Filomat 32:5 (2018), 1629–1638
https://doi.org/10.2298/FIL1805629Z

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Prediction of NOx emission of a power plant boiler based on adaptive simplified T-S model

Wenjie Zhao[a], Baohui Guo[a], Fuchang Chu[a], Yongjie Zhai[a]

[a] *School of Control and Computer Engineering, North China Electric Power University, Baoding 071003, China*

Abstract. The combustion process of power plant boiler often has strong nonlinearity, uncertainty and time-varying nature. Offline models are often difficult to predict the emission of NOx accurately. In this paper, online identification method based on simplified T-S model for prediction of NOx emission is proposed. Firstly, the online subtractive clustering algorithm is used to determine the number of clusters and the clustering center adaptively. After obtaining the parameters, the recursive least squares algorithm is used to identify the parameters of each local model. Finally, the Mackey-Glass chaotic time series and field data of NOx emission of a power plant boiler are predicted by the above algorithm. The results show the effectiveness of the algorithm.

1. Introduction

Nitrogen oxides (NOx) are the products of combustion of power plant boilers, and NOx emissions are closely associated with the working condition of the boilers. Controlling the tunable parameters of the combustion system and optimizing the combustion conditions of boilers can radically cut down the emissions of NOx([1]-[3]). Mathematical model of NOx emissions is the basis for devising the optimization control strategy. Combustion system of power plant boilers is a complex thermal process, featured by nonlinearity and time-varying nature. It is still difficult to model the combustion system of power plant boilers.

In literature [4], nonlinear system identification was realized based on extreme learning machine (ELM) with good effects. Support vector machine ([5],[6]) and neural network[21] have been found extensive applications in the modeling of nonlinear systems. The combination model of gray prediction and BP neural network of thermal power NOx emission is established in [7], and the prediction results are accurate. In addition, the chaos optimization support vector machine model [8] and Bayesian regularization model [9] have also achieved good results. However, the models used are not white box models and they are complex and not fit for the control of field implementation. Literature [10] built the Takagi-Sugeno(T-S) model, which represents the input space with several rules, with each rule corresponding to a local linear model. The output of a global model is obtained by center of gravity defuzzification of local system output.

2010 *Mathematics Subject Classification*. 91D30.
Keywords. Subtractive clustering; online identification; T-S model; NOx emission
Received: 18 September 2017; Accepted: 7 December 2017
Communicated by Hari M.Srivastava
Research supported by the National Key R&D of China (Grant No: 2016YFB0600701) and the Hebei Province Natural Science Foundation of China (No.F2014502059)
Email addresses: zwj12342234@126.com (Wenjie Zhao), 18730271886@163.com (Baohui Guo), 1318061263@qq.com (Fuchang Chu), 13603123517@163.com (Yongjie Zhai)

This modeling method, as a multi-model integration, can well describe nonlinear systems. In addition, the consequent part is linear system, which can be readily analyzed using mature linear system theories. That is why the T-S model has made great progress [11]. In the literature [12], the fuzzy c-regression model is used to find the appropriate rules of the TS fuzzy model. The ant colony optimization algorithm and the orthogonal least squares method are used to obtain the precondition structure and subsequent parameters.

The T-S fuzzy model is established and the results are obtained. However, multi-model systems are generally identified by offline methods [13], which require a large amount of data covering all working condition. Therefore, the model cannot be updated in real-time based on the changes in process object and hardly applies to time-varying systems. Literature [14] proposed a recursive form of clustering algorithm through a simplified representation of fuzzy rules, thus realizing online identification of nonlinear systems. However, this algorithm requires the pre-determined number of clusters, and therefore it is not online identification in a real sense.

We propose a fuzzy identification algorithm by combining with online subtractive clustering ([13],[15]). As the sample input increases, the numbers of clusters and cluster centers can be determined adaptively based on online subtractive clustering. With different rules obtained, the parameters of local model under each rule are identified by using recursive least squares (RLS) algorithm. Finally, the Mackey-Glass chaotic time series and the NOx field data of a power plant boiler are predicted by the above algorithm. The results show the effectiveness of the algorithm.

