



LJMU Research Online

Saif-UI-Allah, MW, Khan, J, Ahmed, F, Salman, CA, Gillani, Z, Hussain, A, Yasin, M, UI-Haq, N, Khan, AU, Bazmi, AA, Ahmad, Z and Hasan, M

Computationally Inexpensive 1D-CNN for the Prediction of Noisy Data of NO_x Emissions From 500 MW Coal-Fired Power Plant

<http://researchonline.ljmu.ac.uk/id/eprint/17830/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Saif-UI-Allah, MW, Khan, J, Ahmed, F, Salman, CA, Gillani, Z, Hussain, A, Yasin, M, UI-Haq, N, Khan, AU, Bazmi, AA, Ahmad, Z and Hasan, M (2022) Computationally Inexpensive 1D-CNN for the Prediction of Noisy Data of NO_x Emissions From 500 MW Coal-Fired Power Plant. *Frontiers in Energy*

LJMU has developed [LJMU Research Online](http://researchonline.ljmu.ac.uk/) for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>



Computationally Inexpensive 1D-CNN for the Prediction of Noisy Data of NOx Emissions From 500 MW Coal-Fired Power Plant

Muhammad Waqas Saif-Ul-Allah¹, Javed Khan¹, Faisal Ahmed^{1*}, Chaudhary Awais Salman^{2*}, Zeeshan Gillani³, Arif Hussain¹, Muhammad Yasin⁴, Noaman Ul-Haq⁴, Asad Ullah Khan^{4,5}, Aqeel Ahmed Bazmi^{1*}, Zubair Ahmad^{6*} and Mudassir Hasan⁷

OPEN ACCESS

Edited by:

Dong Liu,
Nanjing University of Science and
Technology, China

Reviewed by:

Solomon Giwa,
Olabisi Onabanjo University, Nigeria
Baseem Khan,
Hawassa University, Ethiopia

*Correspondence:

Faisal Ahmed
faisalahmed@cuilahore.edu.pk
Aqeel Ahmed Bazmi
abazmi@cuilahore.edu.pk
Chaudhary Awais Salman
chaudhary.awais.salman@mdh.se
Zubair Ahmad
zubair7157@yu.ac.kr

Specialty section:

This article was submitted to
Advanced Clean Fuel Technologies,
a section of the journal
Frontiers in Energy Research

Received: 16 May 2022

Accepted: 20 June 2022

Published: 16 August 2022

Citation:

Saif-Ul-Allah MW, Khan J, Ahmed F,
Salman CA, Gillani Z, Hussain A,
Yasin M, Ul-Haq N, Khan AU,
Bazmi AA, Ahmad Z and Hasan M
(2022) Computationally Inexpensive
1D-CNN for the Prediction of Noisy
Data of NOx Emissions From 500 MW
Coal-Fired Power Plant.
Front. Energy Res. 10:945769.
doi: 10.3389/fenrg.2022.945769

¹Process and Energy Systems Engineering Center-PRESTIGE, Department of Chemical Engineering, COMSATS University Islamabad, Lahore, Pakistan, ²School of Business, Society and Engineering, Mälardalen University, Västerås, Sweden, ³Department of Computer Science, COMSATS University Islamabad, Lahore, Pakistan, ⁴Department of Chemical Engineering, COMSATS University Islamabad, Lahore, Pakistan, ⁵Department of Chemical Engineering, SCME, National University of Science and Technology, Islamabad, Pakistan, ⁶School of Chemical Engineering, Yeungnam University, Gyeongsan, South Korea, ⁷Department of Chemical Engineering, College of Engineering, King Khalid University, Abha, Saudi Arabia

Coal-fired power plants have been used to meet the energy requirements in countries where coal reserves are abundant and are the key source of NOx emissions. Owing to the serious environmental and health concerns associated with NOx emissions, much work has been carried out to reduce NOx emissions. Sophisticated artificial intelligence (AI) techniques have been employed during the past few decades, such as least-squares support vector machine (LSSVM), artificial neural networks (ANN), long short-term memory (LSTM), and gated recurrent unit (GRU), to develop the NOx prediction model. Several studies have investigated deep neural networks (DNN) models for accurate NOx emission prediction. However, there is a need to investigate a DNN-based NOx prediction model that is accurate and computationally inexpensive. Recently, a new AI technique, convolutional neural network (CNN), has been introduced and proven superior for image class prediction accuracy. According to the best of the author's knowledge, not much work has been done on the utilization of CNN on NOx emissions from coal-fired power plants. Therefore, this study investigated the prediction performance and computational time of one-dimensional CNN (1D-CNN) on NOx emissions data from a 500 MW coal-fired power plant. The variations of hyperparameters of LSTM, GRU, and 1D-CNN were investigated, and the performance metrics such as RMSE and computational time were recorded to obtain optimal hyperparameters. The obtained optimal values of hyperparameters of LSTM, GRU, and 1D-CNN were then employed for models' development, and consequently, the models were tested on test data. The 1D-CNN NOx emission model improved the training efficiency in terms of RMSE by 70.6% and 60.1% compared to LSTM and GRU, respectively. Furthermore, the testing efficiency for 1D-CNN improved by 10.2% and 15.7% compared to LSTM and GRU, respectively. Moreover, 1D-CNN (26 s) reduced the training time by 83.8% and 50% compared to LSTM (160 s) and GRU (52 s), respectively. Results reveal that 1D-CNN is more accurate,

more stable, and computationally inexpensive compared to LSTM and GRU on NOx emission data from the 500 MW power plant.

Keywords: NOx prediction, machine learning, 1D-convolutional neural network, LSTM, GRU, coal-fired power plant

INTRODUCTION

Most power plants use a coal-fired boiler (CFB) to generate steam to meet energy requirements at the expense of releasing many harmful gases into our environment. These poisonous gases cause serious environmental problems such as sea level ozone development, forest decline, and health issues such as asthma, lung irritation, and eye cancer (Korpela et al., 2015) that have motivated researchers to investigate it seriously. For example, NOx emissions badly impact our natural ecosystem and life support system in the long run by disturbing the natural air quality (Jonson et al., 2017). Air quality is quantified in terms of the air quality index (AQI) that has a direct impact on the human respiratory system and needs to be properly investigated (Ikram and Yan, 2017). Moreover, in industrial effluents such as those from coal-fired power plant boilers, NOx is emitted directly into the air, destroying air quality and causing serious health issues (Gaffney and Marley, 2009). International Energy Agency has reported a ~4.3% overall increase in coal consumption in power generation industry from 2000 (~23% coal consumption) to 2016 (~27.3% coal consumption) (IEA 2020, 2020). Moreover, a high amount of nitrogen and sulfur compounds in coal leads to NOx and SOx as major pollutant emissions from a CFB. In order to investigate pollutant gases, Continuous Emission Monitoring Systems (CEMS) were employed to monitor NOx emissions in the past (Zhang and Schreifels, 2011). However, CEMS were not cost-effective and required periodic maintenance because of the harsh operating environmental conditions. Therefore, artificial intelligence (AI) models based on algorithms have been investigated, which can be put to work besides NOx measuring and monitoring systems (Lv et al., 2013).

In the past, researchers have employed mathematical models (Ke et al., 2022) and AI algorithms (Ahmed et al., 2015) to predict CFB gas emissions. Amongst the different methods proposed, AI models and mathematical approaches gained ultimate interest (Vo et al., 2019). The formation of NOx from CFB has high complexity and cannot be modeled

mathematically effectively. Some studies have employed mathematical modeling for NOx emissions with some assumptions that ease the calculations and have introduced error methods to predict the accuracy of the model (Ke et al., 2022). Besides, mathematical modeling employs numerical solutions and is found computationally expensive (Vo et al., 2019). In contrast, AI models are inexpensive and require no insight or knowledge of the process. Amongst these AI models, the support vector machine (SVM) and its variant least-squares support vector machine (LSSVM) were found promising for NOx prediction (Zhai et al., 2020). LSSVM prediction quality and computational time were better than most AI models, but it requires well-pre-trained parameters for prediction (Lv et al., 2015). Several methods were employed to estimate the accuracy of the trained model such as root mean square error (RMSE), mean average error (MAE), mean regression error (MRE), and R^2 (Baghban et al., 2019).

