Journal of Power Technologies 100 (1) (2020) 21–31

Computational intelligence approach for NOx emissions minimization in a 30 MW premixed gas burner

Weibo Chen *, Guixiong Liu †

Abstract

Artificial intelligence algorithms have become a research hotspot in attempts to reduce NOx emissions in gas burners through NOx emission modeling and optimizing operating parameters. This paper compres the predictive accuracy of NOx emission models based on LSSVM, SVR and ELM. CGA and three other GA based hybrid algorithms proposed to modify CGA were employed to optimize the operating parameters of a 30MW gas burner in order to reduce NOx emission. The results show that the NOx emission model built by LSSVM is more accurate than that of SVR and ELM. The mean relative error and correlation coefficient obtained by the LSSVM model were 0.0731% and 0.999, respectively. Among the four optimization algorithms, the novel TSGA proposed in this paper showed its superiority over the other three algorithms, excelling in its global searching ability and stability. The LSSVM plus TSGA method is a potential combination for predicting and reducing NOx emission by optimizing the operating parameters for the gas burner on-line.

Keywords: NOx emission modeling, Operating parameters optimization, TSGA, Hybrid algorithm, Gas burner

1. Introduction

Combustion of natural gas is currently an important energy source in thermal power engineering and will be increasingly popular in the future since it is cleaner than coal [1]. However, undesirable emissions will occur if the gas burner is working in inappropriate operation parameters. One significant pollutant impacting the global atmosphere is Nitrogen Oxides (NOx) caused by natural gas combustion. Statistics show that coal combustion emits about 7.7 million tons of NOx, about 70% of the atmospheric emissions of NOx, into the atmosphere, which in China alone leads to costs in excess of $13.3 billion every year. As a main element of air pollution, NOx causes serious respiratory diseases and a range of negative environmental effects. For instance, it is a leading factor in the formation of photochemical smog, which causes severe damage to the organs of many creatures and to crops [2]. Therefore, controling NOx emissions is a worldwide concern, as the utilization of fossil fuels continues to increase and natural gas combustion will have to satisfy increasingly rigorous emission standards in the future.

Many measures can be employed to reduce emission of NOx from gas burners, such as operating parameters optimization, air-staged combustion, microwave-induced NOx reduction, and high-temperature air combustion [3]; [4]; [5]; [6]; [7]. Among them, optimization of operating parameters has demonstrated its potential to reduce NOx emissions in gas burners due to its low cost and high-efficiency advantages over other methods. Operating parameters optimization includes two important and separate steps, i.e. NOx emission modeling and NOx emission optimizing [8] . The first essential problem, which is a prerequisite for the second step, is how to model NOx emission accurately. Once a NOx emission model is built, NOx emissions can be controlled to a certain extent through regulating the inputs of the model. Unfortunately, the concentration of NOx emitted from a gas burner is influenced by multiple operating parameters of the boiler and their relationship is complicated and highly nonlinear [9] . The theoretical model has shown itself to be extremely difficult to determined to date. Therefore, conventional regression methods are not suitable due to their theoretical limitations such as the problem of precise extent and time cost. With advances in computer technology and artificial intelligence (AI), a large amount of literature covers the application of various machine learning algorithms to solve this highly nonlinear modeling problem [10].

During the past thirty years ANN (Artificial Neural Network), a typical machine learning approach, has been developed to maturity in extensive fields of research and has been widely employed to model NOx emission for combustion systems due to its outstanding nonlinear mapping ability [11]; [12]; [13]; [14]; [15]. The radial basis function (RBF), back propagation (BPNN), GRNN (generalized regression neu-

*School of Mechanical Automotive Engineering, South China University of Technology, e-mail
†School of Mechanical Automotive Engineering, South China University of Technology, e-mail

ral network) and time delay neural networks were developed to model pollution emission from coal-fired power plants, internal engine and other energy systems. An extensive and excellent review of artificial intelligence for combustion modeling and pollutant control was made by Kalogirou [16]. ANN models have shown their advantages compared with the conventional regression methods and other methods that are based on complex combustion theory such as CFD (computational fluid dynamics) technique [17]; [18]. However, some insuperable weaknesses occurred with the current application of ANN-based NOx emissions modeling [19]. First, ANN models suffer from some shortcomings including: needs for numerous control parameters, uncertainty in solution (network weights) and danger of overfitting. The parameter calibration procedure is a trial-and-error process, which is time-consuming and stochastic. It seriously restricts the capacity of ANN-based online application. Moreover, the correctness and generalization of the ANN model depend to a great extent on the sufficiency and representativeness of the training data, which takes a large amount of time and money to obtain. Therefore, ANN-based models should be improved though practical use in the energy field.

More recently, another computational intelligence-based method called support vector regression (SVR) was proposed by Vapnik to overcome the drawbacks of ANN and to better solve the highly nonlinear problems [20]; [21]. Another study demonstrates that SVR is a better alternative to the ANN- based model and can predict complex process in many fields [22]; [23]. SVR was also successfully used in NOx modeling [24]. In such studies, the SVR based modeling technique showed a more powerful ability to model the highly nonlinear relationships between the multitudinous operating parameters of a coal boiler and its NOx concentrations. More recently, a modified version of SVM named least squares-SVM (LS-SVM) was introduced by Suykens and Vandewalle. Compared with SVM, LSSVM usually needs less run time and shows more self-adapting ability than SVM because of an added equality constraint-based formulation. As with SVM, LSSVM has also been employed widely to problems in many sectors of industry such as the asphaltene deposition problem [25]. Therefore, LSSVM will be the preferred approach to build the model of NOx emission from gas burners. This model is vitally important as it forms the basis of the following parameters optimization.