2. Representation of simplified T-S model

2.1. T-S model

T-S model[10] is a modeling method that estimates the global nonlinear features by fuzzy reasoning based on local linearization. T-S model consists of several IF-THEN rules. The T-S model of a MISO system can be represented as follows:

$$R^i\text{: IF } x_1 \text{ is } A_1^i, x_2 \text{ is } A_2^i, \ldots, x_m \text{ is } A_m^i, \text{ THEN } y^i = b_0^i + b_1^i x_1 + b_2^i x_2 + \ldots + b_m^i x_m \tag{1}$$

System output is obtained by defuzzification:

$$y = \frac{\sum_{i=1}^{c} \mu^i y^i}{\sum_{i=1}^{c} \mu^i}, \mu^i = \prod_{j=1}^{m} A_j^i(x_j) \tag{2}$$

where $R^i(i=1,2,\ldots,c)$ is the i-th rule; x_i ($i=1,2,\ldots,m$) is the i-th input variable; A_j^i is the fuzzy set over the universe of discourse of input; y_i is the output of the i-th rule; b_m^i is a linear parameter of rule consequent; y is model output; μ^i is the degree of satisfaction of the i-th rule; $A_j^i(x_j)$ is the membership function of x_j with respect to A_j^i , which is usually in the form of a Gaussian function:

$$A_j^i(x_j) = e^{-\frac{(x_j - c_j^i)^2}{\sigma_j^{i2}}} \tag{3}$$

where c and σ are predetermined parameters.

2.2. Simplified T-S model

T-S model requires complex nonlinear dynamic programming to determine fuzzy rules and structure and parameters of the model. Therefore, it is not fit for online identification of the model [16]. By simplifying the antecedent structure of fuzzy rules, we obtain a simplified T-S model [14]. A simplified T-S model of a MISO system has the following form:

$$\text{IF } x \text{ is } (v_i, r_i), \text{THEN } y_i = \theta_i^T z (i = 1, 2, \ldots, q) \tag{4}$$

where $x = [x_1, x_2, ..., x_m]^T$ is the vector of input data; v_i is the cluster center of the i-th data; r_i is the radius of influence of the i-th cluster; y_i is the output of the i-th cluster; $z = [1, x_1, x_2, ..., x_m]^T$, $\theta_i = [a_{i0}, a_{i1}, ..., a_{in}]^T$, θ_i is the consequent parameter of the i-th rule. Identification of cluster center v and radius r in the simplified T-S model is identification of antecedent parameters. Identification of parameter θ in consequent function is identification of consequent parameters. Both are collectively known as identification of the simplified T-S model.

According to equation (4), the system output is

$$y = \frac{\sum_i w_i \theta_i^T z}{\sum_i w_i} \tag{5}$$

where weight coefficient w_i is obtained based on the distance between the input vector and local cluster center v_i,

$$w_i = exp(-\frac{||x(t) - v_i||}{r_i}\gamma) \tag{6}$$

where γ is a sensitivity parameter that reflects the speed of reduction in the degree of membership as the input samples are far away from the center of the data region. When the fresh samples are far away from every cluster, the value of the above equation is small and it has no impact on the output. This method does not use adaptive reasoning, and therefore has no influence on the precision and speed of identification.

3. Online identification algorithm

Training data are usually input continuously in online identification, and they are never fixed data. New data enhance and confirm the information contained in the existing data, while bringing new information into the model. These new data indicate changes in operational conditions or working conditions [11]. This raises the need for online identification of the system, so as to constantly modify the model to better fit the real system.

According to equation (4), identification of antecedent parameters yields (v_i,r_i) and identification of consequent parameters yields θ_i. This is combined with equation (5) and (6) to obtain the model output. On this basis, we will obtain the simplified T-S model of nonlinear system.

3.1. Identification of antecedent parameters of simplified T-S model

3.1.1. Online identification of cluster center v_i

As more operational data are introduced, new rules will be formed from the new information or the existing rules will be modified. The new information contained in the updated data is closely connected to the data that are already collected. The degree of importance of the information is judged based on their spatial proximity.

Subtractive clustering takes all data points as candidate cluster centers and calculates their densities [15]. Density is a measure of the distance from one specific data point to all other samples, and it has the following form [17]:

$$P(x(i)) = \frac{1}{1 + \frac{1}{(N-1)} \sum_{j=1, j\neq i}^{N} (x(i) - x(j))^T (x(i) - x(j))} \tag{7}$$

where $i = 1, 2, ..., N$; $x(i)$ is the data point in dataset X; $X = [x(1), x(2), ..., x(N)]$; N is the number of data points in dataset X.

