Model for Predicting NO_x Emission from Boilers Based on MWOA-LSSVM Integration

Chengang Zhen and Huaiyuan Liu

School of Control and Computer Engineering, North China Electric Power University, Baoding 071000, China

Keywords: NO_x Emission, Whale Optimization Algorithm, LSSVM, Parameters Optimization, Integrated Model

Accurate, reliable prediction of NO_x emission in flue gas is of great significance for operation of power station boilers with low nitrogen emissions. To improve the accuracy of a prediction model, a method for predicting NO_x emission from boilers based on integration of the whale optimization algorithm and least squares support vector machine (MWOA-LSSVM) is proposed in this paper. First, the sample space is divided, and a segmentation logistic chaotic map is then used to initialize the population. The nonlinear adaptive parameters are improved, and quadratic interpolation update position is used to improve the whale optimization algorithm (WOA) by broadening the global exploration ability of the algorithm. The MWOA is used to globally optimize the kernel function width and penalty factor of the LSSVM sub-model in each subspace, yielding the sub-model as an output. Finally, the sub-model output is integrated using the least squares regression, yielding the output from the integrated model. The simulation results show that the MWOA-LSSVM integrated model has stable, high-precision simulation performance compared with other selected prediction models and can provide more accurate predictions of NO_x emissions from boilers.

Introduction

Nitrogen oxide compounds (NO_x) are one of the main components in pollutant emissions generated in coal-fired power plants and have a great impact on the formation of acid rain and fog (Park *et al.*, 2019). With the continuous improvement of mass concentration standards on NO_x emissions from coal-fired power plants in China (Environmental Protection Department, 2014), prediction and control of NO_x emissions from coal-fired plant must be more accurate and faster. However, NO_x emissions from boilers are affected by many factors, such as coal type, unit load, and wind volume ratio; thus, mathematically describing NO_x emissions using simple traditional models is difficult (Lv *et al.*, 2013).

Artificial intelligence algorithms and machine learning are becoming more useful for solving complex nonlinear problems, and many effective methods for predicting NO_x emissions have been derived. At present, the least squares support vector machine (LSSVM) is a machine learning algorithm with good performance and has been widely used for predicting NO_x emissions from power plant boilers (Gu *et al.*, 2010; Li and Wang, 2018). However, the prediction from LSSVM often has large errors. This is because the kernel function width $\boldsymbol{\sigma}^2$ and the penalty factor $\boldsymbol{c}$ influence the prediction accuracy. Unreasonable parameter setting will lead to poor reliability in the model results (Si *et al.*, 2017).

Regarding the parameter optimization problem of LSSVM, Zhang *et al.* (2014) established a model for predicting NO_x emissions and proposed optimizing the parameters σ^2 and c in LSSVM using the fruit fly optimization algorithm (FOA). Gao *et al.* (2012) proposed optimizing the parameters of LSSVM using an improved particle swarm optimization (PSO) algorithm and built a shared model of a boiler combustion system. These super-parameter optimization algorithms can provide more accurate predictions, but they have poor stability and weak search ability.

Mirjalili and Lewis (2016) developed a new population-based meta-heuristic optimization algorithm with high precision and stability, named the whale optimization algorithm (WOA), which simulates the hunting behavior of humpback whales. Based on this, Zhang and Chen (2018) introduced the concept of nonlinear parameters and chaos to improve the global optimization ability and produce greater variations in the initial WOA population. Huang *et al.* (2018) improved the performance of WOA by introducing a cosine control factor and polynomial variation. However, WOA still has problems in that the algorithm can become trapped in a local optimum solution and it has poor global exploration ability, thus further refinement is necessary.

Meanwhile, it is difficult to produce more accurate predictions by simply optimizing the parameters in the LSSVM using a meta-heuristic algorithm due to the limitations of LSSVM. Therefore, a new method for solving the problem is required. The multi-model integrated modeling method provides higher precision and stronger generalization than single-model modeling method (Zhao and Lv, 2016). Zhao and Lv (2016) integrated the LSSVM sub-model using the least squares regression based on membership degree and established an NO_x emission prediction model for a boiler.

Received on January 15, 2019; accepted on April 9, 2019
DOI: 10.1252/jcej.19we004
Correspondence concerning this article should be addressed to H. Liu (E-mail address: 15176260155@163.com).

However, the model implementation process is relatively complicated and the test accuracy is low.

Based on the above problems, the author broadened the global exploration ability of WOA by initializing the population with a segmented logistic chaotic sequence, called the modified whale optimization algorithm (MWOA). The position of whale is updated by introducing modified sinusoidal nonlinear control parameters and quadratic interpolation (QI). The MWOA-LSSVM sub-model was based on partitioning the sample space according to the amount of emitted NO_x. The sub-model output was integrated using least squares regression, yielding the predicted NO_x emission. NO_x emission prediction models such as LSSVM, FOA-LSSVM, and PSO-LSSVM are examined for comparison. Simulation results show that the MWOA-LSSVM integrated model had higher accuracy than the other selected models.

