Journal of Environmental Engineering and Landscape Management
ISSN 1648–6897 / eISSN 1822-4199
2020 Volume 28 Issue 2: 74–87
https://doi.org/10.3846/jeelm.2020.12098

# PREDICTION AND OPTIMIZATION OF A DESULPHURIZATION SYSTEM USING CMAC NEURAL NETWORK AND GENETIC ALGORITHM

Zhiwei KONG[1], Yong ZHANG[1], Xudong WANG[1], Yueyang XU[2], Baosheng JIN[1*]

[1]*Key Laboratory of Energy Thermal Conversion and Control of Ministry of Education, School of Energy and Environment, Southeast University, 210096 Nanjing, China*
[2]*Guo Dian Science and Technology Research Institute Co. Ltd., 210023 Nanjing, China; State Key Laboratory of Clean and Efficient Coal-Fired Power Generation and Pollution Control, 210023 Nanjing, China*

Received 05 November 2019; accepted 17 January 2020

**Highlight**

- Based on 44640 sets of data, a ten-input two-output CMAC model was built for WFGD process with high accuracy.
- Proposed GREA-UDM-CMAC-LNA-GA methodology is logical and effective for model building and optimization of high-dimensional data-generation systems.
- Compared with optimal operation parameters, optimized economic cost with GA was reduced by more than 30% under the same input conditions and restrains.

**Abstract.** In this paper, taking desulphurizing ratio and economic cost as two objectives, a ten-input two-output prediction model was structured and validated for desulphurization system. Cerebellar model articulation controller (CMAC) neural network and genetic algorithm (GA) were used for model building and optimization of cost respectively. In the model building process, the grey relation entropy analysis and uniform design method were used to screen the input variables and study the model parameters separately. Traditional regression analysis and proposed location number analysis method were adopted to analyze output errors of experiment group and predict the results of test group. Results show that regression analyses keep high fit degree with experiment group results while the fitting accuracies for test group are quite different. As for location number analysis, a power function between output errors and location numbers was fitted well with the data of experiment group and test group for $SO_2$. Prediction model was initialized by location number analysis method. Model was validated and cost optimization case was performed with GA subsequently. The result shows that the optimal cost obtained from GA could be reduced by more than 30% compared with original optimal operating parameters under same constraints.

**Keywords:** desulphurization system, cost, optimization, CMAC, GA.

## Introduction

In last several years, under the pressure of environmental issues, many scholars have been committed to the optimization of environmental facilities in thermal power plants. As the most widely used desulphurization method, wet flue gas desulphurization (WFGD) has drawn many attentions absolutely. So far, researchers have built systematic theoretical models, performed detailed numerical simulation, carried out complex simulation experiments and verified the models for the system (Kiil et al., 1998; Warych & Szymanowski, 2001; Gutiérrez et al., 2006; Dou et al., 2009; Wang & Dai, 2018; Liu et al., 2019). In recent years, with the rise and rapid development of intelligent computing, more and more researchers have begun to build mathematical models and optimize the process using operation data.

Proportion Integration Differentiation (PID) control is still an important control method for power stations. However, due to the characteristics of high dimensionality, multivariate, time varying, wide operating conditions and time delay problems, it is difficult for conventional PID control to achieve rapidity and accuracy simultaneously for WFGD process (Villanueva Perales et al., 2010). In this situation, some researchers began to experiment with machine learning methods to help control PH. Wu et al.

*Corresponding author. E-mail: *bsjin@seu.edu.cn*

(2011) studied external recurrent neural network (RNN) for PH control in the absorber. Yang et al. took advantage of Radial Basis Function (RBF) neural network to tune PID parameters online (Yang et al., 2016). Cheng with her fellows applied Back-Propagation (BP) network monitoring the PH of the tank directly to predict desulphurization process and use Levenberg-Marquard (LM) algorithm to improve the BP network (Cheng et al., 2017). They also reproduced the control strategy using RNN neural network like Wu et al. (Cheng & Xie, 2017). Fu et al. (2019) used long short-term memory neural network establishing an efficiency prediction model and got higher efficiency than the least squares support vector machine model and RNN model. Yet, these studies only focus on the traditional parameter PH to ensure sulfur removal rate, which ignored the important parameter of energy consumption or capital cost. Setting the desulphurization cost as first goal and satisfying different sulfur removal rates, Warych and Szymanowski (2002) proposed a systematic cost calculation methodology and used BP net for specific operation parameters' optimization. Guo et al. (2019) used BP net and particle swarm optimization algorithm to model and optimize desulphurization efficiency and comprehensive cost of WFGD system, respectively. In the small quantity of neural network-based optimization studies for WFGD, BP net are in absolute dominance. Although BP net can realize the modelling of WFGD system, it has been widely recognized as having problems of slow learning speed and falling into local minimums easily (D. Jin & Lin, 2012; W. Jin et al., 2000). As for the desulphurization system, it produces a large amount of data all the time, which stings the weakness of BP net. Moreover, lacking of clear guiding theory for the numbers of hidden layers and neurons in the construction process imposes restrictions on the popularization of obtained results. Therefore, model selection is an important part of the research field of WFGD for intelligent computing.

Modelling the structure and function of the cerebellum, CMAC was put forward by Albus for manipulator control problems (Albus, 1975a). Different from other neural networks, CMAC belongs to table reference technology. It uses memory system achieving mapping and has characteristic of partial learning (Albus, 1975b). CMAC has the advantages of fast learning and response, no local minimum problem, being suitable for large amounts of data and possesses fair generalization ability (Albus, 1975b; Ching-Tsan & Chun-Shin, 1996). As can be seen, the characteristics and advantages of CMAC well fit the needs of desulphurization system. Therefore, the application of CMAC in desulphurization system is reasonably expected to achieve satisfying prediction accuracy, which will be further favorable for process optimization. However, the studies on the error bound of CMAC are insufficient (Lin & Chiang, 1997; Lin & Wang, 1996). Physical storage size on the computer required by data is usually the first parameter to be determined (Lin & Wang, 1996; Tamura et al., 2017). This method is feasible for simple models and small data volumes. However, for multi-dimensional and data-generation objects, it is conceivable that the practice of the method will be less effective.

Slow learning speed and falling into local minimums easily of BP net restrain its operation data volume. CMAC can process large volume of data but deficient studies of its error bound limit applications of the method. If an appropriate method can be used to predict the error of CMAC, then the large amount of data will no longer be a disaster, but an advantage for multi-dimensional and data-generation objects. Uniform design method (UDM) has a simple and mature error analysis system, which is incomparable to analyze and predict errors on the basis of limited experiments (Fang et al., 2000). It is more than appropriate to use the UDM to analyze and predict the output error of CMAC for complex systems. In short, using UDM to construct CMAC model for WFGD can not only provide a new mathematical modeling method for the process, but also promote the research of error analysis of CMAC, which will further promote its application in similar objects.

Based on above concepts, distinguishing from existing researches, this paper applies CMAC to build prediction model for WFGD system with desulphurizing ratio and the economic cost two objectives. Data volume is expanded further compared with published studies. A total of 44640 sets of original data are used. To be specific, in the establishment process of prediction model, since the grey relation entropy analysis (GREA) can make full use of the individual information to obtain overall proximity (Liu & Forrest, 2010; Biswas et al., 2014), the method was applied to screen model input variables. To reduce the number of experiments and analyze error effectively, UDM was then applied to design and study the CMAC model parameters. Common linear regression and quadratic linear regression were subsequently used for the analyses of obtained results. At the same time, considering the importance of memory space for CMAC, the size of memory space and output errors were also connected and investigated for error analysis (called as location number analysis method, LNA) innovatively. With the simple logic, mature theoretical system and global optimality, GA is still one of the main optimization tools (Mostofi & Hasanlou, 2017; Armaghani et al., 2018; Munroe et al., 2019). Hence, after the completement of CMAC prediction model building, GA is finally utilized to optimize the cost. A set of logical and scientific analysis methodology was actually proposed and attempted, which follows the process of GREA-UDM-CMAC-LNA-GA. The practice of the methodology will provide reliable experience and reference for many other similar systems with high dimension and large amount of data.

## 1. Research objects and methods

### 1.1. Double-loop wet flue gas desulphurization process

The data analyzed in this paper is from an ultra-supercritical unit in Jiangsu province, China. The unit adopts

WFGD mechanism, launching a double-loop desulphurization tower. The schematic is shown in Figure 1.