The first data point $x(1)$ is taken as the first cluster center and given the value $P_1(x(1))$=1. Here $P_k(k = 1, 2, ...)$ is the density calculated at time k.

At time k, suppose there already exists l cluster centers $v_i(i = 1, 2, ..., l)$. Then the density from the current data point $x(k)$ to the old cluster center v_i is calculated and compared. The density of $x(k)$ is calculated as below:

$$P_k(x(k)) = \frac{1}{1 + \frac{1}{(N-1)} \sum_{j=1}^{k-1} (x(k) - x(j))^T (x(k) - x(j))} = \frac{k-1}{(k-1)(\gamma(k)+1) - 2\eta(k)x(k) + \sigma(k)} \tag{8}$$

where $\eta(k) = \sum j = 1k - 1x^T(j)$,$\gamma(k) = x^T(k)x(k)$,$\sigma(k) = \sum j = 1k - 1x^T(j)x(j)$lies on the right side of equation (2); $\gamma(k)$ and $x(k)$ are calculated from the current data points. There exist the following recursive relations between other parameters:

$$\eta(k) = \eta(k-1) + x^T(k-1) \tag{9}$$

$$\sigma(k) = \sigma(k-1) + x^T(k-1)x(k-1) \tag{10}$$

It can be known from equation (8) to (10) that to calculate $P_k(x(k))$, there is no need to preserve all data before time k. Only by preserving $\eta(k-1)$, $\sigma(k-1)$ and $x(k-1)$ can we derive recursively the current $x(k)$.

For old cluster center v_i, the density at time $k-1$ is calculated from equation (7):

$$P_{k-1}(v_i) = \frac{1}{1 + \frac{1}{(k-2)} \sum_{j=1, j\neq i}^{k-1} (v_i - x(j))^T (v_i - x(j))} \tag{11}$$

Then from the density v_i at time k, the recursive equation can be described as follows:

$$P_k(v_i) = \frac{1}{1 + \frac{1}{(k-1)} \sum_{j=1, j\neq i}^{k} (v_i - x(j))^T (v_i - x(j))} = \frac{(k-1)P_{k-1}(v_i)}{k - 2 + P_{k-1}(v_i) + \zeta(k)P_{k-1}(v_i)} \tag{12}$$

where $\zeta(k) = (v_i - x(k))^T (v_i - x(k))$ is calculated from the current $x(k)$.

Based on the definition of density in equation (7), the density of data points reflects the ability of the data point in describing the surrounding data. The greater the amount of data surrounding this data point, the higher the density and vice versa. Therefore, the following situations are considered when comparing $P_k(x(k))$ and $P_k(v_i)(i = 1, 2, ...l)$:

$$\text{a. IF } \min||x(k) - v_i||_2 > r_1 \text{ and } P_k(x(k)) > \xi, \text{ THEN } v_{l+1} = x(k), l = l + 1(i = 1, 2, ..., l). \tag{13}$$

where $r_1 = (0.3 - 0.5) * \arg\max||x_i - x_j||_2$ $(i \neq j, i, j = 1, 2, ..., N)$, $\xi = 0.5 \max P_{k-1}(v_i)(i = 1, 2, ..., l)$.

It is easy to be seen that this data point has a sufficiently large distance from the nearest cluster center and the density is above the threshold. This indicates that this data point has a stronger ability in describing the surrounding values as compared with the old cluster center. It represents a new working condition.

$$\text{b. IF } \min||x(k) - v_i||_2 \leq r_1 \text{ and } P_k(x(k)) > P_k(v_t), \text{ THEN } v_t = x(k)(i = 1, 2, ..., l). \tag{14}$$

where $t = \arg\min||x(k) - v_i||_2(i = 1, 2, ..., l)$.Apparently, situation (b) indicates that the new data point exhibits a greater describing ability than the nearest cluster center and the new data point is very close to this cluster center. This implies an adaptive transition of the center of working condition along with the operation of the system. If new clusters are formed at this time, there will be too many clusters. To avoid this, the nearest cluster center will be replaced by the new data point, and the parameters of the local parameters are solved after the transition of the working condition.

$$\text{c. IF } P_k(x(k)) \leq P_k(v_t), \text{ THEN } v_i = v_i(i = 1, 2, ..., l). \tag{15}$$

When the density of the new data points is lower than the density of the nearest cluster center, its effect on the clustering result can be neglected and the original clustering is maintained, whatever the distance from the cluster center.