After the model is established, the second step is to find the most suitable combination of operating parameters, which can optimize NOx emission by searching algorithms. In order to satisfy actual on-line application, the optimization algorithms must have the ability to produce high-quality solutions and quickly converge. Since the NOx reduction approach is a multivariable and non-linear problem that needs quite a lot of time to achieve ideal answers, novel methods should be employed in the searching process. In a huge amount of literature, evolutionary algorithms (EAs) have been proved to be sufficiently powerful tools to deal with the problem studied in this paper and can overcome the disadvantage of conventional optimization algorithms.

Genetic algorithm (GA) is one of the typical EAs. GA, which provides an approach to search the optimal solutions, is a calculation model that simulates natural selection based on the theory of evolution and the evolutionary mechanism of genetics. During the GA search progress, the initial population was established randomly and evolved continuously by selection, crossover and mutation. With evolution, high-quality genes are retained and low-quality ones removed. Therefore, the new generations created will improve until they satisfy certain conditions. GA, which has the characteristic of implicit parallelism and good global searching ability, has been widely used in various optimization processes, including combustion optimization [26]. Related research includes improving the yield of diesel and kerosene [27], determining dew point pressure [28] and predicting dissolved calcium carbonate concentration [29]. In these studies, GA has exhibited its powerful ability to optimize the design and operating parameters of a range of industrial equipment. However, GA has also shown some insurmountable shortcomings during the application process. The two major weaknesses are premature convergence and time-consuming process to find the final solution. Therefore, different hybrid algorithms which combine characteristics of GA and other optimization algorithms are proposed in order to adjust the evolution speed and searching range of GA.

SAGA, PSOGA and TSGA are three novel GA-based modified algorithms by adding the theory of SA (simulated annealing algorithm), PSO (particle swarm optimization) and TS (tabu search) into the searching method of GA. These three modified algorithms are able to improve the weaknesses of GA through modifying the search steps. SAGA controls the evolutionary rate of GA by simulating the thermal balance theory of metal annealing process while PSOGA absorbs the crossover and mutation ideas of GA in order to accelerate convergence [30]. And TSGA which extends the search range of GA because of the tabu list searching technique is also a potential algorithm. PSOGA and TSGA have never been employed to reduce NOx emission for gas burner while SAGA has shown its superiority NOx reduction of coal-fired boilers [1]; [31] . Therefore, it is meaningful to introduce these two al-

gorithms into combustion optimization and compared them with SAGA.

2. Description of the burner and data preparation

The experimental installation used in this paper is composed of a 30MW gas burner whose type is RPD80G-EU and its matched boiler. The standard natural gas was employed and experimental data is including the concentrations of NOx and the corresponding operating parameters of the system. Experiment was carried out after the boiler was adjusted to use the standard natural gas suitably so that the boiler can run smoothly. Concentrations of NOx emitted were monitored continuously by a sensitive NOx detector in the flue outlet of the boiler. The experimental data employed in this research were all acquired under stable operating conditions. The operating parameters which are likely to influence NOx emission were recorded by the real-time monitoring system of the gas burner.

There are totally 400 groups of data that were collected after the experiment. Among them, 200 representative cases which have greater diversity with each other were selected for this research. For the selected 200 cases, the concentrations of NOx vary from 13.37 ppm to 45.54 ppm, as shown in Fig. 1. These 200 tests were carried out by changing the boiler load in the range of 7-33 MW, primary air pressure in the range of 0.31–6.5 kPa, secondary air pressure in the range of 0.2–6.8 kPa, and the downstream pressure of gas valve in the range of 13–14 kPa, respectively. The measured NOx concentrations and the corresponding operating parameters were in a large dynamic range which was enough to build a representative model.

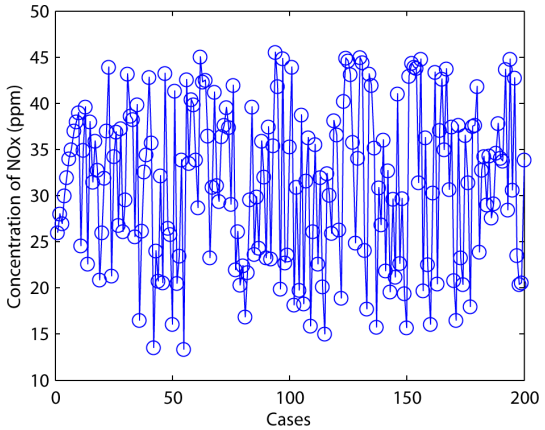

Figure 1: Monitored concentrations of NOx

3. Modeling of NOx emission

3.1. Least squares support vector machine

Support Vector Machine (SVM), introduced by Vapnik and his co-workers, is generally known as a strong mathematical approach to establish an accurate and comprehensive relationship between the variables of a certain mathematical problem.

Least squares-SVM (LS-SVM), introduced by Suykens and Vandewalle, is a modified version of SVM. LS-SVM has wide application in both regression and classification cases. Compared with SVM, LSSVM usually needs less run time and shows more adaptivity. Moreover, LSSVM possesses more characteristic because it employs an equality constraint-based formulation instead of quadratic programming techniques.

In LS-SVMs, the regression is expressed as given below:

$$\begin{array}{ll} & \min_{\omega,\xi,b} J(\omega,\xi)\frac{1}{2}w^T w+\gamma\frac{1}{2}\sum_{j=1}^{M}\xi_j^2 \\ s.t. & \eta_j = w\varphi(V_j)+b+\xi_j \qquad j=1,2,\cdots,M \end{array}$$

where w is the weight of regression. ξ is the random error of the optimization problem and ξ_j(j=1,2,L...M) represents the random error of the *j*th vector γ, which represents the regularization parameter, is a positive number. The nonlinear mapping function is denoted as ψ(g), while the deviation is denoted as *b*.

