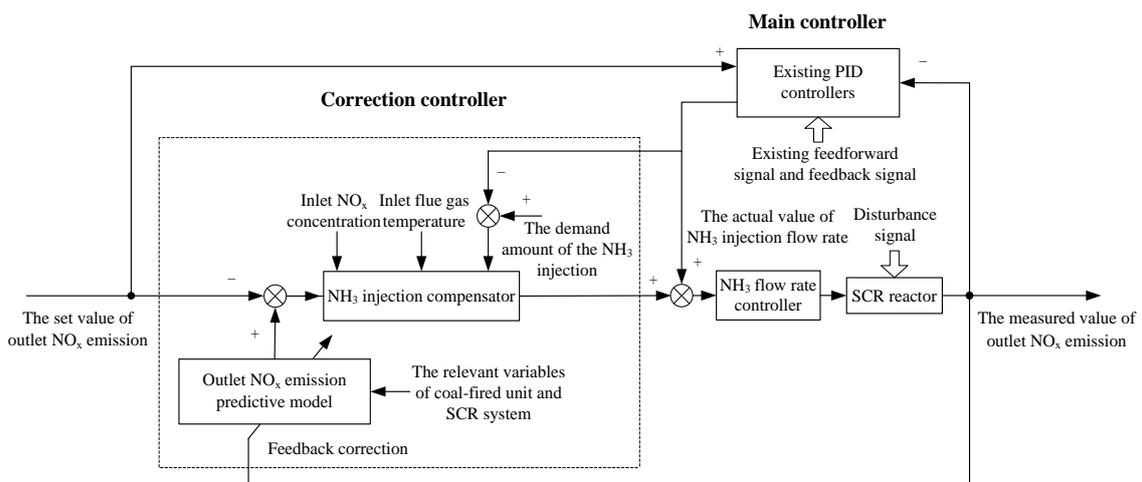


Figure 2. The SCR de-NO_x process control system.

The correction controller includes three parts mainly: outlet NO_x emission predictive model, feedback correction and NH₃ injection compensator (NIC).

- (1) Outlet NO_x emission predictive model

In general, a dynamic data-driven model is constructed by adding the model input, such as the historical input data $x(t-1), \dots, x(t-w+1)$ and output data $y(t-1), \dots, y(t-w+1)$. Therefore, $x(t-\tau)x(t-\tau-1)$ and $y(t-1)$ are selected as the new input variables in this paper. As a kind of model update method, the time difference (TD) algorithm can work out the variable drift and achieve better predictive accuracy. Moreover, the data model based on the TD algorithm does not need frequent reconstruction and parameter updating [21]. The time-delay must be considered in the SCR deNO_x process. Hence, the delay-time difference (DTD) algorithm is proposed. Combining with the DTD method and KPLS model, the dynamic KPLS model, which is named DTD-KPLS, is proposed.

In general, the original input variable sample and output variable sample are used in the regression algorithm. However, $\Delta x(t)$ is calculated between the input variable $x(t)$ at the time of t and the input variable $x(t-i)$ at the time of $t-i$ through the TD method first. Similarly, $\Delta y(t)$ is calculated by the TD method:

$$\Delta x(t) = x(t) - x(t-i) \quad (6)$$

$$\Delta y(t) = y(t) - y(t-i) \quad (7)$$

Then, the relationship between $\Delta x(t)$ and $\Delta y(t)$ is modeled as:

$$\Delta y(t) = f[\Delta x(t)] \quad (8)$$

If new data $x(t')$ is obtained, the TD of input data can be computed as:

$$\Delta x(t') = x(t') - x(t'-i) \quad (9)$$

Thus, the TD of output data can be predicted by the training model as:

$$\Delta y(t') = f(\Delta x(t')) \quad (10)$$

Finally, the predicted output value at time t' is:

$$\hat{y}(t') = \Delta y(t') + y(t'-i) \quad (11)$$

When the DTD method is used, the time-delay τ needs to be considered in Equations (8) and (10). Hence, the training model becomes:

$$\Delta y(t) = f[\Delta x(t-\tau)] = f[x(t-\tau) - x(t-\tau-1)] \quad (12)$$

Hence, the actual predicted output is:

$$\hat{y}(t') = f[x(t'-\tau) - x(t'-\tau-1)] + y(t'-1) \quad (13)$$

In this paper, the estimation of the DTD-KPLS model from the training set is depicted as follows [22]:

1. The relevant variable samples are obtained, and the preprocessing, which contains outlier elimination and data filtering, is performed;
2. The input matrix $\mathbf{X}_{\text{real}} \in \mathbf{R}^{n \times m}$ and the output matrix $\mathbf{Y} \in \mathbf{R}^{n \times 1}$ are confirmed;
3. The time-delay of each input variable $x_i (i = 1, \dots, n)$ is estimated and the phase space reconstruction is made for \mathbf{X}_{real} . Hence, the new input variable matrix \mathbf{X}_{st} is obtained;
4. The $\Delta x(t-\tau)$ and $\Delta y(t-1)$ of new input are calculated, and the $\Delta y(t)$ of output is calculated;
5. The training set $\Delta \mathbf{X}_{\text{tr}}^0$ and $\Delta \mathbf{Y}_{\text{tr}}^0$ are confirmed and $\Delta \mathbf{X}_{\text{tr}}^1$ and $\Delta \mathbf{Y}_{\text{tr}}^1$ (z-score normalized) are obtained;
6. The kernel matrix \mathbf{K}_{tr}^0 of the training set is calculated as:

$$\mathbf{K}_{\text{tr}}^0 = k(\Delta \mathbf{x}_{\text{tr}}, \Delta \mathbf{x}_{\text{tr}}) \quad (14)$$

7. The kernel matrix K_{tr}^0 of the training set is centered as:

$$K_{tr}^1 = \left(I - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) K_{tr}^0 \left(I - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) \quad (15)$$

Here, I is a unit matrix, $\mathbf{1}_n$ is a matrix whose all elements are 1 and whose dimension is n .

8. L is the number of latent variables, i iterates from 1 to L , and initialize the score vector u^i of X_{tr}^1 randomly.
9. The score vector t^i is calculated as:

$$t^i = \frac{K_{tr}^1 u^i}{\|K_{tr}^1 u^i\|} \quad (16)$$

10. The weight vector c^i is calculated as:

$$c^i = \left(Y_{tr}^1 \right)^T t^i \quad (17)$$

11. The score vector u^i is calculated as:

$$u^i = \frac{Y_{tr}^1 c^i}{\|Y_{tr}^1 c^i\|} \quad (18)$$

12. Step (8) to (11) is repeated until that t^i is converged;
13. The matrix K_{tr}^1 and Y_{tr}^1 are reduced as follows: until that t and u are extracted;

$$K_{tr}^{i+1} = \left[I - t^i (t^i)^T \right] K_{tr}^i \left[I - t^i (t^i)^T \right] \quad (19)$$

$$Y_{tr}^{i+1} = Y_{tr}^i - t^i (t^i)^T Y_{tr}^i \quad (20)$$

14. The regression coefficient B is obtained, and the regression model of the training set is obtained as follows:

$$\hat{Y}_{tr} = K_{tr} B = K_{tr} U \left(T^T K_{tr} U \right)^{-1} T^T Y_{tr} \quad (21)$$

Here, T and U are both matrices, which are formed by score vector t and score vector u .

The prediction for the test set is similar to that for the training set, except the computation of the test kernel matrix K_{te}^0 and the centralization of K_{te}^0 :

$$K_{te}^0 = k(\Delta x_{te}, \Delta x_{tr}) \quad (22)$$

$$K_{te}^1 = \left(K_{te}^0 - \frac{1}{n_t} \mathbf{1}_{nt} \mathbf{1}_n^T K_{te}^0 \right) \left(I - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T \right) \quad (23)$$

Here, n_t is the number of the test set. In addition, the kernel function uses the radial basis function (RBF) kernel in this paper.

- (2) Feedback correction

In order to compensate for the prediction error, the feedback correction equation is as follows:

$$\tilde{y}(t+1) = \hat{y}(t+1) + \Delta y(t+1) \quad (24)$$

$$\Delta y(t+1) = \rho \times \Delta y_0(t+1) + (1-\rho) \Delta y(t) \quad (25)$$

$$\Delta y_0(t+1) = y(t) - \hat{y}(t) \quad (26)$$

Here, ρ is 0.3, $\tilde{y}(t)$ is the corrected predicted value at time t , $\hat{y}(t)$ is the predicted value at time t , $y(t)$ is the real value at time t .

Because the SCR de-NO_x process is very complicated, a complete mathematical model of the SCR process has still not been summarized. However, the compensation amount of NH₃ injection Δy_{NH_3} can be calculated according to the nonlinear Equation (27):

$$\Delta y_{\text{NH}_3}(t) = f(\Delta \tilde{y}(t)) \quad (27)$$

Here,

$$\Delta y_{\text{NH}_3}(t) = y'_{\text{NH}_3} - y_{\text{NH}_3}(t) \quad (28)$$

$$\Delta \tilde{y}(t) = \tilde{y}(t) - y_{\text{set}} \quad (29)$$

y'_{NH_3} is the demand amount of NH₃ injection at the current operating state. $y_{\text{NH}_3}(t)$ is the NH₃ injection flow rate at the time of t .