1. WOA and LSSVM

1.1 WOA

WOA (Mirjalili and Lewis, 2016) is a population-based intelligent bionic optimization algorithm that mimics the hunting behavior of humpback whales and is optimized for specific problems. A humpback whale will dive down and then rotate upwards after finding prey while spiraling bubbles around the prey; these bubbles reach the surface of the water in order to attack the prey. **Figure 1** (Mirjalili and Lewis, 2016) shows the hunting behavior of humpback whales. The process involves surrounding prey, attacking prey (development phase), and for searching prey (exploration phase).

During the development phase, the humpback whale regards the current best position as the target prey and then adjusts its current position to achieve local optimum by shrinking encircling mechanism and spiral updating position. The mathematical model for spiraling toward prey is described as follows:

$$D = |C \cdot X^*(t) - X(t)| \tag{1}$$

$$X(t+1) = X^*(t) - A \cdot D \tag{2}$$

Here, t is the current iteration number, X^* is the current optimal position vector, X is the position vector, $|\ |$ is the absolute value operation in Eq. (1), $\cdot$ is the dot product, D is the position measurement parameter vector, and A and C are two control parameter vectors obtained from the following equations:

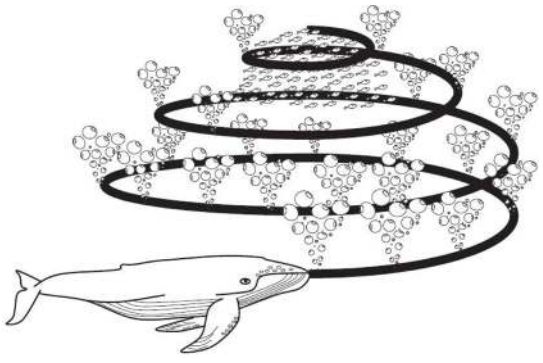

Fig. 1 Hunting behavior (spiral bubble attack) of humpback whales

$$A = 2a \cdot r - a \tag{3}$$

$$C = 2 \cdot r \tag{4}$$

Here, r is a random vector in [0, 1] which is uniformly distributed, and a is the shrinking rate, which decreases linearly from 2 to 0 in successive iterations.

$$a = 2 - 2 \cdot \frac{t}{t_{\text{MaxIter}}} \tag{5}$$

Here, t_{MaxIter} is the maximum number of iterations.

The mathematical description for updating the position spirally is defined in the spiral equation:

$$D' = |X^*(t) - X(t)| \tag{6}$$

$$X(t+1) = D' \cdot e^{bl} \cdot \cos(2\pi l) + X^*(t) \tag{7}$$

Here, D' is the distance between the current position and the optimal position, l is a random number in $[-1, -2]$, and b is a constant that determines the type of logarithmic spiral.

As a humpback whale hunts, shrinking encircling mechanism and spiral updating position is simultaneously carried out. Therefore, both modes are given probability of 1/2.

$$X(t+1) = \begin{cases} X^*(t) - A \cdot D & p_1 < 0.5 \\ D' \cdot e^{bl} \cdot \cos(2\pi l) + X^*(t) & p_1 \geq 0.5 \end{cases} \tag{8}$$

Here, p_1 is a random number in [0, 1].

During the exploration phase, a global search is used to continuously update the optimal position and reach a global optimal solution. The process is described using a mathematical model as follows:

$$D = |C \cdot X_{\text{rand}} - X| \tag{9}$$

$$X(t+1) = X_{\text{rand}} - A \cdot D \tag{10}$$

Here, X_{rand} is a random position vector selected from the current generation. The position is updated by selecting the absolute value of the control parameter vector A. When $|A| < 1$, the local optimal solution is found by shrinking; when $|A| \geq 1$, the global optimal is located by exploration.

WOA still has the disadvantage of easily becoming trapped in a local optimum and poor global searching ability, although it provides better optimization accuracy than some traditional heuristic optimization algorithms. An improved version of WOA is presented in this paper.

1.2 LSSVM

A set of data samples (x_i, y_i) with $i = 1, 2, \ldots, n$ is assumed in the paper. The LSSVM algorithm uses the principle of structural risk minimization to establish optimization problems (Li *et al.*, 2012):

$$\begin{aligned} &\min J = c \cdot \sum_{i=1}^{n} \xi_i^2 + \frac{1}{2}\|w\|^2 \\ &\text{s.t.} \quad y_i = w^{\mathrm{T}} \cdot \varphi(x_i) + b, \quad i = 1, \ldots, n \end{aligned} \tag{11}$$

Here, ξ is a slack variable whose significance is to introduce an outlier, c is a penalty factor, w is a weight vector, b is error, and φ is the kernel function.