The Lagrangian equation is defined as follows:

$$L(w,b,\xi,a)=J(w,\xi)-\sum_{j=1}^{M}\alpha_j w\varphi(V_j)+b+\xi_j-\eta_j$$

where α is the Lagrangian multiplier. If the Lagrangian equation is differentiated with respect to *w*, *b*, α_i and ξ_i, the following relationships are obtained after applying the conditions:

$$\begin{cases} \frac{dL}{dw}=0\rightarrow w=\sum_{j=1}^{M}\alpha_j\varphi(V_j) \\ \frac{dL}{db}=0\rightarrow\sum_{j=1}^{M}\alpha_j=0 \\ \frac{dL}{d\xi_j}=0\rightarrow\alpha_j=\gamma\xi_j \\ \frac{dL}{d\alpha_j}=0\rightarrow w\varphi(V_j)+b+\xi_j-\eta_j=0 \end{cases}$$

After the formula (3) is simplified and the kernel function, which meet the Mercer condition, is defined as:

$$K(V_j, V_k) = \varphi(V_j)\varphi(V_k) \qquad j,k = 1,2,\cdots M$$

the question is converted to solve the unknown variables α and *b*:

$$\begin{bmatrix} 0 & amp;e \\ e^T & amp;\Omega + \gamma_{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \eta \end{bmatrix}$$

$$\begin{bmatrix} 0 & amp;e \\ e^T & amp;\Omega + \gamma_{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \eta \end{bmatrix}$$

where I is the M dimension unit matrix,

$$\begin{gathered} \eta = [\eta_1, \eta_2, \cdots, \eta_M]^T \\ \alpha = [\alpha_1, \alpha_2, \cdots, \alpha_M]^T \\ e = [1, \cdots, 1] \\ \Omega_{jk} = K(V_j, V_k) = \varphi(V_j)\varphi(V_k) \qquad j,k = 1,2,\cdots,M \end{gathered}$$

After α and b is obtained, the following relationship is found as the final result:

$$y(V) = \sum_{j=1}^{M} \alpha_j K(V, V_j) + b$$

Kernel functions K(Vj ,Vk) result in determination of the dot product within a high-dimension feature space where the low-dimension space input data are used without the transfer function φ. The most usual kernel functions are polynomial, Gaussian-like or some particular sigmoids. In this paper, RBF is defined as:

$$K(x_i, x_j) = \exp(-\left\|x_j - x_i\right\|^2 / 2\sigma^2)$$

where σ introduces a positive real number, taken into account as the kernel function.

3.2. Training process of LSSVM

NOx formation is the result of the complex and dynamic combustion process. From the up-to-date literature, there comes a conclusion that only the primary variables, i.e. fuel and airflows, which have a direct impact on the combustion process, couldn't be able to establish a stable and precise model. The data related to chemical reactions, quality of air-to-fuel mixing in boilers, changes in fuel quality, etc. are usually not available. And the calibration situations of boilers which have some degree of uncertainty also influence the NOx emission. These are the reasons that the primary variables alone cannot suffice for an accurate NOx prediction. Some other variables that indirectly describe the calibration situations of boiler such as CO concentration and temperature of the flue gases are considered as the input of LSSVM in order to improve the NOx model. In this respect, the superiority of models including the indirect variables were verified by comparing the accuracy of models with that are only considering the primary variables. Building a LSSVM model includes two steps, training and test. To do this, the total 200 cases were randomly classified into two subsets, the training subset containing 100 cases and the test subset containing another 100 cases. The training subset was used to establish the mathematical relationship between the concentration of NOx and the selected variables. In this paper, the LSSVM toolkit, which is running on MATLAB software and providing comprehensive functions for LSSVR classification, regression, and parameter optimization, was employed to build and train the LSSVM model.

During the training process, the selection of control parameters, including kernel function parameter g, penalty factor C, and loss function parameter ε, must be carefully confirmed because they also have a significant impact on the accuracy of LSSVM model. In this research, grid-search method, in which pairs of (g, C, ε) are tried and the one with the maximum R (correlation coefficient) or the minimum MRE (mean relative error) is picked, was employed to optimize control parameters.

4. Hybrid intelligent algorithms to reduce NOx emission

4.1. Simulated Annealing Genetic Algorithm

The simulated annealing (SA) algorithm, which can find the global minimum of the objective function in probability by random search, is a global optimization method based on the mechanism of metal annealing process. An improved method called the simulated annealing genetic algorithm (SAGA), which combines the simulated annealing and the conventional genetic algorithm, is developed by Mathefound.

First, the initial status *i* which is characterized by fitness of an initial individual is defined as the current status, the energy of which is E_i. Second, the initial individual is imposed with a tiny change on its fitness based on the crossover and mutation of CGA and reaches a new status *j*, the energy of which is E_j. If $E_i < E_j$, the new status *j* is considered as the sig-

nificant status. If $E_i > E_j$, the property of status j has to be justified by the probability that individual in this status considering the influence of thermal motion. The ration of probability for status i to j equals to the ration of Boltzmann factor, as expressed as:

$$p = exp(\frac{E_i - E_j}{kT})$$

where k is the Boltzmann constant. T is the thermal equilibrium temperature.

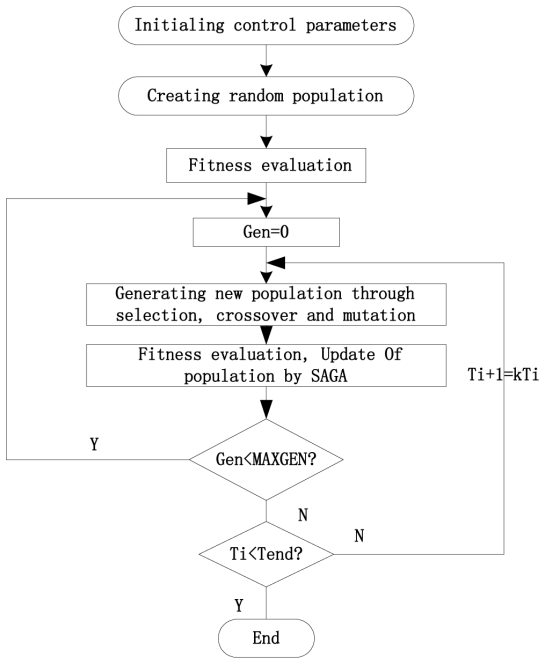

Figure 2: Procedure of SAGA

After that, a random number ξ between 0 and 1 is generated and compared with p. The status j will be accepted as the significant status. Otherwise it will be rejected. The current status is replaced by the significant status. This process is repeated within a given number of generations and at last the system will reach a balanced status with a lower energy. Then the thermal equilibrium temperature T is decreased to a new one according to the following equation:

$$T_{new} = T_{current} \times r \qquad 0 < r < 1$$

where r is the attenuation coefficient of thermal equilibrium temperature.