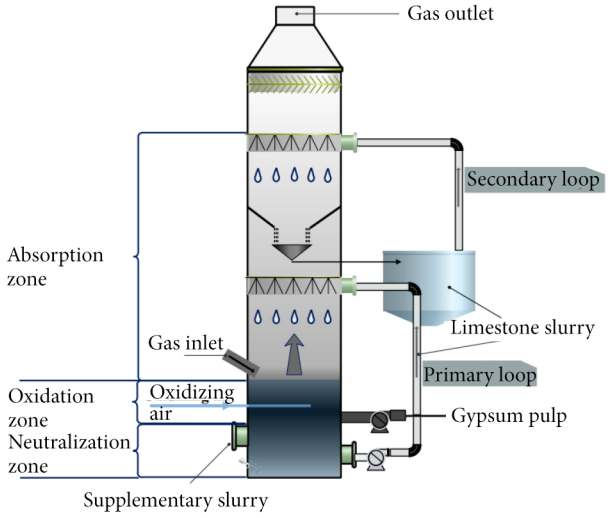


Figure 1. Double-loop WFGD schematic

In this system, when the flue gas enters the desulphurization tower and goes upwards, it meets water and dissolves into it. Then, the dissolved sulfur dioxide releases hydrogen ion, which combines with carbonate to form carbon dioxide. Following, sulfite reacts with calcium ion and oxygen to generate calcium sulfate precipitates. The reactions in these processes are described as R1–R3.

$$SO_2 + H_2O \rightleftharpoons HSO_3^- + H^+ \rightleftharpoons SO_3^{2-} + 2H^+; \quad (R1)$$

$$CO_3^{2-} + H^+ \rightarrow H_2O + CO_2\uparrow; \quad (R2)$$

$$Ca^{2+} + SO_3^{2-} + O_2 \rightarrow CaSO_4\downarrow. \quad (R3)$$

The entire reaction system operates continuously based on the interactions of these reactions. In addition, the double-loop tower is divided into independent upper and lower absorption spaces, which cooperate with and complement each other.

Two traditional indicators are adopted to evaluate the performances of the system, namely desulphurizing ratio ($R_{ds}$) and calcium sulfur ratio ($R_{cas}$), which are defined as followings:

$$R_{ds} = (C_{is}V_{in} - C_{os}V_{out})/C_{is}V_{in}; \quad (1)$$

$$R_{cas} = m_{CaCO_3}/m_{SO_2}, \quad (2)$$

where: $C_{is}$ and $Cos$ represents inlet and outlet concentration of $SO_2$ respectively, mg/m$^3$; $V_{in}$ and $V_{out}$ represents total inlet and outlet volume of flue gas respectively, $10^4$ m$^3$/h; $m$ represents mole number.

### 1.2. CMAC neural network

The parameters that need to be determined in a CMAC model include input and output variables, quantization level ($Q_l$) of each input variable, generalization parameter ($c$) and learning rate (ε). Take the two-dimensional model in Figure 2 as an example to illustrate the principle of CMAC. Under high dimensions, the squares would be hypercubes (Albus, 1975a, 1975b; Ching-Tsan & Chun-Shin, 1996).

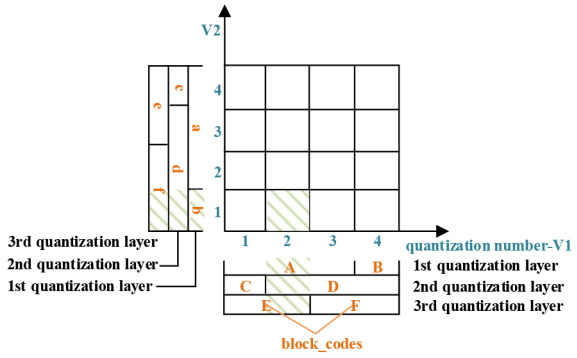


Figure 2. CMAC principle diagram

For every dimension, input data is first quantified by $Q_l$ to quantization numbers. According to different partition modes of quantization layers, quantization numbers are then matched with block codes. Subsequently, block codes of same layer from all dimensions are combined together (Ab, Dd, Ef). The combination of each layer generates feature code. Every feature code corresponds to a memory location, which stores one or more (depends on the number of output variables) weight values. The collection of all possible combinations constitutes the memory space, and weight matric is formulated at the same time. For certain input vector, the sum of the selected weights is the prediction output. The quantization layer number coincides with the value of $c$. Total location number ($ln$) or the size of memory space can be generalized as:

$$l_n = \sum_{i=1}^{c}\prod_{j=1}^{n} Q_{bn(ij)}; \quad (3)$$

$$Q_{bn(ij)} = \left\lceil (Q_{l(j)} - i + 1)/c \right\rceil + \left\lceil (i-1)/c \right\rceil, \ (1 \le i \le c, \ 1 \le j \le n), \quad (4)$$

where: $n$ represents dimension; $Q_{bn}$ represents quantization block number, which is the total number of block codes in certain dimension and certain quantization layer; symbol $\lceil\ \rceil$ refers to round up to an integer.

The calculation of weights ($w$) is usually an iterative process:

$$w' = w + \frac{\varepsilon}{c} \times (y - f(y)), \ (0 < \varepsilon < 1), \quad (5)$$

in which, $f(y)$ means prediction value of the model for output variable, $y$ means actual value of the set. $w'$ represents weight after iteration.

In the model building process, for the ease of comparison with uniform emission standard, the output variables are $Cost$ and $C_{os}$ but not $R_{ds}$, of which the $Cost$ is defined as follows.

$$Cost = (100M_{CaCO_3} + 0.35P_c)/m_{SO_2} \times 600, \quad (6)$$

in which, $M_{CaCO3}$ means the mass of $CaCO_3$, kg; $m_{SO2}$ means mole number of $SO_2$, mol; *Pc* represents power consumption, kW·h, which consists of stirrer power, slurry circulating power, oxygen air power and gypsum discharge pump power. The prices of every unit $CaCO_3$ (t) and auxiliary power (kW·h) are taken as 100 and 0.35 CNY. Multiplying the value by 600 is to avoid operation loss and error. The advantage of defining cost in this way is that different loads can be compared in the same class. In addition, mean relative error (MRE) (Ansari et al., 2018) index is used as training target so as to modify and improve accuracy of model, namely:

$$MRE = \frac{1}{N} \times \sum_{i=1}^{N} \frac{|f(y_i) - y_i|}{y_i}, \quad (7)$$

in which, *N* means total sets of data used. The advantage of taking MRE as indicator lies in that it emphasizes the relative error, which weakens the influence of actual values of data. In this way, the results obtained based on WFGD system can be extended to other similar objects.

### 1.3. Practice of GA in CMAC

General processes of GA are population initialization, fitness calculated, genetic evolution and population convergence (Whitley & Tutorial, 1994; Deb et al., 2002). However, the application of the combination GA and CMAC in industrial field needs some necessary preparation, including:

(1) Determine the values of characteristic parameters to be optimized;

(2) Select similar samples and determine the *maximum* and *minimum* values of the similar samples for the remaining input variables;

(3) Calculate quantization interval length ($Q_{il}$) of each variable based on the extremum of all preprocessed data and $Q_l$, namely:

$$Q_{il(i)} = (MAX_i - MIN_i) / Q_{l(i)}, \ i = 1, 2, ..., 7, \quad (8)$$

in which, $MAX_i$ and $MIN_i$ means the maximum and minimum value of preprocessed data for variable *i*.

(4) Determine new quantization levels $Ql'$, given as:

$$Q_l{}'_{(i)} = \left\lceil ((maximum_i - minimum_i) / Q_{il(i)}) \right\rceil, \ i = 1, 2, ..., 7. \quad (9)$$

Optimization is performed within the extremum values of similar samples, while the data not within the extremum interval of similar samples are not within the optimization range. Therefore, different from the $Q_l$ in the experimental and test groups, the $Q_l'$ here will be greatly reduced.

(5) Determine binary digit numbers ($N_{bd}$). Binary array expression precision should be less than quantization interval length. That is, $2^{N_{bd}}$ should be no less than $Q_l'$. The binary digit numbers of all remaining variables are linked together as individual genes.

(6) Fulfill weight matric, the meaning of which is to fill the possible experienced weight space bound by the minimax of the similar samples under specific characteristic value. The filling principle is that the weight of untraversed location is equal to the average of other weights adjacent to it. In other words, the weight of n-dimensional untraversed hypercube is determined by 2n coplanar hypercubes. The untraversed initial values are set to 1, which is much larger than the maximum of traversed weights. The advantage of the initialization is to avoid false local optimal problems caused by large quantity of small weights, which is beneficial to the next step of optimization.

## 2. Data preparation and preprocessing

### 2.1. Data preparation

The obtained data variables from the power station include unit load (MW), inlet and outlet flue gas parameters: volume ($V_{in}$, $V_{out}$), sulphur concentration ($C_{is}$, $C_{os}$) and oxygen content (%), humidity (%), temperature (°C); liquid tank parameters: slurry density ($\rho_t$, kg/m$^3$), *PH*, liquid level (*H*, m); circulating slurry volume ($V_c$, t/h), volume and density of limestone supplements slurry ($V_s$, t/h and $\rho_s$, kg/m$^3$); current parameters: current of oxidizing air blower ($I_o$, A), stirrer ($I_s$, A), gypsum discharge pump ($I_{gdp}$, A); oxidizing air volume (m$^3$/h), defogging pressure difference (Pa) and some other parameters. Due to the two-stage absorption tower, the primary parameters are suffixed with –1 and the secondary parameters are suffixed with –2, for example, *PH*-1 and *PH*-2.

