

Article

An Interpretable Time Series Forecasting Model for Predicting NO_x Emission Concentration in Ferroalloy Electric Arc Furnace Plants

Youngjin Seol¹, Seunghyun Lee¹, Jiho Lee¹, Chang-Wan Kim² , Hyun Su Bak³, Youngchul Byun³ and Janghyeok Yoon^{1,*}

¹ Department of Industrial Engineering, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, Seoul 05029, Republic of Korea; syj6151@konkuk.ac.kr (Y.S.); ant6836@konkuk.ac.kr (S.L.); oscar@neopons.com (J.L.)

² School of Mechanical Engineering, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, Seoul 05029, Republic of Korea; goodant@konkuk.ac.kr

³ Particulate Matter Research Center, Research Institute of Industrial Science and Technology (RIST), Gwangyang 57801, Republic of Korea; hsbak@rist.re.kr (H.S.B.); ycbyun@rist.re.kr (Y.B.)

* Correspondence: janghyoon@konkuk.ac.kr

Abstract: Considering the pivotal role of ferroalloys in the steel industry and the escalating global emphasis on sustainability (e.g., zero emissions and carbon neutrality), the demand for ferroalloys is anticipated to increase. However, the electric arc furnace (EAF) of ferroalloy plants generates substantial amounts of nitrogen oxides (NO_x) because of the high-temperature combustion processes. Despite the substantial contributions of many studies on NO_x prediction from various industrial facilities, there is a lack of studies considering the environmental condition of the EAF in ferroalloy plants. Therefore, this study presents a deep learning model for predicting NO_x emissions from ferroalloy plants and further can provide guidelines for predicting NO_x in industrial sites equipped with electric furnaces. In this study, we collected various historical data from the manufacturing execution system of electric furnaces and exhaust gas systems to develop a prediction model. Additionally, an interpretable artificial intelligence method was employed to track the effects of each variable on the NO_x emissions. The proposed prediction model can provide decision support to reduce NO_x emissions. Furthermore, the interpretation of the model contributes to a better understanding of the factors influencing NO_x emissions and the development of effective strategies for emission reduction in ferroalloys EAF plants.

Keywords: NO_x emission; electric arc furnace; deep learning; explainable artificial intelligence

MSC: 68T05; 68T07; 68T37



Citation: Seol, Y.; Lee, S.; Lee, J.; Kim, C.-W.; Bak, H.S.; Byun, Y.; Yoon, J. An Interpretable Time Series Forecasting Model for Predicting NO_x Emission Concentration in Ferroalloy Electric Arc Furnace Plants. *Mathematics* **2024**, *12*, 878. <https://doi.org/10.3390/math12060878>

Academic Editor: Xiang Li

Received: 15 February 2024

Revised: 11 March 2024

Accepted: 13 March 2024

Published: 16 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The escalating global emphasis on sustainability, such as zero emissions, has led to changes in various industries, including the steel industry [1]. A wide range of policies, strategies, protocols, and interventions related to emission reductions for specific air pollutants have been implemented globally [2]. Particularly, in the Republic of Korea, owing to increasingly stringent environmental regulations, government agencies have installed sensors in stacks for telemonitoring and regulating factories that emit environmental pollutants [3]. Additionally, the demand for ferroalloy, which is an essential raw material for the steel industry, is anticipated to increase not only because of its importance in manufacturing steel but also because of evolving production technologies aimed at reducing emissions [1]. Ferroalloys are iron alloys with a high proportion of one or more elements, such as manganese (Mn), aluminum (Al), and silicon (Si), which enhance the characteristics of steel and cast iron or serve essential functions in the manufacturing process [4]. Although

ferroalloy production using electric arc furnaces (EAFs) results in lower emissions than steel production using blast furnaces, the ferroalloy production process still generates a considerable amount of NO_x emissions, a substantial portion of which originates from EAFs themselves [5]. As NO_x is a major contributor to air pollution and a factor that affects human health [6], there has been an increasing focus on research to predict and reduce NO_x emissions from various facilities [7–9]. However, despite the high demand for ferroalloys and the significant NO_x emissions associated with their production, there is a lack of research on predicting NO_x emissions from EAFs, which are the primary sources of NO_x at production sites. In this study, we developed a deep learning-based time series prediction model to predict NO_x emissions from electric furnaces. For this purpose, data were collected from electric furnace and exhaust gas equipment at a ferroalloy production site. Furthermore, to interpret the deep learning-based prediction model, we used interpretable artificial intelligence techniques to identify the variables that have a significant impact on NO_x prediction in the electric furnace environment.

Predicting NO_x emissions in EAFs can facilitate the reduction in emissions through both pre-management, by adjusting key operating variables in production facilities, and post-management, by enhancing denitrification facilities for efficient NO_x removal from exhaust gases. However, predicting NO_x emissions in EAFs is challenging, owing to the severe internal environment and complex combustion reactions [10,11]. EAFs produce various exhaust gases and particulate matter during high-temperature combustion. Additionally, the inside of the furnace chimney is exposed to a hot, humid environment containing a mixture of various gases. Owing to these conditions, it is difficult to accurately predict the NO_x concentration within EAFs. There are two main approaches to predicting NO_x emissions [9,12]. The first one is the mechanism-based calculation approach, which involves various parameters and empirical formulas for heat transfer, combustion, and turbulence [12–15]. However, this approach to NO_x prediction requires various assumptions and time to simulate the combustion process and predict NO_x emissions, making it challenging to model a combination of various factors inside EAFs [9,12,16].