In this paper, the TD-KPLS model is used to establish the nonlinear model of the NIC, which is shown in Equation (27). The difference between the TD-KPLS model and DTD-KPLS model is that the samples of the TD-KPLS model are not reconstructed. In the model, the deviation $\Delta \tilde{y}(t)$ between the corrected DTD-KPLS model output $\tilde{y}(t)$ and the set value of outlet NO_x mission y_{set} is used as the input variables, and the compensation amount of NH₃ injection $\Delta y_{\text{NH}_3}(t)$ is used as the output variable. In addition, input variables also include inlet flue gas temperature and inlet NO_x concentration, which reflect the changing trend of the boiler load to a certain extent.

The adverse effects brought by the inertia characteristic of the SCR system are eliminated by feed-forward compensation. Hence, the NIC can compensate for the NH₃ injection in time to ensure that the outlet NO_x emission is stable at the set value.

In the SCR de-NO_x system, the NH₃ injection valve can be seen as a linear part, so the algorithm design and parameter setting are performed on the basis of that. However, the NH₃ injection valve has a certain dead zone and saturation zone, which shows nonlinearity. Therefore, in this paper, the middle curve in the on and off stroke is taken as the fitting curve between the NH₃ injection valve opening and the NH₃ injection flow rate.

In order to avoid valve saturation and action frequently, the following constraints of the NH₃ injection valve need to be considered:

$$U_{\text{Vmin}} \leq \Delta u_{\text{V}}(t+1|t) + u_{\text{V}}(t) \leq U_{\text{Vmax}} \quad (30)$$

$$u_{\text{Vmin}} \leq \Delta u_{\text{V}}(t+1|t) \leq u_{\text{Vmax}} \quad (31)$$

Here, U_{Vmax} and U_{Vmin} are the upper and lower limits for the NH₃ valve opening. u_{Vmax} and u_{Vmin} are the upper and lower limits for the increment of the NH₃ valve opening. $\Delta u_{\text{V}}(t+1|t)$ represents the increment of the NH₃ valve opening from the next time $t+1$ to the current time t .

3. The Delay Estimation and Variable Selection Method

3.1. Delay Estimation

Aiming at nonlinearity and large lag in SCR de-NO_x process, this paper proposes the k-nearest neighbor MI(KNN_MI) method to estimate the time-delay of the input variable and the effect of the historical data lengths on KNN_MI was taken into account. The KNN_MI method can be summarized as follows [8].

The KNN_MI of variable X and Y is:

$$\text{MI}(X, Y) = \psi(k) - \langle \psi(\mathbf{n}_x + 1) + \psi(\mathbf{n}_y + 1) \rangle + \psi(n) \quad (32)$$

Here, $\mathbf{n}_x(i)$ is the sample number x_j whose distance from x_i is less than $\varepsilon_i/2$ strictly, $\varepsilon_i/2$ denotes the distance from x_i to its k -th neighbor, and $\mathbf{n}_y(i)$ is calculated similarly, $i \in [1, \dots, n]$. $\Psi(x)$ is the

digamma function, $\Psi(x) = \Gamma(x)^{-1}d\Gamma(x)/dx$. It meets the recursion equation $\Psi(x + 1) = 1/x + \Psi(x)$ and $\Psi(1) \approx -0.5772156$. The symbol $\langle \dots \rangle$ denotes the mean of all variables in it.

The KNN_MI of high dimensional variables $(X_1, X_2, \dots, X_m, Y)$ is:

$$\text{MI}(X_1, X_2, \dots, X_m, Y) = \psi(k) - \langle \psi(\mathbf{n}_{x_1}) + \dots + \psi(\mathbf{n}_{x_m}) + \psi(\mathbf{n}_y) \rangle + m\psi(n) \quad (33)$$

Because of the time-delay between each input variable x_i and the output variable y is different, the phase space of x_i is reconstructed by embedding with time-delay $\tau_i \in [\min(\tau_i), \max(\tau_i)]$ ($\min(\tau_i)$ and $\max(\tau_i)$ are resolved by field experience). In addition, MI value has a relation to the length w of x'_i . Therefore, the time-delay τ_i and length w are calculated at the time of t as follows:

$$\begin{aligned} \max_{\tau_i=\tau'_i, w_i=w'_i} \text{MI}([x_i(t-\tau_i-w_i+1), \dots, x_i(t-\tau_i-1), x_i(t-\tau_i)]^T, [y(t-w_i+1), \dots, y(t-1), y(t)]^T) \\ \text{s.t. } \min(\tau_i) \leq \tau_i \leq \max(\tau_i), \tau_i + w_i < T_{\max}; i \in [1, m] \end{aligned} \quad (34)$$

Here, T_{\max} is the maximum reaction time when the SCR system responds. Equation (34) involves constrained multivariable nonlinear optimization. If the number of input variables is m , $2m$ parameters need to be obtained. Thus, within the above constraints, a global searcher maximizes the objective function by particle swarm optimization (PSO) algorithm. Hence, the optimal $\tau' = [\tau'_1, \tau'_2, \dots, \tau'_m]$ and $w' = [w'_1, w'_2, \dots, w'_m]$ are obtained. By estimating the time-delay τ'_i of input variable x_i , the reconstructed input matrix X_{st} is assumed as follows:

$$X_{\text{st}} = \begin{bmatrix} x_{.1}(t-\tau'_1-n+1) & \cdots & x_{.i}(t-\tau'_i-n+1) \\ \vdots & \ddots & \vdots \\ x_{.1}(t-\tau'_1-1) & \cdots & x_{.i}(t-\tau'_i-1) \\ x_{.1}(t-\tau'_1) & \cdots & x_{.i}(t-\tau'_i) \end{bmatrix} \quad (35)$$

3.2. Variable Selection

The subset S , which has m input variables, equals to set (X_1, X_2, \dots, X_m) . The MI between S and Y is defined as:

$$\text{MI}(S, Y) = \text{MI}(X_1, X_2, \dots, X_m, Y) \quad (36)$$

Information entropy represents a measure of uncertainty. If the added variable d_i is irrelevant, the change of uncertainty is less. Therefore, the change rate of KNN_MI, which shown in Equation (37) is larger, when the added variable d_i is relevant:

$$R_{\Delta\text{MI}}(S + d_i, Y) = \frac{\Delta\text{MI}(S, d_i, Y)}{\text{MI}(S, Y)} = \frac{\text{MI}(X_1, X_2, \dots, X_m, d_i, Y) - \text{MI}(X_1, X_2, \dots, X_m, Y)}{\text{MI}(X_1, X_2, \dots, X_m, Y)} \quad (37)$$

The high-dimensional KNN_MI and the impact on original $\text{MI}(S, Y)$ are considered in the above equation. Therefore, $R_{\Delta\text{MI}}(S + d_i, Y)$ can be used to select variable subset S in forward-selection (step (2)–step (5)). In addition, the one-dimensional KNN_MI between any two variables in the variable subset S can be used to delete redundant variables in backward selection (step (6)–step (7)). Based on the change rate of KNN_MI (KNN_MI_CR), the algorithm for the variable selection is designed as follows:

- (1) One-dimensional $\text{MI}(X_i, Y)$ between each input variable X_i and output variable Y are calculated, $i = 1, \dots, m$;
- (2) Based on one-dimensional KNN_MI values, the input variables are sorted from max to min. The most relevant variable which corresponds to the maximum KNN_MI value is selected as the initial variable set S ;

- (3) Based on the most relevant variable and the sorting result of other input variables, $R_{\Delta MI}(S + d_i, Y)$, which relates to the other input variable d_i and the output variable Y , is calculated in sequence by Equation (37);
- (4) The calculated $R_{\Delta MI}(S + d_i, Y)$ are sorted from min to max. The most relevant variable d_i , which corresponds to the minimum $R_{\Delta MI}(S + d_i, Y)$ is selected. If $R_{\Delta MI}(S + d_i, Y) < \alpha$ is satisfied, $S = S \cup d_i$. Here, α is a threshold value;
- (5) Step (3) and step (4) are repeated until all input variables were selected. The relevant variable subset S is finally obtained;
- (6) Backward selection: The one-dimensional KNN_MI between any two variables in S is calculated, and then the combination of redundant variables which has the max KNN_MI value is selected.
- (7) The redundant variables are eliminated by combining with the one-dimensional KNN_MI in step (2), and the optimal relevant variable subset S is finally obtained.

4. The Benchmark Dataset Experiment

In order to verify the proposed variable selection method, the Friedman dataset and Housing dataset [23] are used to experiment, and the predictive accuracy of the KPLS model is evaluated by the root mean square error (RMSE).

4.1. Friedman Dataset

Friedman function:

$$Y = 10 \sin(\pi X_1 X_2) + 20(X_3 - 0.5)^2 + 10X_4 + 5X_5 + \varepsilon \quad (38)$$

Here, the input variable $X_i (1 \leq i \leq 10)$ is a uniform distribution between 0 and 1, ε is the Gaussian noise with a variance of 1. The variables from X_6 to X_{10} are irrelevant variables. The variable $X_{11} = 0.5X_1 + \varepsilon$ is a redundant variable. The sample number is 500. The number of training sets is 338. The number of the test set is 162.

First, this paper uses the proposed variable selection method (KNN_MI_CR) to select variables. In Section 3.2, the one-dimensional MI(X_i, Y) calculated in step (1) is shown in Figure 3.