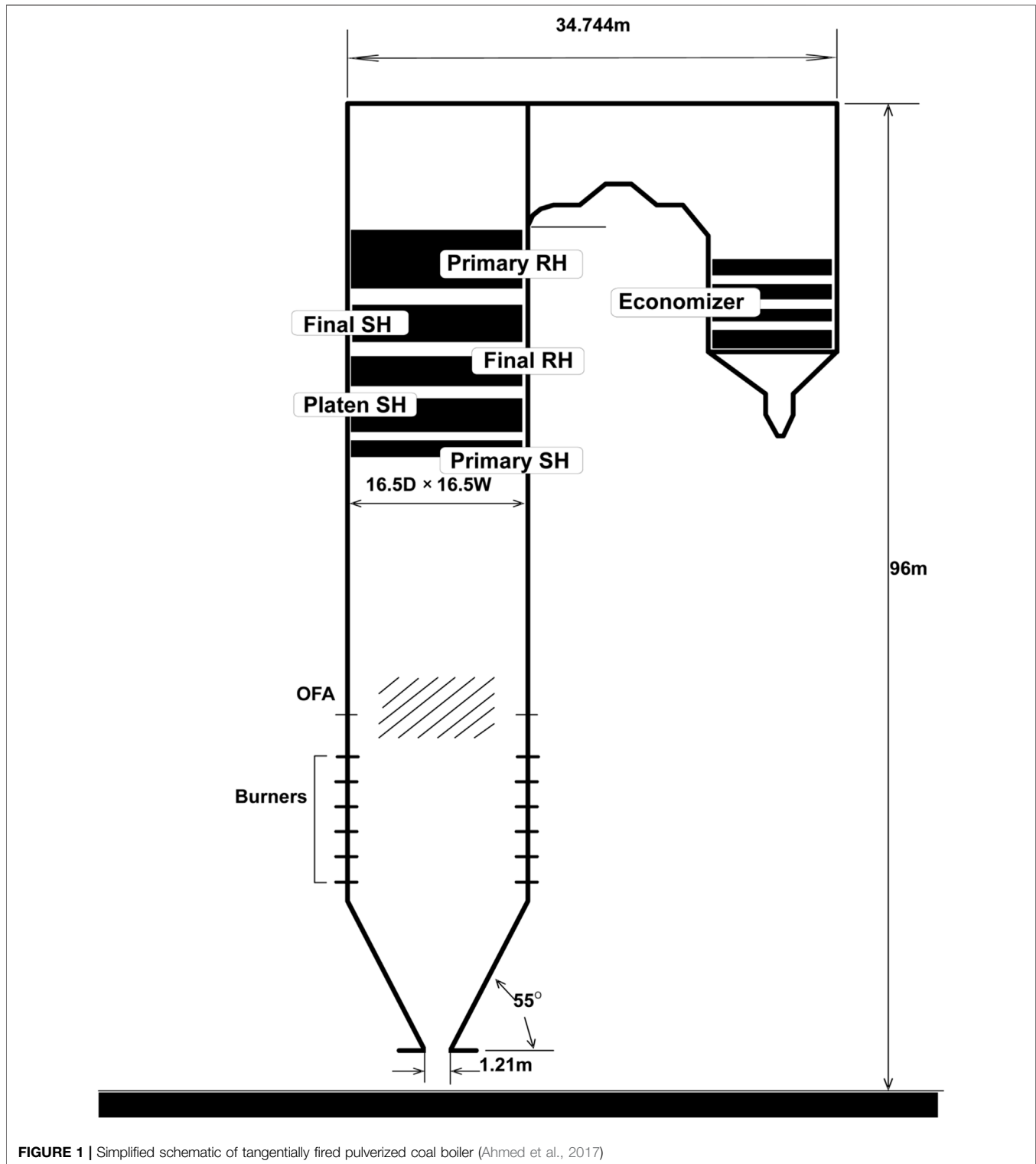
More sophisticated artificial neural networks (ANN) have evolved, such as single-layer neural networks, shallow neural networks, and deep neural networks (DNN), which are more reliable than other AI techniques. ANN consists of layers of neurons conceived from the human neural system (Jain et al., 1996). A single-layer neural network has a relatively simple architecture with one input layer and one output layer. A shallow neural network adds one hidden layer between the input and output layer. DNN have a relatively very complex architecture because of multiple hidden layers between the input and output layers (Miikkulainen et al., 2019). The current era has converged its focus on DNN because of their versatility and precise performance. Like other AI models, DNN also demands hyperparameter training from historical data from the equipment. Several studies (Adams et al., 2020; Shin et al., 2020; Spinelli et al., 2020; Yang et al., 2020) have employed different DNN architectures and proved the prediction efficiency. DNN with more complex geometry, if providing enough training data to properly train its hyperparameters (weights, bias), can be employed for NOx emission prediction effectively (Spinelli et al., 2020). Deep hybrid neural networks have also been developed and investigated for CFB gas emissions and control (Hu et al.,

TABLE 1 | Data collected for modeling experiment

Variable name		Units
Coal feeder flow rate	$i_1, i_2, i_3, i_4, i_5, i_6$	t/h
Mill differential pressure	$i_7, i_8, i_9, i_{10}, i_{11}, i_{12}$	A
Boiler outlet flue gas temperature	i_{13}	°C
Flue gas O ₂ concentration	i_{14}, i_{15}, i_{16}	%
Flue gas CO concentration	i_{17}	%
FGD in SOx	i_{18}	%
RH out temperature	i_{19}	°C
Power generated	i_{20}	MW
SCR average temperature	i_{21}, i_{22}	°C

TABLE 2 | Coal properties.

Variable name	Value	Units
Top total moisture	16.97	%
Intrinsic moisture	9.3	%
Ash	9.7	%
Bottom volatile compounds	37.05	%
Top volatile compounds	37.05	%
Sulphur contents	0.6725	%
Calorific value	5532.19	kJ/kg
Granularity	2	Millimeters



2020). A partially supervised DNN regression-based model has been successfully investigated for airborne hyperspectral imagery for improved spatial soil properties prediction (Ou et al., 2021). Moreover, larger and more sophisticated DNN models require large hardware resources for training, making them very

computationally expensive. A variant of DNN, a bidirectional gated recurrent neural network unit, has also been developed and applied for early prediction of the AQI and investigated the effect of temporal and spatial correlation on model performance (Zhang K. et al., 2020). Various studies have proposed and

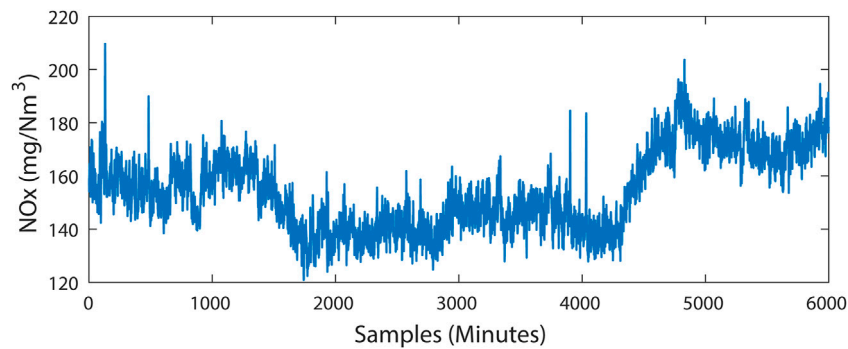


FIGURE 2 | Boiler NOx emissions concentration.