The data selection period is August 2018. The data is recorded in minute and totals 44,640 sets. The changes of traditional parameter load are shown in Figure 3.

### 2.2. Data preprocessing

Raw data is always incomplete and inconsistent, which not only brings difficulty to data analysis directly, but also gives rise to barriers to conclusions. To improve the quality of the data to be mined, the raw data should firstly be preprocessed. This article illustrates the data preprocessing process through the case of slurry density. Raw data and processed data of the slurry density is shown in Figure 4. Black and red dots represent raw and processed data respectively.

As can be seen from the Figure 4, although raw data can basically reflect changes in trends, it has distinct features of high noises and multiple interruption intervals, which brings great resistance to subsequent model building and data analysis process. To address the problems, the characteristic values of interruption intervals were first obtained, including the length of intervals, variable values of starting and ending points and extreme values. Then, similar intervals were found for each interruption interval. The similar interval refers to close extreme values, the close span of variable values, and the same or opposite

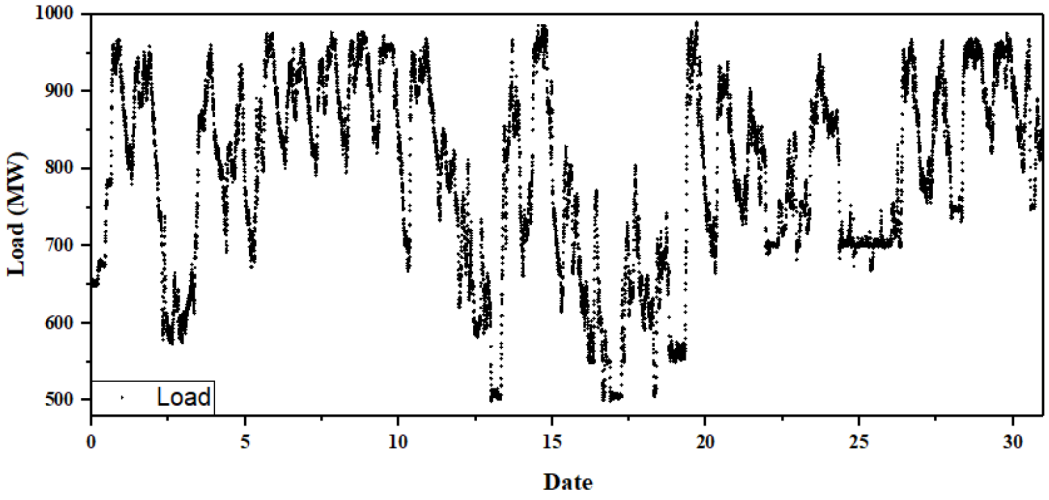


Figure 3. Load changes

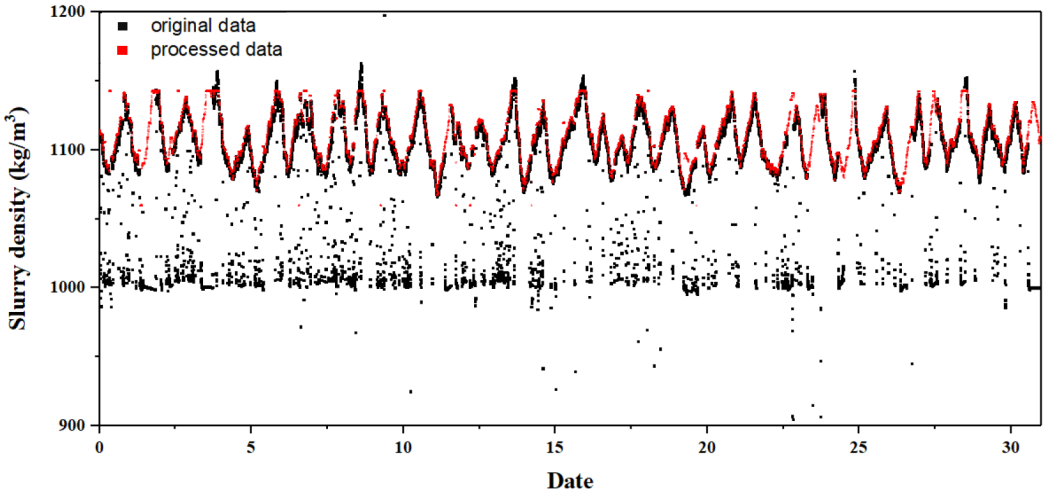


Figure 4. Raw and processed data of slurry density

trend. Subsequently, interpolation or resampling method was used to fill the discontinuous intervals. After that, fit and filter the overall data samples. In the fitting and filtering process, the Savitzky-Golay filter fitting method was adopted (Schafer, 2011). The sliding window width is set as 201, while the rank of sliding curve equation is 3 in the case.

After preprocessing, the data can be applied to the next data modeling and optimization process. Portion feature values of processed data are shown in Table 1.

## 3. CMAC prediction model

### 3.1. Experimental design

GREA is generally considered as an objective and effective system analysis method and it is widely used in multi-factor correlation analysis (Deng, 1982, 1989; Liu et al., 2011). Here, this method was used to analyze the correlation of cost with variables. Based on the results, variables of *load*, $C_{is}$, $V_{in}$, $I_o$, $\rho_s$, *H*-1, *PH*-1, $\rho_{t-1}$, $V_{s-1}$ and $V_c$ were adopted to constitute the model inputs. As introduced

Table 1. Maximin values of processed data

| Variables | Load | $C_{is}$ | $V_{in}$ | $V_c$ | $\rho_{t-1}$ | $I_s$ | $I_o$ |
|---|---|---|---|---|---|---|---|
| MAX | 990.4 | 2917 | 299.9 | 31.14 | 1143 | 874.2 | 248.4 |
| MIN | 498.1 | 658.5 | 162.8 | 0.5500 | 1060 | 772.7 | 0.2360 |
| Variables | $I_{gdp}$ | *PH*-1 | *H*-1 | $V_{s-1}$ | $V_{s-2}$ | $\rho_s$ | $C_{os}$ |
| MAX | 203.3 | 7.880 | 12.62 | 63.75 | 74.43 | 1400 | 2.978 |
| MIN | 0.7100 | 2.930 | 9.170 | 0 | 0 | 1236 | 49.89 |

in 2.2, the $Q_l$ of the ten parameters, accompanied with generalization parameter *c*, total 11 parameters need to be determined. *Cost* and $C_{os}$ are employed as output variables normally. After preliminary testing, it is appropriate to take ε as 0.34. All 44640 sets of data are in service.

### 3.1.1. Experiment group

For multi-factor and multi-level experiments, uniform design method has incomparable advantages on saving time and experimental materials (Fang, 1994; Fang et al., 2000; Li et al., 2004). According to the principle of uniform design method, 13 levels are necessary for 11 variables (factors). Since U* table generally has less uniformity discrepancy, $U^*_{12}{}^{11}$ table was selected. The specific data is shown in Table 2.

Regression analysis theory is the common method applied to the analysis of the results of uniform design (Li et al., 2003). The linear regression analysis was firstly used for the analysis, of which the basic formula is expressed as:

$$Y = b_0 + \sum_{i=1}^{11} b_i x_i \,, \tag{10}$$

where, *x* represents each variable, *Y* refers to output MRE. Based on unique calculation principles of UDM (Fang, 1994), the obtained results of linear regression analysis are shown in Table 3.

In the formulas, the values of $b_0$ are the actual, and the rest values are all enlarged by $10^5$ times. From the results, the fitting formulas of *Cost* and $C_{os}$ have certain similarities: (1) The positive and negative properties of their coefficients are consistent; (2) Both coefficients of parameter *c* are much larger than those of other ten parameters. On the one hand, the results illustrate the importance of *c* for CMAC. On the other hand, it is proved that the MRE index can indeed weaken the influence of sample data, which is beneficial to the promotion of obtained conclusion.

Subsequently, quadratic linear analysis was also used for error analysis, of which the basic formula is shown in Formula 11.

$$Y = b_0 + \sum_{i=1}^{11} b_i x_i + \sum_{\substack{i=1,j=1 \\ i\neq j}}^{i,j<=11} b_{ij} x_i x_j + \sum_{i=1}^{11} x_i^2 \,. \tag{11}$$

Using quadratic stepping method of quadratic linear analysis, the obtained error prediction formulas of cost *Yc* and $SO_2$ *Ys* are:

$$Yc = 0.5279 + 1.208\times10^{-3} x_{11} + 10^{-9} \times (-7.720 x_1 x_3 - 2.850 x_2 x_7 - 6.697 x_4 x_5); \tag{12}$$

$$Ys = 0.3458 + 0.02491 x_{11}{}^2 \times 10^{-4} + (3.721 x_3 x_4 - 1.066 x_8 x_{10}) \times 10^{-8}. \tag{13}$$

For the convenience of comparison, the fitting results of the two regression methods and actual vaues of experimental group are drawn in Figure 5 intuitively.