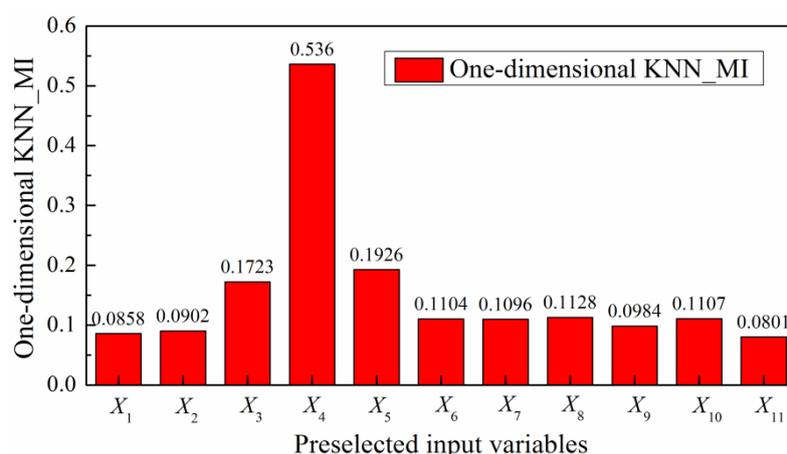


Figure 3. One-dimensional mutual information (MI)(X_i, Y) between each preselected input variable and output variable.

In Figure 3, if the variable selection is only based on one-dimensional KNN_MI, X_8 and X_{10} are mistakenly selected, X_1 and X_2 are mistakenly eliminated.

In step (2), the most relevant variable X_4 , which corresponds to the maximum one-dimensional KNN_MI value, is selected as the initial variable set S . Hence, the initial variable set S is X_4 .

High dimensional KNN_MI and its change rate, which is calculated in step (4), are shown in Figure 4. There, $+X_5$ represents $MI(X_4, X_5, Y)$, $+X_3$ represents $MI(X_4, X_5, X_3, Y)$ and so on. The threshold value α is 0.1. Therefore, variables from X_6 to X_{10} are eliminated, and the relevant variable subset $S = (X_4, X_5, X_3, X_2, X_{11}, X_1)$.

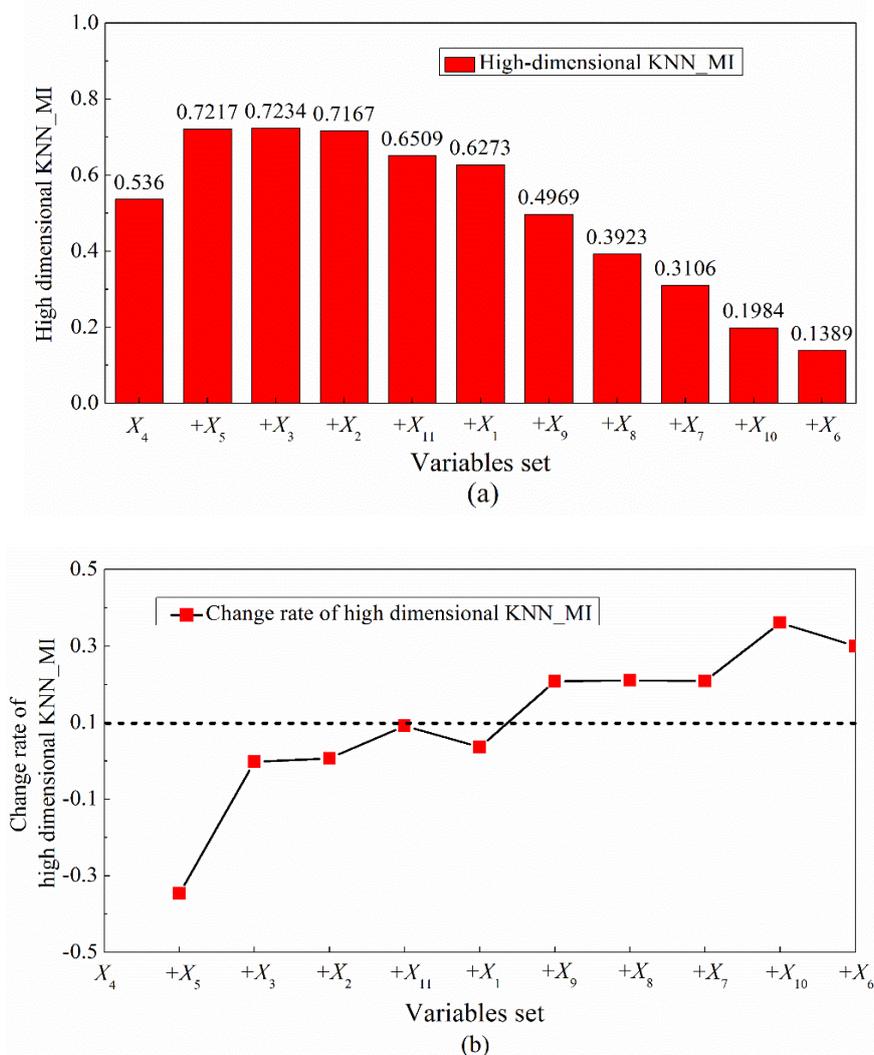


Figure 4. High dimensional k-nearest neighbor mutual information (KNN_MI) (a) and its change rate (forward selection) (b).

In Figure 3, $MI(X_1, Y)$, $MI(X_2, Y)$ and $MI(X_{11}, Y)$ are relatively close. In addition, $MI(X_3, Y)$ and $MI(X_5, Y)$ are also relatively close. Therefore, the relevant variable subset S could have a redundant variable. One-dimensional KNN_MI between any two variables in S is shown in Figure 5. It is obvious that $MI(X_1, X_{11})$ is the maximum value. In Figure 3, $MI(X_1, Y) > MI(X_{11}, Y)$. Therefore, X_{11} is a redundant variable. Finally, optimal relevant variable subset $S = (X_4, X_5, X_3, X_2, X_1)$. This is consistent with Equation (38).

Second, the comparison among variable selection methods and the test results of the KPLS model are shown in Tables 1 and 2.

In Tables 1 and 2, it can be seen that BIF, mRMR, MIFS, and CMIM have not selected all valid input variables, and the CMIM method selects redundant variable X_{11} in the input variable subset. Compared with the other four variable selection methods, the KNN_MI_CR method selects all relevant input variables. Therefore, the KNN_MI_CR method proposed in this paper can select the input variable subset effectively.

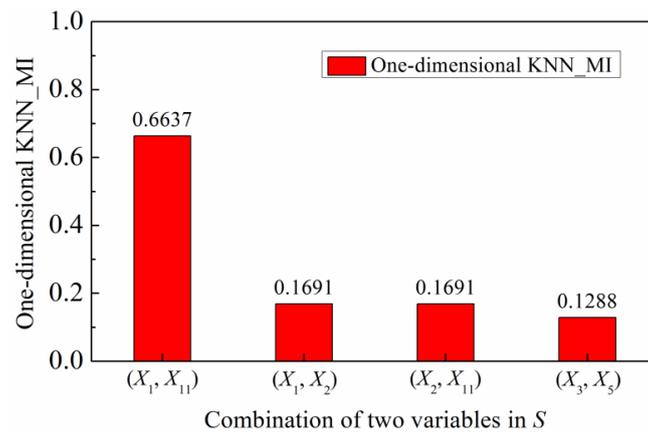


Figure 5. One-dimensional KNN_MI between any two variables in S (backward selection).

Table 1. The variable selection result and model test result for the Friedman dataset (no redundant variable).

Variable Selection Method	Variable Selection Result	RMSE
All variables	$X_1 \sim X_{10}$	1.4966
MIFS [11]	X_4, X_5, X_3, X_1, X_9	1.2019
mRMR [12]	X_4, X_5, X_3, X_1, X_9	1.2019
CMIM [13]	X_4, X_5, X_3, X_7, X_8	1.2433
BIF [24]	$X_4, X_5, X_3, X_{10}, X_9$	1.2079
KNN_MI_CR	X_4, X_5, X_3, X_2, X_1	1.1260

Table 2. The variable selection result and model test result for the Friedman dataset (contain redundant variable X_{11}).

Variable Selection Method	Variable Selection Result	RMSE
All variables	$X_1 \sim X_{11}$	1.6222
MIFS [11]	X_4, X_5, X_3, X_1, X_9	1.2043
mRMR [12]	X_4, X_5, X_3, X_1, X_9	1.2043
CMIM [13]	$X_4, X_5, X_3, X_{11}, X_8$	1.1843
BIF [24]	$X_4, X_5, X_3, X_8, X_{10}$	1.3810
KNN_MI_CR	X_4, X_5, X_3, X_2, X_1	1.1272

4.2. Housing Dataset

The Housing dataset is used to predict Boston house price (defined as the output variable Y) by statistic values (defined as input variables X_1, \dots, X_{13}), which represent 13 attributes in each house. The number of training sets is 338. The number of the test set is 168. The comparison of different variable selection methods and test results of the KPLS model are shown in Table 3.

It can be seen from the results that the predictive accuracy of the model is highest by using the KNN_MI_CR method. Therefore, the KNN_MI_CR method contributes to establishing a more effective prediction model.

Table 3. The variable selection result and model test results for the Housing dataset.