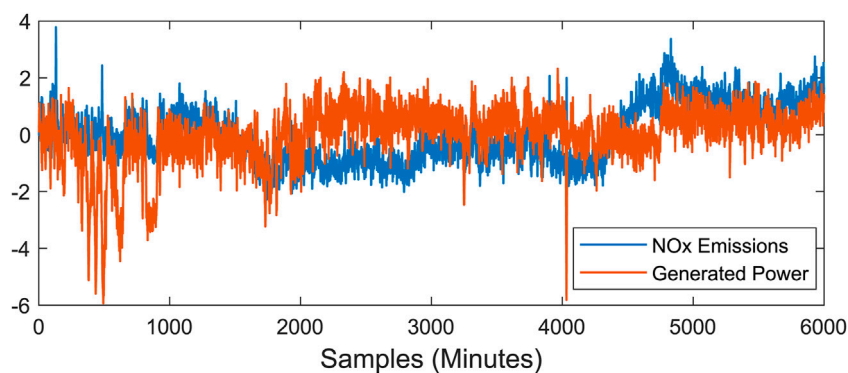


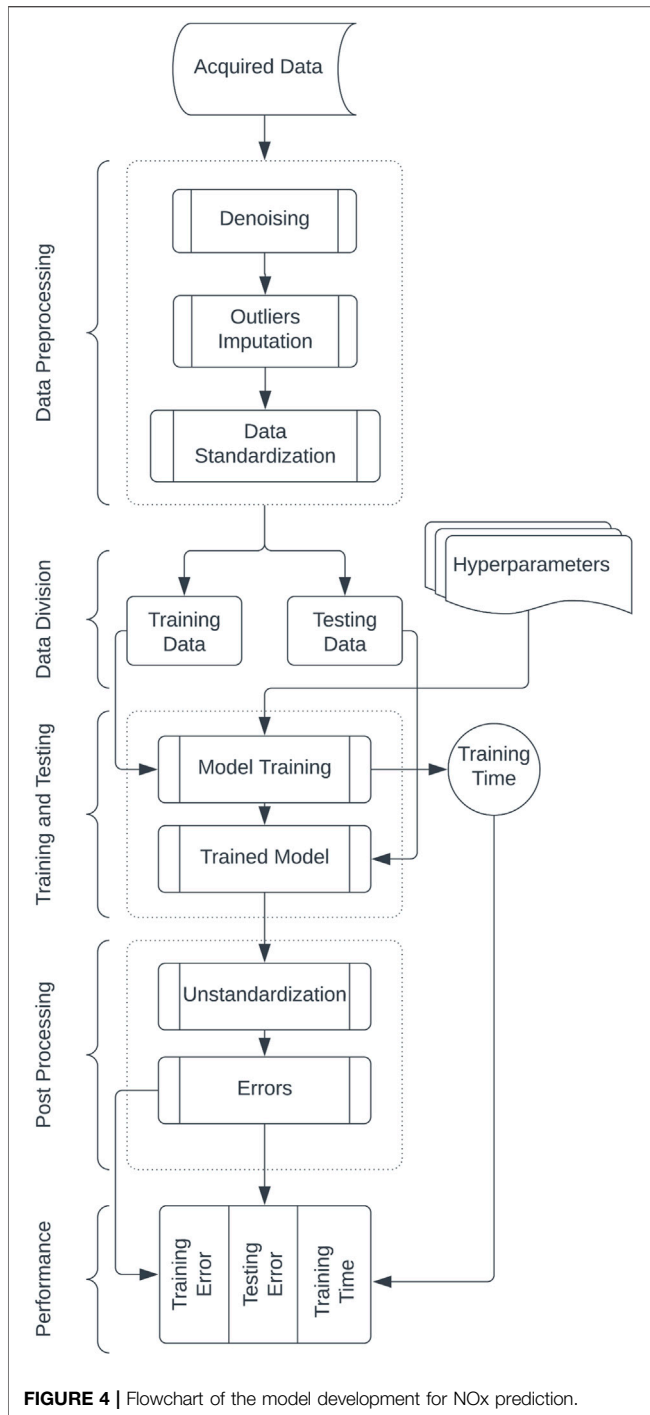
FIGURE 3 | Variation in unit load with NOx emissions.

employed various DNN models for complex processes such as boiler NOx emissions and prediction (Song et al., 2022; Wang et al., 2022). Recently, a variant of DNN S3LX has been proposed for boiler NOx prediction, and the result was shown with $\sim 4.28\%$ error (Zhang Y. et al., 2020). The experimental work has revealed the promising performance of the proposed DNN model compared to the multi-layer perceptron model and support vector regression model. A study has employed a long short-term memory (LSTM) DNN variant for boiler NOx prediction (Yang et al., 2020). A novel deep bidirectional learning machine model has been employed successfully and showed good generalization capability to improve the efficiency and NOx emissions for the 300 MW plant boiler (Li et al., 2017). Moreover, a convolution operation-based neural network has also been proposed for boiler NOx emissions prediction from real plant data (Spinelli et al., 2020). The study showed the excellent performance of the model and proved promising results against LSTM and the other DNN model proposed.

Training the DNN model requires high-quality data with quality variables. Sometimes, technical glitches and other technical interferences reduce data quality with missing values. A study has put forward different sophisticated mechanistic approaches to deal with missing data (Kang, 2013). Moreover, a non-iterative neural-like framework has also been investigated

for filling missing data and has shown promising results (Tkachenko et al., 2019). Also, a study proposed a modified DNN-based convolutional imputation method and has shown excellent efficiency (Spinelli et al., 2020). Moreover, the outliers in the data might reduce the data quality that should be treated accordingly. A study reported the best possible ways to handle unnecessary outliers to improve data quality for model training and testing (Bortolotti, 2018). Moreover, another study suggested the use of statistical preprocessing of the obtained data that help in building a good model (Al-jabery et al., 2020).

Previous studies had proposed DNN models to predict NOx emissions from coal-fired power plant boilers. However, there is a need to investigate DNN-based models that are accurate and computationally inexpensive. This study investigated recurrent neural network (RNN) variants, such as LSTM and GRU, and a newly developed one-dimensional convolutional operation-based neural network (1D-CNN) for boiler NOx emissions prediction for real plant data. The performance of the investigated models has been evaluated using training RMSE, training time, and testing RMSE. This study also considered modeling simple ANN and AI models such as linear regression (LR) and SVM for comparison. This study aims to bring forth the outperformed model that has a good generalization ability and is computationally inexpensive.



DATA PREPARATION

Generally, power plant operational data are logged using distributive control systems (DCS). For this study, the collected data for the 500 MW power plant contained 55 process variables containing NOx as a dependent variable. Every independent variable was plotted, and the variables that were not correlated with boiler NOx emissions concentration were dropped. However, by visualizing

correlations amongst all collected features, 22 input variables, as shown in **Table 1**, and one output variable (NOx) were selected for modeling.

Process Description

For this study, a 500 MW power plant was investigated, and NOx data were collected for tangentially fired pulverized coal boiler as shown in **Figure 1**. The coal is stacked in silos conveyed to the crushing mill for size reduction through a conveyor belt. Furthermore, the boiler consists of 24 burners and six layers. Before entering the mill, which transports the pulverized coal to the boiler, the primary air is preheated in the preheater. Secondary air is forced to draft through the preheater that entered the furnace to burn the coal. Combustion gas and ash are produced after the combustion. NOx was removed in a selective catalytic reactor (SCR), and the remaining NOx was captured in a dust collector.

In a CFB, coal is the core factor that has a direct impact on NOx emissions. The boiler was equipped with a coal analyzer, and coal properties were analyzed per minute. The properties of the coal are listed in **Table 2**.

Every variable of the boiler data was visually analyzed to estimate the normal boiler operation condition without disturbance. As the boiler NOx emissions are highly non-linear (**Figure 2**), non-stationarity in data was also considered. For this study, 6,000 samples per minute were selected for NOx emissions modeling.

The selected 6,000 samples were then analyzed carefully for noise and outliers as they can drastically impact the model performance. The data were recorded per minute, constituting almost 4 days of data. The variables given in **Table 1** were employed as input variables for the NOx emission modeling. For visual inspection, data were first standardized using **Eq. 1**. Visual inspection revealed that load (MW) was correlated with NOx emissions, as plotted in **Figure 3**. It can be noted that NOx emissions became higher when the load increased and became lower when the load decreased. After de-noising and removing outliers, the data were standardized using **Eq. 1** to remove the mean and include unit variance for effective model training:

$$\text{standardization} = \frac{x_i - \mu}{\sigma} \quad (1)$$

Initially, the outliers in the data were visually inspected in each variable. As deleting a sample from time series might cause potent information to be lost, outliers were imputed using the missing value imputation technique. Saif-ul-Allah et al. compared different missing data imputation techniques and reported project to model plan (PMP) as outperforming method for PM2.5. This study employed PMP as an imputation technique and used deep learning methods LSTM, Bi-LSTM, and GRU to predict PM2.5 concentration in Guangzhou city, China (Saif-ul-Allah et al., 2022). Moreover, Fast Fourier Transformation (FFT) was employed to analyze the frequency component of each variable with respect to time. The high-frequency component with a low power value was then removed using a bandpass filter.

TABLE 3 | Performance metrics results of LSTM, GRU, and 1D-CNN with varying hyperparameters.

Models	Hyperparameters		Performance metrics		
	Hidden units		Training RMSE	Testing RMSE	Training time (s)
LSTM	LSTM-20		4.52	9.99	56 s
	LSTM-40		3.96	7.75	80 s
	LSTM-60		3.78	7.6	110 s
	LSTM-80		3.75	7.13	160 s
	LSTM-100		3.67	7.94	180 s
GRU	GRU-20		3.55	9.52	25 s
	GRU-40		3.06	8.71	40 s
	GRU-60		2.76	7.6	52 s
	GRU-80		2.59	8.52	100 s
	GRU-100		2.46	8.5	110 s
1D-CNN	Number of filters	Filter size	Training RMSE	Testing RMSE	Training time
	100	5	2.05	8.76	10
	200	5	1.96	8.7	13
	300	5	1.9	7.85	17
	400	5	1.72	7.02	24
	500	5	1.1	6.4	26
	600	5	1.65	9.4	31

