ANN-weighted Distance Grey Wolf Optimizer for NOx Emission Optimization in Coal Fired Boilers of a Thermal Power Plant

Mahmood Raphiyoddin Shaphiyoddin Malik
BS Abdur Rahman Crescent Institute of Science and Technology, Chennai, India
rafimalik@gmail.com (corresponding author)

Viswambaran Saraswati Priya
BS Abdur Rahman Crescent Institute of Science and Technology, Chennai, India
priya@crescent.education

Received: 15 May 2024 | Revised: 29 May 2024 | Accepted: 31 May 2024

ABSTRACT
This research work suggests the application of predictive modeling for the Nitrogen Oxide (NOx) emissions from the 210 MW pulverized boiler that burns coal. In order to lower the NOx emissions in the flue gas, it is necessary to optimize various operational parameters during combustion, including oxygen in flue gas, various damper opening positions, air-distribution system, nozzle tilt, and the temperature of the flue gas outlet. Information gathered from variable parametric field tests was used to create an Artificial Neural Network (ANN) model. The ANN model was trained to predict NOx emissions based on the parameters of coal combustion. The trained ANN with its biases and weights in the form of arithmetical equations was given as a fitness function to the weighted distance Grey Wolf Optimizer (GWO) to improve operating conditions for decreased NOx emissions.

Keywords-thermal power plant, air pollution, weighted distance grey wolf optimizer, artificial neural networks

I. INTRODUCTION

Coal is the main energy source used in India. It constitutes over 60% of the commercial energy needs, being responsible for generating 70% of the electricity. Burning coal in thermal power stations emits harmful pollutants into the atmosphere, including suspended particulate matter, sulfur oxides (SOx), carbon oxides (COx), and nitrogen oxides (NOx). Power plants use different methods to reduce and control these emissions to make energy production cleaner and safer [1]. These pollutants contribute to acid rain and climate change, with SO2, NOx, and CO2 being the main culprits. Efforts have been made to reduce emissions by employing technologies such as electrostatic precipitators to remove fly-ash and flue gas and desulfurization plants. Improving power generation efficiency is a successful technique to lessen CO2 emissions. However, reducing NOx emissions remains a challenge, which encourages more research in this area. NOx is produced when coal and air are combined in coal-fired power plants. There are primary-based (combustion modification), and secondary based (flue gas treatment) methods for reducing NOx emissions. Secondary techniques involve adding reagents like urea or ammonia in the flue gas. Primary techniques include the use of low NOx burners, modern boiler control and operation system. However, implementing low-NOx burners can negatively impact other operational parameters, necessitating the need for advanced operational and control systems [2]. Experimental work combined with modeling has been used to evaluate the effectiveness of reducing NOx emissions through oxygen-enhanced combustion. NOx emissions can be reduced by replacing a little amount of oxygen. Precise oxygen addition by burners results in increased coal devolatilization and quick ignition, reducing the distance between the flame and the burner while creating a concentrated fuel-rich area nearby. To decrease NOx emissions, it is crucial to regulate coal combustion variables such as the flame temperature, the ratio of fuel to oxygen, and the duration of fuel exposure during combustion [3].

Attempts have been made to model and optimize the parameters of the boiler to minimize the emissions. Conventional as well as meta-heuristic methods, such as Artificial Intelligence (AI) techniques, have been created for nonlinear problems. Among these techniques, Artificial Neural Networks (ANNs) have been effective in non-statistical model building and can achieve high accuracy. They consist of nodes (neurons) connected through weights. They are effectively combined with optimization algorithms when trained on experimental data. ANNs consist of input, hidden layer(s), and an output layer. They are able to process information about
various problems, including highly non-linear or stochastic, non-differentiable, or discontinuous [4].

In this research, the combination of ANN and optimization algorithms is used to optimize NOx emission. This work seeks to establish the connection between the operating variables and the flue gas NOx emissions through the parametric field experiments. The operational conditions observed during the full-load operation of the 210 MW thermal power plant were utilized to illustrate the capability of ANNs to model the attributes of NOx emissions in pulverized coal combustion. The ANN model reveals the correlation between NOx combustion features, coal attributes, and combustion conditions.