α_i is defined as a Lagrange multiplier, and the Lagrange method is used to solve the optimization problem in Eq. (11), which is structured as follows:

$$L(w,b,\xi,\alpha) = \frac{1}{2}c\cdot\sum_{i=1}^{n}\xi_i^2+\frac{1}{2}\|w\|^2-\sum_{i=1}^{n}\alpha_i\{w^{\mathrm{T}}\cdot\varphi(x_i)+b+\xi_i-y_i\} \quad (12)$$

The Karush–Kuhn–Tucker (KKT) optimality condition is defined as follows:

$$\begin{cases}\dfrac{\partial L}{\partial w}=0\rightarrow w=\sum_{i=1}^{n}\alpha_i\cdot\varphi(x_i)\\ \dfrac{\partial L}{\partial b}=0\rightarrow\sum_{i=1}^{n}\alpha_i=0\\ \dfrac{\partial L}{\partial \xi_i}=0\rightarrow\dfrac{1}{c}\alpha_i=\xi_i\\ \dfrac{\partial y}{\partial \alpha_i}=0\rightarrow y_i=w^{\mathrm{T}}\cdot\varphi(x_i)+b+\xi_i\end{cases} \quad (13)$$

$K(x_i, y_i)=\langle\varphi(x_i), \varphi(y_i)\rangle$ is the definition of the kernel function. Equation (14) is obtained by Eq. (13). Taking $\sum_{i=1}^{n}\alpha_i=0$ in Eq. (14) yields:

$$y_i=\sum_{j=1}^{n}(\alpha_j, K(x_i, y_i))+b+\frac{1}{2c}\alpha_i \quad (14)$$

$$\begin{bmatrix}0 & 1 & 1 & \cdots & 1\\ 1 & K(x_1,x_1)+\frac{1}{2c} & K(x_1,x_2) & \cdots & K(x_1,x_n)\\ 1 & K(x_2,x_1) & K(x_2,x_2)+\frac{1}{2c} & \cdots & K(x_2,x_n)\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 1 & K(x_n,x_1) & K(x_n,x_2) & \cdots & K(x_n,x_n)+\frac{1}{2c}\end{bmatrix}\cdot\begin{bmatrix}b\\ \alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}=\begin{bmatrix}0\\ y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix} \quad (15)$$

The regression parameters $[b\,\alpha_1\,\alpha_2\ldots\alpha_n]$ in the model are obtained by solving Eq. (15). The decision function is defined as follows:

$$f(x)=\sum_{i=1}^{n}\alpha_i K(x_i,y_i)+b \quad (16)$$

The Gaussian radial basis (RBF) function is chosen as the kernel function, and σ is the width of the kernel function:

$$K(x_i,x_j)=e^{-\|x_i-x_j\|^2/\sigma^2} \quad (17)$$

2. Modified WOA Algorithm

2.1 Population initialization based on chaos theory

The degree of diversity in the initial population has great significance on convergence of a population-based optimization algorithm. The initial population in WOA is too random and lacks regularity (Zhang and Chen, 2018). Chaos provides randomness and ergodicity. A piecewise logistic chaotic map has better nonlinear characteristics can be used to generate a chaotic sequence without disturbing the calculation. Based on this, a piecewise logistic chaotic map was used to initialize the population in the WOA. The specific expression is as follows:

$$X(t+1)=\begin{cases}4\mu X(t)(0.5-X(t)), & 0\le X(t)\le 0.5\\ 1-4\mu X(t)(0.5-X(t))(1-X(t)), & 0.5\le X(t)\le 1\end{cases} \quad (18)$$

2.2 Nonlinearization of the control parameter *a*

The transition between the development phase and exploration phase in WOA proceeds according to the control vector A. When the $|A|$ is small, the local optimization ability is strong, and when the $|A|$ is large, the global optimization ability is strong. The key to obtaining good performance such as accuracy depends on whether the algorithm can achieve perfect coordination between the two phases. Therefore, the parameter a determined from A plays a crucial role in this process. While a decreases linearly from 2 to 0 in WOA, it cannot accurately reflect and adapt to the complex nonlinear search process. The global searching ability of the algorithm must be improved in order to avoid premature convergence when solving complex, high-dimensional problems (Long *et al.*, 2018). Therefore, it is necessary to define a as a nonlinear function of time. A sine function is used to provide some nonlinear decay over successive iterations:

$$a=2\cdot\sin\left(\frac{\pi}{2}\cdot\frac{t}{t_{\mathrm{MaxIter}}}+\frac{\pi}{2}\right) \quad (19)$$

2.3 Addition of a nonlinear adaptive parameter

The inertia weight has a significant impact on the exploration ability of a meta-heuristic algorithm (Guo *et al.*, 2017). The global exploration effect of the algorithm is broader when the value of the nonlinear adaptive is large. Local exploration of the algorithm is more effective when the value is small. Therefore, a nonlinear adaptive parameter is introduced in the development phase of the WOA to speed up convergence and broaden global searching ability:

$$X(t+1)=\begin{cases}X^*(t)-\omega(t)\cdot A\cdot D, & p_1<0.5\\ \omega(t)\cdot D'\cdot e^{bl}\cdot\cos(2\pi l)+X^*(t), & p_1\ge 0.5\end{cases} \quad (20)$$

Considering the synchronous nonlinear variation with the control parameter a, ω is defined as follows:

$$\omega(t)=2\cdot\sin\left(\frac{\pi}{2}\cdot\frac{t}{t_{\mathrm{MaxIter}}}+\frac{\pi}{2}\right)-1 \quad (21)$$

The algorithm initially has a larger step size and stronger global searching ability. As the number of iterations tends to t_{MaxIter}, ω gradually converges to 0, the step size of algorithm decreases, and local searching is enhanced such that accuracy is ensured.