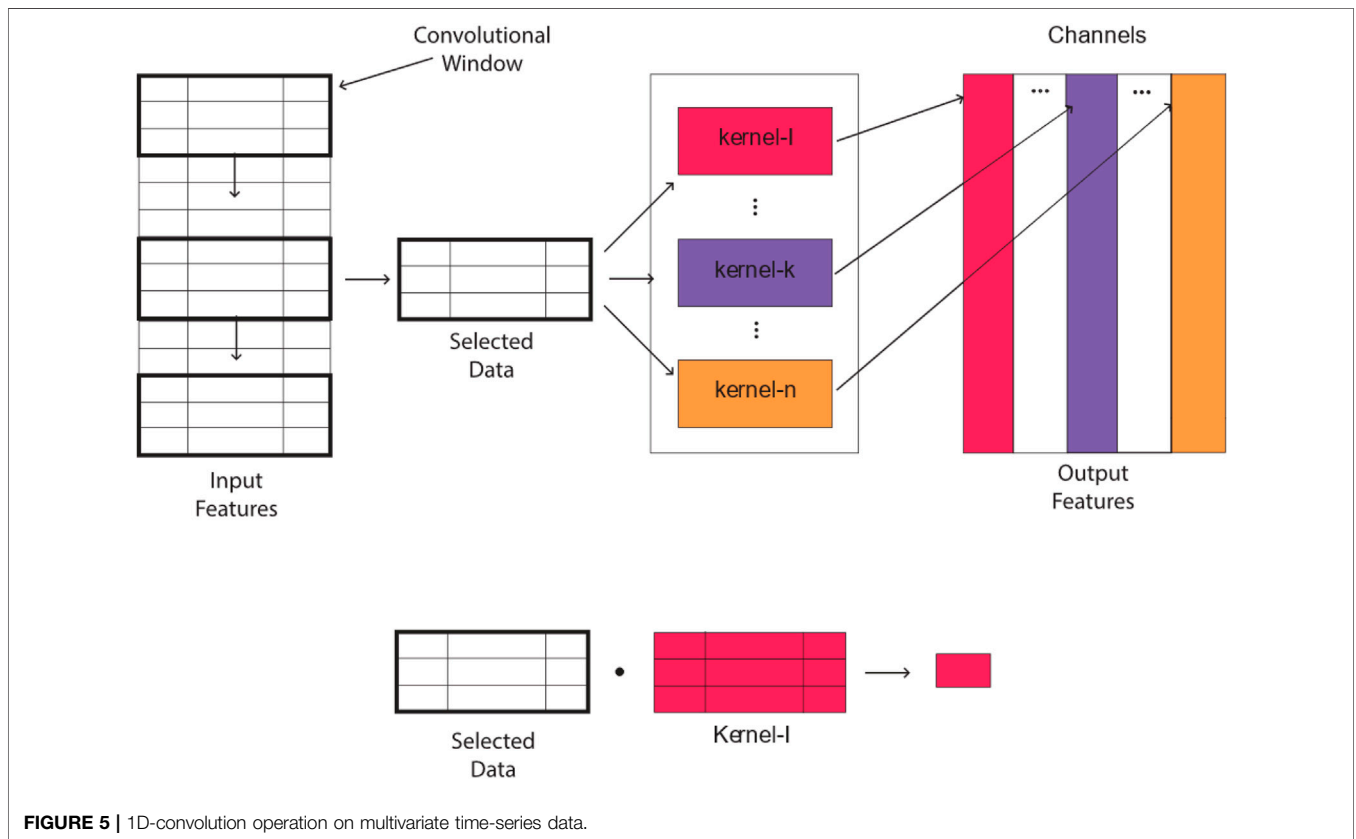


FIGURE 5 | 1D-convolution operation on multivariate time-series data.

MODELING AND SIMULATION

NOx Prediction Modeling

Boiler NOx prediction modeling with high dimensional data containing noise is very complicated. Visual inspection of the input data resulted in a good correlation with NOx emissions concentration. However, the correlation amongst the input data itself produces redundant information that makes NOx

prediction modeling more challenging and complex. A detailed flowchart of modeling strategy applied in this study is shown in **Figure 4**.

LR, SVM, and ANN Modeling

This study also considered modeling traditional AI models such as LR, SVM, and ANN. A simple LR model can be represented as follows:

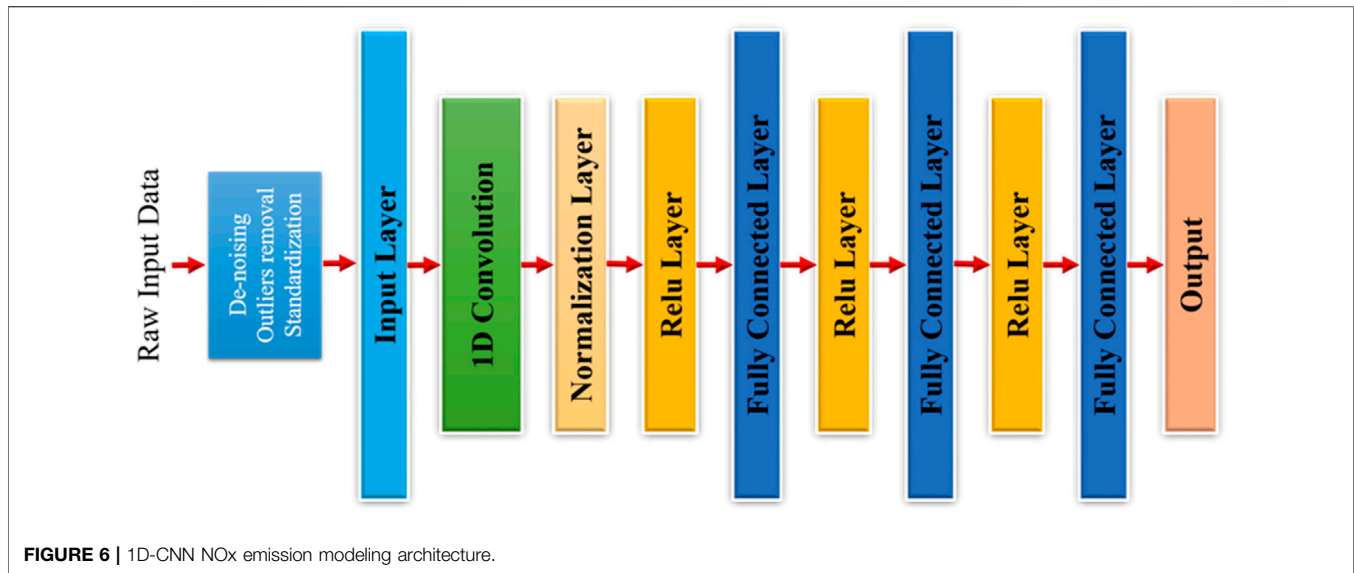


TABLE 4 | Comparison of AI models for NOx emissions test data.

Models	Performance metrics		
	RMSE	MAE	MAPE
LR	11.35	9.08	0.0570
SVM	11.42	9.17	0.0573
ANN	10.75	8.14	0.0506
LSTM-80	7.13	5.62	0.0349
GRU-60	7.6	5.99	0.0365
1D-CNN-500	6.4	5.06	0.0315

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_n X_{in} + \varepsilon_i, \quad (2)$$

where y_i is the output variable (NOx), β_0 is a constant term, β_k represents model coefficient, X_{ij} is the i th observation in j th input variable, and ε_i is the noise term.

The SVM model used the linear kernel with the kernel scale. Training the SVM model searches for suitable values of hyperparameters such as kernel scale, box constraint, and epsilon that better correlate the input and output variable. Using random search algorithm, SVM hyperparameters were tuned and the values of kernel scale, box constraint and epsilon were found to be 4.2, 12.40 and 1.24, respectively. On the contrary, ANN was devised based on the human nervous system and tried to learn like humans with experience. The training data are used to train the model. Moreover, multiple neurons in ANN, which are small mathematical models, propagate information from one node to another. Those neurons are associated with the weights and bias matrices learned while training ANN. For ANN training, hyperparameters such as the number of layers, activation, and lambda were tuned using a random search algorithm, finding three hidden layers, sigmoid activation, and 0.33 lambda as tuned hyperparameters. Also, for ANN training, the learning rate of 0.01 with 1,000 iterations was used to adjust weights and biases for hidden layers.