Variable Selection Method	Variable Selection Result	RMSE
All variables	$X_1 \sim X_{13}$	9.9566
MIFS [11]	$X_{13}, X_6, X_{11}, X_1, X_4$	9.6594
mRMR [12]	$X_{13}, X_{10}, X_{12}, X_6, X_7$	9.8252
CMIM [13]	$X_{13}, X_{10}, X_2, X_6, X_4$	10.5333
BIF [24]	$X_{13}, X_{11}, X_6, X_3, X_5$	9.7156
KNN_MI_CR	$X_{13}, X_{10}, X_4, X_3, X_2$	9.4404

5. The Field Data Experiment for SCR de-NO_x Process

5.1. Field Data and Preselected Input Variables

The inlet NO_x concentration and NH₃ injection flow rate reflect the NH₃/NO_x molar ratio, which influences the de-NO_x efficiency and the NH₃ escape directly. The NH₃ injection flow rate is controlled to match up the change of boiler load by the NH₃ valve. Inlet O₂ content influences outlet NO_x concentration and de-NO_x efficiency directly. Moreover, the change in boiler load often influences the inlet flue gas flow rate. Then, the flue gas temperature is changed through heat exchange. The inlet flue gas temperature influences the reaction rate and catalyst activity.

In a word, the outlet NO_x emission is related to many factors. In addition, the changes in boiler load, coal quality, and combustion condition could cause a large fluctuation in inlet NO_x concentration. Therefore, preselected input variables and the output variable are shown in Table 4.

Table 4. The preselected input variables and output variable for SCR de-NO_x process.

Variable	Parameter	Unit
Preselected input variables	Boiler load/ x_1	MW
	Total coal feed rate/ x_2	t/h
	Total air volume / x_3	t/h
	Inlet NO _x concentration/ x_4	mg/m ³
	Inlet O ₂ content/ x_5	%
	NH ₃ valve opening/ x_6	%
	NH ₃ injection flow rate/ x_7	kg/h
	Inlet flue gas temperature/ x_8	°C
	Inlet flue gas flow rate/ x_9	km ³ /h
Output variable	Outlet NO _x emission/ y	mg/m ³

Assuming that coal quality is invariable, the selected field data cover steady state and variable state. One-dimensional linear interpolation is carried out on the measured NO_x concentration at the process of blowback. Therefore, 2000 samples were collected, and the sampling period is 10 s. Due to the length of this paper, the samples of part preselected input variables and output variables are shown in Figure 6.

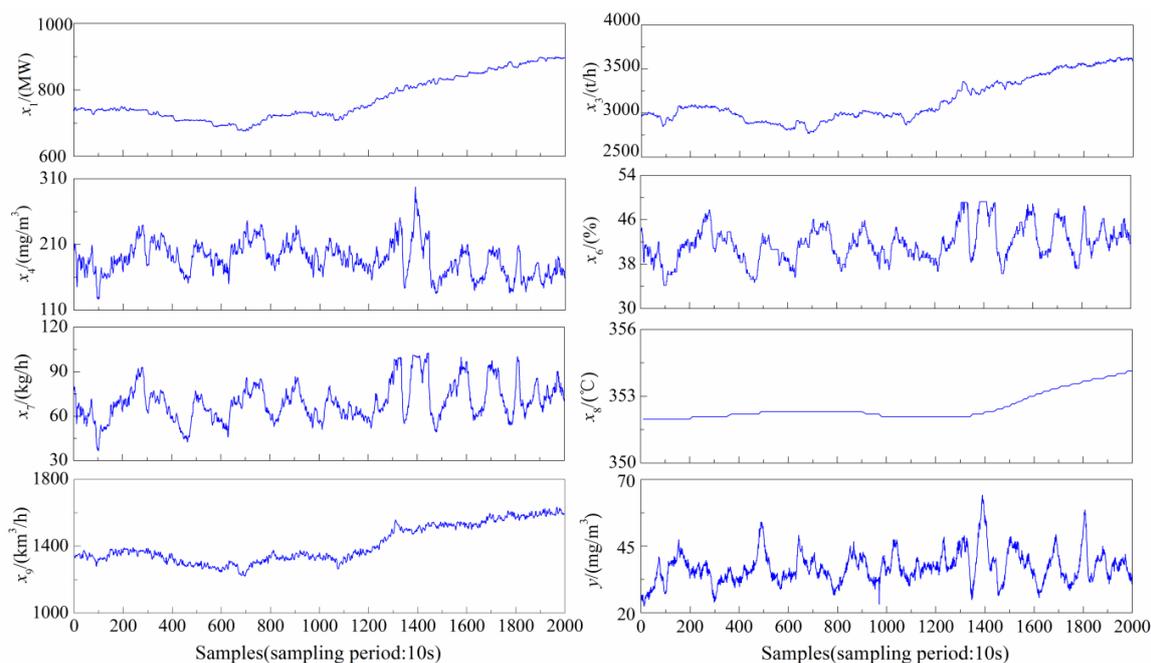


Figure 6. The samples of part preselected input and output variables.

5.2. Result Analysis of Delay Estimation and Variable Selection

In general, the maximum time-delay of SCR reaction which includes NO_x analysis is 120~400 s, and the maximum time-delay of boiler load which influences inlet NO_x concentration is 600 s, so the range of τ is (20,60) from x_1 to x_3 , and (1,40) from x_4 to x_9 . T_{\max} in Equation (32) is 1200 s, so the range of w is (1,120). The time-delay of each input variable is different-in-different operation point. The samples from 200 point to 500 point at 800 MW operating state are taken as an example to analyze the delay estimation and variable selection methods. The size of the SCR reactor is 11.67 m \times 13.95 m \times 12.6 m. The optimal fitness curves of the PSO algorithm for all preselected input variables are shown in Figure 7, and the results are shown in Table 5.

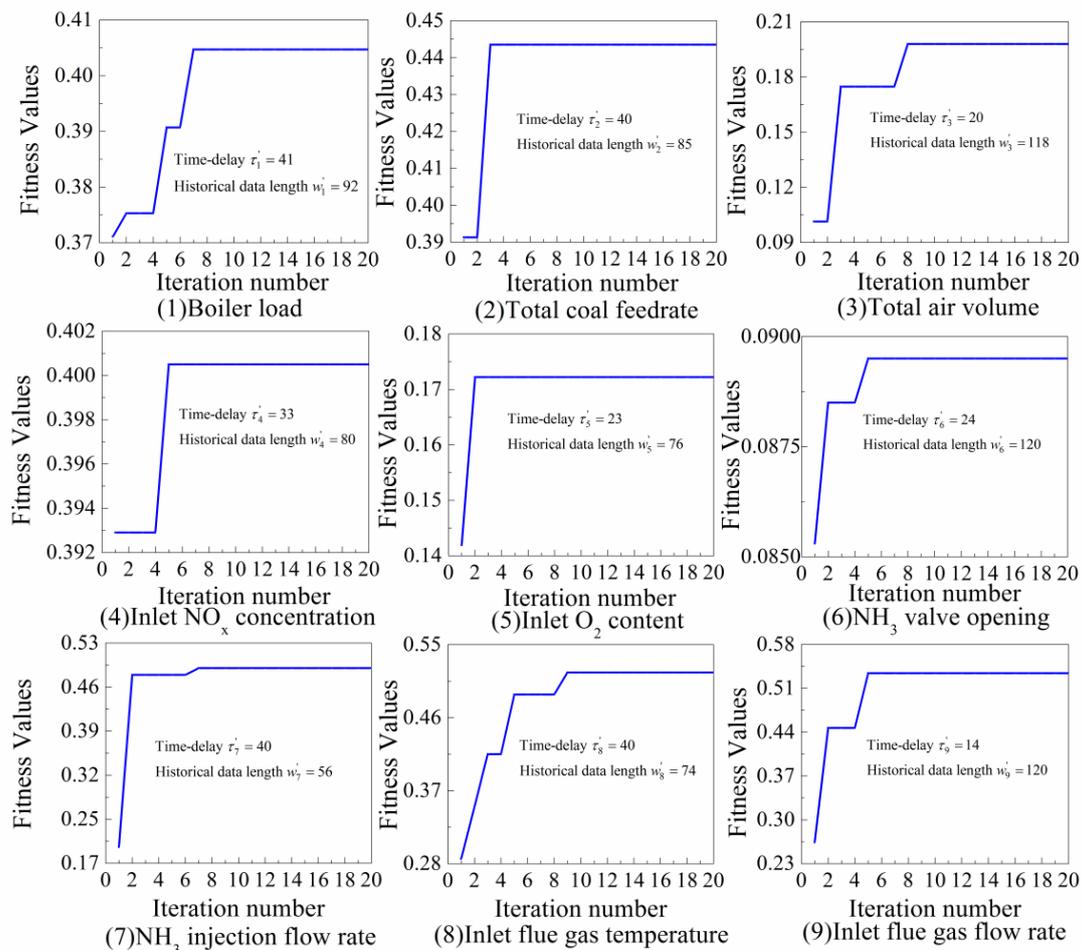


Figure 7. The fitness curve of particle swarm optimization (PSO) algorithm.

Table 5. The results of delay estimation.