When T tends to be 0, the optimal result is obtained. The SA algorithm, which overcomes the essential deficiency of CGA, enhances the ability of the searching process to get rid of the local optimal solution.

The SAGA implements the genetic crossover and mutation of Conventional Genetic Algorithm by temperature controlling of Simulated Annealing Algorithm. By combining the advantages of CGA and SA, a more reasonable solution is expected. The procedure of SAGA method is shown in Fig.2.

4.2. PSO- Genetic Algorithm

The particle swarm optimization (PSO), which has the ability of fast convergence, is an evolutionary computation technique developed by Kennedy and Eberhart based on the population behavior. PSO is initialized with a population of random candidate solutions. Each particle in the population is a point in an N -dimension space (N is the dimension of the space, which equals the number of control variables), which is a potential solution to the optimization problem. Every single particle is assigned a randomized velocity and is iteratively moved through the problem space. The quality of each particle is evaluated by the fitness value which is calculated by a given function. The PSO algorithm has the feature of rapid convergence which can improve the searching efficiency if combined with GA. The PSO-GA is a hybrid algorithm which has a better performance on converging speed and the ability of global searching. Let x and v denote a particle position (solution) and its corresponding travelling velocity in a search space, respectively. P_{best} represents the best previous position of one particle, while the best position of the whole population is denoted as g_{best}.

In each iteration, particles firstly modify their velocities and positions according to P_{best} and g_{best} as shown in the following formulations:

$$\begin{gathered} v_{t+1} = \omega v_t + c_1 r_1(t)(P_{best} - x_t) \\ +c_2 r_2(t)(g_{best} - x_t) \\ x_{t+1} = x_t + v_{t+1} \\ w(t) = \begin{cases} \frac{w_{max} - w_{min}}{q}(t-1) + w_{min}, iter \leq 0 \\ w_{min}, iter > 0 \end{cases} \end{gathered}$$

where v_i represents the current velocity of the particle at the i-th iteration while x_i is the current position of the particle at iteration i.

Next, the crossover and mutation progress will be implemented on the particles according to the thought of CGA as the following equations (12) and (13).

$$\begin{aligned} x_{i1} &= x_{i1}(1-b) + x_{i2}b \\ x_{i2} &= x_{i2}(1-b) + x_{i1}b \end{aligned}$$

where b is a random number between 0 and 1.

$$a_i = \begin{cases} a_i + (a_i - a_{max}) \cdot f(g), r \geq 0.5 \\ a_i + (a_{min} - a_i) \cdot f(g), r < 0.5 \\ f(g) = r_2(1 - g/G_{max})^2 \end{cases}$$

where a_i represents the ith variable of x while a_{max} and a_{min} are the upper and lower bound of a_i respectively. r_2 is a random number and the current number of iterations is denoted as g. The random number between 0 and 1 is represented by r while the largest number of evolution is described by G_{max} .

When the PSO-GA was employed to solve the optimization problem in this paper, the specific procedures were as follows:

Step 1: Initialize population in a random fashion.

Step 2: Measuring the fitness of individuals in population. The fitness is defined as the reciprocal of the concentration of NOx which is calculated with the LSSVM model developed above.

Step 3: Particles modify their velocities and positions according to P_{best} and g_{best}.

Step 4: The crossover and mutation is implemented on population.

Step 5: Update the generation $t=t+1$, return to Step 2.

4.3. Tabu Search Genetic Algorithm

The tabu search (TS) algorithm, introduced by Glover [32] , is a classical meta-heuristic which searches optimal solutions in the adjacent domains of candidate solutions created randomly to solve combinatorial optimization problems. The TS can be seen as a hill-climbing algorithm that tries to prevent local optimality by allowing non-improving moves. The process of the TS can be briefly described as follows [33]. First TS algorithm starts with a group of random initial solutions and iteratively improves the quality of the solution by searching all the candidates in the adjacent domains of that solution. These possible solutions in adjacent domains can be created as the values which have a distance in a certain range with the current solution. Using tabu list, the neighboring solutions are evaluated based on the objective function and selected even they don't have an ideal performance in the current stage. The best candidate solution so far is put on the tabu list, so that it cannot be operated for some number of iterations in order to avoid repeating searching. Through designing the size of the tabu list, the number of iterations that a candidate solution is set tabu can be controlled. The large enough size of the list is required in order to avoid circulation and guaranteeing the searching intensity. What's more, an aspiration criterion must be set to cope with the special situation. If the algorithm finds that a solution which is still in the tabu list creates an objective value much better than the current best-known solution, then this solution is allowed to be operated. When the searching progress reaches a preset stopping criterion, the algorithm stops and the final solution is obtained from the best candidate solution in the list. Since GA shows its weaker ability of hill-climbing, an elevated genetic algorithm which compounds the TS algorithm is proposed.