Modern power plants now employ advanced combustion optimization software, utilizing meta-heuristic methods like Genetic Algorithms (GAS). This software simultaneously accomplishes two critical goals: it minimizes NOx emissions while optimizing boiler efficiency by fine-tuning the damper settings for both secondary air and detached overfire air [5]. This research introduced a backpropagation neural network model as an objective function, enabling real-time simultaneous optimization of boiler efficiency and NOx emissions. This approach enhances operational efficiency and reduces environmental impact by minimizing harmful emissions, representing a significant technological advancement in power plant operations.

The current paper emphasizes the effectiveness of Weighted Distance Grey Wolf Optimizer (WDGWO) in conjunction with ANNs to identify the maximum combustion conditions for minimizing NOx emissions. These optimization methods are employed to explore the input space of the neural network model, aiming to minimize NOx emissions. The study implements many ANN models utilizing MATLAB R2015a.

II. EXPERIMENTAL DATA

Model tests were conducted in a 210 MW large furnace that was part of a dry bottom boiler with tangential firing. There are six primary and seven secondary air burners and combustion air nozzles in the furnace positioned at every furnace corner. These nozzles have the ability to tilt in the vertical direction, allowing for a range of approximately 30° upward and downward from the horizontal axis. During the experiments, burner levels A to F were activated under rated load conditions. In the scenario of burning bituminous coal in a power plant, coal pulverisers play a crucial role by supplying air-coal mixtures to the burners according to their respective grades. This was facilitated by employing the tangential firing system, with the burner layout and furnace dimensions illustrated in Figure 1. A comprehensive analysis of NOx emission characteristics was conducted through 35 tests on the boiler, varying primary air, secondary air distribution, boiler load, and nozzle tilting angles. Focusing on full load conditions (210 MW), 14 data sets were meticulously chosen from this extensive dataset to train and test an ANN, exclusively addressing NOx emissions. Throughout these experiments, coal fineness remained consistent, and NOx and O2 concentrations at the boiler outlet were continuously monitored, ensuring precise data collection and analysis [6].

![Diagram illustrating the dimensions and burner placement of the furnace.](image)

The NOx concentrations presented in this work were measured under dry gas circumstances and represent average values over extended periods of continuous operation. Table I shows the test results under full load conditions. Table II provides a summary of the actual recorded NOx emissions under full load conditions.

**Fig. 1.** Diagram illustrating the dimensions and burner placement of the furnace.
A. **ANN Modeling**

Many studies have up to this point concentrated on using ANNs to model pollutant emissions. The architecture of ANNs emulates the way that the human brain learns. It comprises interconnected neurons, categorized into 3 different groups: neurons in the input, output, and hidden layers. While the output layer neurons are where the data is displayed, the input layer neurons feed the neural network with the input data. Scalar functions known as weights establish the connections between neurons and are essential to the network's ability to learn. The backpropagation procedure, which introduces a variety of data input and output to the system, is frequently used to train ANNs. This incoming data is processed by each hidden layer and output layer neuron by multiplying it by the corresponding weight and then by a transfer function [7].

The network learning method comprises repeatedly iterating weight adjustments to decrease the discrepancy between the experimentally measured response and the anticipated response from the ANN model. The various input variables to be considered are like damper opening positions, anticipated response from the ANN model. The various input variables to be considered are like damper opening positions, anticipated response from the ANN model. The varied input data and output to the system, is frequently used to train ANNs. The incoming data is processed by each hidden layer and output layer neuron by multiplying it by the corresponding weight and then by a transfer function [7].

The network learning method comprises repeatedly iterating weight adjustments to decrease the discrepancy between the experimentally measured response and the anticipated response from the ANN model. The various input variables to be considered are like damper opening positions, tilting of burners, flue gas temperature at outlet, oxygen percentage.

There are nine input layer nodes with the first bias node connected to ten nodes in the hidden layer. As a result, there are 90 values of weight \( W_{ij} \) and 10 values of biases \( b_j \) between the input and the hidden layers. On the hidden layer, “tansig” transfer function was used to calculate the sum of the 90 weighted inputs and 10 biases:

\[
Z_j = f'(W_{ij} X_i + b_j), i = 1 to 9; j = 1 to 10 \quad (1)
\]

where \( X_i \) represents the nine inputs of input layer, \( Z_j \) the ten outputs of the hidden layer, \( b_j \) the 10 biases of the hidden layer, \( f' \) is the “tansig” transfer function, and \( W_{ij} \) represents the weights from the input layer \( i \) to the hidden layer \( j \).