From Figure 5, both linear regression and quadratic stepping method have high fit degree with the actual results no matter *Cost* or *Cos*. Correlation coefficients $R^2$ are

Table 2. Experiment group values of 11 variables

| | Load | $C_{is}$ | $V_{in}$ | $V_c$ | $\rho_{t\text{-}1}$ | $I_o$ | *PH*-1 | *H*-1 | $V_{s-1}$ | $\rho_s$ | *c* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2500 | 1800 | 1800 | 2200 | 1950 | 3800 | 2200 | 2300 | 3700 | 2600 | 545 |
| 2 | 3000 | 2600 | 2100 | 3400 | 3200 | 5600 | 1600 | 1800 | 2500 | 2400 | 515 |
| 3 | 3500 | 3400 | 2400 | 4600 | 1200 | 3500 | 2300 | 2600 | 1300 | 2200 | 485 |
| 4 | 4000 | 4200 | 2700 | 1900 | 2450 | 5300 | 1700 | 2100 | 4000 | 2000 | 455 |
| 5 | 4500 | 5000 | 1700 | 3100 | 3700 | 3200 | 2400 | 1600 | 2800 | 1800 | 425 |
| 6 | 5000 | 5800 | 2000 | 4300 | 1700 | 5000 | 1800 | 2400 | 1600 | 1600 | 395 |
| 7 | 5500 | 1400 | 2300 | 1600 | 2950 | 2900 | 2500 | 1900 | 4300 | 2700 | 365 |
| 8 | 6000 | 2200 | 2600 | 2800 | 950 | 4700 | 1900 | 2700 | 3100 | 2500 | 335 |
| 9 | 6500 | 3000 | 1600 | 4000 | 2200 | 2600 | 2600 | 2200 | 1900 | 2300 | 305 |
| 10 | 7000 | 3800 | 1900 | 1300 | 3450 | 4400 | 2000 | 1700 | 4600 | 2100 | 275 |
| 11 | 7500 | 4600 | 2200 | 2500 | 1450 | 2300 | 2700 | 2500 | 3400 | 1900 | 245 |
| 12 | 8000 | 5400 | 2500 | 3700 | 2700 | 4100 | 2100 | 2000 | 2200 | 1700 | 215 |

Table 3. Results of linear regression

| | $b_0$ (×1) | $b_1$ (×$10^{-5}$) | $b_2$ (×$10^{-5}$) | $b_3$ (×$10^{-5}$) | $b_4$ (×$10^{-5}$) | $b_5$ (×$10^{-5}$) |
|---|---|---|---|---|---|---|
| Cost | 0.8263 | –4.454 | –0.3215 | –3.796 | –0.6901 | –0.9360 |
| $SO_2$ | 0.8563 | –5.609 | –0.8170 | –8.718 | –1.373 | –2.327 |
| | $b_6$ (×$10^{-5}$) | $b_7$ (×$10^{-5}$) | $b_8$ (×$10^{-5}$) | $b_9$ (×$10^{-5}$) | $b_{10}$ (×$10^{-5}$) | $b_{11}$ (×$10^{-5}$) |
| Cost | 0.1460 | –0.4380 | 2.340 | 0.6901 | 1.286 | 74.23 |
| $SO_2$ | 0.1666 | –0.4997 | 5.8179 | 1.3728 | 3.268 | 93.48 |

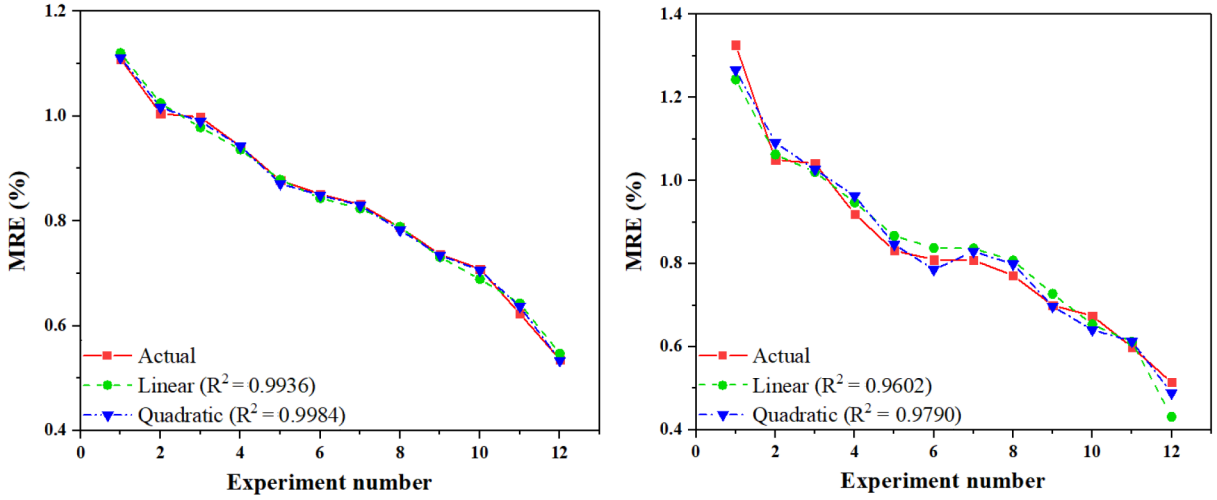


Figure 5. Experiment results and analysis: (a) cost (b) $SO_2$

always greater than 0.96. This embodys the effectiveness of error analysis system of UDM. Comparatively speaking, quadratic analysis model has a little bit better accuracy. However, this does not mean that quadratic analysis is more valuable than linear regression prediction. Quadratic analysis has much more variables' combinations than linear regression analysis. It supposed to own a higher fitting accuracy in theory. The superiority of the two methods requires a further comparison of predictive capabilities.

### 3.1.2. Test group

In order to further verify the prediction ability of the formulas obtained by the experiment group, test group was then conducted. The logic of the parameters for test group is unified level from $U^*_{12}{}^{11}$ uniform table. The specific data is shown in Table 4. Data, learning rate and other parameters remain the same with experiment group.

Prediction formulas obtained in 4.1.1 were further used to predict MREs of test group. Both actual and predicted values are drawn in Figure 6.

As can be seen from Figure 6, for cost, linear regression formula still has high accuracy with the actual results, although it does not completely follow the trends of actual results. In terms of $SO_2$, prediction errors of linear regression formula have an increasing deviation with actual data. The maximum even reached 8.6%. Whether the error will continue to increase or not is inevitably questioned. The straightforward explanation for this phenomenon is that the coefficients of some variables are a little bigger, which can be further attributed to the use of lesser levels of uniform design table. Generally speaking, for UDM, the more levels are, the stronger independence between experiments will be, and the obtained formulas are naturally more predictive. It can be expected that a more accurate result would be obtained by improving levels (Fang et al., 2000). As the reason for why there is no similar problem of deviation for cost maybe ascribed to data. Since cost is calculated rather than measured, it is relatively easier to speculate.

Table 4. Test group values of 11 factors

| | Load | $C_{is}$ | $V_{in}$ | $V_c$ | $\rho_{t-1}$ | $I_o$ | $PH$-1 | $H$-1 | $V_{s-1}$ | $\rho_s$ | $c$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 8000 | 5800 | 2700 | 4600 | 3700 | 5600 | 2700 | 2700 | 4600 | 2700 | 545 |
| 2 | 7500 | 5400 | 2600 | 4300 | 3450 | 5300 | 2600 | 2600 | 4300 | 2600 | 515 |
| 3 | 7000 | 5000 | 2500 | 4000 | 3200 | 5000 | 2500 | 2500 | 4000 | 2500 | 485 |
| 4 | 6500 | 4600 | 2400 | 3700 | 2950 | 4700 | 2400 | 2400 | 3700 | 2400 | 455 |
| 5 | 6000 | 4200 | 2300 | 3400 | 2700 | 4400 | 2300 | 2300 | 3400 | 2300 | 425 |
| 6 | 5500 | 3800 | 2200 | 3100 | 2450 | 4100 | 2200 | 2200 | 3100 | 2200 | 395 |
| 7 | 5000 | 3400 | 2100 | 2800 | 2200 | 3800 | 2100 | 2100 | 2800 | 2100 | 365 |
| 8 | 4500 | 3000 | 2000 | 2500 | 1950 | 3500 | 2000 | 2000 | 2500 | 2000 | 335 |
| 9 | 4000 | 2600 | 1900 | 2200 | 1700 | 3200 | 1900 | 1900 | 2200 | 1900 | 305 |
| 10 | 3500 | 2200 | 1800 | 1900 | 1450 | 2900 | 1800 | 1800 | 1900 | 1800 | 275 |
| 11 | 3000 | 1800 | 1700 | 1600 | 1200 | 2600 | 1700 | 1700 | 1600 | 1700 | 245 |
| 12 | 2500 | 1400 | 1600 | 1300 | 950 | 2300 | 1600 | 1600 | 1300 | 1600 | 215 |