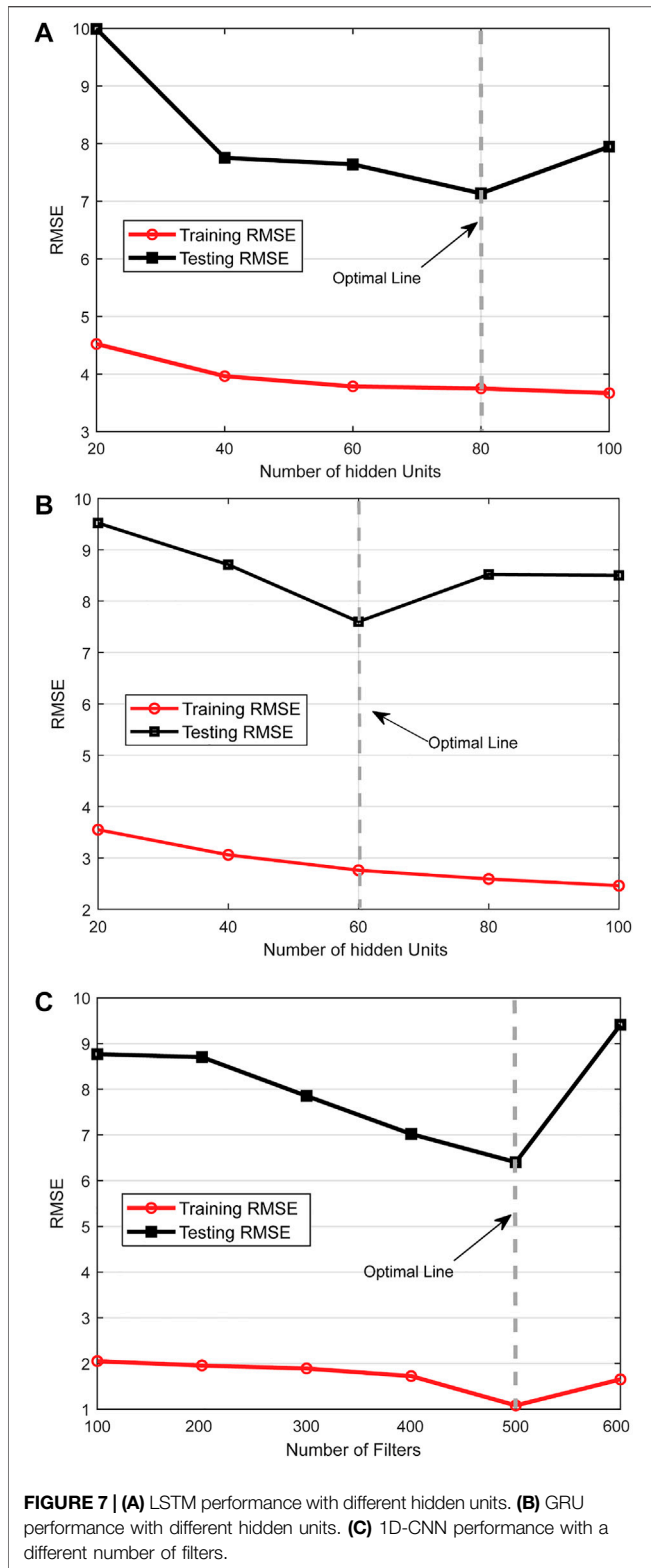
LSTM, GRU, and 1D-CNN Modeling

Besides these regression techniques, more advanced time-series RNN models such as LSTM and gated recurrent unit (GRU) were also modeled in this study. Very few studies have incorporated the 1D-CNN technique for boiler NOx emissions. This study considered modeling 1D-CNN for boiler NOx emission prediction.

The LSTM model was trained for NOx prediction with one input layer, one hidden layer, one fully connected layer, and one output layer. Also, a dropout layer was used after the LSTM layer to avoid overfitting the model. The activation function used for the hidden layer was soft-sigmoid, and the output layer was simple linear regression. In the case of LSTM, the number of hidden units is an important hyperparameter that drastically affects the model efficiency. In this study, LSTM was modeled using five different values of hidden layer neurons such as 20, 40, 60, 80, and 100. Moreover, GRU was also modeled for NOx emission prediction with the same layer arrangements as LSTM. Also, five different values of hidden layer neurons such as 20, 40, 60, 80, and 100 were employed to model GRU (Table 3).

CNN has already shown its superiority for 2D images (Tang et al., 2022). For one-dimensional data, 1D-CNN can be modeled for time-series multivariate NOx data. Moreover, 1D-CNN was found computationally very efficient and able to deal with time-series data. The 1D-convolution operation for time-series one-dimensional data is shown in Figure 5. The selected window size convolves over the time-series data and applies kernel on the selected window. The kernel matrix is the weight matrix whose values are learned while training the network. The dot product of the selected window and the kernel matrix essentially provides the scalar values.

The 1D-CNN model architecture contained one input layer, one hidden layer, and one output regression layer, as shown in Figure 6. Different numbers of filters (100, 200, 300, 400, 500, and 600) of size five were used in the 1D-convolutional layer to



convolve with time sequence data with padding mode “as same.” After the 1D-convolutional layer, one normalization layer was added to normalize the output of the convolutional layer. Moreover, the normalized data were then processed with the

ReLU layer. Performance measuring criteria such as training RMSE, testing RMSE, and training time were compared for LSTM, GRU, and 1D-CNN NOx emission models. This study aims to report the model that outperforms with less testing RMSE and training time.

Performance Metrics

After training, performance measuring criteria RMSE, MAPE, and MAE were employed to quantify training and testing prediction performance using Eqs 3–5:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - pred_i)^2}, \quad (3)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - pred_i}{x_i} \right|, \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - pred_i|. \quad (5)$$

RESULTS AND DISCUSSION

Performance Comparison of LR, SVM, and ANN

AI techniques such as LR, SVM, and ANN were modeled for NOx emissions, and testing data prediction performances were reported and compared (Table 4). It can be seen from the table that LR reported the testing data prediction RMSE, MAE, and MAPE of 11.35, 9.08, and 0.0570, respectively. On the contrary, SVM with its tuned hyperparameters reported poorer performance than LR and reported testing data prediction RMSE, MAE, and MAPE of 11.42, 9.17, and 0.0573. Moreover, ANN hyperparameters were also tuned, and the training was performed with different learning rates and iterations. With three hidden layers, sigmoid activation, and 0.33 lambda, ANN correlated, the input features with NOx concentrations were better than LR and SVM at the learning rate of 0.01 with 1,000 training iterations. ANN improved the testing data prediction results compared to LR and SVM and reported RMSE (10.75), MAE (8.14), and MAPE (0.0506).

LSTM Model for NOx Emissions

The LSTM model was independently trained with 20, 40, 60, 80, and 100 hidden layer neurons. Parameters such as training RMSE, testing RMSE, and training time were recorded; each time model was trained to compare the model’s efficiency. RMSE and training time results have been tabulated in Table 3, where the fourth row shows that the LSTM model performed better in terms of testing RMSE with 80 hidden layer neurons.

The training and testing RMSEs against the LSTM models with 20, 40, 60, 80, and 100 hidden units have been plotted in Figure 7A. The minimum testing RMSE was found at 80 hidden units, which is the optimal value. After increasing the number of neurons from the optimal value, the LSTM model overfitted and lost the prediction generalizability, as can be seen in Figure 7A. However, with fewer neurons compared to the optimal value, the model could not learn from the training data; hence, LSTM was

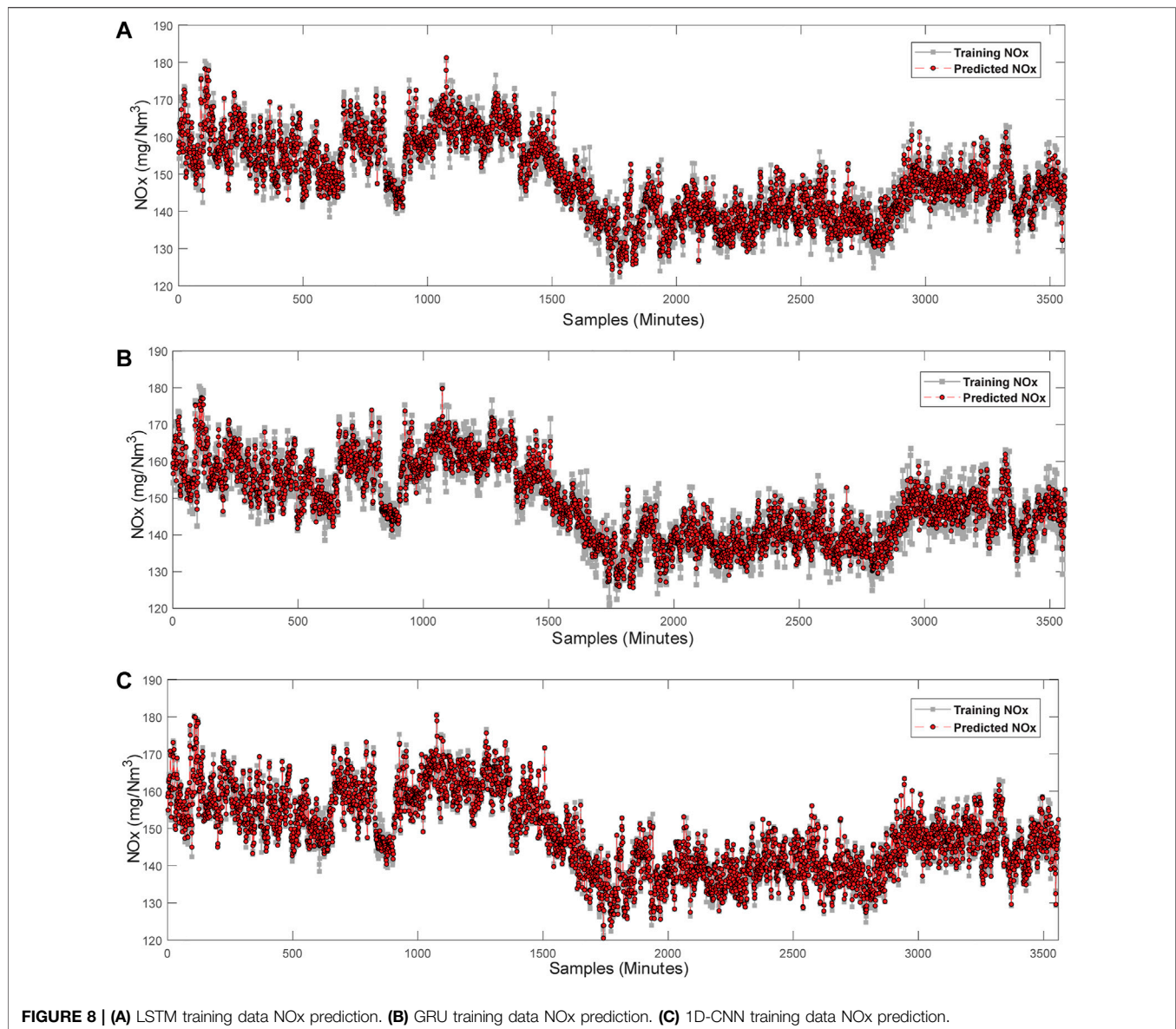


FIGURE 8 | (A) LSTM training data NO_x prediction. **(B)** GRU training data NO_x prediction. **(C)** 1D-CNN training data NO_x prediction.

not generalized enough to perform well. Moreover, an increase in hidden layer neurons effectively decreased the training RMSE, but testing RMSE increased accordingly, representing that model started overfitting beyond optimal value.