2.4 Improvement based on QI

In order to further improve the global search ability of WOA and solve large-scale global optimization problems,

a QI method is used in WOA to improve the search ability of algorithm. In recent years, QI has been widely used in meta-heuristic optimization algorithms (Singh and Agrawal, 2016; Gupta *et al.*, 2017). QI uses three vectors selected in an n-dimensional space to find the minimum value on a quadratic curve, called the second crossing. The three selected vectors in WOA are the current optimal position vector $\boldsymbol{X}^*=(x_1, x_2,\ldots, x_n)$ and two population random vectors $\boldsymbol{Y}=(y_1, y_2,\ldots, y_n)$ and $\boldsymbol{Z}=(z_1, z_2,\ldots, z_n)$. The new solution vector $\boldsymbol{X}'=(x_1', x_2',\ldots, x_n')$ is calculated using Eq. (20):

$$x_i'=\frac{1}{2}\cdot\frac{(y_i^2-z_i^2)\cdot f(X^*)+(z_i^2-x_i^2)\cdot f(Y)+(x_i^2-y_i^2)\cdot f(Z)}{(y_i-z_i)\cdot f(X^*)+(z_i-x_i)\cdot f(Y)+(x_i-y_i)\cdot f(Z)} \quad (22)$$

Here, $f(X^*)$, $f(Y)$, and $f(Z)$ are the fitness of X^*, Y, and Z, respectively, and i is the i-th dimension.

QI is used in the development phase in the WOA to enhance the development ability and ensure population differentiation. In this way, WOA includes two methods in the development phase: quadratic crossover and spirally updating position. A uniformly distributed parameter is used to control the two modes. Spiral position updating is calculated with Eq. (20) when the probability $\boldsymbol{p}_2$ is less than 0.8. Otherwise, the position is updated by using the quadratic crossover, as shown in Eq. (17). Therefore, Eq. (20) is updated as follows:

$$X(t+1)=\begin{cases}X^*(t)-\omega(t)\cdot A\cdot D, & p_1<0.5\\ \omega(t)\cdot D'\cdot e^{bl}\cdot\cos(2\pi l)+X^*(t), & p_1\geq 0.5,\ p_2<0.8\\ \text{Eq. (17)}, & p_1\geq 0.5,\ p_2\geq 0.8\end{cases} \quad (23)$$

3. MWOA-LSSVM Model

3.1 Data preparation

3.1.1 Selection of input variables The experimental data was derived from the 1099.3 t/h forced circulation boiler (Alstom, France). Input variables are confirmed according to the formation mechanism analysis of NO_x emission in flue gas (Wang and Zhou, 2002; Zhao *et al.*, 2015).

There are three factors that affect NO_x formation: nitride content in coal, oxygen content in coal, and the combustion situation.

Nitrogen compounds in coal react with a large amount of oxygen to produce a large amount of NO_x. Oxygen content in flue gas is used to measure whether there is adequate combustion, which has an important impact on the emission of NO_x.

Variation in the unit load directly affects the temperature of the furnace and oxygen concentration, thus affecting NO_x emission. Therefore, NO_x concentration shows a positive trend as the unit load changes.

Coal with higher volatile content and lower calorific value is easier to ignite. Because the combustion of pulverized coal consumes a significant amount of oxygen, a hypoxic environment forms in the later stage of combustion, which has a great impact on the production of NO_x.

In a boiler combustion system, primary wind mainly carries pulverized coal into the furnace and provides volatile combustion. Secondary wind provides an amount of air required for pulverized coal combustion in the furnace. Tertiary wind has a strong mixing effect on combustion and provides the oxygen required in the burnout stage. The primary wind volume is determined by the primary wind pressure, while the other wind volume is determined by the dampers.

The unit load, volatile matter of coal, low calorific value of coal (LCV), primary wind pressure (P_{A_SET}), opening degree of secondary wind dampers A, B and C (S_{E_A}, S_{E_B}, S_{E_C}), upper and lower tertiary wind damper opening degree (S_{R_U}, S_{R_D}), and oxygen content in flue gas ($\rho(O_2)$) were used as input variables. The output variable is NO_x emission. 105 sets of data in different steady-state conditions were obtained after a large number of orthogonal experiments (Zhen and Liu, 2018). Some of these data are shown in **Table 1**.

3.1.2 Data processing All data sample vectors are normalized as follows:

$$x'=\frac{x-x_{\min}}{x_{\max}-x_{\min}} \quad (24)$$

Here, $\boldsymbol{x}$ is the initial sample vector, $\boldsymbol{x}_{\min}$ is the minimum of the vector, $\boldsymbol{x}_{\max}$ is the maximum of the vector, and $\boldsymbol{x}'$ is the processed sample vector.