After the above online fuzzy clustering, the models structural parameters and clustering result have the following relationship:

$$n = l; c_i = v_i(i = 1, 2, ..., l) \tag{16}$$

3.1.2. Online identification of radius r_i

The radius of each input region is calculated after obtaining the data input center [16]:

Step 1. Initialize the radius, $r_i = 0(i = 1, 2, ..., n)$.

Step 2. For the new data point $x(t)$, calculate its distance to each cluster center to obtain the nearest input center v_i:

$$\|x(t) - v_s\| = \min \|x(t) - v_i\|(i = 1, 2, ..., n) \tag{17}$$

Step 3. Update the radius of input region r_s with v_s as the cluster center:

$$r_s = \max\{\|x(t) - v_s\|, r_s\} \tag{18}$$

Step 4. Repeat step 2 and 3 to obtain the radius of each cluster $r_i(i = 1, 2, ..., n)$.

3.2. Identification of consequent parameters in the simplified T-S model

Since the consequent part of the fuzzy model is local linear model, the consequent parameters can be identified using RLS method.

From the above clustering process there is $(v_i, r_i)(i = 1, 2, ..., n)$. For the new data point $(x(t), y(t))$,

a. IF $\|x(t) - v_i\| \le r_i$

$$\theta_i(k+1) = \theta_i(k) + P_i(k+1)x(t)[y(t) - x^T(t)\theta_i(k)], P_i(k+1) = P_i(k) - \frac{P_i(k)x(t)x^T(t)P_i(k)}{1 + x^T(t)P_i(k)x(t)} \tag{19}$$

b. IF $\|x(t) - v_i\| > r_i$

$$\theta_i(k+1) = \theta_i(k), P_i(k+1) = P_i(k) \tag{20}$$

c. $k = k + 1, (k = 1, 2, ..., N)$. The new data points are input continuously, and steps a and b are implemented to obtain the parameters of each rule.

The determination of initial values of $\theta(0)$ and $P(0)$ in the recursive process is a key problem. Generally the value of $\theta(0)$ is taken as zero or even smaller. Here $P(0) = I$, where α is a sufficiently large real number and $\alpha = 10^6$.

To sum up, the online identification algorithm proposed in this study consists of the following steps:

Step 1. Input data point $\{x(1), y(1)\}$, and initialize each value. Take the first data sample $x(1)$ as the first cluster center and give it the value $P_1(x(1)) = 1$. Initialize each cluster radius, $r_i = 0(i = 1, 2, ..., n)$. Set $\theta(0)$ to zero, $P(0) = \alpha I$, $\alpha = 10^6$. I is the identity matrix.

Step 2. Continuously input data point $\{x(t), y(t)\}$ (t=2,3,...).

Step 3. Calculate the density of the input data points using equation (8). Calculate the density of each cluster center using equation (12).

Step 4. Implement equation (13)-(15) and update the clustering rules.

Step 5. Calculate the cluster radius r_i using equation (17) and update the cluster center v_i using equation (18).

Step 6. Update consequent parameter θ_i using equation (19) or (20).

Step 7. Calculate the degree of membership of the data point using equation (6).

Step 8. Calculate system output using equation (5).

Step 9. Turn to step 2.

Online identification of nonlinear system is thus realized following the above steps.

4. Simulation example

4.1. Prediction of chaotic time series

Mackey and Glass[18] first reported the chaos phenomenon in the time delay system in 1977. The time-delay chaotic system began to draw an increasing attention after that. However, the fact that the time-delay system has infinite-dimensional state space makes the study of time-delay system very difficult. Here Mackey-Glass chaotic time series are predicted to verify the effectiveness of the proposed algorithm. Chaotic time series can be estimated using delay differential equation [19]:

$$\dot{x} = \frac{\alpha x(t-\tau)}{1+x^{\gamma}(t-\tau)} - \beta x(t) \tag{21}$$

Here $\alpha = 0.2$,$\beta = 0.1$,$\gamma = 10$, and τ is the only tunable parameter. Chaos phenomenon occurs with the above equation when $\tau \geq 17$ (see Figure 2).