The second approach is a data-driven method that establishes the relationships between NO_x emissions and output variables based on data [9,10]. Compared with the first approach, this data-driven approach does not need to solve complex equations [12]. In this regard, many studies have employed data-driven methods to predict NO_x emissions by applying a deep belief network (DBN) [9], artificial neural network (ANN) [17], extreme learning machine (ELM) [8], and long short-term memory (LSTM) [7] to various facilities such as coal-fired boilers, cement precalcining kilns, and industrial waste incinerators [7,17,18]. Although these data-based approaches have proven to be effective in predicting NO_x emissions in diverse applications, applying data-based approaches to the EAFs of ferroalloy production facilities has certain limitations. First, these approaches were designed without considering the specific environment of the EAFs in ferroalloy production facilities. EAFs have unique environmental conditions compared to other high-temperature industrial settings, including higher temperatures from electric arcs (commonly reach 2000 °C) and reliance on electrical energy instead of fossil fuels [19,20]. Unlike the constant conditions found in other combustion processes, EAFs involve a dynamic process for producing products. This includes adding raw materials, removing slag, and adjusting alloy compositions even during operation. In addition, the absence of a data collection system to gather the necessary information from ferroalloy production systems presents another challenge. Identifying the data that are essential for the accurate prediction of NO_x emissions from EAFs remains unclear. Therefore, further research is required to develop NO_x emission prediction models that account for the unique conditions within EAFs. Second, many previous studies have employed machine learning-based models for NO_x prediction, focusing only on predictive performance analysis. For the successful collaboration between experts and machine learning technology, the key factor is interpretability [21]. This ensures that behaviors of the model and predictions are understandable to humans, facilitating further application for EAFs. Accordingly, for NO_x prediction and reduction,

it is necessary to analyze the critical factors for NO_x prediction and to understand the behaviors of the predictive models.

This study conducts a data-driven NO_x emission prediction that is suitable for the exhaust gas generation mechanism in the EAFs of ferroalloy production facilities. Additionally, the study employs interpretable artificial intelligence (AI) algorithms to identify variables that contribute significantly to NO_x emissions, thereby proposing an interpretable model that considers the characteristics of EAFs to predict NO_x emissions. Therefore, this study provides guidance for constructing NO_x prediction systems and data collection systems for EAFs and can also be utilized to support the installation and operation of denitrification equipment for NO_x reduction by providing NO_x emissions from EAFs and furnishing data on NO_x emissions from EAFs. It also offers insights into the factors influencing these emissions, facilitating the environmental management and efficient control of NO_x emissions during ferroalloy production.

2. Background

2.1. Operation and Gas Exhaustion Process of the Electric Arc Furnace

EAFs are primarily used to produce ferroalloys. The ferroalloy production process consists of raw material transportation, raw material pretreatment, electric furnace melting, refining, and casting. In EAFs, raw materials such as iron scrap are melted and refined; subsequently, oxidizing slag is produced to remove impurities from the molten pool [20]. An EAF is a sealed structure in which raw materials are charged and electrodes are inserted after sealing with a cover. As illustrated in Figure 1, the melting of the raw materials begins with an arc discharge. Electrical energy is supplied through the graphite electrodes, creating a powerful electric arc between the electrodes and raw materials. This intense electric arc, with its strong voltage, serves as the primary heat source for melting the material. The internal temperature of the furnace is controlled by adjusting the position of the electrodes. During the melting process, EAFs emit thermal NO_x and other gases because of the high temperatures generated, and NO_x emission is predominantly concentrated in this melting process.

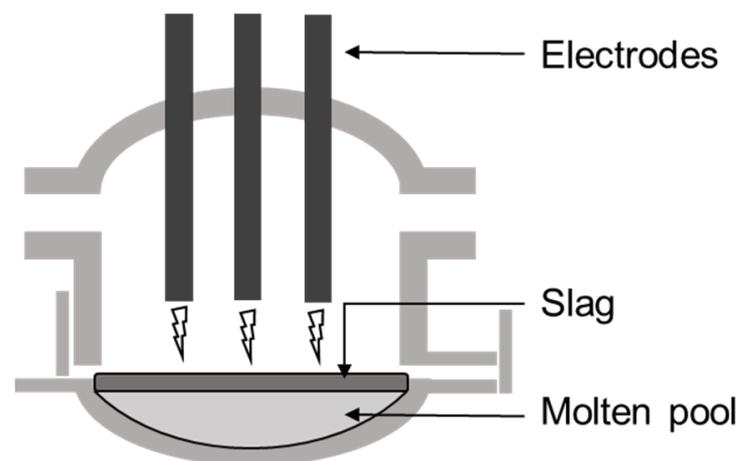


Figure 1. Schematic of an electric arc furnace.