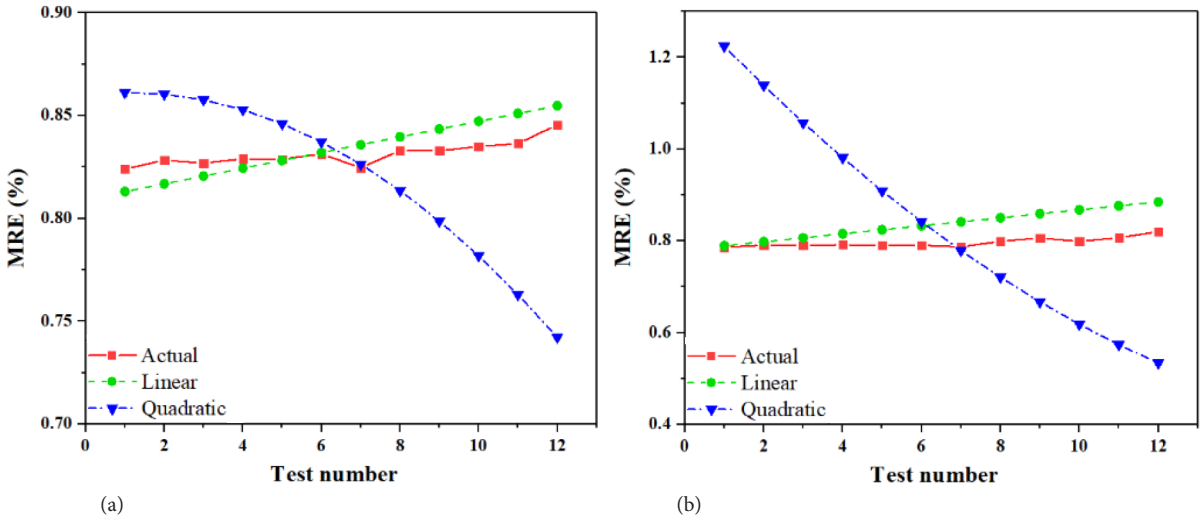


Figure 6. Prediction results of test group: (a) cost (b) $SO_2$

Even worse, prediction formulas of quadratic stepping method have completely wrong trends for both cost and $SO_2$. This demonstrates that the regression formulas are only the fitting of the experimental data, but not implys a deep physical relationship. The reason why the method is able to achieve such a high consistency with the results of experiment group simply lies on that it has more compound modes of variables. Attempts to use other quadratic stepping methods yielded even worse results. It is no doubt that quadratic linear analysis is not suitable for the case.

Quadratic linear regression has no error prediction ability while linear regression has progressively larger prediction errors. In order to obtain more accurate fitting curve for $SO_2$, new analysis approaches must be explored based on the current results. It is explained in more detail in part 4.2 that why more accurate error prediction model of $SO_2$ is needed.

### 3.1.3. Location number analysis method

Considering the fact that CMAC belongs to table reference techniques, the size of the "table" is the reflection of detailed degree of mapping and determines the space to store information (Tao et al., 2002). Prediction error was related to location number tentatively. Location numbers are calculated by Formula 3. The numbers of previous experiment and test group are shown in Table 5.

Based on the results of experiment group, the relationship between prediction error of $SO_2$ and location number is fitted as:

$$Ys_{\ln} = 10.35 l_n^{-0.09290}. \quad (14)$$

Prediction results of the two groups are presented in Figure 7.

From the Figure 7, LNA prediction results have the validity for trends, but often have apparent deviation in comparison with other prediction results for experimental group data. However, for test group, the prediction results of fitting formula were highly coincident with actual values. It is easy to find out that the location numbers of test group were generally close to $10^{12}$, so the basic idea is that the location number prediction formula has high accuracy in the nearby area. Reviewing the diagram (a) in Figure 7, location number prediction results also have high consistency at some other areas, for example, at number 2, 3, 4, 8, 11, 12. Obviously, the least squares principle is working. Generally, location number fitting formula keeps trends correctness in wide span of location number and maintains high precision in specific areas simultaneously.

Location number curve fitting produces satisfactory prediction results. This conforms the importance of the size of the "table". From the results, the error and location

Table 5. Location numbers of experiment and test group

| | Experiment group | Test group |
|---|---|---|
| 1 | $1.361\times10^{10}$ | $1.009\times10^{12}$ |
| 2 | $4.649\times10^{10}$ | $9.972\times10^{11}$ |
| 3 | $4.907\times10^{10}$ | $1.037\times10^{12}$ |
| 4 | $2.044\times10^{11}$ | $1.037\times10^{12}$ |
| 5 | $3.001\times10^{11}$ | $1.002\times10^{12}$ |
| 6 | $4.790\times10^{11}$ | $9.888\times10^{11}$ |

| | Experiment group | Test group |
|---|---|---|
| 7 | $5.363\times10^{11}$ | $9.481\times10^{11}$ |
| 8 | $1.305\times10^{12}$ | $8.842\times10^{11}$ |
| 9 | $2.722\times10^{12}$ | $8.9288\times10^{11}$ |
| 10 | $1.081\times10^{13}$ | $8.3615\times10^{11}$ |
| 11 | $2.278\times10^{13}$ | $7.670\times10^{11}$ |
| 12 | $1.539\times10^{14}$ | $7.163\times10^{11}$ |

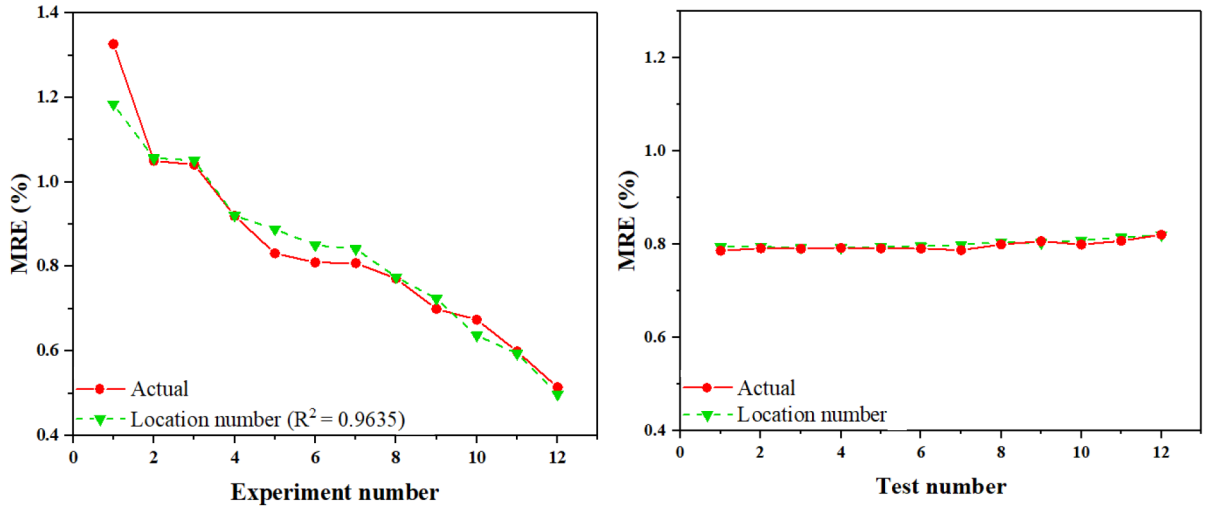


Figure 7. LNA prediction results: (a) experiment group (b) test group

Table 6. Determined model parameters

| Load | $C_{is}$ | $V_{in}$ | $V_c$ | $\rho_{t-1}$ | $I_o$ | $PH$-1 | $H$-1 | $V_{s-1}$ | $\rho_s$ | $c$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 2500 | 1400 | 1600 | 1300 | 950 | 2300 | 1600 | 1600 | 1300 | 1600 | 259 |

number are in relationship of power function. However, this result is based on the current range of location number. Whether it is practical in a wider range can be further studied in the field of CMAC error analysis. As for this paper, a more credible $SO_2$ error prediction formula has been obtained.

Comparing the results of above methods comprehensively, quadratic linear analysis has the highest consistency with experiment results and the worst predictions for test group. Linear regression analysis has relatively poorer similarity with experimental data but is able to maintain validity for prediction. Location number analysis method keeps trends correctness in wide space, and meanwhile maintains high precision in areas.