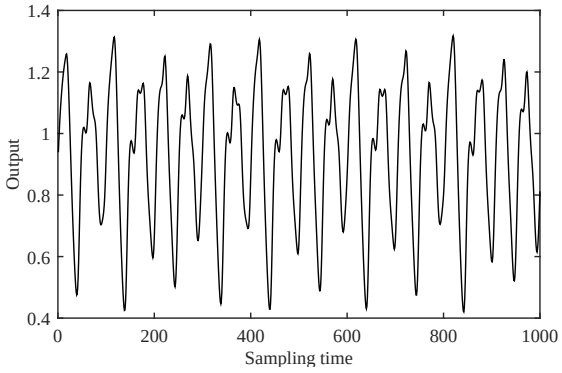

Figure 1: Phase diagram of Mackey-Glass chaotic time series at τ=17.

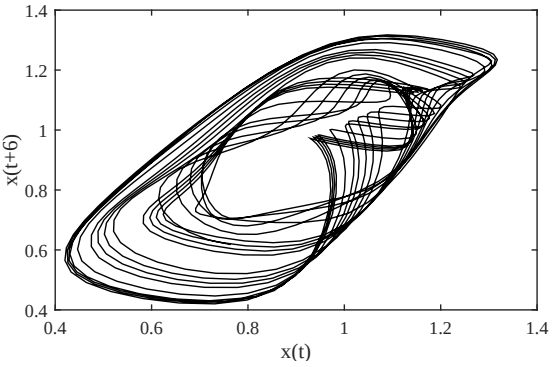

Figure 2: Phase diagram of the system at τ=17.

Chaotic time series are approximated using the following model:

$$x(t+1) = f(x(t), x(t-6), x(t-12), x(t-18)) \tag{22}$$

From equation (22), 3000 data points are generated. The first 2500 data points are the training data and the last 500 data points are the testing data. Identification accuracy is measured using the performance indicator below:

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(y(t) - \hat{y}(t))^2 \tag{23}$$

Figure 3 is the comparison between the predicted output and the actual output. It can be seen that the proposed algorithm has a high prediction accuracy.

The above model yields the following cluster centers:

$$[v_1, v_2, v_3, v_4, v_5]^T = \begin{bmatrix} 1.1157 & 0.9343 & 0.9696 & 0.8544 \\ 1.1323 & 0.9496 & 0.9630 & 0.8991 \\ 0.9590 & 1.0414 & 1.1521 & 1.0632 \\ 1.0517 & 0.9335 & 0.9672 & 0.7310 \\ 0.9104 & 0.7458 & 0.7158 & 0.9217 \end{bmatrix}$$

Cluster radius is:

$$[r_1, r_2, r_3, r_4, r_5] = \begin{bmatrix} 0.7475 & 0.7956 & 0.7126 & 0.7202 & 0.7033 \end{bmatrix}$$

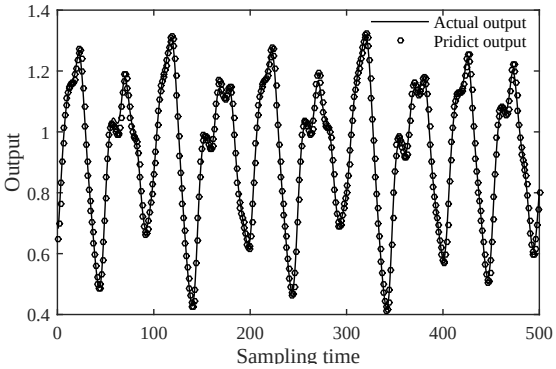

Figure 3: Comparison between the predicted output(dashed line) and the actual output (solid line).