In this study, we collected data from EAFs and exhaust gas emission facilities at a ferroalloy production site in South Korea. Figure 2 illustrates the exhaust gas emissions generated in the EAFs considered in this study. By melting the raw materials, EAFs generate a significant amount of thermal NO_x. Subsequently, the dust duct captures and collects the exhaust gases and fine particulates generated in the preceding processes. The exhaust gases are then directed to a semi-dry reactor (SDR), where water is injected to control the temperature of the exhaust gas. In the SDR, further treatment, such as desulfurization, occurs to remove additional pollutants from the exhaust before release.

The role of SDRs extends beyond temperature control to actively reduce the concentrations of various harmful substances in exhaust gas. Bag filters remove particulate pollutants (e.g., dust) but not gaseous pollutants (e.g., NO_x and SO_x). This phase primarily focuses on eliminating fine particulates from the gas stream. An induced draft fan (IDF) is operated to expel gases from outside the stack using pressure. The IDF creates a suction effect that ensures the efficient and effective discharge of gases, thereby minimizing the emission of untreated or partially treated exhaust gases into the environment.

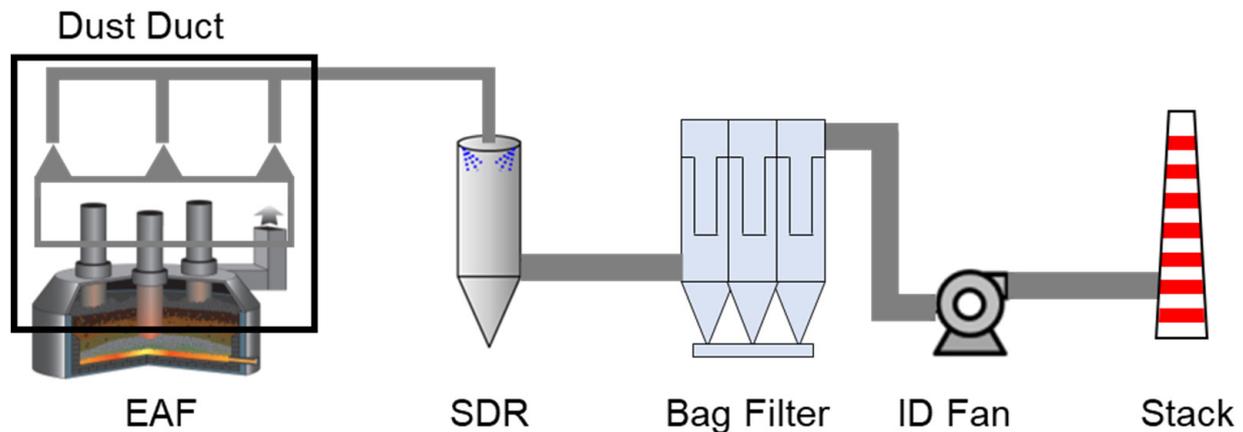


Figure 2. Schematic of the NO_x emission process from the electric arc furnace to the chimney.

2.2. Data-Driven NO_x Emissions Prediction Research

Owing to the complex mechanism of NO_x emissions from facilities involving high-temperature processes, such as coal-fired power plants, research has been conducted on predicting NO_x emissions from exhaust gases (Table 1). Research has been conducted to predict NO_x emissions by utilizing computational fluid dynamics (CFD) simulations to generate data on flow, temperature, and chemical reactions within the furnace. Faravelli et al. [22] proposed a method for predicting NO_x emissions from gas/oil boilers by utilizing CFD to obtain data on flow, temperature, and stoichiometry within the furnace. They simplified the conditions with an ideal reactor network which is interconnected and perfectly stirred or plug flow reactors to predict NO_x emissions using a detailed kinetic scheme. Likewise, Lv et al. [23] utilized CFD simulations to generate 3D NO_x spatial distribution data and applied extreme learning machine modeling for accurate predictions of NO_x distribution in the furnace. This study partitioned data based on NO_x generation mechanisms for enhancing model accuracy and provided a detailed approach for NO_x prediction in furnace environments. However, a potential limitation of these studies is the requirement for fluid dynamics experts to effectively use CFD, requiring expertise in handling diverse parameters and empirical formulas for heat transfer, combustion, and turbulence specific to each facility's environmental conditions. This complexity can pose challenges for implementation in areas with limited research, such as EAFs, due to the variability in environmental conditions across different facilities.

Compared to the challenges associated with CFD studies, research utilizing data-driven methods has been conducted to establish the relationship between operational variables and NO_x generation, thereby enabling the prediction of NO_x emissions from various facilities with less difficulty. Wang, Ma, Wang, Li, and Zhang [9] proposed a method for data acquisition and NO_x emission prediction in coal-fired power plants using DBN-based models utilizing historical operating data. Tang, Wang, Chai, Cao, Ouyang, and Li [8] proposed an autoencoder ELM model to predict NO_x emission concentrations from a coal-fired boiler. In their study, an autoencoder was utilized to extract hidden features from the variables of operational data, and an ELM model was then applied to predict NO_x emissions from the hidden features. Zhang, Wang, Shao, Duan, and Hou [17] utilized an ANN to predict NO_x in cement precalcining kilns and a genetic algorithm to

search for optimal operation parameters to achieve the lowest concentration of nitrogen oxide emissions.