### 3.2. Model initialized

Now that the error prediction formulas have been settled. The cost and $SO_2$ share the same structure of CMAC and have similar output precision. Then, the next step is to determine the allowable error line. The average sulfur removal rate of preprocessed data is 0.9887, and $SO_2$ concentration of chimney inlet is 17.87 mg/m$^3$. If the allowable deviation is 1%, that is, if the sulfur removal rate is reduced by 1%, the average concentration of export $SO_2$ would be 34.64 mg/m$^3$, which meets the national $SO_2$ ultra-low emissions of 35 mg/m$^3$ standard in China (Zhang et al., 2019). Therefore, model parameters are determined as the prediction error of $SO_2$ is 1%.

As can be seen from Figure 6a, location number fitting formula of $SO_2$ has high prediction accuracy nearby 1% error. Therefore, the prediction formula is adopted for model determination. The $Q_l$ of 10 input parameters are fixed as minimum values from uniform design table, and final model parameters were determined by only adjusting the value of $c$. Calculated $c$ should be no more than 259. Hence, the model parameters are acquired as listed in Table 6.

Using the constructed CMAC model, the cost and $SO_2$ error are 0.9741% and 0.9812% respectively. Output $SO_2$ error is close to the objective error. The results demonstrated that this dual-output model already has high accuracy. There is no need to further adjust parameters. So far, CMAC prediction model has finally been initialized.

### 3.3. Model validation

Model validation is an essential step in model building process. In order to verify the above determined model parameters, model validation was then conducted. The model training was performed with randomly selected 95% of the preprocessed data, and the correctness test was then proceeded with the remaining 5% (2232 sets). Except data volume, the remaining parameters remain the same with experiment and test group. The prediction results are drawn in Figure 8.

It needs to be noted that each output of CMAC is the sum of $c$ weights. Since the amount of data is much smaller than weight space and the validation model was built with 95% of data, it is inevitable that the locations are not traversed for the remaining 5% data.

As can be seen from Figure 8, the results of cost and $SO_2$ share similar data. the MRE is approximately an oblique straight line with not traversed number, which is the embodiment of the CMAC prediction principle. This shows that large prediction errors are mainly caused by

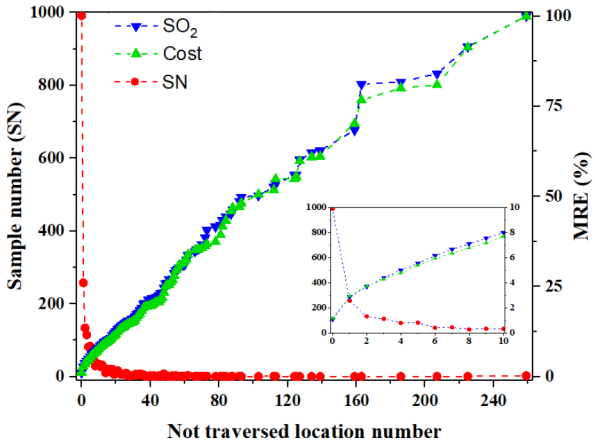


Figure 8. Model validation results

non-ergodicity. Then, the MRE of $SO_2$ and cost for 991 sets of completely traversed data are 1.118% and 1.204% respectively, which are larger than the design error. This may be due to the lack of generalized traversed data resulting in insufficient generalization performance. In optimization process, all sets of data would be in service. At that time, if 5% of data was randomly selected, the predicted MRE would be 0.9741% and 0.9812% in theory. That means the error curve has no influence on optimization process. Therefore, although the errors are larger than the design values, it is not caused by inappropriate CMAC model parameters, but by non-ergodicity. Hence, the model parameters are finally determined.

In addition, the importance of data volume to the construction of the CMAC model can be clearly perceived. Although it is unrealistic to increase the amount of data to completely fill the weight matrix, when the amount of data increases, the number of corresponding weights within the range of each variable will be greatly improved under different working conditions. In that case, even if some locations are not traversed, predictions of weights can be filled by appropriate methods. Therefore, it is worthwhile to enlarge the data volume in future work.

A total of 44640 sets of data were used in the construction of the prediction model. However, it still faces the problem of small data volume, which shows the ability of the CMAC in mathematical modeling of large amounts of data (Ching-Tsan & Chun-Shin, 1996). This characteristic is obviously very important especially for systems that produce large amounts of data. Compared with the most published studies based on small data volume, the CMAC model proposed in the paper seems more practical.

## 4. GA optimization model

One of the significances of model building is to optimize process. After the preparation progress as described in section 2.3, binary array can establish contact with the actual CMAC model input data through minimax values. Optimization process can then be conducted through general GA steps. In the case, characteristic parameters are determined as *Load*, $C_{is}$ and $V_{in}$, while similarity interval are selected as *Load* ±5, $C_{is}$ ±50, $V_{in}$ ±15.

As can be seen form Figure 3, load was in rapid transition, which greatly limits the number of similar samples. With more similar samples, the range of parameters values would be wider, the results would be more global and less likely affected by deviation points. Hence, the horizontal interval near the 25th is the best period for the selection of samples to be optimized. Operating condition of (*Load*, $C_{is}$, $V_{in}$) = (701.73, 1262.03, 237.93) is selected as optimization objective. The results of the preparation process are shown in Table 7.

Table 7. GA preparation processed results

| Variables | Maximum | Minimum | $Q_{il}$ | $Q_l'$ | $N_{bd}$ |
|---|---|---|---|---|---|
| $V_c$ | 18.69 | 1.587 | 0.0235 | 728 | 10 |
| $\rho_{t-1}$ | 1143 | 1075 | 0.0874 | 778 | 10 |
| $I_o$ | 74.13 | 69.09 | 0.1079 | 47 | 6 |
| *PH*-1 | 7.064 | 3.192 | 0.0031 | 1251 | 11 |
| *H*-1 | 12.55 | 9.807 | 0.0022 | 1273 | 11 |
| $V_{s-1}$ | 39.93 | 0 | 0.0490 | 794 | 10 |
| $\rho_s$ | 1400 | 1240 | 0.1026 | 1563 | 11 |

After completion of the preparatory phase, specific evolutionary rules need to be developed. In the process of genetic evolution, the population keeps 200, and population initialization was generated by random binary array. Individual fitness was calculated by Formula 15:

$$fitness = (300 / Cost)^2 . \quad (15)$$

In the sixth step of the preparation phase, most of the vacant weights are initially assignment with large values and then determined by iterations, so these weights tend to be much bigger, which makes it difficult to meet the required sulfur removal rate. In order to avoid this situation, a strict evolutionary rule is adopted as: The top 10 percent of the fitness samples (first 20 samples) were identified as strong biological samples. The next 20 samples are positioned as medium biological sample. Half of the offspring are completely generated by combination of strong parent samples randomly. The other half progenies are generated equally by strong-medium and medium-medium combination. Gene intersections are randomly selected. Before and after the intersections are copies of the parental genes separately. The rate of gene mutation is set to 5%, and the mutation point is also random.

Due to the fast transitions of working conditions during the operation of the desulphurization system and the discontinuous operation of the pump, it is inevitable that some obtained values are smaller than the normal. From the Formula 6, cost is the superposition of factors. Hence, in these situations, the calculated fitness would be larger than actual values, which would naturally affect the

process of genetic variation and ultimately affect the optimization results. In order to avoid the influence of abnormal data on the optimization results, the optimization was proceeded with improvement of $R_{cas}$ constraint. Before the selection of strong and medium samples, population is screened firstly and these with $R_{cas}$ < 0.9 are deleted. To ensure the characteristic of global optimization, a relatively conservative *Rcas* is selected.

In the process of optimization, CMAC model was constructed based on the determined parameters in Table 6. All 44640 sets of data were put into service. Learning rate also kept 0.34. After several optimization experiments and comparative analysis of the results, the optimal parameters and their fitness curves were obtained. The optimal fitness in each generation is shown in Figure 9.

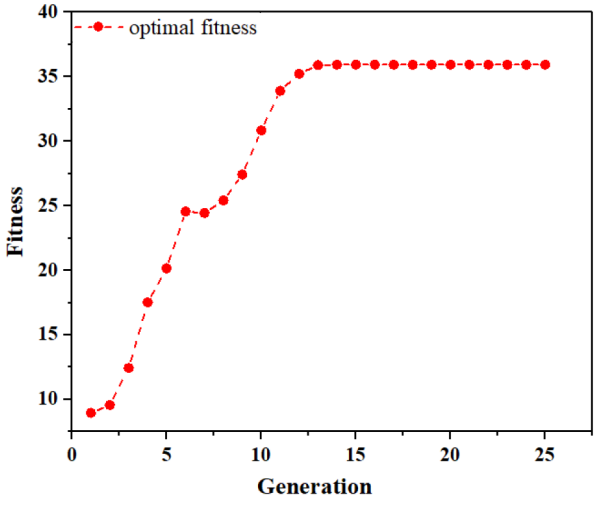


Figure 9. Optimal fitness curve

As can be seen, the optimal fitness in the first ten generations is rising rapidly except the short pause in the 6th and 7th generations. This might be because the local optimum was encountered. Besides, after 13 generations of genetic evolution, the optimal fitness reached a stable value of 35.94, which means the optimum operating condition has been found under the input condition.