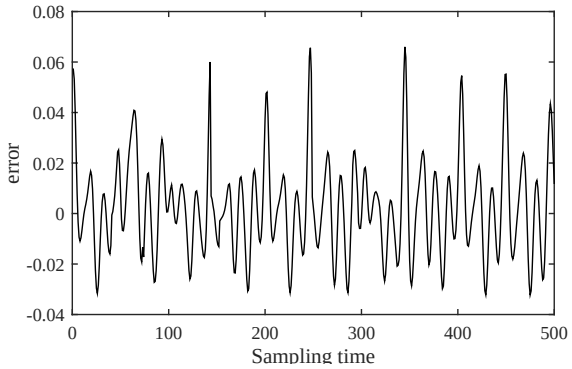

Figure 4: Error of model prediction.

There is no need to know the number of clusters under the adaptive clustering algorithm. The final number of clusters is reasonable and the prediction accuracy is high, which verifies the effectiveness of the proposed algorithm.

4.2. NOx prediction of power plant boiler

4.2.1. Brief description of the boiler

No. 2 boiler at a power plant was tested. It was a 1099.3t/H forced circulation boiler. The combustion system is shown in Figure 5 [20]. The boiler had a double-arch furnace with W-shaped flame and 36 direct flow burners in double arches. The flames were ejected from the two rows of burners to the center of the furnace bottom in an oblique direction. The flames bent upwards to form a W-shape at the water-cooled hopper.

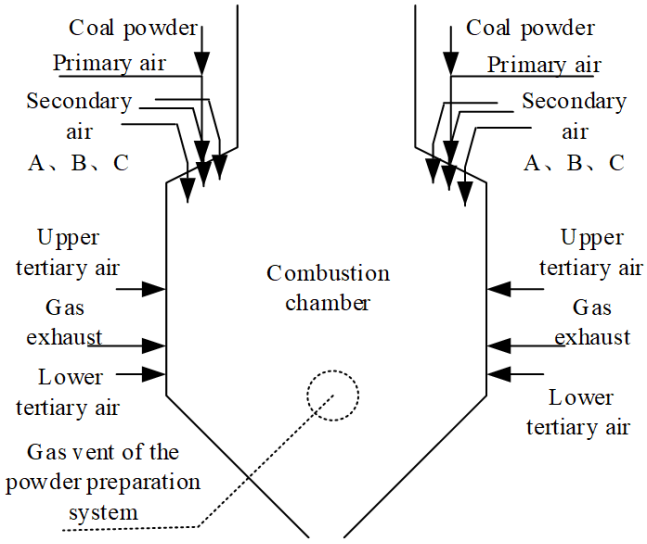

Figure 5: Combustion system of the boiler

Orthogonal test was performed for the boiler, and 105 groups of data on steady-state working conditions were obtained. According to the generation mechanism of NOx, the following variables were chosen as the input data: unit load *LOAD*, combustion value *Q*, fuel volatility *V*, primary air pressure *PA*, opening

Table 1: Steady-state working condition of the boiler

data	$LOAD/MW$	$Q/MJkg^{-1}$	$V/\%$	PA/kPa	$SE_A/\%$	$SE_B/\%$	$SE_C/\%$	$SR_U/\%$	$SR_D/\%$	$O_2/\%$	NOx/mgm^{-3}
1	237.68	23.1	9.27	3.46	41	7	42	20	50	3.675	659.18
2	238.76	23.11	8.85	3.46	41	7	42	20	50	3.445	599.43
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
40	291.53	19.23	9.67	3.5	80	20	55	5	50	1.585	616.70
41	292.04	23.64	8.68	3.64	94	36.5	67	65	75	1.76	664.70
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
80	347.84	21.7	9.97	3.57	53	22	43	33	33	1.805	824.05
81	348.3	21.02	9.27	4.13	67	30	55	33	57	1.16	708.77
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 2: Comparison of prediction accuracy of different models

Performance indicator	Literature [20]	Our study
$MAE(mg/m^3)$	24.1337	16.5108
$MRE(\%)$	0.0337	0.0249

of three layers of secondary air baffles SE_A, SE_B and SE_C (upper, middle and lower), opening of two layers of tertiary air baffles SR_U and SR_D, oxygen content in flue gas O_2. The output variable was NOx emission. Some of the experimental data are shown in Table 1.

4.2.2. Model identification

The first 90 groups of experimental data were used as training data, and the last 15 groups as the testing data. The effectiveness of the model was verified through experiments, and two indicators of prediction accuracy were used.