According to the level of output NO_x, the sample data

Table 1 Experimental data for the boiler in steady conditions

Data	Input										Output
	LOAD [MW]	Volatile [%]	LCV [%]	P_{A_SET} [kPa]	S_{E_A} [%]	S_{E_B} [%]	S_{E_C} [%]	S_{R_U} [%]	S_{R_D} [%]	$\rho(O_2)$ [%]	NO_x [mg·m^{-3}]
1	237.68	9.27	23.1	3.46	41	7	42	20	50	3.68	659.18
2	238.17	9.15	22.3	2.94	41	7	42	20	50	3.70	643.72
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
39	289.78	8.62	23.71	3.96	46	27.5	43	65	55	2.43	776.06
40	290.01	8.99	23.32	3.96	46	36.5	55	65	35	2.15	812.79
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
72	327.16	9.19	23.46	3.85	53	22	55	45	57	2.89	987.34
73	327.43	9.30	23.22	3.85	81	38	43	33	45	2.42	965.22
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

is divided into three data subspaces D_l, D_m, and D_h as the sample set in the sub-model. The specific division rules are as follows:

$$\begin{aligned} D_l &= \{X \mid 0 \le y_i < 0.47\} \\ D_m &= \{X \mid 0.47 \le y_i < 0.64\} \\ D_h &= \{X \mid 0.64 \le y_i \le 1\} \end{aligned} \tag{25}$$

Here, the input of data space $D = [D_l, D_m, D_h]$ is $\boldsymbol{X} = [x_1, x_2, \ldots, x_{10}]$, and $\boldsymbol{y}_i$ ($i = 1, 2, \ldots, n$) is NO_x emission.

3.2 Modeling process

After enhancing the global search ability of WOA, an integrated MWOA-LSSVM model is proposed to improve the accuracy of NO_x prediction. First, the data space is divided according to the level of NO_x emission value by Eq. (25). Then, the LSSVM sub-model in each subspace is globally optimized by MWOA. Finally, the sub-models are integrated using least squares method to obtain the output. The specific modeling steps are as follows:

1) The sample set D is normalized;

2) Define MWOA: the five parameters (variable number ***dim***, maximum iteration number $\boldsymbol{t}_{\text{MaxIter}}$, number of whales ***SearchAgents_no***, variable lower limit ***lb***, and variable upper limit ***ub***) in the WOA;

3) Initialize the population: the whale population is initialized using Eq. (18) to increase the population difference.

4) Define the fitness function: the position of the whale is used to represent the parameters of the LSSVM model σ^2 and c, namely $X(i, 1)$ and $X(i, 2)$, respectively. Therefore, $X(i, 1)$ and $X(i, 2)$ are substituted into LSSVM to predict the test sample, and the root mean square error (RMSE) between the predicted value y_i' and true value y_i is taken as the fitness function in WOA. ***RMSE*** is defined as follows:

$$RMSE = \sqrt{\frac{1}{n\sum_{i=1}^{n}(y_i' - y_i)^2}} \tag{26}$$

5) MWOA is used to determine the best σ^2_{best} and c_{best} values;

6) σ^2_{best} and c_{best} are used as parameters in LSSVM to determine NO_x emission in each subspace.

7) The integrated MWOA-LSSVM model is defined in Eq. (27) using least squares regression, where $\hat{Y}$ is the input and y is the output:

$$y(x) = \sum_{i=1}^{d} p_i h_i(x) \tag{27}$$

Equation (27) is written in matrix form $Y = \hat{Y} \cdot W$, where $W = (H^{\mathrm{T}}H)^{-1}H^{\mathrm{T}}y$, and Y is the output from the integrated model.

Table 2 shows the pseudo code for execution of MWOA-LSSVM.

3.3 Defining parameters

The parameters optimization algorithm significantly affect the performance of the optimization method. Generally speaking, the optimization parameters are determined based on experience or results from tests.

Population size and number of search agents: a smaller population can fully search the solution space, avoid overfitting, and requires less computing time. Usually 20–40 individuals are used.

Variable number *dim*: refers to the number of optimization variables. σ^2 and c are parameters to be optimized, thus $dim = 2$ in this paper.

Maximum number of iterations t_{MaxIter}: this value is based on an analysis of a specific problem. If t_{MaxIter} is too small, the algorithm may not converge to the true optimum. If t_{MaxIter} is too large, the computation time may be unnecessarily long.

Range of parameters: the upper and lower bounds are determined by previous experience or from experiments.

Learning factor: usually c_1 and $c_2 = 2$ in PSO.

4. Results and Discussion

One hundred five sets of sample data were divided into training and test sets. The first 95 sets of sample data were used for training and the last 10 sets were used for testing. The test sample is not involved in training. The WOA parameters were $dim = 2$, $SearchAgents\_no = 30$, $t_{\text{MaxIter}} = 100$, $lb = [0.01, 0.001]$, and $ub = [500, 500]$. All simulations were performed on a PC with an Intel Core i5-4590 3.3 GHz processor with 4 GB RAM. All programs were implemented with the M scripting language in MATLAB 2015b.