At the same time, the parameters of lowest cost satisfying the ultra-low emissions standard were found in similar samples and compared. The optimized results using GA (condition 1) and optimum operating parameters in original data (condition 2) were listed in Table 8.

The cost and $R_{ds}$ of the two optimum conditions are respectively 50.04, 0.9905 and 71.88, 0.9883. From the two sets of data, most data of condition 1 have decreased to some extent while $V_{s-1}$ increases due to the low level of $\rho_{t-1}$. From the perspective of cost, the reduction of $V_c$ and $\rho_{t-1}$ means savings of actual materials. Meanwhile, the reductions of $V_c$ and $I_o$ represent a reduction in power consumption. From the Formula 6, they together lead to a cut down of cost. In fact, the success should be attributed to the deeper reason, the excellent generalization and prediction ability of CMAC (Zhou et al., 2018). Because of CMAC's extraordinary generalization ability, untraversed operating conditions gain their predicted values, and optimization can be carried out smoothly. In addition, this result also reflects the global optimization of genetic algorithms (Horton et al., 2018).

From the definite optimization case, it can be seen that the model is expected to be able to give appropriate adjustable parameters such as $V_c$, $V_{s-1}$ in time under different conditions. This will provide guidance for actual engineering operation and help the system to ensure a higher sulfur removal rate with lower cost, thereby achieving the effect of energy conservation. Of course, the above results are based on the data and the specific object, an absorber of 1000 MW plant with a double-loop desulphurization process. For other different objects or desulfurizing towers with different design loads and structures even same object with different data, the model dimensions or input variables of the prediction model may be different. However, no matter how objects and data change, the entire method proposed in the article has been proved to be correct and feasible.

The article is just a tentative application of CMAC neural network for modeling systems with large amount of data. There are details that can be further studied and improved. Specifically, the data used in the article is monthly data. It is still small compared with the research object which generates data all the time. It is worth studying on larger volume of data in future research. What's more importantly, the characteristics of the research subject may be gradually changing, which means the closer the data is to the current, the higher the value, and conversely, the longer the data is, the lower the value. Therefore, it is meaningful to weight the prediction accuracy of different data period to make the prediction model closer to the current object.

## Conclusions

Using 44640 sets of data, a ten-input two-output CMAC model was built for WFGD system based on GREA and UDM. Compared with traditional regression analysis, location number analysis method was attempted to predict error and determined model parameters. Optimization of cost was further conducted using GA. From the results, some conclusions can be drawn as followings:

Table 8. The optimal parameters of two methods

| | Load | $C_{is}$ | $V_{in}$ | $V_c$ | $\rho_{t-1}$ | $I_o$ | *PH*-1 | *H*-1 | $V_{s-1}$ | $\rho_s$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 701.73 | 1262.03 | 237.93 | 10.21 | 1080.78 | 69.95 | 4.89 | 11.40 | 6.77 | 1399.61 |
| 2 | 701.40 | 1289.69 | 237.71 | 10.55 | 1106.51 | 72.61 | 4.76 | 11.44 | 6.24 | 1400.00 |

1) From the high accuracy of the prediction model and the significant energy saving of optimization case, it can be demonstrated that CMAC is suitable for mathematical modeling of such multi-dimensional and data-generation system. The GREA-UDM-CMAC-LNA-GA methodology is logical and effective for model building and optimization.
2) The power function between error of $SO_2$ and location number and its accurate predicted values of test group prove the correctness of the location number analysis method. It is much more convenient and effective to determine model parameters of CMAC compared with regression analysis.
3) With mature theoretical basis, it is promising to use GA to optimize parameters. In the optimization case, CMAC's optimal cost is reduced by more than 30% compared with regression analysis under same constrains, which fully shows feasibility and effectiveness of GA.

The table reference principle of CMAC is the full extraction and utilization of the information of historical data. The successful implement of CMAC in WFGD system provides a new feasible and reliable solution for industrialization to reduce economic cost. Besides, the use of MRE index makes the analysis results in this paper have reference value for similar systems.

## Funding

This work has been financially supported by the National Key R&D Program of China (2018YFC1901200).

## Author contributions

Baosheng Jin and Yueyang Xu conceived the study and were responsible for the design and development of the data analysis. Yueyang Xu were responsible for data collection and analysis. Zhiwei Kong, Yong Zhang and Xudong Wang were responsible for data interpretation. Zhiwei Kong wrote the first draft of the article.

## Disclosure statement

The authors declare no competing financial interest.

## References

Albus, J. S. (1975a). A new approach to manipulator control: The Cerebellar Model Articulation Controller (CMAC). *Journal of Dynamic Systems, Measurement, and Control*, *97*(3), 220–227. https://doi.org/10.1115/1.3426922

Albus, J. S. (1975b). Data storage in the Cerebellar Model Articulation Controller (CMAC). *Journal of Dynamic Systems, Measurement, and Control*, *97*(3), 228–233. https://doi.org/10.1115/1.3426923

Ansari, H. R., Zarei, M. J., Sabbaghi, S., & Keshavarz, P. (2018). A new comprehensive model for relative viscosity of various nanofluids using feed-forward back-propagation MLP neural networks. *International Communications in Heat and Mass Transfer*, *91*, 158–164. https://doi.org/10.1016/j.icheatmasstransfer.2017.12.012

Armaghani, D. J., Hasanipanah, M., Mahdiyar, A., Majid, M. Z. A., Amnieh, H. B., & Tahir, M. M. (2018). Airblast prediction through a hybrid genetic algorithm-ANN model. *Neural Computing and Applications*, *29*(9), 619–629. https://doi.org/10.1007/s00521-016-2598-8

Biswas, P., Pramanik, S., & Giri, B. C. (2014). Entropy based grey relational analysis method for multi-attribute decision making under single valued neutrosophic assessments. *Neutrosophic Sets and Systems*, *2*, 102–110.

Cheng, H., & Xie, J. (2017, October). Study on the application of recurrent fuzzy neural network in PH control system of absorption tower. In *2017 Chinese Automation Congress (CAC)* (pp. 5962–5966). IEEE. https://doi.org/10.1109/CAC.2017.8243850

Cheng, H., Cui, L., & Li, J. (2017, October). Application of improved BP neural network based on LM algorithm in desulfurization system of thermal power plant. In *2017 Chinese Automation Congress (CAC)* (pp. 5917–5920). IEEE. https://doi.org/10.1109/CAC.2017.8243841

Ching-Tsan, C., & Chun-Shin, L. (1996). CMAC with general basis functions. *Neural Networks*, *9*(7), 1199–1211. https://doi.org/10.1016/0893-6080(96)00132-3

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. M. T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, *6*(2), 182–197. https://doi.org/10.1109/4235.996017

Deng, J. L. (1982). Control problems of grey systems. *Systems & Control Letters*, *1*(5), 288–294. https://doi.org/10.1016/S0167-6911(82)80025-X

Deng, J. L. (1989). Introduction to grey system theory. *The Journal of Grey System*, *1*(1), 1–24.

Dou, B., Pan, W., Jin, Q., Wang, W., & Li, Y. (2009). Prediction of $SO_2$ removal efficiency for wet flue gas desulfurization. *Energy Conversion and Management*, *50*(10), 2547–2553. https://doi.org/10.1016/j.enconman.2009.06.012

Fang, K. T. (1994). *Uniform design and uniform design table*. Science Press.