The network architecture consists of a hidden layer comprising 10 nodes connected to a single node in the output layer. Within the intermediate layer, bridging the hidden and output layers, there are 10 sets of weight values, arranged in 10 rows, along with a singular bias value. The ‘purelin’ transfer function serves as the mechanism to compute the weighted summation of all 10 inputs \( W_j \), adding the bias \( b_j \) to determine the output layer's result. Equation (2) represents the output layer’s overall weights and bias value.

\[
Y = f^p(W_j * Z_j + b_j), j = 1 to 10 \quad (2)
\]

where \( Y \) is the output, i.e. NOx emissions, \( Z_j \) represents the 10 inputs of the hidden layer, \( f^p \) is the “purelin” transfer function of the output layer, \( b_j \) is the bias of the output layer, and \( W_j \) represents the weights from the hidden layer \( j \) to the output layer.

The ANN underwent training using data sourced from Tables I and II. During this training process, the network’s parameters were fine-tuned, and weight connections were established between the input and the hidden layers, and between the hidden and the output layers. This training aimed to minimize disparities between the measured values and the network's predictions. These results were pivotal in establishing a correlation between operational parameters and NOx emissions under constant full-load conditions in the boiler, enhancing our understanding of their relationship.

B. **Grey Wolf Optimizer**

The GWO draws inspiration from grey wolves’ coordinated hunting techniques [8]. Because they adhere to a very strict social hierarchy, these wolves hunt their prey effectively. They are separated into four groups in the hierarchy: alphas, betas, omegas, and deltas. An omega occupies the lowest position in the hierarchy, while alphas enjoy the highest. Since they are the strongest members of the pack, the alphas are the ones who command the others. These wolves possess the ability to pinpoint the target, and when they do, the entire pack charges in to attack [9]. The three major stages of hunting are: 1) Locating and pursuing, 2) surrounding and pestering the target before it ceases its motion, and 3) assaulting the target. The developed mathematical model for prey hunting and attacking follows.

The pack's position is updated by:

\[
\vec{D} = |\vec{C} \times \vec{X}_p(t) - \vec{X}(t)| \quad (3)
\]

\[
\vec{X}(t + 1) = \vec{X}_p(t) - \vec{A} \times \vec{D} \quad (4)
\]

where \( \vec{X} \) signifies a grey wolf’s location vector, representing the wolf’s position or location within a mathematical model or algorithm, \( t \) represents the present iteration, \( \vec{X}_p \) corresponds to prey’s location vector, and \( \vec{A} \) and \( \vec{C} \) represent coefficient vectors employed in various calculations.

Computation of vectors \( \vec{A} \) and \( \vec{C} \) proceeds as outlined below:

\[
\vec{A} = 2\vec{a} \times r_1 - \vec{a} \quad (5)
\]

\[
\vec{C} = 2\vec{a} \quad (6)
\]

The element \( \vec{a} \) undergoes a linear decline from 2 to 0 with the advancement of iterations. Additionally, \( r_1 \) and \( r_2 \) represent stochastic vectors. Alpha, beta, and delta positions are leveraged to synchronise the entire wolf pack’s position as they
get closer to and converge on their prey during the hunting process. Equations (7)-(13) describes the working method:

\[ D_a = |C_1 \times X_a - X| \]  
(7)

\[ D_\beta = |C_2 \times X_\beta - X| \]  
(8)

\[ D_\delta = |C_3 \times X_\delta - X| \]  
(9)

\[ X_1 = X_a - A_1 \times D_a \]  
(10)

\[ X_2 = X_\beta - A_1 \times D_\beta \]  
(11)

\[ X_3 = X_\delta - A_1 \times D_\delta \]  
(12)

\[ X(t + 1) = \frac{x_1 + x_2 + x_3}{3} \]  
(13)