Input Variables	KNN_MI	Optimal Result of Time-Delay	Time-Delay
x_1	$\text{MI}(x_1, x_3)$	41	$410 + 330 = 740$ s
x_2	$\text{MI}(x_2, x_3)$	40	$400 + 330 = 730$ s
x_3	$\text{MI}(x_3, y)$	20	$200 + 330 = 530$ s
x_4	$\text{MI}(x_4, y)$	33	330 s
x_5	$\text{MI}(x_5, y)$	23	230 s
x_6	$\text{MI}(x_6, y)$	24	240 s
x_7	$\text{MI}(x_7, y)$	40	400 s
x_8	$\text{MI}(x_8, y)$	40	400 s
x_9	$\text{MI}(x_9, y)$	14	140 s

It can be seen from Figure 7, each preselected variable reaches the optimal global solution before the 10th iteration, and the optimal parameter τ' and w' are obtained. The optimal result multiplies the sampling period to obtain the time-delay of the preselected variable, which is shown in Table 5.

The effect of time-delay needs to be considered for the variable selection in SCR de-NO_x process. Therefore, based on the delay estimation result, the phase space reconstruction on each input variable is performed first, and then the KNN_MI_CR method is used to select a suitable variable.

It can be seen from Table 6 that if the model samples are reconstructed by phase space, the result of variable selection is different. The DTD-KPLS model, whose variables are selected by the KNN_MI_CR method, has the highest predictive accuracy. Although the variable selection result is the same, the predictive accuracy of the DTD-KPLS model whose samples are reconstructed is improved in the BIF and CMIM method. In addition, the sequence of the selected input variable is different, and the model accuracy is different, which indicates that the consideration of high-dimensional KNN_MI is conducive to the selection of suitable input variables.

Table 6. The variable selection result and test result for SCR de-NO_x process.

Method	Samples		Reconstructed Samples	
	Result	RMSE (TD-KPLS)	Result	RMSE (DTD-KPLS)
All input variable	$X_1 \sim X_9$	4.4561 mg/m ³	$X_1 \sim X_9$	4.2619 mg/m ³
MIFS [11]	X_3, X_4, X_6	3.2071 mg/m ³	X_3, X_4, X_8	3.2089 mg/m ³
mRMR [12]	X_3, X_5, X_4	3.2101 mg/m ³	X_3, X_5, X_2	3.2168 mg/m ³
CMIM [13]	X_3, X_7, X_4	3.2061 mg/m ³	X_3, X_7, X_4	3.2003 mg/m ³
BIF [24]	X_3, X_4, X_7	3.3069 mg/m ³	X_3, X_4, X_7	3.2207 mg/m ³
DISR [25]	X_3, X_4, X_7	3.3069 mg/m ³	X_3, X_4, X_2	3.2130 mg/m ³
KNN_MI_CR	X_3, X_5, X_4	3.2101 mg/m ³	X_3, X_7, X_4	3.2003 mg/m ³

5.3. Analysis of Dynamic Model

In general, dynamic modeling is realized by adding the historical input data $x(t-1), \dots, x(t-w+1)$ and output data $y(t-1), \dots, y(t-w+1)$ as a new input of the model. The 2000 prediction samples are used to analyze model accuracy by using different input variables. The results are shown in Table 7.

Table 7. The predictive accuracy of DT-KPLS and DTD-KPLS model using different input variables.

Model	Input Variables	Dimension of Input Variable	RMSE
DT-KPLS	$x(t)$	3	3.2268 mg/m ³
	$y(t-1), x(t)$	4	3.2166 mg/m ³
	$x(t-1), x(t)$	6	3.2054 mg/m ³
	$x(t-1), y(t-1), x(t)$	7	3.2061 mg/m ³
DTD-KPLS	$x(t-\tau)$	3	3.2139 mg/m ³
	$y(t-1), x(t-\tau)$	4	3.2082 mg/m ³
	$x(t-\tau-1), x(t-\tau)$	6	3.2035 mg/m ³
	$x(t-\tau-1), y(t-1), x(t-\tau)$	7	3.2003 mg/m ³

It can be seen from Table 7 when input variables are $x(t-\tau-1), y(t-1)$ and $x(t-\tau)$, the predictive accuracy of the model is highest. In addition, because the model input is historical input data before the current time t , the DTD-KPLS model can predict the current value of the output variable in advance accurately.

When parameter optimization is used in model control strategies, a large number of calculations need to be avoided. In this paper, the DTD-KPLS model has only one RBF kernel parameter. The predictive accuracy of the DTD-KPLS model in different kernel parameters is analyzed. The results are shown in Figure 8.

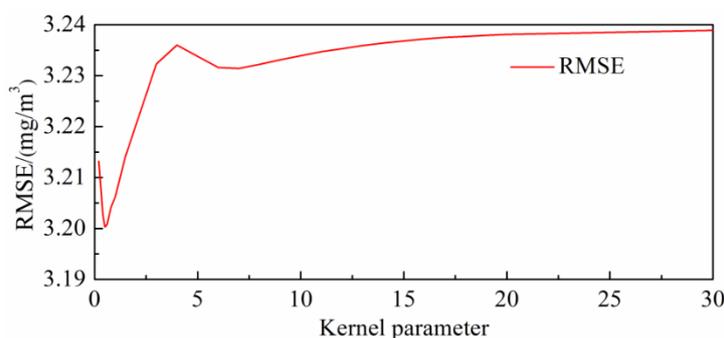


Figure 8. The predictive accuracy of delay–time difference (DTD)–kernel partial least squares (KPLS) model with different kernel parameters.

The change rate of the RMSE value was no more than 1%, as shown in Figure 8, which indicates that the predictive accuracy was not sensitive to the kernel parameter. Therefore, the DTD–KPLS model could adopt fixed parameters to avoid a large number of calculations in parameter optimization. In this paper, the single-step prediction time was 1 s, so it could meet the demand for control.

In order to compensate for the prediction error, this paper proposes a feedback correction strategy for the DTD–KPLS model. The predicted value and the corrected predicted value are shown in Figure 9.

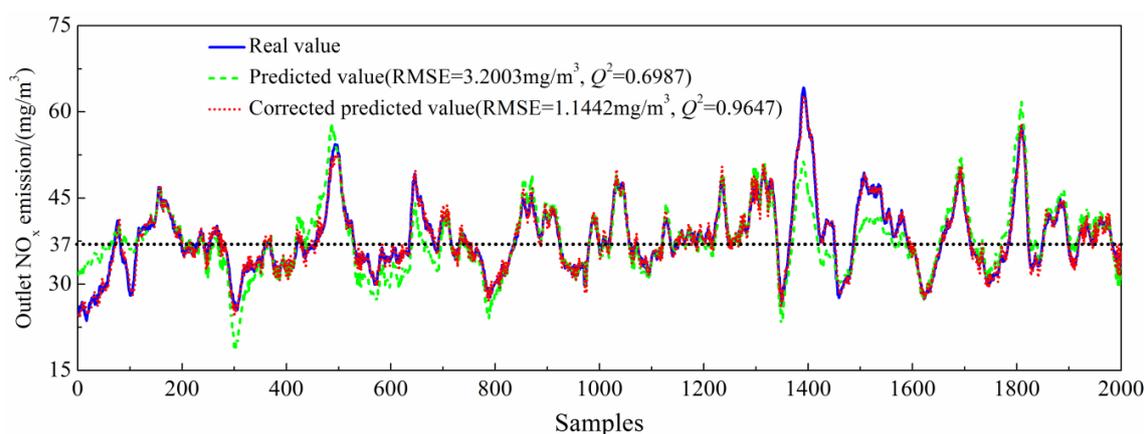


Figure 9. The predicted value and the corrected predicted value for DTD–KPLS model.

It can be seen from Figure 9, and the DTD–KPLS model shows high predictive accuracy with a low RMSE of 3.2933 mg/m³ and a high determinate coefficient (Q^2) of 0.7011 in the whole state. A large amount of NO_x was produced in a variable state. Hence the predictive accuracy was lower at the peak of the outlet NO_x emission curve. After feedback correction, the predicted values were closer to the real value. Steady state and variable state were covered in the model samples, so the DTD–KPLS model had strong self-adaptability.

5.4. Analysis of Control Effect

Two kinds of NH₃ injection control modes have mainly existed in China. One is fixed NH₃/NO_x molar ratio control, and the other one is fixed outlet NO_x value control. In this paper, the fixed outlet NO_x value control was analyzed. First, the NO_x flow rate was determined by the inlet NO_x concentration and the flue gas flow rate, and the preset molar ratio was calculated. Then, the NH₃/NO_x molar ratio was corrected based on the measured value and the set value of outlet NO_x emission. The amount of NH₃ injection was determined by the corrected NH₃/NO_x molar ratio and the calculated NO_x flow rate. Finally, the fixed outlet NO_x emission was achieved.

In order to obtain the demand amount of NH₃ injection y'_{NH_3} at the current operating state, the transfer function from the NH₃ injection to the decreased amount of outlet NO_x emission was investigated through a dynamic characteristic test as follows:

$$W(s) = -\frac{1.73e^{-21s}}{10236s^2 + 196s + 1} \tag{39}$$

Then, the control structure diagram set up in MATLAB/SIMULINK is shown in Figure 10.