Fang, K. T., Lin, D. K., Winker, P., & Zhang, Y. (2000). Uniform design: Theory and application. *Technometrics*, *42*(3), 237–248. https://doi.org/10.1080/00401706.2000.10486045

Fu, J., Xiao, H., Wang, T., Zhang, R., Wang, L., & Shi, X. (2019, July). Prediction model of desulfurization efficiency of coal-fired power plants based on long short-term memory neural network. In *2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)* (pp. 40–45). IEEE. https://doi.org/10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00030

Guo, Y., Xu, Z., Zheng, C., Shu, J., Dong, H., Zhang, Y., Weng, W., & Gao, X. (2019). Modeling and optimization of wet flue gas desulfurization system based on a hybrid modeling method. *Journal of the Air & Waste Management Association*, *69*(5), 565–575. https://doi.org/10.1080/10962247.2018.1551252

Gutiérrez Ortiz, F. J., Vidal, F., Ollero, P., Salvador, L., Cortés, V., & Gimenez, A. (2006). Pilot-plant technical assessment of wet flue gas desulfurization using limestone. *Industrial & Engineering Chemistry Research*, *45*(4), 1466–1477. https://doi.org/10.1021/ie051316o

Horton, P., Jaboyedoff, M., & Obled, C. (2018). Using genetic algorithms to optimize the analogue method for precipitation prediction in the Swiss Alps. *Journal of Hydrology*, *556*, 1220–1231. https://doi.org/10.1016/j.jhydrol.2017.04.017

Jin, D., & Lin, S. (Eds.). (2012). *Advances in computer science and information engineering* (Vol. 1). Part of the Advances in Intelligent and Soft Computing book series (AINSC, Vol. 168). Springer. https://doi.org/10.1007/978-3-642-30126-1

Jin, W., Li, Z. J., Wei, L. S., & Zhen, H. (2000, August). The improvements of BP neural network learning algorithm. In *WCC 2000-ICSP 2000. 2000 5th International Conference on Signal Processing Proceedings. 16th World Computer Congress 2000* (Vol. 3, pp. 1647–1649). IEEE. https://doi.org/10.1109/ICOSP.2000.893417

Kiil, S., Michelsen, M. L., & Dam-Johansen, K. (1998). Experimental investigation and modeling of a wet flue gas desulfurization pilot plant. *Industrial & Engineering Chemistry Research*, *37*(7), 2792–2806. https://doi.org/10.1021/ie9709446

Li, R., Lin, D. K. J., & Chen, Y. (2004). Uniform design: Design, analysis and applications. *International Journal of Materials and Product Technology*, *20*(1–3), 101–114. https://doi.org/10.1504/IJMPT.2004.003915

Li, J. F., Liao, H., Normand, B., Cordier, C., Maurin, G., Foct, J., & Coddet, C. (2003). Uniform design method for optimization of process parameters of plasma sprayed TiN coatings. *Surface and Coatings Technology*, *176*(1), 1–13. https://doi.org/10.1016/S0257-8972(03)00019-7

Lin, C., & Chiang, C. (1997). Learning convergence of CMAC technique. *IEEE Transactions on Neural Networks*, *8*(6), 1281–1292. https://doi.org/10.1109/72.641451

Lin, C. C. D., & Wang, H. P. B. (1996). Performance analysis of rotating machinery using enhanced cerebellar model articulation controller (E-CMAC) neural networks. *Computers & Industrial Engineering*, *30*(2), 227–242. https://doi.org/10.1016/0360-8352(95)00168-9

Liu, L., Deng, Q., Zheng, C., Wang, S., Wang, J., & Gao, X. (2019). An insight into electrostatic field effects on SO3 adsorption by CaO with $CO_2$, $SO_2$ and $H_2O$: A DFT approach. *Aerosol and Air Quality Research*, *19*(10), 2320–2330. https://doi.org/10.4209/aaqr.2019.04.0206

Liu, S., & Forrest, J. Y. L. (2010). *Grey systems: Theory and applications*. Springer Science & Business Media.

Liu, S., Cai, H., Cao, Y., & Yang, Y. (2011, October). Advance in grey incidence analysis modelling. In *2011 IEEE International Conference on Systems, Man, and Cybernetics*. Anchorage, AK (pp. 1886–1890). IEEE. https://doi.org/10.1109/ICSMC.2011.6083947

Mostofi, N., & Hasanlou, M. (2017). Feature selection of various land cover indices for monitoring surface heat island in Tehran city using Landsat 8 imagery. *Journal of Environmental Engineering and Landscape Management*, *25*(3), 241–250. https://doi.org/10.3846/16486897.2016.1223084

Munroe, S., Sandoval, K., Martens, D. E., Sipkema, D., & Pomponi, S. A. (2019). Genetic algorithm as an optimization tool for the development of sponge cell culture media. *In Vitro Cellular & Developmental Biology-Animal*, *55*(3), 149–158. https://doi.org/10.1007/s11626-018-00317-0

Schafer, R. (2011). What is a Savitzky-Golay filter? *IEEE Signal Processing Magazine*, *28*(4), 111–117. https://doi.org/10.1109/MSP.2011.941097

Tamura, T., Eguchi, M., Qiao, M., & Ohmori, H. (2017). Diesel engine combustion control with triple fuel injections based on cerebellar model articulation controller (CMAC) in feedback error learning. In *The Proceedings of the International Symposium on Diagnostics and Modeling of Combustion in Internal Combustion Engines*, 2017.9 (p. C110). The Japan Society of Mechanical Engineers. https://doi.org/10.1299/jmsesdm.2017.9.C110

Tao, T., Lu, H. C., & Su, S. F. (2002). Robust CMAC control schemes for dynamic trajectory following. *Journal of the Chinese Institute of Engineers*, *25*(3), 253–264. https://doi.org/10.1080/02533839.2002.9670700

Villanueva Perales, A. L., Gutiérrez Ortiz, F. J., Vidal Barrero, F., & Ollero, P. (2010). Using neural networks to address nonlinear pH control in wet limestone flue gas desulfurization plants. *Industrial & Engineering Chemistry Research*, *49*(5), 2263–2272. https://doi.org/10.1021/ie9007584

Wang, P., & Dai, G. (2018). Synergistic effect between spraying layers on the performance of the WFGD spray column. *Asia-Pacific Journal of Chemical Engineering*, *13*(6), e2266. https://doi.org/10.1002/apj.2266

Warych, J., & Szymanowski, M. (2001). Model of the wet limestone flue gas desulfurization process for cost optimization. *Industrial & Engineering Chemistry Research*, *40*(12), 2597–2605. https://doi.org/10.1021/ie0005708

Warych, J., & Szymanowski, M. (2002). Optimum values of process parameters of the "wet limestone flue gas desulfurization system". *Chemical Engineering & Technology*, *25*(4), 427–432. https://doi.org/10.1002/1521-4125(200204)25:4<427::AID-CEAT427>3.0.CO;2-X

Whitley, D., & Tutorial, A. G. A. (1994). A genetic algorithm tutorial. *Statistics and Computing*, *4*, 65–85. https://doi.org/10.1007/BF00175354

Wu, Y., Shen, L. E., & Zhang, L. (2011). Study on nonlinear pH control strategy based on external recurrent neural network. *Procedia Engineering*, *15*, 866–871. https://doi.org/10.1016/j.proeng.2011.08.160

Yang, Z. K., Liu, C. Y., Song, X. L., Song, Z. Y., & Wang, Z. S. (2016, July). Application of RBF neural network PID in wet flue gas desulfurization of thermal power plant. In *2016 International Conference on Machine Learning and Cybernetics (ICMLC)* (Vol. 1, pp. 301–306). Jeju, South Korea. IEEE. https://doi.org/10.1109/ICMLC.2016.7860918

Zhang, Y., Liu, Q., Wei, H., Du, Z., & Zhu, Y. (2019, October). Study on economy of flue gas ultra-low emission in coal-fired power. In *IOP Conference Series: Earth and Environmental Science* (Vol. 349, No. 1, 012017). IOP Publishing. https://doi.org/10.1088/1755-1315/349/1/012017

Zhou, D., Shi, M., Chao, F., Lin, C. M., Yang, L., Shang, C., & Zhou, C. (2018). Use of human gestures for controlling a mobile robot via adaptive CMAC network and fuzzy logic controller. *Neurocomputing*, *282*, 218–231. https://doi.org/10.1016/j.neucom.2017.12.016

# APPENDIX

| **Nomenclature** | | | |
|---|---|---|---|
| $c$ | generalization parameter | $Q_l$ | quantization level |
| $C_{is}$ | sulphur content of inlet flue gas | $Q_n$ | quantization number |
| $C_{os}$ | sulphur content of outlet flue gas | $R_{cas}$ | calcium sulfur ratio |
| $\varepsilon$ | learning rate | $R_{ds}$ | desulphurizing ratio |
| $H$ | liquid level height | $V_c$ | circulating slurry volume |
| $I_{gdp}$ | current of gypsum discharge pump | $V_{in}$ | volume of inlet flue gas |
| $I_o$ | current of oxidizing air blower | $V_{out}$ | volume of inlet flue gas |
| $I_s$ | current of Stirrer | $V_s$ | supply slurry volume |
| $l_n$ | location number | $w$ | weight |
| $m$ | mole | **Abbreviations** | |
| $M$ | mass | CMAC | cerebellar model articulation controller |
| $n$ | dimension | GA | genetic algorithm |
| $N_{bd}$ | binary digit number | GREA | grey relation entropy analysis |
| $P_c$ | power consumption | LNA | location number analysis |
| $\rho_s$ | supply slurry density | MRE | mean relative error |
| $\rho_t$ | slurry density of the tank | UNM | uniform design method |
| $Q_{il}$ | quantization interval length | WFGD | wet flue gas desulphurization |