4.1 Evaluation criterion and NO_x prediction using MWOA-LSSVM

The root mean square error (RMSE), average relative error δ_{PMRE}, average absolute error δ_{PMAE}, and relative error δ_{MRE} are the error criteria used to evaluate the performance of the model. RMSE is used to measure the deviation between the predicted value and the true value, thus it is more suitable as a fitness function for parameter optimization. δ_{PMRE}, δ_{PMAE}, and δ_{MRE} can better reflect the actual prediction error, thus o they are more suitable for analyzing the results. Therefore, δ_{PMRE}, δ_{PMAE}, and δ_{MRE} are introduced to evaluate the accuracy in this section in order to better describe the performance of the proposed model.

$$\delta_{\text{PMRE}} = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{y_i} \tag{28}$$

$$\delta_{\text{PMAE}} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{29}$$

$$\delta_{\text{MRE}} = \frac{\hat{y}_i - y_i}{y_i} \tag{30}$$

Here, n is the number of sample data, $\hat{y}_i$ is the predicted output from the model, and y_i is the true value.

In order to verify that MWOA has broader global searching ability, MWOA and WOA are used to directly optimize

Table 2 Pseudocode for the MWOA-LSSVM algorithm

	Pseudo code of MWOA-LSSVM integration algorithm
01	Normalize the sample set D by Eq. (24);
02	Divide the sample space into D_l, D_m and D_h according to the level of output NO_x value;
03	Set *dim*, $t_{MaxIter}$, *SearchAgents_no*, *lb*, and *ub*;
04	Establish sub-model based on data set which is a sample subspace;
05	**while** $t < t_{MaxIter}$ **do**
06	**for1** 1 to n **do**
07	Let $n = SearchAgents\_no$ and initialize the whale population $X = \{x_1, x_2, ..., x_i, ..., x_n\}$ by Eq. (18). Let $t = 1$ and calculate fitness $f(x_i)$ of each whale individual x_i, the current optimal individual fitness value *Leader_score*, and its position *Leader_pos*;
08	**end for1**
09	Repeat 03–06 twice to get the population Y, Z, and their optimal fitness and position;
10	Calculate a using Eq. (19);
11	**for2** 1 to n **do**
12	Update A, C, b, l, p_1, and p_2
13	**if1** $p_1 < 0.5$
14	**if2** $\|A\| < 1$
15	Update position vector by Eq. (20);
16	**elseif2** $\|A\| \geqq 1$
17	Select a random position vector X_{rand};
18	Update position vector by Eq. (10);
19	**end if2**
20	**else if1** $p_1 \geqq 0.5$
21	**if3** $p_2 < 0.8$
22	Update position vector by Eq. (20);
23	**else if3** $p_2 \geqq 0.8$
24	Update position vector by Eq. (22);
25	**end if3**
26	**end if1**
27	**end for2**
28	Check and fix duplicate individuals to ensure that each search vector is valid;
29	Update *Leader_pos* if there is a better location;
30	$t = t + 1$;
31	**end while**
32	Return *Leader_score*, *Leader_pos*;
33	Substitute *Leader_pos* into LSSVM to predict the test sample;
34	Record the output of the sub-model;
35	Return to 04 until sub-models are established on all sample subspaces;
36	Synthesize the output of each subspace by Eq. (27);
37	Generate prediction graph and calculate the error.

the kernel function width σ^2 and penalty factor c in LSSVM without data space division. LSSVM was used for training in order to build the NO_x emission prediction model and predict emissions. The progress of parameter optimization is shown in **Figure 2**.

Figure 2 shows that MWOA and WOA provide the same precision after a sufficient number of iterations, but MWOA requires more iterations to converge. MWOA has wider search scope compared with WOA, but the wider search scope and the faster search time are contradictory, thus the search time must be increased in order to search a broader range. It is difficult for MWOA to fall into a local optimum and it can be used for optimization over a lager range in order to enhance the performance and generalization of the algorithm. It is difficult for a simple parameter optimization algorithm to enhance the prediction accuracy of LSSVM due to the limitations of the LSSVM and the data set. Therefore, the integrated MWOA-LSSVM model is proposed for solving this problem. The integrated model is used to estab-