The two indicators of prediction accuracy were calculated as follows:

$$MRE = \frac{1}{N}\sum_{t=1}^{N}\frac{|y(t)-\hat{y}(t)|}{y(t)} \tag{24}$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|y(t)-\hat{y}(t)| \tag{25}$$

where $y(t)$ is actual output; $\hat{y}(t)$ is predicted output. The prediction result is shown in Figure 6.

It can be seen from the model that the proposed model has a high accuracy and provides a good prediction of NOx emission. Table 2 is the comparison of accuracy indicators between different algorithms.

As shown above, the proposed algorithm has a high stability, as well as strong adaptive learning ability and generalization performance. Compared with LS-SVM in literature [14], our model is a white box model, which is featured by higher generalization performance and accuracy and simple structure. Our model is more suitable for online prediction and field applications.

5. Conclusion

We propose an online identification method for nonlinear systems based on simplified T-S model. This algorithm can achieve online identification of complex nonlinear systems and its effectiveness is verified through the prediction of Mackey-Glass chaotic time series and field data of NOx emission of a power plant boiler. The online recursive form of subtractive clustering is applied to fuzzy modeling. Data are divided into regions adaptively and local linear model is established. Compared with black box models such as neural network and LS-SVM, the consequent part of the proposed algorithm is a linear model, which is fit for the control of field implementation. Our algorithm has fewer rules than T-S fuzzy model and can

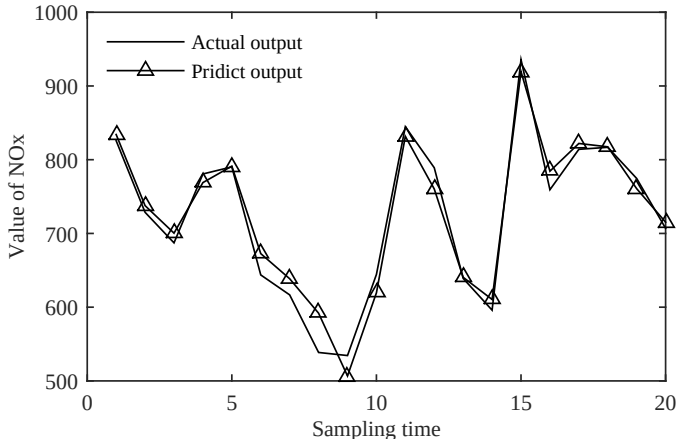

Figure 6: Prediction result using the proposed algorithm

determine the number of clusters in an online, adaptive manner. These features make our algorithm highly suitable for real-time field modeling.