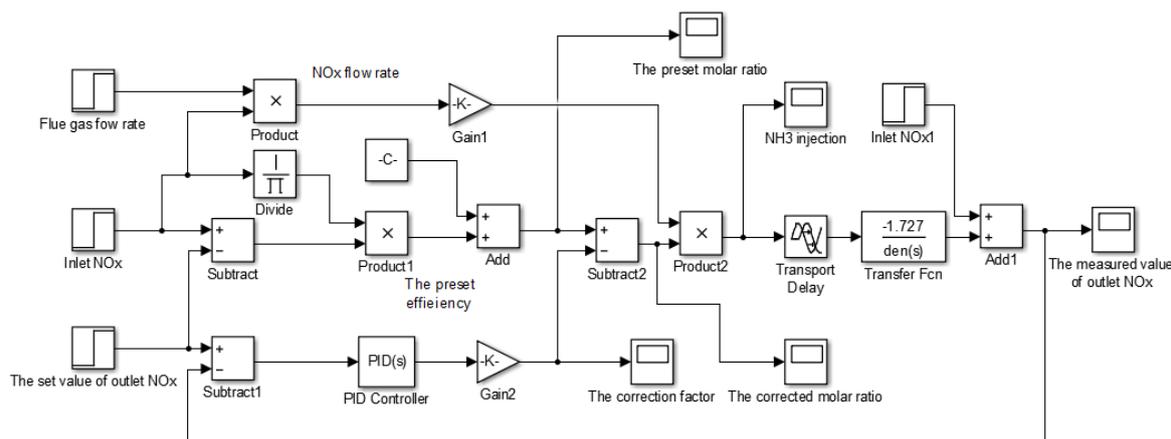


Figure 10. The control structure diagram of fixed outlet NO_x value.

In the field data in Figure 6, the inlet NO_x and the flue gas flow rate fluctuated from 135~230 mg/m³ to 120~160 km³/h. During the simulation, the inlet NO_x, the flue gas flow rate and the set value of outlet NO_x were set to 160 mg/m³, 1300 km³/h and 37 mg/m³, which were consistent with the field data. PID tuning program was used to get optimal PID parameters. The results are shown in Figures 11–13.

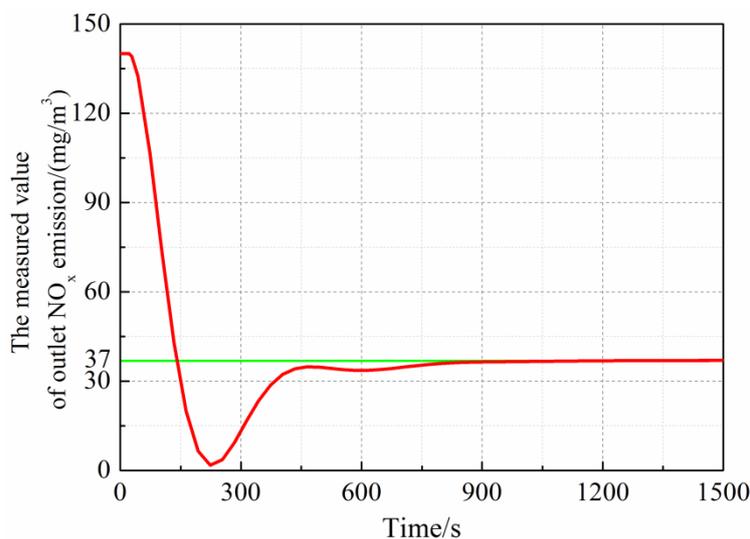


Figure 11. The proportion integration differentiation (PID) control effect of outlet NO_x.

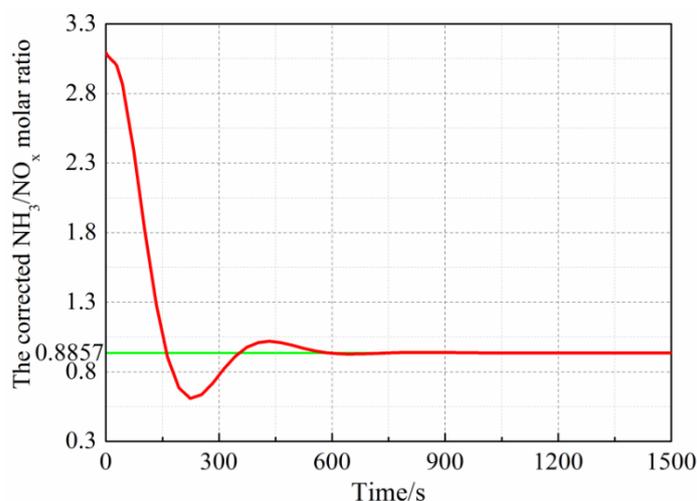


Figure 12. The curve of corrected NH_3/NO_x molar ratio.

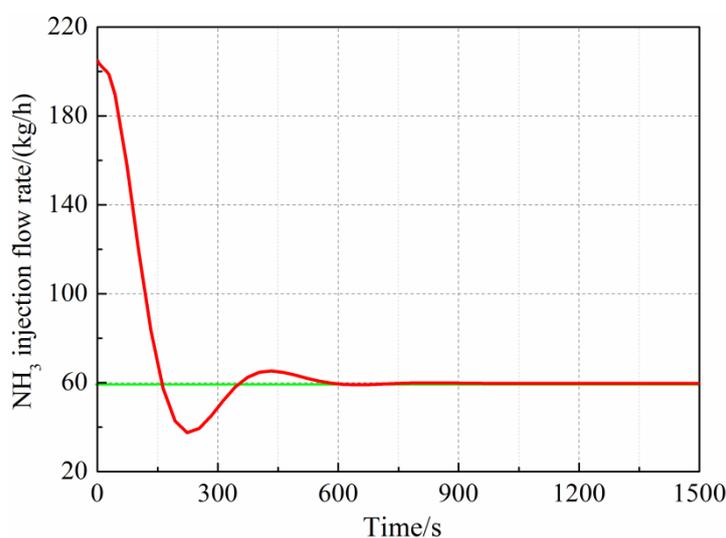


Figure 13. The curve of NH_3 injection flow rate.

If the inlet NO_x value and the set value of outlet NO_x were fixed, as the deviation between them decreased, the PID controller reduced the actual NH_3/NO_x molar ratio; thereby, the amount of NH_3 injection was reduced gradually. After the system was stabilized, the outlet NO_x emission, the NH_3 injection flow rate and the NH_3/NO_x molar ratio were stable at 37 mg/m^3 (set value), 60 kg/h and 0.8857 , which were close to the average value of field data. So the demand amount of NH_3 injection y'_{NH_3} corresponding to the set value of outlet NO_x emission was 59.6 kg/h , and this method could accurately simulate the fixed outlet NO_x control mode. The control process lasted for nearly 1000 s before the system became stable. Hence, the existing system could not meet real-time control because of the long adjustment time.

The predicted value of the TD-KPLS model was taken as the output of NIC in each model update. The input variables included the deviation between the corrected predicted NO_x emission value and the set value of outlet NO_x emission, inlet flue gas temperature and inlet NO_x concentration. The output variable was the amount of NH_3 injection compensation. The output of the NIC is shown in Figure 14.

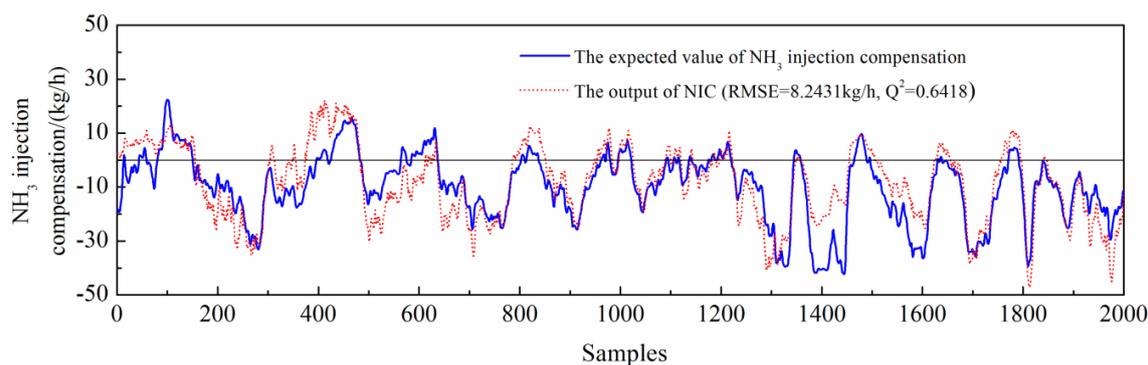


Figure 14. The output of NH_3 injection compensator (NIC).

Figure 14 shows that the output of the NIC was consistent with the trend of the expected value of NH_3 injection compensation. Therefore, the predictive accuracy of TD-KPLS was high. In addition, the predicted value was a mostly negative value indicating that the amount of NH_3 injection was reduced by 11% after adding the NIC. Assuming that the price of NH_3 water was about 3000 RMB per ton, the data used in this paper was taken as an example. After the application of the NIC, the cost of de- NO_x could be saved by 531.2 RMB per day.

Because the SCR de- NO_x system had large inertia and had a large delay in NO_x analysis; hence, the following analysis started from the 100th point. The control effect of outlet NO_x emission is shown in Figure 15.