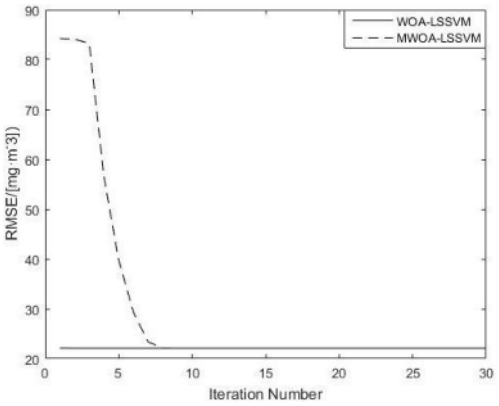

Fig. 2 RMSE from MWOA and WOA over successive Iterations

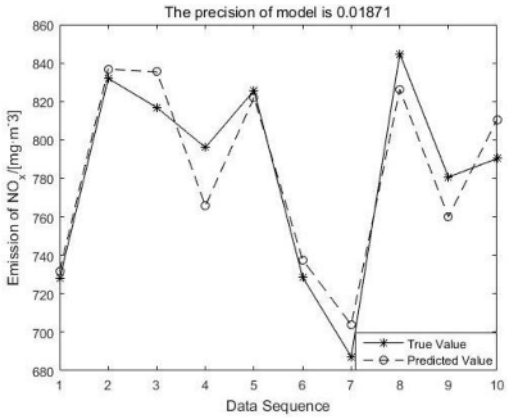

Fig. 3 Comparison between the measured and predicted values using MWOA-LSSVM

Table 3 Comparison between parameters from different optimization models

Model	LSSVM	FOA-LSSVM	PSO-LSSVM	WOA-LSSVM
σ^2	57.15	48.19	203.12	126.45
c	8245.92	7.24	140.62	349.31

lish a model after the sample data space is divided, and the output is shown in **Figure 3**. Figure 3 shows that the true and predicted values are highly fitted, and the error for the test sample is small ($\delta_{PMRE}=1.87\%$ and $\delta_{PMAE}=14.643$). The MWOA-LSSVM method can sufficiently predict NO_x emissions from a utility boiler.

4.2 Comparison of predictions from selected models

In order to verify the prediction accuracy of MWOA-LSSVM integration model is higher than that from other prediction models, the fruit fly optimization algorithm (FOA) and particle swarm optimization algorithm (PSO) were used to optimize the kernel function width σ^2 and penalty factor c in LSSVM for comparison. The results for σ^2 and c from different optimization algorithms after parameter optimization are listed in **Table 3**.

The NO_x emission prediction model can be established using the parameters in Table 3, yielding output values from different prediction models. The relative error δ_{MRE} values with different prediction models are compared in **Figure 4**. Figure 4 shows that the δ_{MRE} with WOA-LSSVM is distributed over a narrower range compared to the other models.

The output results with testing data with different prediction models are compared in **Figure 5**, and the accuracy is compared in **Table 4**. Figure 5 and Table 4 show that the MWOA-LSSVM model has higher prediction accuracy than the other selected models because this model used the least squares method to significantly improve the prediction accuracy compared with the single model.

Case studies were presented in the literature (Zhao *et al.*, 2015; Zhao and Lv, 2016; Zhen and Liu, 2018). Most of these introduced the fuzzy concept to improve the prediction accuracy, which often complicates the modeling process.

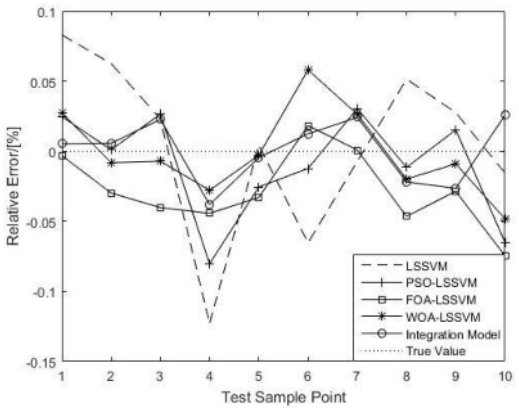

Fig. 4 Comparison between of relative error in NO_x emission with different models

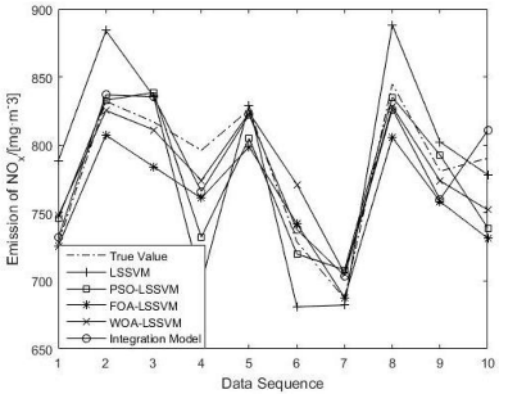

Fig. 5 Comparison between the true and predicted values with different models

Table 4 Precision comparison between different prediction models

Model	δ_{PMRE}/[%]	δ_{PMRE}/[mg·m^{-3}]
LSSVM	4.63	36.24
FOA-LSSVM	3.19	25.67
PSO-LSSVM	2.93	22.85
WOA-LSSVM	2.35	17.93
Integration Model	1.87	14.64

Compared with the aforementioned studies, the proposed model does not require computation of a high-dimensional membership matrix. Therefore, the proposed model can provide high prediction accuracy while simplifying implementation.

Conclusion

In this paper, an NO_x emission prediction model based on MWOA-LSSVM is proposed. For WOA, the global searching ability has been improved by modifying the population initialization method through introducing a control parameter, and an algorithm for updating the position. The accuracy of the prediction model is higher when sub-models

are integrated using the least squares method. Compared with other prediction models such as LSSVM, FOA-LSSVM, and PSO-LSSVM, the proposed MWOA-LSSVM integration model provides more accurate prediction than the other selected models.