Table 1. Prior NOx prediction studies in facilities with combustion process.

Facility	Data	Prediction Method	Reference
Gas/oil-fired boiler	Fluent-based simulation data	Computer fluid dynamics, ideal reactor network	Faravelli, Bua, Frassoldati, Antifora, Tognotti, and Ranzi [22]
Coal-fired boiler	Fluent-based simulation data	Computer fluid dynamics, extreme learning machine	Lv, Zhao, Cao, and Shen [23]
Coal-fired boiler	Historical operation data, fluent-based simulation data, and experimental data	Deep belief network	Wang, Ma, Wang, Li, and Zhang [9]
Coal-fired boiler	Historical operation data	Auto-encoder, extreme learning machine	Tang, Wang, Chai, Cao, Ouyang, and Li [8]
Cement precalcining kiln	Historical operation data	Artificial neural network	Zhang, Wang, Shao, Duan, and Hou [17]
Coal-fired boiler	Historical operation data	Recurrent neural network	Safdarnejad, Tuttle, and Powell [24]
Coal-fired boiler	Historical operation data	Long short-term memory network	Yang, Wang, and Li [7]
Diesel engine	World harmonized transient cycle (WHTC) emission test data	Convolutional neural network, long short-term memory network	Shen, Wang, Wang, Zeng, Yu, and He [25]
Coal-fired boiler	Historical operation data	Random forest algorithm, lightweight convolutional neural network	Wang, Peng, Cao, Zhou, Fan, Li, and Huang [12]
Coal-fired boiler	Historical operation data	Convolutional neural networks, channel Attention mechanism	Li et al. [26]

However, previous studies utilizing ANN-based models have certain limitations. This is because they do not utilize the temporal dynamics of the operating variables in facilities, which can contribute to NOx emissions. As the combustion of an EAF is a dynamic process, EAFs' working conditions are correlated with historical NOx emissions. Given that manufacturing execution systems (MESs) and telemonitoring systems (TMSs) store dynamic time series data, previous time series data can be leveraged to develop prediction models. Safdarnejad et al. [24] developed a dynamic data-driven model for a coal-fired utility boiler to estimate NOx and CO emissions simultaneously, utilizing recurrent neural networks to capture time series characteristics of the data. Yang, Wang, and Li [7] focused on using LSTM networks to model the relationship between the operational parameters and NOx emissions in a 660 MW boiler. To enhance NOx emissions prediction in diesel engine transient environments, Shen et al. [25] proposed a prediction model based on a hybrid neural network architecture that combines the feature extraction capabilities of a convolutional neural network (CNN) with the time series prediction proficiency of LSTM networks. In addition to models considering the temporal dynamics of the operating variables in facilities, research has been conducted to modify the characteristics and purposes of prediction in facilities or enhance the performance of existing models. To improve the efficiency of the denitrification process in power plants, Wang, Peng, Cao, Zhou, Fan, Li, and Huang [12] proposed a modeling method using a random forest algorithm for the dimensionality reduction in input data and a lightweight CNN. In their study, satisfactory NOx predictive performance was obtained. A lightweight CNN is preferred over a high-performance CNN, which requires numerous parameters and floating-point operations. Lightweight CNN could offer the advantage of efficient computation and reduced complexity, making them more suitable for real-time NOx emission prediction tasks in coal-fired boilers. Li et al.'s [25] study presents a CNN-based model for the

accurate prediction of NO_x emissions from a coal-fired power plant boiler. An attention mechanism was integrated into the CNN-based model, with the attention module focusing on the interdependencies between channels in the input feature maps to capture important information in latent space.

Though previous studies have proposed data-driven NO_x prediction methods for facilities with combustion systems, research on predicting NO_x emissions from EAFs is still lacking. Consequently, data-acquisition systems tailored for NO_x prediction in EAF environments are lacking. An EAF generates extremely high temperatures to melt raw materials, and owing to the characteristics of the molten pool during ferroalloy production, noise is generated when measuring the exhaust gases emitted during ferroalloy production. As the gas trapped beneath the slag layer in the molten pool and the collapse of charged raw materials can lead to sudden explosions and a rapid increase in NO_x emissions, it is necessary to smooth the NO_x emission values before utilizing them as training data for the prediction model. In addition, exhaust gases in ferroalloy production facilities motion at high speeds in hot and humid environments. In such an environment, data collected by IoT sensors in a pipe may contain noise, owing to various factors. Thus, data preprocessing techniques are required to construct training data for the prediction mode by smoothing the noise. To smooth out noise or outliers, a Kalman filter is used to estimate the current state from past measurements and correct outlier data based on the distribution of the given data.

2.3. Interpretable Prediction Models

Despite the contributions of previous studies to the prediction of NO_x emissions in various combustion processes, an interesting yet unexplored angle still exists. In the case of deep-learning-based prediction models, numerous studies have focused on performance analysis, making it difficult to track the impact of input variables on NO_x emissions. As deep learning-based predictions rely solely on black-box models with undisclosed internal mechanisms, experts in decision making have experienced challenges in utilizing these predictive models [27].