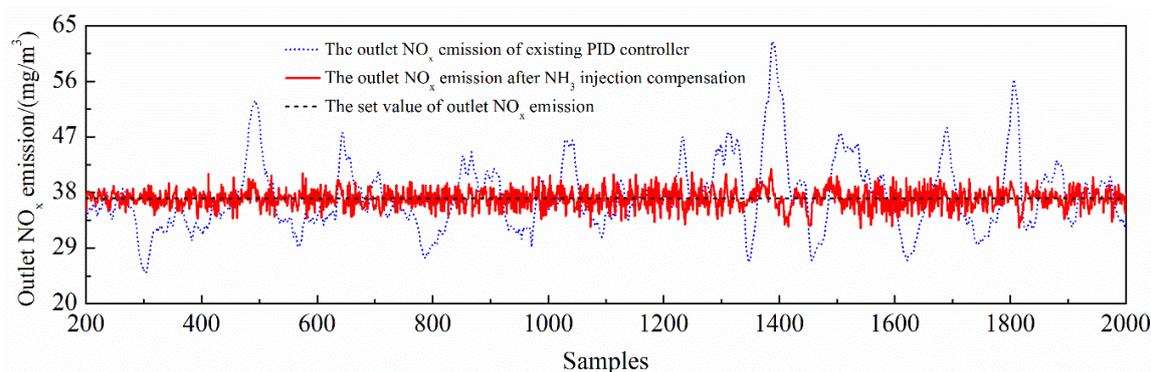


Figure 15. The control effect of outlet NO_x emission.

Figure 15 shows the control effect of the existing controller was not ideal, and the outlet NO_x emission produced drastic fluctuation when the load changes. For example, the outlet NO_x emission exceeded the national NO_x emission standard (50 mg/m^3) at the 500, 1400 and 1800 points. In contrast, after the NIC was added, the outlet NO_x emission whose fluctuation ranged in $\pm 5 \text{ mg/m}^3$ was stable near the set value. Therefore, the outlet NO_x emission was far lower than the national NO_x emission standard.

It can be seen from Figure 16 that the minimum value of the de- NO_x efficiency after adding the NIC (76.31%) was higher than the minimum value of the de- NO_x efficiency of the existing controller (73.79%). The average de- NO_x efficiency was 80.12% by using the existing controller, and the average de- NO_x efficiency was 80.15% after adding the NIC. This indicated that the de- NO_x efficiency was basically unchanged after adding the NIC.

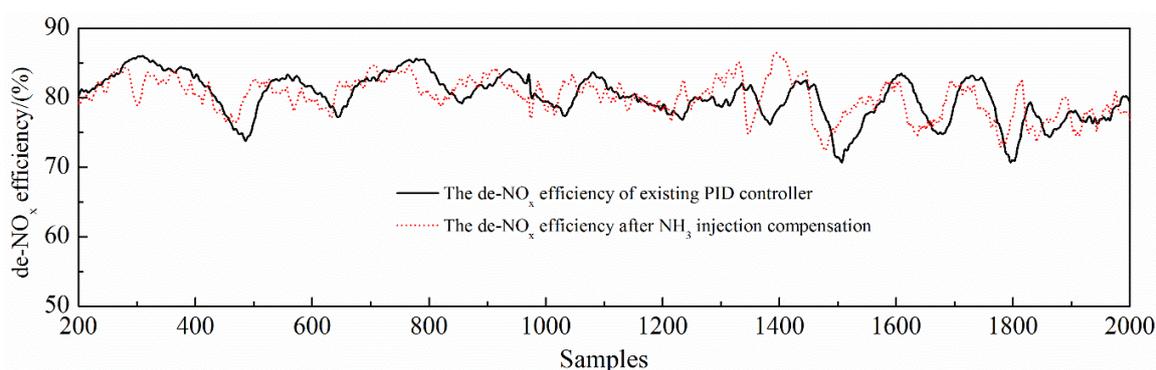


Figure 16. The de-NO_x efficiency of the existing controller and the de-NO_x efficiency after NH₃ injection compensation.

In order to calculate the NH₃ escape further, the following formula was introduced.

$$p(t) = \frac{(W_m(t) - W_c(t)) \times 10^6}{17 \times V_{gas}(t)} \quad (40)$$

Here, t represents the sampling time; W_m is the measured NH₃ injection flow rate, kg/h; W_c is the NH₃ injection consumption, kg/h; V_{gas} is the inlet flue gas flow rate, m³/h; p is the NH₃ escape, μg/m³.

Because the compensation amount of NH₃ injection in this paper was negative value mostly, the consumption of NH₃ injection was regarded unchanged basically. Figure 17 was the comparison between the NH₃ escape of the existing controller and the NH₃ escape after NH₃ injection compensation.

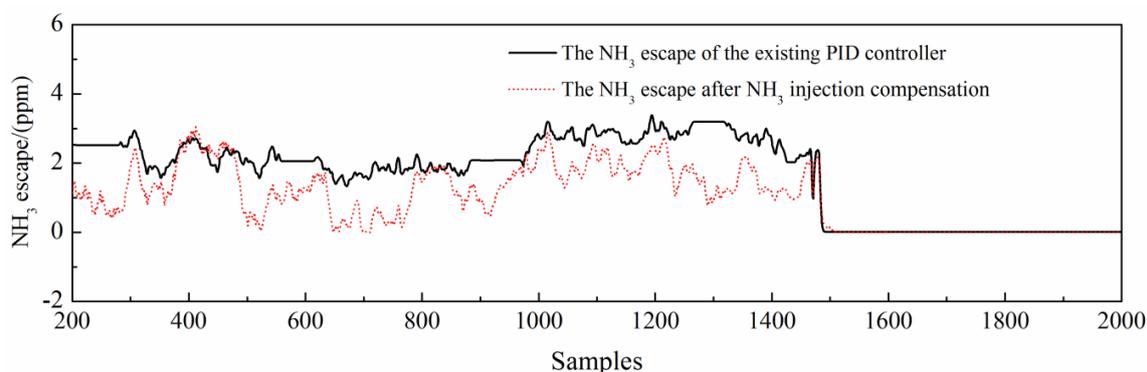


Figure 17. The NH₃ escape of the existing controller and the NH₃ escape after NH₃ injection compensation.

In Figure 17, after the NIC was added, although the average of de-NO_x efficiency did not change, the NH₃ escape was reduced by 39% after NH₃ injection compensation.

Actually, in order to avoid actuator saturation, the NH₃ injection valve opening was restrained by conditions (23) and (24) in the NH₃ injection control. The NH₃ valve opening of the existing controller and NH₃ valve opening after NH₃ injection compensation are shown in Figure 18.

Due to the lag of NO_x analysis, the NH₃ control valve was prevented from being over-opened or over-closed. Therefore, when the NH₃ control valve was in automatic mode, the NH₃ valve opening was limited from $U_{Vmin} = 20\%$ to $U_{Vmax} = 50\%$. In Figure 18, the NH₃ valve opening of the existing controller reached 50% from 1300 points to 1500 points, which indicated that the actuator reached saturation in automatic control. When the NIC was used, the average of the NH₃ valve opening was reduced, thereby actuator saturation was avoided.

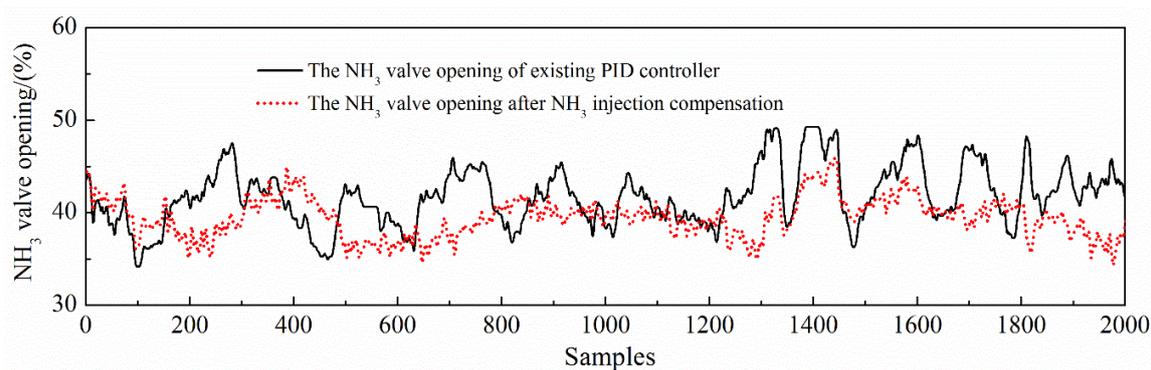


Figure 18. The NH_3 valve opening of the existing controller and the NH_3 valve opening after NH_3 injection compensation.

The upper limitation $u_{V_{\max}}$ and the lower limitation $u_{V_{\min}}$ of the increment of NH_3 valve opening were 2% and -2%, respectively. Figure 19 shows that the increment of NH_3 valve opening after NH_3 injection compensation was within limitation. Therefore, the NH_3 valve was protected, and the control accuracy was improved.

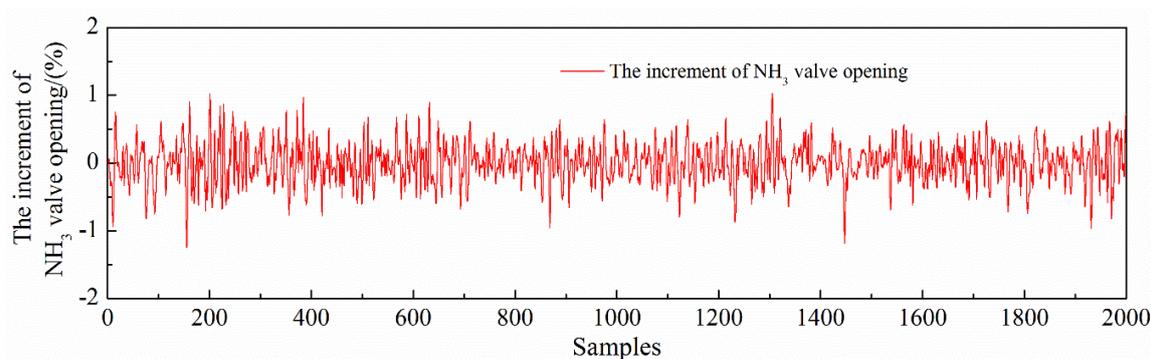


Figure 19. The increment of NH_3 valve opening after NH_3 injection compensation.