C. Weighted Distance Grey Wolf Optimizer

WDGWO represents an adapted iteration derived from GWO. Unlike the original GWO, which relies on the simple average of the three best positions within the wolf pack to update their locations, WDGWO incorporates a weighted approach. This variation is designed to improve the optimization process. WDGWO enhances the original GWO approach by introducing weighted best positions of the pack leaders. These weights \( w_i \) are computed using coefficient vectors \( (A_1, C_1) \) in each iteration, altering the location update equation accordingly. This strategy demonstrates particular efficacy in optimizing intricate problems. It refines the GWO algorithm by incorporating weighted distances and provides a more sophisticated approach to problem-solving, making it valuable in tackling intricate optimization challenges [10]. Equations (14) and (15) describe the process, and Algorithm-1 provides a detailed explanation of the algorithm and the archive update strategy.

\[ w_1 = A_1 \times C_1, \quad w_2 = A_2 \times C_2, \quad w_3 = A_3 \times C_3 \]  
(14)

\[ X(t + 1) = \frac{w_1 x_1 + w_2 x_2 + w_3 x_3}{w_1 + w_2 + w_3} \]  
(15)

Algorithm-1: Pseudo Code of WDGWO

1. Initialize Iteration Count (MaxIter)
2. Initialize Size of the Pack (NG)
3. Initialize Grey Wolf Population
4. Initialize \( \tilde{a}, \tilde{A}, \tilde{C} \)
5. Initialize archive size to 100
6. Evaluate fitness of each grey wolf \( f_i(X) \)
7. Compute \( X_a \) = the first best grey wolf
8. Compute \( X_\beta \) = the second best grey wolf
9. Compute \( X_\delta \) = the third best grey wolf
10. While \( t \leq \) MaxIter do
11. Update \( \tilde{a}, \tilde{A}, \tilde{C} \)
12. Update \( X_a, X_\beta, X_\delta \)
13. Calculate weights \( w_i \)
14. Update position vector \( X(t + 1) \)
15. Evaluate fitness \( f_i(X) \)

The process commences with various initializations, including setting parameters such as the maximum number of iterations (1000), pack size (20), and employing a random Gaussian distribution strategy to initialize the location of the entire pack. The search begins after immediately recording the top three wolves. During the search process, the fitness of each wolf will be evaluated, and the best locations of the top three wolves will be updated and documented. The search process is further improved by the updated best grey wolves. To preserve the Pareto dominated solution, an archive of size 100 is kept. In order to keep archive's size constant, the archive crowding distance concept is used if the solution exceeds the limit. When compared to its counterparts, WDGWO has an exceptional effect on the reduction of pollutants. These outcomes are primarily attributable to the weighted distances built into the original GWO, which computes distances among wolves with a basic linear average, potentially trapping the pack in local minima. The WDGWO, an improved GWO variant, offers significantly enhanced results [11].

The primary goals of boiler combustion optimization are to assist operators in utilizing coal more efficiently and with reduced emissions. Hence, the learned parameters of the network, including weights and biases, in a trained backpropagation ANN, were utilized to formulate the objective function for NOx optimization [12]. Algorithms for optimization can be combined with the ANN model. An ANN can be used to create a fitness function that illustrates the connection of NOx emissions and the input operating parameters. The ability of ANNs to map non-linear input-output relationships quickly, makes them a suitable choice for serving as the fitness function in optimization techniques. The best operating parameters can be found by using these optimization algorithms’ searching capabilities. WDGWO was utilized to optimize the NOx emissions in this research.

The proposed methodology integrates WDGWO algorithm and predictive modeling, incorporating an ANN model to estimate NOx emissions. By refining the optimization techniques of WDGWO and incorporating innovative approaches, the methodology aims to contribute significantly to a more sustainable and efficient resolution of reducing NOx emission from the complex environmental challenges within the thermal power plants.

IV. RESULTS AND DISCUSSION

A. Modeling of ANN

Performance evaluation employed metrics like Sum of Square Error (SSE), Mean Square Error (MSE), and Determination Coefficients \( R^2 \) [14]. Figure 2 displays the
measured and network-predicted NOx emission concentrations for both trained and new data. The comparison indicates that the trained network is performing reasonably well in its predictions. This suggests the model's effectiveness in generalizing to new data. The trained network attained the maximum $R^2$ value of 0.99994 and the lowest SSE value of 0.02004 by using a trial-and-error method. The evaluation of network's performance was carried out using 2 new test data and 12 trained test data. Figure 2 illustrates the comparison among the predicted and the observed values, all of which were evaluated under full-load conditions. Figure 3 shows the regression coefficient of the network.