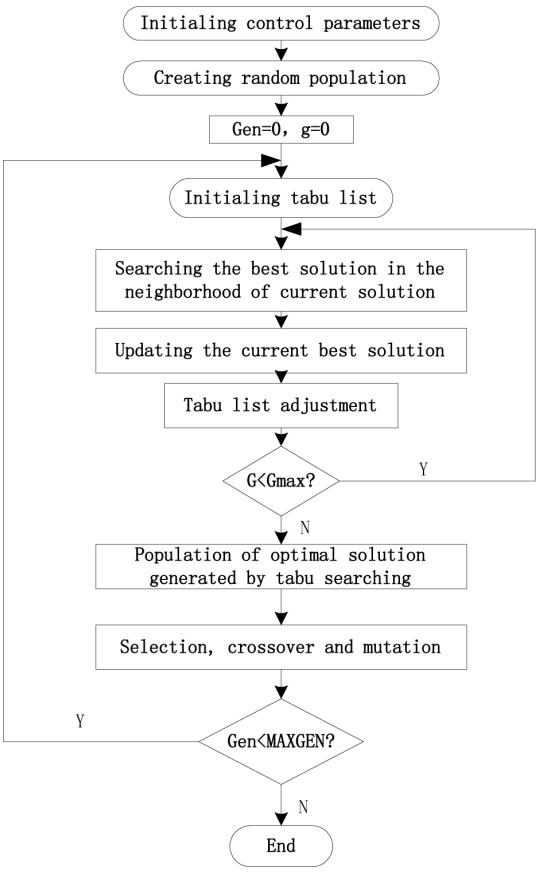

Figure 3: Procedure of TSGA

The procedure of TSGA method is shown in Fig.3. After the initial population is generated, the TS searching process is executed firstly in order to optimize the initial solution. Since all target variables are coded in terms of real number in this paper, the domain of candidate solutions is created as the following equation:

$$S(x) = \{s | s = x + ud \qquad s, x \in X\}$$

where s and x represent the new and original individual respectively. X is the discrete space of variables while d represents the fixed distance of the displacement. The u is used to control both the direction and level of the movement which can be certain or random number or a function decline with iterative times. In this paper, the exponential function which decreases with the iteration k_i increases is used as u:

$$u = exp(\frac{k_i - k_{max}}{R})$$

where k_{max} represents the max iterations while R is a coefficient which controls the descent rate of search range. The superiority on the optimization efficiency of this type of u will be discussed in the following research.

The crossover and mutation will follow the TS algorithm and the whole progress will continuously cycle until certain condiction is satisfied.

5. Results and discussion

5.1. Modeling of NOx emission

As pointed out in section 3.2, παραμετερς (γ, ˆ, ε) have a significant influence on the predictive performance of the LSSVM model. The selection of (g, ˆ, ε) must be completed prior to obtaining a desired LSSVM model. The grid-search employed in the present study was implemented on a desktop computer with a 2.33 GHz CPU and 8.0 G DRAMs under Windows 7. It took nearly 1 min of CPU time to get the best (g, C, ε) with an MRE of 1.59%. The optimal pairs of (g, C, ε) were (2373.7, 3.262, 0.167).

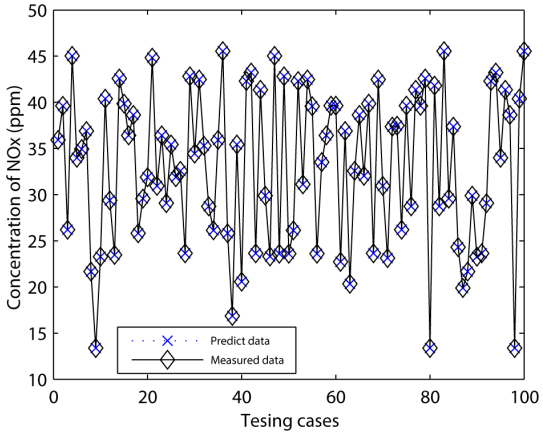

Figure 4: Measured and predicted NOx concentrations by LSSVM

With the optimized parameters, the LSSVM model of NOx emission was trained using the 100 training cases. After that the performance of the trained LSSVM model was evaluated by the remained 100 cases as the testing subset employed. Training a LSSVM model from the data set consisting of 100 samples may typically take several seconds. The comparison between the predicted and the measured NOx concentration is shown in Figure 4. The solid line represents the measured NOx concentrations, while open circles represent the predicted NOx concentrations from the LSSVM model.

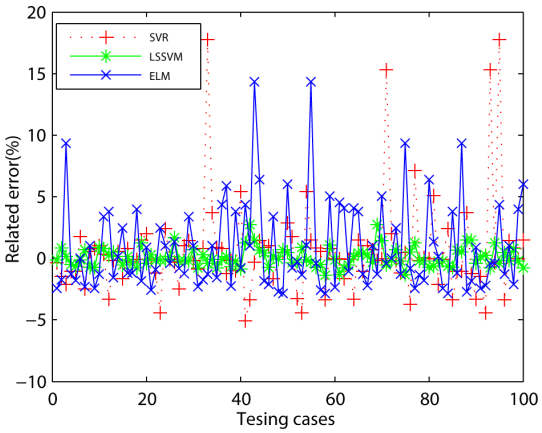

Figure 5: Relative errors producted by SVR, LSSVM and ELM

The ELM and SVR model, which were widely used models to predict NOx emissions, were compared with LSSVM in order to verify the superiority of the LSSVM model. The ELM model was optimized using the conventionally employed "trial-and-error" method in which control parameters were carefully tuned. The ELM model which has a single hidden layer with 12 neurons was chosen. The learning rates of the whole net and layers were both set as 0.2. The TRAINGD method was selected as the training method for the layers. The tansig and purelin function were chosen as the transfer function of the hidden layer and output layer respectively. The SVR model was built by the LIBSVM software package. The RBF was also used as the kernel function while the parameters (C, γ) were optimized by 10-fold cross validation. The final parameters used were (5990, 10).