The prediction performance of NO_x using the LSTM-80 model with training and testing data is presented in **Figure 8A** and **Figure 9A**. The trained LSTM model provided a good prediction of the NO_x with training data. The training and testing RMSEs were 3.75 and 7.13, respectively. Compared to ANN, LSTM-80 decreased the prediction errors by 33.6% RMSE, 30.9% MAE, and 30.02% MAPE (**Table 4**).

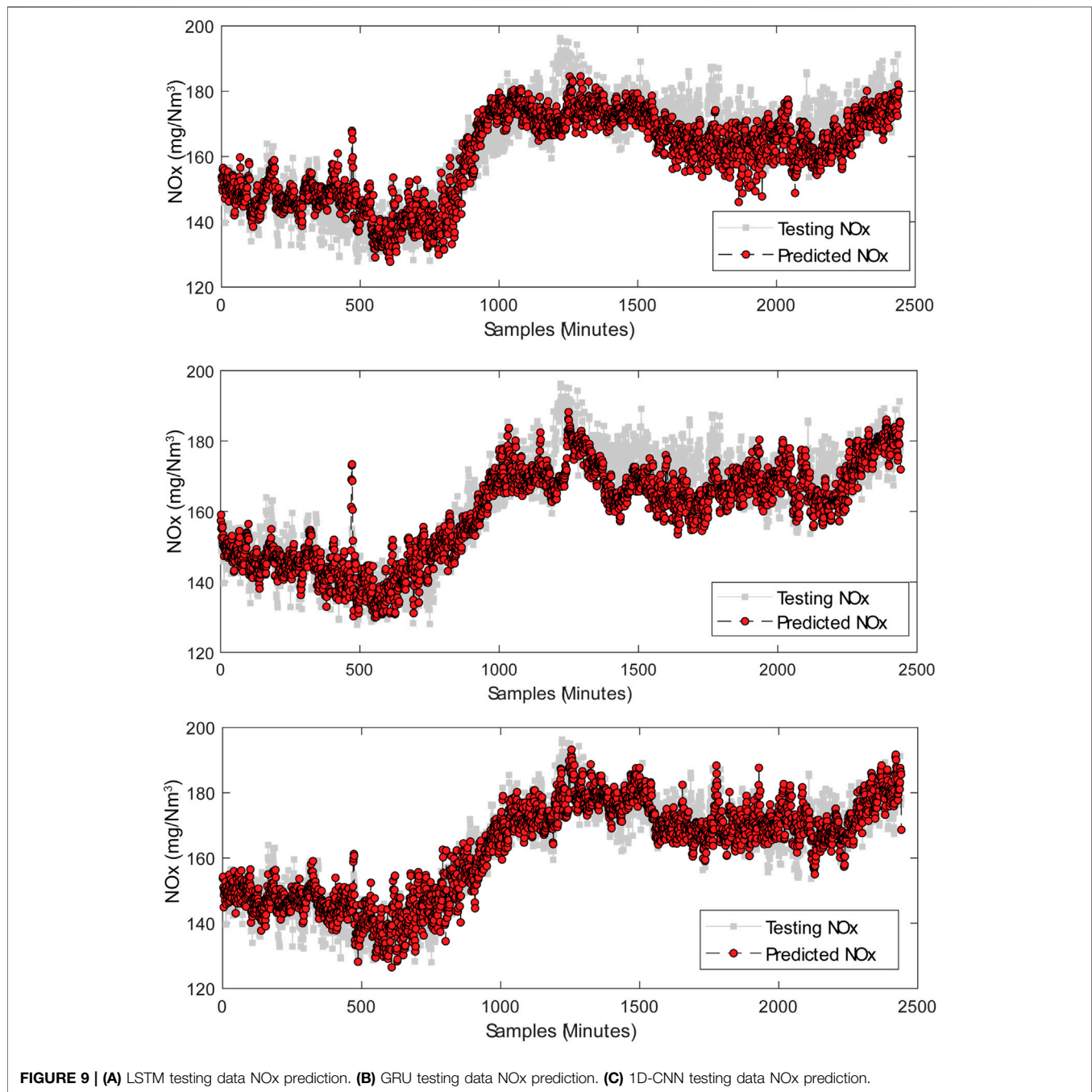
Moreover, **Table 3** also reports the training time for every LSTM run for 20, 40, 60, 80, and 100 hidden layer neurons. LSTM training with 20 hidden layer neurons took 56 and 80 s for 40, 110 s for 60, 160 s for 80, and 180 s for 100 hidden layer neurons. Furthermore, with the increase in hidden layer neurons, the

training time increased, depicting intensive computation while training.

GRU Model for NO_x Emissions

GRU model was also trained independently with 20, 40, 60, 80, and 100 hidden layer neurons. Moreover, while modeling, GRU started overfitting at 80 hidden layer neurons. The results in **Table 3** show that the GRU model performed better with 60 neurons giving the least testing RSME of 7.6. With the increase in the number of hidden layer neurons, the GRU training RMSE reduced, and testing RMSE increased, indicating model overfitting beyond optimal value, as shown in **Figure 7B**.

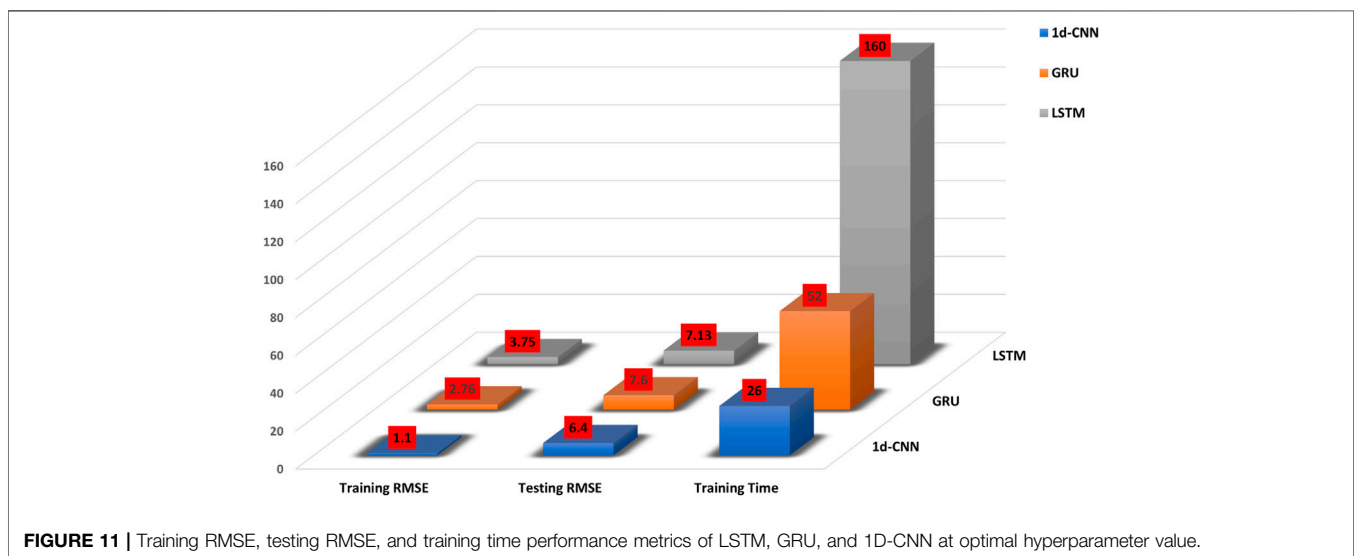
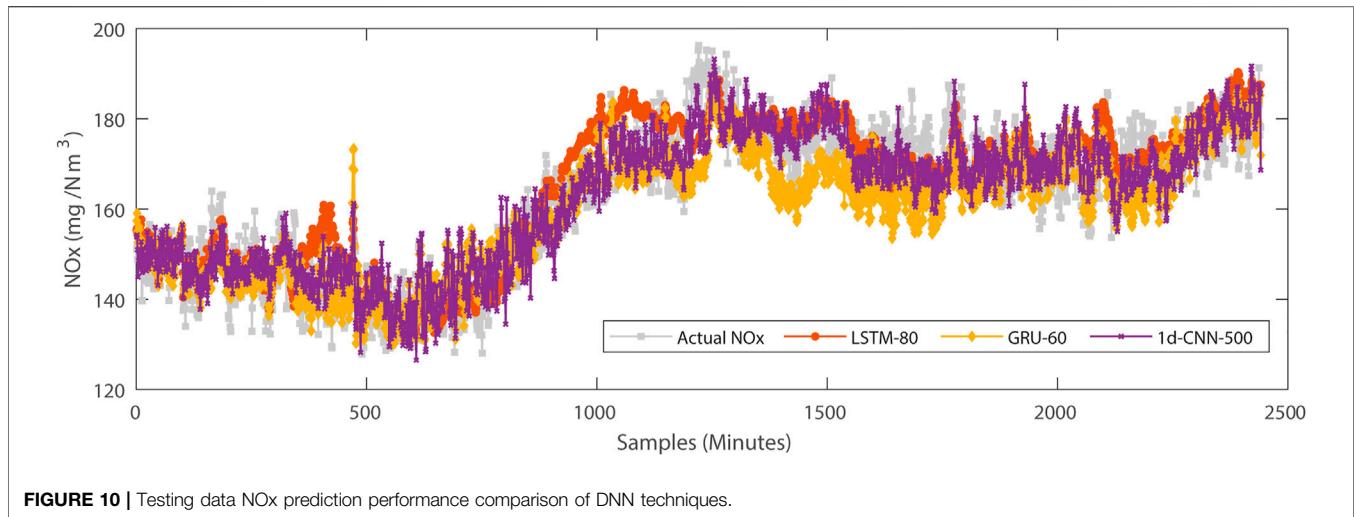
The prediction efficiency of the GRU model on training and testing data is presented in **Figures 8B, 9B**, showing that the model could predict the NO_x on training data effectively with an