In all, the proposed control strategy had a good control effect. Through the NH_3 injection was adjusted by NIC continually, the controller could decrease the NH_3 injection and decrease the NH_3 escape as much as possible while ensuring that the high de- NO_x efficiency and NO_x emission meets the national NO_x emission standard. Therefore, the operation cost was reduced, and secondary pollution caused by the incomplete reaction of NH_3 was avoided.

6. Conclusions

A dynamic KPLS model incorporated with variable selection and delay estimation was proposed for the prediction of outlet NO_x emission in advance. The KNN_MI method was used to estimate the time-delay, and the effect of historical data lengths on the KNN_MI value was considered. Bidirectional search, which is based on the change rate of KNN_MI (KNN_MI_CR), was used for variable selection, which avoids using a single evaluation strategy and considers the effect of high-dimensional KNN_MI. Finally, the DTD update method and feedback correction strategy are proposed, so the fixed parameter is used in the final model. Therefore, the final model could be identified offline by using field data and have high predictive accuracy.

The proposed control system includes the main controller and the correction controller. The main controller is the existing PID controller. The combination of NIC and the outlet NO_x emission model is used as the correction controller. According to the results of the field data experiment, the proposed control strategy has several findings as follows:

- (1) In the aspect of the control effect, the outlet NO_x emission meets the national NO_x emission standard, whose fluctuation range in ± 5 mg/m³ is stable near the set value, and the system maintains high de-NO_x efficiency; the NH₃ escape is reduced by 39% after NIC.
- (2) Compared with the conventional PID control, the proposed control strategy not only protects the NH₃ injection valve but also reduces the operation cost because of the reduction of NH₃ injection;
- (3) The reduction of NH₃ escape avoids operating accidents and secondary pollution.

The potential disadvantage of the proposed control strategy is that the calculation of the demand amount of NH₃ injection y'_{NH_3} at the current operating state involves a dynamic characteristic test. For different operating states, y'_{NH_3} will be different. If the operating state changes frequently, it will not be easy to determine y'_{NH_3} accurately.

Author Contributions: L.L. and Z.D. conceived the research. L.Y. and Ming Sun participated in the analysis of the data and writing the initial manuscript. L.Y., M.S. and J.L. revised the manuscript and adjusted the data presentation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Hebei Province, China. (Grant No. E2018502111) and the Fundamental Research Funds for the Central Universities, China. (Grant No. 2017MS187).

Acknowledgments: Authors acknowledge the Guodian Taizhou Generation Co, Ltd., Taizhou, China, for providing field data support to make the article for its quality in presentation and execution.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yan, L.Q.; Dong, Z.; Jia, H.; Huang, J.N. Dynamic inferential NO_x emission prediction model with delay estimation for SCR de-NO_x process in coal-fired power plants. *R. Soc. Open Sci.* **2020**, *7*, 191647. [[CrossRef](#)] [[PubMed](#)]
2. Vignesh, R.; Ashok, B. Critical interpretative review on current outlook and prospects of selective catalytic reduction system for De-NO_x strategy in compression ignition engine. *Fuel* **2020**, *276*, 117996. [[CrossRef](#)]
3. Zambrano, D.; Tayamon, S.; Carlsson, B.; Wigren, T. Identification of a discrete-time nonlinear Hammerstein-Wiener model for a selective catalytic reduction system. In Proceedings of the American Control Conference, San Francisco, CA, USA, 29 June 29–1 July 2011; pp. 78–83. [[CrossRef](#)]
4. Hui, P.; Ozaki, T.; Toyoda, Y.; Shioya, H.; Nakano, K.; Haggan-Ozaki, V.; Mori, M. RBF-ARX model-based nonlinear system modeling and predictive control with application to a NO_x decomposition process. *Control Eng. Pract.* **2004**, *12*, 191–203. [[CrossRef](#)]
5. Si, F.; Romero, C.E.; Yao, Z.; Schuster, E.; Xu, Z.; Morey, R.L.; Liebowitz, B.N. Optimization of coal-fired boiler SCRs based on modified support vector machine models and genetic algorithms. *Fuel* **2009**, *88*, 806–816. [[CrossRef](#)]
6. Krijnsen, H.C.; Bakker, R.; Kooten, W.E.J.V.; Calis, H.P.A.; Verbeek, R.P.; Bleek, C.M.V.D. Evaluation of fit algorithms for NO_x emission prediction for efficient deNO_x control of transient diesel engine exhaust gas. *Ind. Eng. Chem. Res.* **2000**, *39*, 2992–2997. [[CrossRef](#)]
7. Gao, J.B.; Tian, G.H.; Sornioti, A.; Karci, A.E.; Palo, R.D. Review of thermal management of catalytic converters to decrease engine emissions during cold start and warm up. *Appl. Therm. Eng.* **2017**, *147*, 177–187. [[CrossRef](#)]
8. Kraskov, A.; Stögbauer, H.; Grassberger, P. Estimating mutual information. *Phys. Rev. E.* **2004**, *69*, 066138. [[CrossRef](#)]
9. Paninski, L. Estimation of entropy and mutual information. *Neural Comput.* **2003**, *15*, 1191–1253. [[CrossRef](#)]
10. Stanišić, D.; Jorgovanović, N.; Popov, N.; Čongradac, V. Soft sensor for real-time cement fineness estimation. *ISA Trans.* **2015**, *55*, 250–259. [[CrossRef](#)]
11. Battiti, R. Using mutual information for selecting features in supervised neural net learning. *IEEE Trans. Neural Netw.* **1994**, *5*, 537–550. [[CrossRef](#)]
12. Peng, H.; Long, F.H.; Ding, C. Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal.* **2005**, *27*, 1226–1238. [[CrossRef](#)]

13. Fleuret, F. Fast binary feature selection with conditional mutual information. *J. Mach. Learn. Res.* **2004**, *5*, 1531–1555. [[CrossRef](#)]
14. Ludwig, O.; Nunes, U.; Araújo, R.; Schnitman, L.; Lepikson, H.A. Applications of information theory, genetic algorithms, and neural models to predict oil flow. *Commun. Nonlinear Sci.* **2009**, *14*, 2870–2885. [[CrossRef](#)]
15. Chiang, C.J.; Kuo, C.L.; Huang, C.C.; Lee, J.Y. Model predictive control of SCR aftertreatment system. In Proceedings of the 5th IEEE Conference on Industrial Electronics and Applications (ICIEA), Taichung, Taiwan, 15–17 June 2010; pp. 2058–2063. [[CrossRef](#)]
16. Nakamoto, M.; Kokubo, T.; Kamito, A.; Shimizu, K. Cascade Control Using GPC and LQR for a NO_x Decomposition Process of a Thermal Power Plant. *Trans. SICE* **2000**, *36*, 962–971. [[CrossRef](#)]
17. Zhang, Y.; Shen, J.; Li, Y.; Wu, X. Cascade constrained DMC-PID control for SCR denitrification System. In Proceedings of the 34th Chinese Control Conference, Hangzhou, China, 28–30 July 2015; pp. 7906–7909. [[CrossRef](#)]
18. Peng, H.; Gui, W.H.; Shioya, H.; Zou, R. A predictive control strategy for nonlinear NO_x decomposition process in thermal power plants. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **2006**, *36*, 904–921. [[CrossRef](#)]
19. Peng, H.; Nakano, K.; Shioya, H. Nonlinear predictive control using neural nets-based local linearization ARX model—Stability and industrial application. *IEEE Trans. Control. Syst. Technol.* **2006**, *15*, 130–143. [[CrossRef](#)]
20. Wu, X.; Shen, J.; Sun, S.Z.; Li, Y.G.; Lee, K.Y. Data-driven disturbance rejection predictive control for SCR denitrification system. *Ind. Eng. Chem. Res.* **2016**, *55*, 5923–5930. [[CrossRef](#)]
21. Kaneko, H.; Funatsu, K. Maintenance-free soft sensor models with time difference of process variables. *Chemom. Intell. Lab. Syst.* **2011**, *107*, 312–317. [[CrossRef](#)]
22. Rosipal, R.; Trejo, L.J. Kernel partial least squares regression in reproducing kernel hilbert space. *J. Mach. Learn. Res.* **2001**, *2*, 97–123. [[CrossRef](#)]
23. Francois, D.; Rossib, F.; Wertza, V.; Verleysen, M. Resampling methods for parameter-free and robust feature selection with mutual information. *Neurocomputing* **2007**, *70*, 1276–1288. [[CrossRef](#)]
24. Jain, A.K.; Duin, R.P.W.; Mao, J.C. Statistical Pattern Recognition: A Review. *IEEE Trans. Pattern Anal.* **2000**, *22*, 4–37. [[CrossRef](#)]
25. Meyer, P.E.; Schretter, C.; Bontempi, G. Information-Theoretic Feature Selection in Microarray Data Using Variable Complementarity. *IEEE J.-STSP* **2008**, *2*, 261–274. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).