Interpretable artificial intelligence methods are processes that provide interpretability in a form understandable by humans, based on the explainability of how a model works [28]. They can be classified based on the complexity of the model into post hoc and intrinsic approaches [29]. The intrinsic approach involves models that are naturally interpretable due to their simple structure (e.g., decision trees, linear SVMs). On the other hand, the post hoc approach is applied after the model has been trained, focusing on the analysis and interpretation of the model's behavior. LIME (local interpretable model-agnostic explanations) and SHAP (SHapley Additive exPlanations) are well-known methods, offering insights into how the model makes its predictions. Both approaches are model-agnostic and can be utilized across various models. LIME focuses on local explanations, offering insights into the interpretation process for specific data points, but it has limitations in providing global interpretations and consistency in the contribution of input variables [30]. SHAP similarly allows for an understanding of individual contributions to predictions across the entire dataset, but this approach can offer a broader analysis of model predictions, such as feature importance [31]. Consequently, it has been utilized across various domains for its comprehensive insights into model behavior [32,33].

Considering the ability to provide model-agnostic interpretations and both global and local explanations [21,31], therefore, this study utilizes SHapley Additive exPlanations (SHAP) to uncover the inner workings of a machine learning model for time series data to predict NO_x emissions from EAFs. This study constructs a model that reflects the relationships between input variables over time and employs preprocessing techniques specific to the features of EAFs to build the training dataset. Additionally, interpretable AI is utilized to analyze the impact of the input variables on NO_x emission predictions.

3. Methodology

3.1. Kalman Filter-Based Smoothing Algorithm

Owing to the extreme environment in the chimney, the NOx data collected by the sensors often contain noise. To address this issue, a Kalman-filter-based smoothing algorithm is introduced to mitigate sensor noise, remove outliers, and enhance the quality of the collected data to train the prediction model [34]. Kalman filtering is a method for estimating the state of a dynamic system [35,36]. It predicts the next state based on the current state and subsequently updates the predicted state using new measurements. The mathematical model can be expressed as follows:

X_k is the state vector representing the system's state at time k .

Y_k is the measurement at time k .

Q is the process noise variance.

R is the measurement variance.

P_k is the error covariance matrix at time k .

K_k is the Kalman gain at time k .

The Kalman gain adjusts the confidence between the current prediction and observed data, thereby determining the optimal state correction. Therefore, a higher Kalman gain value places more trust in the observed data and less emphasis on prediction, allowing the Kalman filter to estimate and predict the system state more accurately. The state variables of the system are estimated using measured data. The measurement data sequence is used as the input to estimate the state of the system, and the Kalman filter-based smoothing algorithm is performed as follows (Algorithm 1):

$$\hat{X}_{k+1|k} = X_k \quad (1)$$

$$P_{k+1|k} = P_k + Q \quad (2)$$

$$e_k = Y_{k+1} - \hat{X}_{k+1|k} \quad (3)$$

$$K_k = \frac{P_{k+1|k}}{P_{k+1|k} + R} \quad (4)$$

$$X_{k+1} = \hat{X}_{k+1|k} + K_k \cdot e_k \quad (5)$$

$$P_{k+1|k} = (I - K_k) \cdot P_{k+1|k} \quad (6)$$

Algorithm 1. Kalman filter-based smoothing algorithm

Input: X_k, Y_{k+1}, P_k

Output: X_{k+1}, P_{k+1}

Prediction:

(a) State prediction (Equation (1)).

(b) Error covariance prediction (Equation (2)).

Update:

(c) Innovation (Equation (3))

(d) Kalman gain (Equation (4))

(e) State update (Equation (5))

(f) Error covariance update (Equation (6))

3.2. NOx Emission Prediction

3.2.1. Long Short-Term Memory Network

NOx emissions in EAFs represent a time series issue because of the relationship between the past operating conditions and the current state. Since NOx emissions during combustion in an EAF is a non-linear and complex process [15], variables and temporal factors should be considered. Given the time series nature of NOx emissions in EAFs, an LSTM neural network-based model is adopted (Figure 3). Owing to the ability of the LSTM network to remember long-term dependencies, it can capture patterns in emission data [7],

and it has been increasingly utilized in various time series prediction domains [7,34,37]. LSTM networks employ a unique architecture that uses structures known as gates to regulate a value called the cell state (C). The cell state acts as the memory for the network, which is crucial for retaining and carrying relevant information throughout the data sequence. The ability of LSTM networks to use gates to regulate cell states is crucial, and this mechanism allows the network to selectively retain or discard information, thereby enhancing its efficiency in analyzing time series data. The forget gate in LSTM networks uses a sigmoid function to assess previous outputs (h_{t-1}) and current inputs (i_t), determining which past information to retain or discard from the cell state. The input gate in the LSTM network updates cell state C_t . It employs a sigmoid function to identify which elements of the current input are significant and identifies a tangent hyperbolic function to generate a vector of new candidate values, \tilde{C}_t . These elements are integrated to update C_t using essential new information. The output gate determines the final output h_t by filtering the cell state \tilde{C}_t through a tangent hyperbolic function and then multiplying it by the output of the sigmoid function. This selectively updates h_t with the relevant information from C_t . Each gate in the LSTM network operates according to the following formula:

$$\text{The forget gate : } f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{7}$$

$$\text{The input gate : } i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{8}$$

$$\text{Alongside } \tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{9}$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tilde{C}_t \tag{10}$$

$$\text{The output gate : } o_t = \sigma(W_O \cdot [h_{t-1}, x_t] + b_o) \tag{11}$$

$$h_t = o_t \otimes \tanh(C_t) \tag{12}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{13}$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{14}$$