RMSE of 2.76. Moreover, on optimal hyperparameter values, GRU improved the performance in terms of training time, that is, 52 s, compared to the LSTM model (160 s) but failed to outperform the LSTM in terms of the NOx prediction and reported a bit higher testing RMSE of 7.6 (Table 3). However, compared to ANN, GRU-60 could still report improved performance. GRU-60 decreased the prediction error compared to ANN by 29.3% RMSE, 26.4% MAE, and 27.8% MAPE.

1D-CNN Model for NOx Emissions

For time-series one-dimension data, RNN and its other variants have played an important role in prediction and classification tasks. However, the major drawback with LSTM and GRU was the computational expensiveness. While experimenting on 1D-CNN for modeling NOx emissions, it was found computationally lighter than LSTM and GRU. Moreover, it conclusively showed excellent time-series multi-feature data modeling ability. However, like LSTM and GRU, 1D-CNN also required a few



hyperparameters and found the number of convolving filters most effective. 1D-CNN modeling of NO_x emissions was performed with different values of the number of filters, such as 100, 200, 300, 400, 500, and 600, independently, and the results are reported in **Table 3**. The convolving filters extract features from the selected window, and those features were used to adjust the weights of hidden layers. 1D-CNN NO_x modeling experiment with 500 filters showed the best result in terms of training RMSE, testing RMSE, and training time. However, it was noticed that 1D-CNN stopped learning from the extracted features if the number of filters increased from 500. With the increase in the number of filters from the optimal value, the model was not learning from the extracted features effectively, and training RMSE was observed to increase, hence reporting poor testing data prediction performance, as shown in **Figure 7C**.

Training and testing results with 500 filters are presented in **Figures 8C, 9C**. The comparison with other models showed that

1D-CNN outperformed the LSTM and GRU model with training RMSE 1.1 and testing RMSE of 6.4 with CFB NO_x data. **Table 3** also presents the training time taken by each model in the experiment. The 1D-CNN model outperformed LSTM and GRU in terms of training time. The best results of the LSTM and GRU experiment in terms of testing RMSE showed training times of 160 and 52 s, respectively. The 1D-CNN experiment revealed that it outperformed the LSTM and GRU models for NO_x prediction in terms of not only testing RMSE but also training time, that is, 26 s.

Overall comparison of LSTM, GRU, and 1D-CNN on testing data prediction performance can be analyzed in **Figure 10**, highlighting that 1D-CNN outperformed the other models on testing data NO_x predictions with testing RMSE of 6.4. **Figure 10** shows that at 1,150–1,220 sample points, LSTM and GRU went largely out of the original NO_x emissions trend. At this time interval, samples of input variables seemed normal, and no

issues were detected. However, at the same time interval, samples of NO_x were a bit noisy that LSTM and GRU unable to predict properly. At the same time interval of 1,150–1,220, 1D-CNN could estimate the NO_x trend accurately. At the optimal value of hyperparameters, 1D-CNN improved the prediction performance on testing data by 10.2% compared to LSTM and 15.7% compared to GRU, as shown in **Figure 11**. For complete training at optimal hyperparameter value, LSTM and GRU took 160 and 52 s, respectively. On the contrary, at the optimal hyperparameter value, 1D-CNN took only 26 s for training, was computationally inexpensive, and decreased the training time by 84% and 50%, respectively, compared to LSTM and GRU (**Figure 11**).

The outperforming 1D-CNN-500 model decreased the testing data prediction error compared to ANN by 40.4% RMSE, 37.8% MAE, and 37.7% MAPE. Overall, **Table 4** compares the performances of the more advanced NO_x emissions models such as LSTM-80, GRU-60, and 1D-CNN-500 with other AI techniques such as LR, SVM, and ANN. **Table 4** suggests that, with noisy and highly non-linear NO_x data, LR, SVM, and ANN could not perform well. On the contrary, 1D-CNN-500 outperformed LSTM-80 and GRU-60 (**Tables 3, 4**).

CONCLUSION

Acquired time-series NO_x emission data were modeled, and the modeling behavior between NO_x emission concentrations and other process variables was explored using different modeling techniques such as LR, SVM, ANN, LSTM, GRU, and 1D-CNN. Training and testing RMSEs of the LSTM, GRU, and 1D-CNN models were investigated while varying the relevant hyperparameter to attain an optimal value. Testing RMSE, MAE, and MAPE of LSTM and GRU models at optimal hidden layer neurons were lower than those of LR, SVM, and ANN. On the contrary, 1D-CNN decreased testing RMSE by 10.2% and 15.7% compared to LSTM and GRU, respectively. Moreover, 1D-CNN reduced the training time (26 s) by 83.8% and 50% compared to LSTM (160 s) and GRU (52 s), respectively. The comparative analysis concluded that 1D-CNN, with optimal hyperparameters, was very computationally inexpensive and accurate compared to LR,

REFERENCES

- Adams, D., Oh, D.-H., Kim, D.-W., Lee, C.-H., and Oh, M. (2020). Prediction of SO_x-NO_x Emission from a Coal-Fired CFB Power Plant with Machine Learning: Plant Data Learned by Deep Neural Network and Least Square Support Vector Machine. *J. Clean. Prod.* 270, 122310. doi:10.1016/j.jclepro.2020.122310
- Ahmed, F., Cho, H., Kim, J.-K., Seong, N., and Yeo, Y.-K. (2015). Prediction of NO_x Emission from Coal Fired Power Plant Based on Real-Time Model Updates and Output Bias Update. *J. Chem. Eng. Jpn./JCEJ* 48, 35–43. doi:10.1252/jcej.13we326
- Ahmed, F., Kim, J.-K., Khan, A., Park, H., and Yeo, Y. (2017). A Fast Converging and Consistent Teaching-Learning-Self-Study Algorithm for Optimization: A

SVM, ANN, LSTM, and GRU with acquired NO_x data. The promising results of 1D-CNN suggest that it can be used in other similar fields in the future for robust and computationally inexpensive predictions. An operation engineer in the field can benefit from the proposed 1D-CNN modeling strategy without using hardcore hardware resources for taking quick and accurate predictions that can help engineers control NO_x emissions.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors without undue reservation.

AUTHOR CONTRIBUTIONS

Conceptualization: FA and CS; methodology and software: MS-U-A, JK, and ZG; formal analysis: AH, MY, and NU-H; investigation: MS-U-A and JK; writing—original draft preparation: MS-U-A, JK; project administration: FA and AK; writing—review and editing: ZA and MH; supervision: FA and AB. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was funded by the Higher Education Commission Pakistan (HEC) under the National Research Program for Universities (NRPU), 2017, Project Number 8167. This work was also supported by the Deanship of Scientific Research at King Khalid University, Abha, KSA, through General Research Project under grant number (R.G.P. 411 2/189/43).

ACKNOWLEDGMENTS

The authors would like to appreciate the contributions of Muhammad Abdul Qyum (Department of Petroleum and Chemical Engineering, Sultan Qaboos University, Muscat, Oman) for his valuable input to the publication of this article.