The capabilities of LSSVM, SVR and ELM model could be revealed through direct comparison between their optimized models. Figure 5 compares the modeling error of the LSSVM, SVR and ELM models case by case. The maximum modeling error of the LSSVM model was 2.729%, presented by the 33th case in the testing subset. All the 100 testing cases had mod-

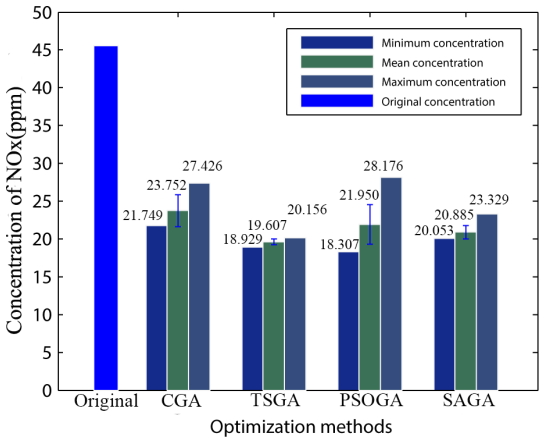

Figure 6: Original NOx concentration and the optimized results based on CGA, TSGA, PSOGA and SAGA

eling error less than 5%. The mean modeling error and the correlation factor were 0.0731% and 0.999, respectively. As a whole, the predicted values showed a rather good agreement with the measured values. The maximum modeling errors of the SVM and ELM models were 17.768% and 14.358%, respectively both presented by the 43th case. Using SVR, a total of 91% (91 cases) of testing cases had modeling error less than 5%. The mean modeling error and the correlation factor were 0.560% and 0.995, respectively. However, the percentage of testing cases had modeling error less than 5% reached 88% (88 cases) when the ELM model was used. The mean modeling error and the correlation factor were also increasing to 0.876 % and 0.881, respectively. The modeling error of 100 cases in the LSSVM was smaller than that both in the SVR model and ELM model. And the correlation coefficient generated by LSSVM is the highest among the three models. It is hence concluded that the LSSVM is the optimal modeling method to build the accurate predictive model of NOx emission from a 30MW.

5.2. NOx reduction by GA, SAGA, PSOGA and TSGA

The control parameters of all the algorithms have significant influence on the performance of optimization procedure. The "trial-and-error" procedures were employed for all the four algorithms presented to obtain the suitable control parameters. Part of the control parameters for the three optimization algorithms are listed in Table 1. The population sizes for the four algorithms were all set to 50 to make a fair comparison. And real number encoding is used as the coding method for all the four algorithms.

It is interesting to compare the performance of the four optimization algorithms mentioned about. Because the 94-th case of a total of 200 cases has the maximum NOx concentration of 45.541 ppm, the four optimization algorithms are conducted on this case. All the algorithms were set with a population of 50 individuals and 100 generations. Figure 6 shows the statistical treatment of the original concentration of NOx and the optimized results based on CGA, SAGA, PSOGA and TSGA. Optimize process of each algorithm was implemented for 50 times and the average value, minimum value and value of the optimized results were calculated and studied due to the randomness of the optimization process of the four algorithms. The standard deviations of 50 results created by the four algorithms were represented by the error bars. As shown in Fig.6, the optimized results for NOx concentration of all the four optimization algorithms become much lower than the original concentration of NOx. This confirms that all the four algorithms have the different degrees of ability to reduce the NOx emission for the 30MW gas burner. However, the strength of the ability to reduce the NOx emission of the four optimization algorithms must be evaluated and compared in order to make the best choice. There are mainly three indicators to describe the optimization capability including the rate of convergence, global searching ability and the stability of searching. The results shown in Fig.6 indicate that the CGA obtained the worst solution among the four algorithms. The average value of optimized NOx con-

Algorithms	Control parameters
CGA	Coding length:12; crossover probability:0.7; mutation probability:0.7; recombine probability:0.7
SAGA	Coding length:12; crossover probability:0.7; mutation probability:0.7; number of generations per temperature phase:100; Initial temperature:100K; end temperature:0K; descending coefficient:0.8.
PSOGA	Coding length:12; crossover probability:0.7; mutation probability: 0.3
TSGA	Coding length:12; crossover probability:0.7; mutation probability:0.7; number of iterations per generation:20; length of taboo list:13.

Table 1: Control parameters of four optimization algorithms

centration found by the CGA is 23.75 ppm which is the highest one and the standard deviation of CGA results which is only less than that of PSOGA is 2.132 ppm, showing that the CGA is the worst in global searching ability and the stability of searching. This conclusion which corresponds well to other research means CGA is very easy to fall into the locally optimal solution and there is also a huge difficulty in CGA to provide stable solutions than other algorithms for the optimization problem in this paper. In contrast, TSGA obtained both the lowest average value and the standard deviation of optimized NOx concentrations, which are 19.607 ppm and 0.418 ppm respectively, among the four algorithms. Therefore, according to the evaluation criteria above within a given iterations, the TSGA is the most effective algorithms because it provides the highest quality and more stable solutions for the optimization problem in this paper. In the 94-th case, the TSGA successfully reduced 38.39% of the NOx concentration. At the meantime, the second lowest values of average concentration and standard deviation of NOx concentration which are 20.885 ppm and 0.860 ppm respectively were obtained by SAGA while PSOGA got the third lowest value of average concentration which is 131.25 ppm and the highest standard deviation of 2.624 ppm respectively. The PSOGA increases the unsteadiness of the searching progress on the contrary.