References

[1] Peng Tan, Ji Xia, Cheng Zhang, etc, Modeling and optimization of NOX emission in a coal-fired power plant using advanced machine learning methods, Energy Procedia 61 (2014) 377–380.
[2] T.K.SaiK.A.Reddy, Measurement and Control of NOx emissions using Soft Computing in a Thermal power Plant, Proceedings of 2014 RAECS UIET Punjab University Chandigarh (2014) 06–08.
[3] Zhao W J, Zhao G, Lv M, et al, Fuzzy optimization control for NOx emissions from power plant boilers based on nonlinear optimization1, Journal of Intelligent and Fuzzy Systems 29 (2015) 2475–2481.
[4] Grzeidak E, Vargas J A R, Alfaro S C A, Online neuro-identification of nonlinear systems using Extreme Learning Machine,International Joint Conference on Neural Networks, (2016) 2923–2930.
[5] Marconato, Anna and Boni, Andrea and Petri, Dario and Schoukens, Johan, Identification of weakly nonlinear systems based on Support Vector Machines,Instrumentation and Measurement Technology Conference, 2009. I2MTC '09. IEEE, (2009) 122–127.
[6] Filho, Orlando Donato Rocha and Serra, Ginalber Luiz De Oliveira, Online identification based on instrumental variable evolving neuro-fuzzy model for stochastic dynamic systems, 2016 IEEE International Conference on Fuzzy Systems (2016) 9–16.
[7] Zhao Y, Zhou J G, Prediction of NOx Emissions from Thermal Power Plants in China Based on a Grey Support Vector Machine Combination Model, Research of Environmental Sciences 24 (2011) 489–496.
[8] Zhou J G, Liang H T, Peng Q, et al, Prediction of the NOx Emissions from Thermal Power Plant Based on Chaotic Optimized SVM, East China Electric Power 39 (2011) 1335–1339.
[9] Yang F. Prediction of NOx Emissions Based on Principal Components and Bayesian Regularization, Thermal Power Generation 39 (2010) 24–27.
[10] Takagi T, Sugeno M. Sugeno, Fuzzy Identification of Systems and its Applications to Modeling and Control,IEEE Transactions on Systems Man and Cybernetics 15 (1985) 116–132.
[11] Angelov P P, Filev D P, An Approach to Online Identification of Takagi-Sugeno Fuzzy Models, IEEE Transactions on Systems Man and Cybernetics Part B Cybernetics A Publication of the IEEE Systems Man and Cybernetics Society 34 (2004) 484–498.
[12] S. H. Tsai and Y. W. Chen, A Novel Fuzzy Identification Method Based on Ant Colony Optimization Algorithm, in IEEE Access 4 (2016) 3747–3756.
[13] Tianhong Pan, Zhefkuang Xue, Shaoyuan Li, An Online Multi-model Identification Algorithm Based on Subtractive Clustering , Acta Automatica Sinica 35 (2009) 220–224.
[14] Jianqin Chen, Yugeng Xi, On-Line Identification of Nonlinear Systems Using Fuzzy Model, Acta Automatica Sinica 24 (1998) 90–94.
[15] Karimoddini A, Salahshoor K, Fatehi A, et al. A new approach for online fuzzy identification by potential clustering including rule reduction,Control Conference. IEEE (2015) 747–754.
[16] Jianhong Lv, Jianqin Chen, Zhiyuan Liu, A Study On Fuzzy Rules Based and Nonlinear Models For Thermal Processes,Proceedings of The Chinese Society for Electrical Engineering 22 (2002) 132–137.
[17] Chiu S L. Fuzzy Model Identification Based on Cluster Estimation,Journal of Intelligent and Fuzzy Systems Applications in Engineering and Technology 2 (1994) 267–278.
[18] Mackey M C, Glass L, Oscillation and chaos in physiological control systems, Science 197 (1977) 287–289.
[19] W. Xie and X. x. Zhang, An improved online self-organizing dynamic fuzzy neural network for nonlinear dynamic system identification, 2016 UKACC 11th International Conference on Control (CONTROL), Belfast, (2016) 1–6.

[20] Wenjie Zhao, Meng Lv, NOx emission modeling of coal-fired boiler based on multi LS-SVM ensemble model, Journal of Electronic Measurement and Instrumentation 30 (2016) 1037–1044.
[21] Yu W, Li X,Some new results on system identification with dynamic neural networks,IEEE Transactions on Neural Networks 12 (2001) 412–417.
[22] Ghoul Y., Ibn Taarit K. and Ksouri M., Online identification of continuous-time systems with multiple unknown time delays based on optimization approaches from sampled data, 2016 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), Hammamet, Tunisia, (2016) 204-208.
[23] K. S. Arysheva, Fuzzy logic and the mountain clustering, 2014 International Conference on Mechanical Engineering, Automation and Control Systems (MEACS), Tomsk, (2014) 1–4.
[24] C. T. Lin, N. R. Pal, S. L. Wu, Y. T. Liu and Y. Y. Lin, An Interval Type-2 Neural Fuzzy System for Online System Identification and Feature Elimination, in IEEE Transactions on Neural Networks and Learning Systems 26 (2015) 1442–1455.
[25] Sugeno M, Yasukawa T, A fuzzy-logic-based approach to qualitative modeling, IEEE Transactions on Fuzzy Systems 1 (1993) 7.
[26] Salahshoor K, Gharibshaiyan S, Online multivariable identification of a nonlinear distillation column using an adaptive Takagi-Sugeno fuzzy model, Cybernetics and Intelligent Systems, 2008 IEEE Conference on. IEEE, (2008) 527–532.
[27] Chen J Q, Lu J H, Chen L J, An on-line identification algorithm for fuzzy systems , Fuzzy Sets and Systems 64 (1994) 63-72.
[28] Bezdek, James C. Fuzzy models for pattern recognition , Institute of Electrical and Electronics Engineers, 1992.
[29] Angelov P P, Evolving Rule-Based Models: A Tool for Design of Flexible Adaptive Systems, 2002.