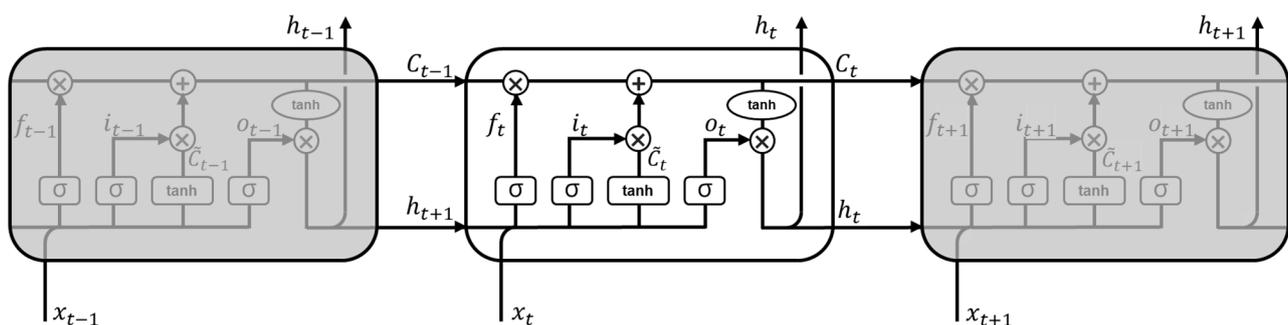


Figure 3. Structure of LSTM.

C_t denotes the state of the LSTM cells at time t , and h_t denotes the output of the unit at time t . W denotes the weight parameter metrics. f_t , i_t , and o_t denote the forget, input, and output gates and state vector at time t . \otimes represents element-wise multiplication. When applied to EAFs, utilizing their strengths in learning the sequence of features [7,25,34,37], the LSTM network can offer advantages in improving the accuracy of predicting NOx emissions typically associated with the operations of these furnaces.

3.2.2. Delay Time Determination

At Korean ferroalloy production sites, TMSs are commonly used to measure the NOx concentrations in EAFs [3]. Throughout the processes of NOx generation, detection, and control, numerous parameters are monitored using the MES, as depicted in Figure 2. However, these parameters are not measured simultaneously by the sensors, which leads to inherent delays in data acquisition. In addition, the combustion processes in EAFs involve complex reactions that occur over time and can influence NOx emissions. Given that changes in the variables within the process do not immediately affect the NOx emissions, it is necessary to select an appropriate delay time between the variables. This helps determine the suitable length of the sequence to be input into the LSTM model, which is suitable for processing and predicting events with intervals and delays in a time series [38]. The delay time selection method based on mutual information (MI) focuses on identifying the most effective sequence length from the operation variables to predict NOx emissions. Tang, Wang, Chai, Cao, Ouyang, and Li [8] determined the delay time between each feature and NOx emission concentration using the MI method. This is achieved by maximizing the combined MI between the input features and target variable. MI is an information theory measure that quantifies the amount of information obtained from one random variable by observing another [39]. MI is frequently used to evaluate the dependence or correlation between variables, capturing insights that traditional regression analyses may not reveal. Here, MI serves as a metric for measuring the extent to which one variable informs another, thereby indicating their level of interdependence. $X = [x_1, x_2, \dots, x_n]$, and n is the number of samples in dataset X . $H(X)$ represents the information entropy of random variable x . The probability distribution of x_i is $p(x_i)$. $H(X, Y)$ is the joint entropy of X and Y . The probability density functions of x and y are $p(x)$ and $p(y)$. The degree of correlation between the two random variables can be expressed by the MI as follows [40]:

$$H(X) = -\sum_{i=1}^n p(x_i) \log p(x_i) \quad (15)$$

$$I(X; Y) = H(X) + H(Y) - H(X, Y) \quad (16)$$

$$I(x; y) = \sum_{xy} \sum_{x \in X} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \quad (17)$$

To determine the delay time, it is varied starting from one step, and the time step that yields the highest MI is selected. By analyzing the MI between the variables and the NOx emission concentration, it is possible to determine the maximum feasible delay time for all input variables.