Acknowledgement

The present project is supported by Fundamental Research Funds for the Central Universities (Grant Nos. 2016MS143, 2018ZD05), Natural Science Foundation of Beijing (Grant No. 4182061).

Literature Cited

Environmental Protection Department; "General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China. Emission Standard of Air Pollutants for Boiler: GB 13271-2014," *Beijing: China Environmental Science Press*, **3**, (2014)

Gao, F., Y.-J. Zhai, Y. Zhuo, P. Han and Y. Lu; "Combustion Optimization for Utility Boilers Based on Sharing LSSVM Model," *Journal of Chinese Society of Power Engineering*, **32**, 928–933 (2012)

Gu, Y.-P., W.-J. Zhao and Z.-S. Wu; "Combustion Optimization for Utility Boiler Based on Least Square Support Vector Machine," *Proceeding of the Chinese Society for Electrical Engineering*, **30**, 91–97 (2010)

Guo, Z.-Z., P. Wang, Y.-F. Ma, Q. Wang and C.-Q. Gong; "Whale Optimization Algorithm Based on Adaptive Weight and Cauchy Mutation," *Microelectronics & Computer*, **34**, 20–25 (2017)

Gupta, K., K. Deep and J.-C. Bansal; "Improving the Local Search Ability of Spider Monkey Optimization Algorithm Using Quadratic Approximation for Unconstrained Optimization," *Comput. Intell.*, **33**, 210–240 (2017)

Huang, Q.-B., J.-X. Li, C.-N. Song, C.-H. Xu and X.-F. Lin; "Whale Optimization Algorithm Based on Cosine Control Factor and Polynomial Mutation," *Control and Decision*, 1–9 https://doi.org/ (2018)

Li, H.-Z., S. Guo, H.-R. Zhao, C.-B. Su and B. Wang; "Annual Electric Load Forecasting by a Least Squares Support Vector Machine with a Fruit Fly Optimization Algorithm," *Journal of Energies*, **5**, 4430–4445 (2012)

Li, Y.-B. and D.-F. Wang; "Application of an Improved LSSVM in Combustion Modeling and Optimization of Utility Boilers," *Journal of Chinese Society of Power Engineering*, **38**, 258–264 (2018)

Long, W., J.-J. Jiao, X.M. Liang and M.-Z. Tang; "An Exploration-Enhanced Grey Wolf Optimizer to Solve High-Dimensional Numerical Optimization," *Eng. Appl. Artif. Intell.*, **68**, 63–80 (2018)

Lv, Y., J.-Z. Liu, T.-T. Yang and W.-Y. Sun; "NOx Emission Characteristic Modeling Based on Feature Extraction Using PLS and LS-SVM," *Chinese Journal of Scientific Instrument*, **34**, 2418–2424 (2013)

Mirjalili, S. and A. Lewis; "The Whale Optimization Algorithm," *Adv. Eng. Software*, **95**, 51–67 (2016)

Park, J.-H., J.-W. Ahn, K.-H. Kim and Y.-S. Son; "Historic and Futuristic Review of Electron Beam Technology for the Treatment of SO_2 and NO_x in Flue Gas," *Chem. Eng. J.*, **355**, 351–366 (2019)

Si, G.-Q., S.-W. Li, J.-Q. Shi and Z. Guo; "Least Square Support Vector Machine Parameter Optimization Method Using Improved Fruit Fly Optimization Algorithm," *Journal of Xi'an Jiaotong University*, **51**, 14–19 (2017)

Singh, D. and S. Agrawal; "Self Organizing Migrating Algorithm with Quadratic Interpolation for Solving Large Scale Global Optimization Problems," *Appl. Soft Comput.*, **38**, 1040–1048 (2016)

Wang, C.-C. and H.-G. Zhou; "The Influence of Boiler Structural Parameters Upon its NOx Emission Capacity," *Thermal Power Generation*, **1**, 11–13 (2002)

Zhang, Y. and F. Chen; "A Modified Whale Optimization Algorithm," *Computer Engineering*, **44**, 208–213 (2018)

Zhang, Z.-X., B.-M. Sun, J. Xing and Y. Li; "Fruit Fly Optimization Algorithm Based High Efficiency and Low NOx Combustion Modeling for a Boiler," *Thermal Power Generation*, **43**, 19–30 (2014)

Zhao, W.-J., G. Zhao, M. Lv and J.-J. Zhao; "Fuzzy Optimization Control for NOx Emissions from Power Plant Boilers Based on Nonlinear Optimization," *J. Intell. Fuzzy Syst.*, **29**, 2475–2481 (2015)

Zhao, W.-J. and M. Lv; "Modeling of NOx Emission of Boiler Based on Multiple LS-SVM Integrated Model," *Journal of Electronic Measurement and Instrument*, **30**, 1037–1044 (2016)

Zhen, C.-G. and H.-Y. Liu; "Prediction on NOx Emission of Coal-Fired Boiler Based on Multi-Model Clustering Ensemble," *Thermal Power Generation*, 1–8 https://doi.org/10.19666/j.rlfd.201808160 (2018)