- Case Study of Tuning of LSSVM Parameters for the Prediction of NO_x Emissions from a Tangentially Fired Pulverized Coal Boiler. *J. Chem. Eng. Jpn./JCEJ* 50, 273–290. doi:10.1252/jcej.16we002
- Al-Jabery, K. K., Obafemi-Ajayi, T., Olbricht, G. R., and Wunsch, D. C., II (2020). “2 - Data Preprocessing,” in *Computational Learning Approaches to Data Analytics in Biomedical Applications*. Editors K. K. Al-Jabery, T. Obafemi-Ajayi, G. R. Olbricht, and D. C. Wunsch II (Academic Press).
- Baghban, A., Sasanipour, J., Habibzadeh, S., and Zhang, Z. (2019). Sulfur Dioxide Solubility Prediction in Ionic Liquids by a Group Contribution - LSSVM Model. *Chem. Eng. Res. Des.* 142, 44–52. doi:10.1016/j.cherd.2018.11.026
- Bortolotti, R. (2018). “Tutorial L - Data Prep 2-3: Outlier Handling,” in *Handbook of Statistical Analysis and Data Mining Applications*. Editors R. NISBET, G. MINER, and K. YALE. Second Edition (Boston: Academic Press).

- Gaffney, J. S., and Marley, N. A. (2009). The Impacts of Combustion Emissions on Air Quality and Climate - from Coal to Biofuels and beyond. *Atmos. Environ.* 43, 23–36. doi:10.1016/j.atmosenv.2008.09.016
- Hu, X., Niu, P., Wang, J., and Zhang, X. (2020). Multi-objective Prediction of Coal-Fired Boiler with a Deep Hybrid Neural Networks. *Atmos. Pollut. Res.* 11, 1084–1090. doi:10.1016/j.apr.2020.04.001
- IEA (2020). *World Energy Outlook*. Paris: IEA. Available at: <https://www.iea.org/reports/world-energy-outlook-2020>.
- Ikram, M., and Yan, Z. J. (2017). Statistical Analysis of the Impact of AQI on Respiratory Disease in Beijing: Application Case 2009. *Energy Procedia* 107, 340–344. doi:10.1016/j.egypro.2016.12.169
- Jain, A. K., Jianchang Mao, M., and Mohiuddin, K. M. (1996). Artificial Neural Networks: a Tutorial. *Computer* 29, 31–44. doi:10.1109/2.485891
- Jonson, J. E., Borken-Kleeefeld, J., Simpson, D., Nyiri, A., Posch, M., and Heyes, C. (2017). Impact of Excess NO X Emissions from Diesel Cars on Air Quality, Public Health and Eutrophication in Europe. *Environ. Res. Lett.* 12, 094017. doi:10.1088/1748-9326/aa8850
- Kang, H. (2013). The Prevention and Handling of the Missing Data. *Korean J. Anesthesiol.* 64, 402–406. doi:10.4097/kjae.2013.64.5.402
- Ke, X., Engblom, M., Yang, H., Brink, A., Lyu, J., Zhang, M., et al. (2022). Prediction and Minimization of NOx Emission in a Circulating Fluidized Bed Combustor: A Comprehensive Mathematical Model for CFB Combustion. *Fuel* 309, 122133. doi:10.1016/j.fuel.2021.122133
- Korpela, T., Kumpulainen, P., Majanne, Y., and Häyrynen, A. (2015). Model Based NOx Emission Monitoring in Natural Gas Fired Hot Water Boilers. *IFAC-PapersOnLine* 48, 385–390. doi:10.1016/j.ifacol.2015.12.409
- Li, G.-Q., Qi, X.-B., Chan, K. C. C., and Chen, B. (2017). Deep Bidirectional Learning Machine for Predicting NOx Emissions and Boiler Efficiency from a Coal-Fired Boiler. *Energy Fuels* 31, 11471–11480. doi:10.1021/acs.energyfuels.7b01415
- Lv, Y., Liu, J., Yang, T., and Zeng, D. (2013). A Novel Least Squares Support Vector Machine Ensemble Model for NOx Emission Prediction of a Coal-Fired Boiler. *Energy* 55, 319–329. doi:10.1016/j.energy.2013.02.062
- Lv, Y., Yang, T., and Liu, J. (2015). An Adaptive Least Squares Support Vector Machine Model with a Novel Update for NOx Emission Prediction. *Chemom. Intelligent Laboratory Syst.* 145, 103–113. doi:10.1016/j.chemolab.2015.04.006
- Miikkulainen, R., Liang, J., Meyerson, E., Rawal, A., Fink, D., Francon, O., et al. (2019). “Chapter 15 - Evolving Deep Neural Networks,” in *Artificial Intelligence in the Age of Neural Networks and Brain Computing*. Editors R. KOZMA, C. ALIPPI, Y. Choe, and F. C. Morabito (Academic Press).
- Ou, D., Tan, K., Lai, J., Jia, X., Wang, X., Chen, Y., et al. (2021). Semi-supervised DNN Regression on Airborne Hyperspectral Imagery for Improved Spatial Soil Properties Prediction. *Geoderma* 385, 114875. doi:10.1016/j.geoderma.2020.114875
- Saif-Ul-Allah, M. W., Qyyum, M. A., Ul-Haq, N., Salman, C. A., and Ahmed, F. (2022). Gated Recurrent Unit Coupled with Projection to Model Plane Imputation for the PM2.5 Prediction for Guangzhou City, China. *Front. Environ. Sci.* 9. doi:10.3389/fenvs.2021.816616
- Shin, S., Lee, Y., Kim, M., Park, J., Lee, S., and Min, K. (2020). Deep Neural Network Model with Bayesian Hyperparameter Optimization for Prediction of NOx at Transient Conditions in a Diesel Engine. *Eng. Appl. Artif. Intell.* 94, 103761. doi:10.1016/j.engappai.2020.103761
- Song, M., Xue, J., Gao, S., Cheng, G., Chen, J., Lu, H., et al. (2022). Prediction of NOx Concentration at SCR Inlet Based on BMIFS-LSTM. *Atmosphere* 13. doi:10.3390/atmos13050686
- Spinelli, I., Scardapane, S., and Uncini, A. (2020). Missing Data Imputation with Adversarially-Trained Graph Convolutional Networks. *Neural Netw.* 129, 249–260. doi:10.1016/j.neunet.2020.06.005
- Tang, S., Zhu, Y., and Yuan, S. (2022). A Novel Adaptive Convolutional Neural Network for Fault Diagnosis of Hydraulic Piston Pump with Acoustic Images. *Adv. Eng. Inf.* 52, 101554. doi:10.1016/j.aei.2022.101554
- Tkachenko, R., Mishchuk, O., Izonin, I., Kryvinska, N., and Stoliarchuk, R. (2019). A Non-iterative Neural-like Framework for Missing Data Imputation. *Procedia Comput. Sci.* 155, 319–326. doi:10.1016/j.procs.2019.08.046
- Vo, N. D., Oh, D. H., Hong, S.-H., Oh, M., and Lee, C.-H. (2019). Combined Approach Using Mathematical Modelling and Artificial Neural Network for Chemical Industries: Steam Methane Reformer. *Appl. Energy* 255, 113809. doi:10.1016/j.apenergy.2019.113809
- Wang, X., Liu, W., Wang, Y., and Yang, G. (2022). A Hybrid NOx Emission Prediction Model Based on CEEMDAN and AM-LSTM. *Fuel* 310, 122486. doi:10.1016/j.fuel.2021.122486
- Yang, G., Wang, Y., and Li, X. (2020). Prediction of the NO Emissions from Thermal Power Plant Using Long-Short Term Memory Neural Network. *Energy* 192, 116597. doi:10.1016/j.energy.2019.116597
- Zhai, Y., Ding, X., Jin, X., and Zhao, L. (2020). Adaptive LSSVM Based Iterative Prediction Method for NOx Concentration Prediction in Coal-Fired Power Plant Considering System Delay. *Appl. Soft Comput.* 89, 106070. doi:10.1016/j.asoc.2020.106070
- Zhang, K., Thé, J., Xie, G., and Yu, H. (2020a). Multi-step Ahead Forecasting of Regional Air Quality Using Spatial-Temporal Deep Neural Networks: A Case Study of Huaihai Economic Zone. *J. Clean. Prod.* 277, 123231. doi:10.1016/j.jclepro.2020.123231
- Zhang, X., and Schreifels, J. (2011). Continuous Emission Monitoring Systems at Power Plants in China: Improving SO2 Emission Measurement. *Energy Policy* 39, 7432–7438. doi:10.1016/j.enpol.2011.09.011
- Zhang, Y., Li, Y., Zhang, X., and Zheng, S. (2020b). Prediction Method of NOx from Power Station Boilers Based on Neural Network. *J. Circuits, Syst. Comput.*, 2150097. doi:10.1142/s0218126621500973

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Saif-Ul-Allah, Khan, Ahmed, Salman, Gillani, Hussain, Yasin, Ul-Haq, Khan, Bazmi, Ahmad and Hasan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.