3.2.3. NOx Emission Prediction Model Development

This study develops a model to predict future NOx emissions using a sequence of data comprising 19 variables, including NOx emissions. To capture the trend of previously emitted NOx levels, NOx emissions are utilized as predictive variables. The performance of the NOx prediction model is assessed using quantitative performance evaluation metrics. The mean absolute percentage error (MAPE) measures the average percentage error between predicted and actual values. The R-squared (R^2) score, or the coefficient of determination, indicates how well the predicted values fit the actual data, with a score of 1 representing a perfect fit. The mean squared error (MSE) quantifies the average of the squares of errors and measures the variance of the prediction errors. The mean absolute error (MAE) measures the average magnitude of errors between the predicted and actual values without considering direction.

$$MAPE = \frac{100}{n} \sum_{i=1}^n \frac{y_i - \hat{y}_i}{\hat{y}_i} \quad (18)$$

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (19)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (20)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (21)$$

3.3. Interpretation of the NOx Emissions Prediction

Although machine-learning-based models have been adopted in various domains, their black-box nature, which enables powerful predictions, presents a key impediment in that AI-based systems often lack interpretability and need interpretable machine learning [27]. To address the lack of interpretability of complex and nonlinear machine-learning-based models, the post hoc interpretation method employs a model-agnostic method to explain how certain features contribute to predictions and the model's behavior [21]. Among the various interpretation methods, SHAP is a widely used framework for interpreting the predictions of machine learning models based on the Shapley value of the conditional expectation of a model [41,42].

SHAP evaluates the feature importance using additive feature attribution methods, as illustrated in Equation (22).

$$g(z') = \phi_0 + \sum_{i=1}^M \phi_i z'_i \quad (22)$$

Let f be the original predictive model to be explained and g be the explanation model. Where $z' \in \{0, 1\}^M$ is a coalition vector that indicates whether the i th feature is present (=1) or absent (=0), M is the number of features, $\phi_i \in \mathbb{R}$ is the importance value of the i th feature, and ϕ_0 is the baseline outcome without any feature. Specifically, SHAP identifies the importance of each feature as a change in the expected model prediction when conditioning on that feature and explains how to change from the base value $E[f(z)]$ to the current output $f(x)$. SHAP averages the ϕ_i values across all possible ordering. Hence, when defining $f_x(S) = E[f_x|x_s]$ for a subset of features (S), the SHAP value (ϕ_i) is expressed as in Equation (23).

$$\phi_i = \sum_{S \subseteq \{x_1, \dots, x_m\} \setminus \{x_i\}} \frac{|S|!(M - |S| - 1)!}{M!} (f_x(S \cup \{x_i\}) - f_x(S)) \quad (23)$$

where $f_x(S \cup \{x_i\})$ and $f_x(S)$ are the model prediction with and without the i th feature. SHAP is an additive feature attribution method when ϕ_0 equals $f_x(\emptyset)$, representing the baseline prediction with no features. The original model's prediction for each sample is equal to the sum of all the feature SHAP values. Thus, the SHAP values indicate the contribution of each feature to the predictions of the model.

Calculating the precise SHAP value poses a challenge due to the necessity of evaluating each potential feature subset, resulting in exponential computational complexity [21]. Therefore, we utilized deep SHAP, a method that aggregates SHAP values calculated for individual network components to derive SHAP values for the entire network [42,43]. Using deep SHAP, we obtained the SHAP values for each feature. The absolute SHAP value of the i th feature for the j th time-steps is expressed as in $|\phi_{i,j}|$, and the SHAP value of the i th feature ϕ_i is the average of $\phi_{i,j}$.

4. Analyses and Results

4.1. Data Preparation

Data were collected from the MES and TMS of EAFs, and 18,834 data points were collected from 1 May 2023 to 7 July 2023. Among them, 17,422 data points were used for training and 1412 for validation; 1412 data points were collected from 12 July to 17 July for the test dataset. Given that the model based on deep learning demands numerous variables and substantial data, long-term observation and data collection are essential. However, due to the nature of the data collected by sensors, there can be gaps, and there may be times when data are not collected due to operational schedules. Therefore, for research purposes, it is crucial to collect long-term data without gaps across many variables. To

build the training dataset for model learning, the NO_x emission measurement data were smoothed using a Kalman filter-based smoothing algorithm. Increasing the value of the measurement variance (R) gives more weight to noise in the observed data. In this case, the Kalman filter is less influenced by the predicted values, and the resulting graph is smoother, following the volatility of the observed data more closely. In this study, R was set as 10^2 . Increasing the value of the process noise variance (Q) results in greater uncertainty in the system. In this case, the Kalman filter considered the predictions to be more uncertain. Consequently, the graph maintains higher volatility in the predictions. In this study, Q was set as 5^2 . The initial error covariance matrix (P) affected the initial state prediction of the Kalman filter. Increasing the initial error covariance matrix (P) value increases the uncertainty of the initial prediction, resulting in a larger initial prediction error. In this study, hyperparameters R, P, and Q were selected as trials and errors. We set the sum of the R and Q so as not to exceed the actual variance of NO_x, which is 17.75^2 . To smooth the fluctuations in the graph, R was maintained larger than Q. P was set to 0.7, based on the initial measurements' difference, which was approximately 1.3. The smoothed NO_x data are shown in Figure 4. We selectively illustrated key examples, as visualizing all data points would obscure this effect.