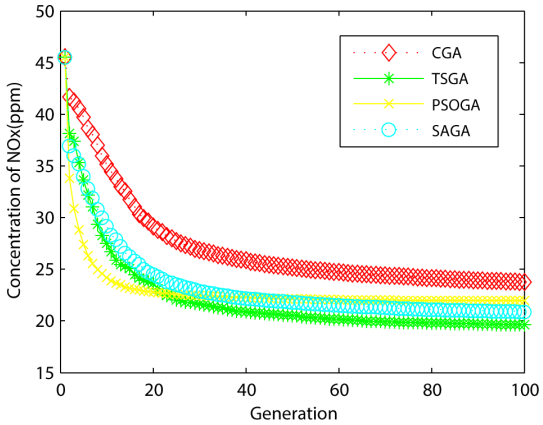

Figure 7: Convergence curves of four algorithms

The rate of convergence can be observed by comparing the convergence curves based on the four algorithms, which is shown in Fig. 7. The searching process can also reflect the global searching ability of the four algorithms. Each convergence curve shown in Fig.7 is the mean curve of 50 curves obtained by 50 times repeated trials for every algorithm. The red curve reveals that the CGA has the lowest rate of convergence and falls into a local optimum after the 50-th generation. This is because the rigorous selection criteria of CGA decreases the diversity of population significantly and finally leads to low convergence speed and premature of searching process. The searching process of CGA is easy to stick in one of the local optima due to the limitation of the worse individuals. And the tardiness of crossover and variation slows down the convergence rate. In contrast, the other three searching process are significantly different from that of CGA. The PSOGA has the fastest convergence speed before the 20-th generation because of the high-efficiency way of particles update. However, it still falls into a local optimum after the 20-th iteration caused by the selection mechanism of GA. The optimization procedure of SAGA has a better performance than that of PSOGA. The final optimal solution of SAGA is lower although the convergence rate is slower than that of PSOGA. This is because the looser selection criteria controlled by the high thermal equilibrium temperature during the early iteration of SAGA process avoids eliminating the worse individuals too quickly. In this way, the genetic diversity of population is maintained so the premature of searching process is prevented. However, it seems that the hill climbing ability of SAGA is not strong enough to get the final optimal value. The reason is the thermal equilibrium temperature decreases and tends to be zero during the last stage, so the SAGA weakens the characteristic of Simulated Annealing Algorithm and becomes similar to the CGA. Therefore, the weakness of CGA which search the final optimal solution only by crossover and variation is still occurred so it needs much more time to converge to the best value. The convergence curve of TSGA displays the best searching performance among the four algorithms as the faster convergence rate than that of SAGA and the lowest value of NOx concentration obtained. This is because TSGA introduces the tabu searching technique which searches the adjacent domains of every individual in order to maintain the genetic diversity of population at every step of iteration. What's more, the searching range of adjacent domains is controlled according to the iterations. During the early stage, the wide adjacent domains are searched in order to cover the candidate solutions as many as possible. And the adjacent domains become so narrow which improves the climbing ability of crossover and variation to converge to the best solution more quickly at the last stage.

Consequently, all the hybrid algorithms in different extent improve the deficiency of the CGA which causes higher probability to fail when CGA is employed to find the extreme value of the optimization problem. Considering the three evaluation criteria above comprehensively, TSGA seems to be a best choice because its best global searching ability and stability of search-

ing in the four algorithms and the convergence rate only slower than that of PSOGA. In the future, the influence of different functions which control the range of neighborhood on the performance of TSGA should be further researched.

6. Conclusions

In the current study, a complex and nonlinear model of NOx concentration from a 30MW gas burner was built by LSSVM model which was optimized by grid-search to predict the concentration of NOx. And the superiority of LSSVM model was verified through comparing the predicted results of the widely used ELM model and SVR, respectively. The results show that the MRE and correlation coefficient of the proposed LSSVM model are 0.0731% and 0.999, which are much better than those of the ELM model and SVR. Using the LSSVM model as the objective function, the CGA and three other GA-based hybrid algorithm including the SAGA, PSOGA and SAGA were employed to optimize the operating parameters for the gas burner to reduce the concentration of NOx. Their optimization capabilities were evaluated and compared. There comes a conclusion that all the four optimization algorithms can optimize the operating parameters and finally reduce the concentration of NOx for the 30MW gas burner. All the hybrid algorithms in different extent improve the deficiency of the CGA which has insurmountable disadvantages to fail in finding the globally optimal solution. The TSGA is preferable because its best global searching ability and stability of searching in the four algorithms and the convergence rate only slower than that of PSOGA. This is because TSGA introduces the tabu searching technique and controls the searching range of adjacent domains using a function. In summary, the LSSVM plus TSGA method is the best combination among all the methods involved in this paper when employed to predict and reduce NOx emission by optimizing the operating parameters for the gas burner on-line.

Acknowledgements

This work was partially supported by the Innovation Foundation Program (No. 2013J4400223) and Business Incubator Project (No. 2013J4200015) from Technology and Information Bureau of Guangzhou City, China.

References

1.González, M.A., Oddi, A., Rasconi, R., and Varela, R. (2015) Scatter search with path relinking for the job shop with time lags and setup times. *Computers & Operations Research*, **60**, 37–54.

2.Xie, K., Li, W., and Zhao, W. (2010) Coal chemical industry and its sustainable development in China. *Energy*, **35** (11), 4349–4355.

3.Wu, S. (2002) AIR-SURROUNDING-FUEL(ASF) PULVERIZED COAL BIAS COMBUSTION TECHNOLOGY WITH LOW Nox EMISSION. *Chinese Journal of Mechanical Engineering*, **38** (01), 108.

4.BARTOSZEK, F.E., VÁSQUEZ, L.R., HE, W., and CHANG, J.S. (1998) Removal of NOxfrom Flue Gas by Reburning with Plasma Activated Natural Gas: Review and Economics. *Combustion Science and Technology*, **133** (1-3), 161–190.

5.Seputro, S., and Sudarmanta, B. (2019) Numerical Investigation of Over Fire Air (OFA) Effect on Flow Characteristics NOx Combustion and Emission in a 600 Mw Tangentially Fired Pulverized Coal Boiler. *IPTEK The Journal of Engineering*, **5** (2).

6.Hiramatsu, T., Najafabadi, H.M., Kobayashi, N., and Itaya, Y. (2012) Non-Catalytic Decomposition of NO by Microwave Plasma Induced at Atmospheric Pressure in an Activated Coke Bed. *KAGAKU KOGAKU RONBUNSHU*, **38** (6), 397–402.