B. Optimization Results of WDGWO

The rapid and efficient nature of the search process makes ANNs suitable for use as fitness functions. Optimization algorithms can leverage this searching capability to identify optimal operating parameters effectively. To pinpoint the most advantageous parameters for minimizing NOx emissions, the WDGWO optimization technique was applied to find the best solution for (2). Figure 4 outlines the flowchart of the proposed ANN-WDGWO approach. Key parameters in this optimization effort involve factors like the amount of flue gas oxygen, nozzle tilt angle, flue gas temperature as it exits, and the positions of the secondary air burner damper openings. It is important to note that when optimizing combustion, certain input factors like the feeder opening size for the pulveriser, the airflow rate through the pulveriser, and the pressure in the wind box are typically considered constant and not adjustable.

However, if necessary, the same procedure can also be used to improve these properties. Figure 3 shows the NOx emissions computed by the WDGWO over a number of iterations. The displayed graphs vividly illustrate the search process's progressive nature, showcasing an impressively rapid convergence rate. Table III presents the optimal operating parameters for full-load conditions, encompassing secondary air burner damper openings, nozzle tilting, flue gas outlet temperature, and flue gas oxygen content, all associated with reduced NOx emissions. These lowered NOx emission values at full load are a testament to the optimization's success, as the results align closely with the experimental data, reflecting an effective and efficient approach.

C. Comparison with Existing Approaches

The GA approach, the Sequential Quadratic Programming (SQP) approach, and the suggested WDGWO approach were juxtaposed and the operating parameters were carefully chosen in order to assure the fairness of the comparison (Table IV).

The objective of this section is to confirm the effectiveness of the WDGWO algorithm in reducing NOx emissions. So, two optimized scenarios from GA and SQP were selected, both achieving a NOx emission level of 205 ppm. It was assessed
how well the developed ANN-WDGWO method performs compared to other existing approaches. In addition to the configuration parameters mentioned above, the operational parameters optimized by the three methods are likewise identical and are based on the same ANN models. The same laptop (2.6 GHz, 4GB RAM) was used for all simulation studies. Table IV lists the optimization outcomes for the three algorithms.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Opening Position of the Damper (%)</th>
<th>Tilting of Burner</th>
<th>Flue Gas Temperature (Low Temp Super Heater Outlet)</th>
<th>Oxygen Content</th>
<th>NOx Emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>WDGWO</td>
<td>AB 82 BC 29 CD 33 DE 68 EF 36 FF 24</td>
<td>Degrees</td>
<td>°C</td>
<td>%</td>
<td>ppm</td>
</tr>
<tr>
<td>GA</td>
<td>100 84</td>
<td>1 1 1 0</td>
<td>30</td>
<td>400</td>
<td></td>
</tr>
<tr>
<td>SQP</td>
<td>100 84</td>
<td>1 1 1 0</td>
<td>30</td>
<td>400</td>
<td></td>
</tr>
</tbody>
</table>

These findings unequivocally validate the dominance of the WDGWO algorithm over the GA and SQP mechanism. The optimization results of NOx emissions for GA, SQP, and WDGWO are 205 ppm, 205 ppm, and 191 ppm, respectively. It is evident that the WdGWO method significantly improves the quality of optimization by 6.8% compared to GA and SQP.

V. CONCLUSION

This paper addresses the combined approach of ANN and Weighted Distance Grey Wolf Optimization (WDGWO) algorithm for predicting and optimizing NOx emissions from flue gas in order to minimize them. The results show significant improvement with reduction of 6.8% in NOx emissions by WDGWO as compared to GA and SQP under a variety of operating situations while burning various types of coal. The optimum operating parameters can be attained in combination with meta-heuristic techniques to reduce NOx emissions. However, in order to attain low NOx combustion conditions, the neural network model must be trained using new experimental data if the coal type and operational parameters change during normal operations.

REFERENCES


www.etasr.com Maliki Priya: ANN-weighted Distance Grey Wolf Optimizer for NOx Emission Optimization in Coal …