7.ZHANG, X. (2009) Effects of Stereo-staged Combustion Technique on NOx Emmision Charactisctics. *Chinese Journal of Mechanical Engineering*, **45** (02), 199.

8.Zheng, L., Zhou, H., Wang, C., and Cen, K. (2008) Combining Support Vector Regression and Ant Colony Optimization to Reduce NOx Emissions in Coal-Fired Utility Boilers. *Energy & Fuels*, **22** (2), 1034–1040.

9.Kowalczyk, B. (2014) Model of an ANSALDO V94.2 gas turbine from Lublin Wrotków Combined Heat and Power Plant using GateCycle™ software. *Journal of Power Technologies*, **94**.

10.Isaza, C.V., Sarmiento, H.O., Kempowsky-Hamon, T., and LeLann, M.-V. (2014) Situation prediction based on fuzzy clustering for industrial complex processes. *Information Sciences*, **279**, 785–804.

11.Li, G., Niu, P., Wang, H., and Liu, Y. (2014) Least Square Fast Learning Network for modeling the combustion efficiency of a 300WM coal-fired boiler. *Neural Networks*, **51**, 57–66.

12.Li, G., Niu, P., Ma, Y., Wang, H., and Zhang, W. (2014) Tuning extreme learning machine by an

improved artificial bee colony to model and optimize the boiler efficiency. *Knowledge-Based Systems*, **67**, 278–289.

13.Sainlez, M., and Heyen, G. (2013) Comparison of supervised learning techniques for atmospheric pollutant monitoring in a Kraft pulp mill. *Journal of Computational and Applied Mathematics*, **246**, 329–334.

14.Li, G., Niu, P., Liu, C., and Zhang, W. (2012) Enhanced combination modeling method for combustion efficiency in coal-fired boilers. *Applied Soft Computing*, **12** (10), 3132–3140.

15.Rusinowski, H., and Stanek, W. (2010) Hybrid model of steam boiler. *Energy*, **35** (2), 1107–1113.

16.Kalogirou, S.A. (2003) Artificial intelligence for the modeling and control of combustion processes: a review. *Progress in Energy and Combustion Science*, **29** (6), 515–566.

17.Lee, J.W., Park, Y.J., Kim, I.T., and Lee, K.W. (2008) Clinical Results After Application of Bevacizumab in Recurrent Pterygium. *Journal of the Korean Ophthalmological Society*, **49** (12), 1901.

18.Yin, C., Rosendahl, L., Clausen, S., and Hvid, S.L. (2012) Characterizing and modeling of an 88 MW grate-fired boiler burning wheat straw: Experience and lessons. *Energy*, **41** (1), 473–482.

19.Lopes, C., and Perdigão, F. (2008) Event Detection by HMM SVM and ANN: A Comparative Study, in *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, pp. 1–10.

20.Smrekar, J., Potočnik, P., and Senegačnik, A. (2013) Multi-step-ahead prediction of NOx emissions for a coal-based boiler. *Applied Energy*, **106**, 89–99.

21.Guo, M.L., Li, D.J., Du, C.B., Jia, Z.H., Qin, X.Z., Chen, L., Sheng, L., and Li, H. (2012) Prediction of the Busy Traffic in Holidays Based on GA-SVR, in *Advances in Intelligent and Soft Computing*, Springer Berlin Heidelberg, pp. 577–582.

22.Lu, Y., and Roychowdhury, V. (2007) Parallel randomized sampling for support vector machine (SVM) and support vector regression (SVR). *Knowledge and Information Systems*, **14** (2), 233–247.

23.Hong, W.-C., Dong, Y., Chen, L.-Y., and Wei, S.-Y. (2011) SVR with hybrid chaotic genetic algorithms for tourism demand forecasting. *Applied Soft Computing*, **11** (2), 1881–1890.

24.Wei, Z., Li, X., Xu, L., and Cheng, Y. (2013) Comparative study of computational intelligence approaches for NOx reduction of coal-fired boiler. *Energy*, **55**, 683–692.

25.Chamkalani, A., Zendehboudi, S., Bahadori, A., Kharrat, R., Chamkalani, R., James, L., and Chatzis, I. (2014) Integration of LSSVM technique with PSO to determine asphaltene deposition. *Journal of Petroleum Science and Engineering*, **124**, 243–253.

26.Wu, J., Shen, J., Krug, M., Nguang, S.K., and Li, Y. (2011) GA-based nonlinear predictive switching control for a boiler-turbine system. *Journal of Control Theory and Applications*, **10** (1), 100–106.

27.Zhou, H., Lu, J., Cao, Z., Shi, J., Pan, M., Li, W., and Jiang, Q. (2011) Modeling and optimization of an industrial hydrocracking unit to improve the yield of diesel or kerosene. *Fuel*, **90** (12), 3521–3530.

28.Ahmadi, M.A., and Ebadi, M. (2014) Evolving smart approach for determination dew point pressure through condensate gas reservoirs. *Fuel*, **117**, 1074–1084.

29.Ahmadi, M.-A., Bahadori, A., and Shadizadeh, S.R. (2015) A rigorous model to predict the amount of Dissolved Calcium Carbonate Concentration throughout oil field brines: Side effect of pressure and temperature. *Fuel*, **139**, 154–159.

30.Ahmadi, M.A., Zahedzadeh, M., Shadizadeh, S.R., and Abbassi, R. (2015) Connectionist model for predicting minimum gas miscibility pressure: Application to gas injection process. *Fuel*, **148**, 202–211.

31.Sels, V., Coelho, J., Dias, A.M., and Vanhoucke, M. (2015) Hybrid tabu search and a truncated branch-and-bound for the unrelated parallel machine scheduling problem. *Computers & Operations Research*, **53**, 107–117.

32.Glover F, L.M. (1997) *Tabu search*, Kluwer Academic Publishers.

33.Gendreau M, P.J.-Y. (2010) *Handbook of metaheuristics*, Springer.