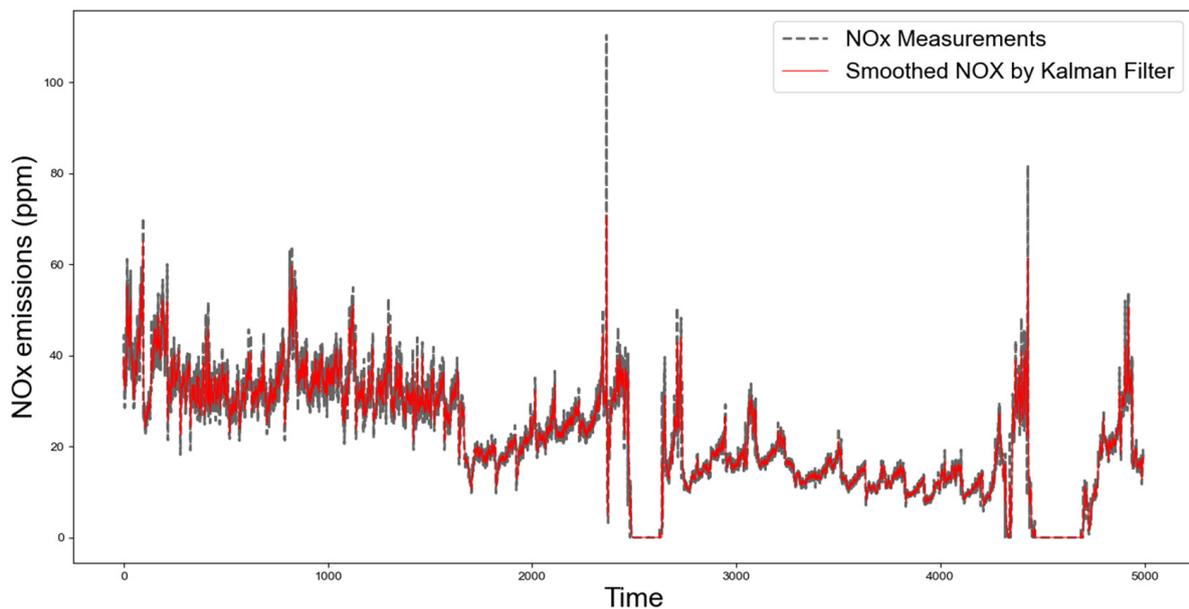


Figure 4. The denoised results of the NO_x data using the Kalman filter-based smoothing algorithm.

4.2. NO_x Emission Prediction

A comparison of the MI between NO_x emissions and variables was conducted to determine the appropriate sequence length for the prediction model input. As shown in Table 2, each variable had a range of delay times from 1 to 6 steps, and each step was 5 min. This design, resulting in a maximum delay time of 30 min, was influenced by regulatory standards mandating emissions monitoring over 30 min intervals in Korea.

To capture the changes in each variable over time, we selected six steps as inputs for the prediction model. The number of units was chosen from the range [64, 128, 256, 512], and the numbers of LSTM and dense layers were varied to identify the optimal number of units that yielded the highest performance. The output of an LSTM layer is a high-dimensional feature vector that cannot be directly used to predict a single NO_x emission value. Therefore, a dense layer was employed, wherein each input node was connected to every output node. This setup transformed the LSTM layer's output into a single, predictive NO_x emission value. After analyzing the performance evaluation metrics in the pilot experiments, two LSTM layers and one dense layer were used (Figure 5), and the optimal units for each layer were determined as follows: LSTM1 (128), LSTM2 (64), and

dense layer (64). As illustrated in Figure 6, the red line representing the predicted values from the model closely followed the dotted line representing the actual NOx emissions. The alignment of these two lines suggests that the model can effectively predict NOx emissions. Figure 7 shows a scatter plot of the prediction models from the test data, where each dot represents an individual prediction against the actual value. The linear fit line indicates the trajectory of the predicted value, and the perfect prediction line in dashed red represents the ideal points at which the predicted values would match the actual values. The 95% prediction band indicates the area in which 95% of the predicted values lie, thus demonstrating the consistency of the model. A narrow 95% prediction band signifies concentrated, accurate predictions within the confidence interval, reflecting a model’s consistent output. Conversely, a wide band indicates greater uncertainty and dispersed predictions. The performance of the prediction model is shown in Table 3, and a comparison analysis was conducted to observe the effects of the presence of previous NOx emissions and temporal factors. This analysis revealed that incorporating the previous NOx emissions and temporal factors yielded better results, as reflected by the improved performance metrics. The ‘Model without NOx’ did not utilize the previous NOx emissions values, indicating that incorporating past NOx emissions data is indeed valuable. The ‘Model with only NOx’ showed satisfactory performance. It seems that including NOx as a feature is crucial. ‘Linear Regression’, ‘Deep Neural Network (DNN)’, ‘Gradient Boosting Regression’, and ‘Random Forest Regression’ employed the same variables as the proposed model. However, due to the nature of their models, they did not incorporate the temporal aspect. This table demonstrates the effectiveness of using a model capable of reflecting temporal elements and leveraging previous NOx emissions data.

Table 2. Highest MI according to the delay time of each feature.

Data Description	Mutual Information	Delay Time (5 min)	Data Description	Mutual Information	Delay Time (5 min)
Electrode Depth-A	0.2646	3	Dust Duct Temperature-A	0.8403	1
Electrode Depth-B	0.2283	1	Dust Duct Temperature-B	0.7489	1
Electrode Depth-C	0.2538	1	Dust Duct Temperature-C	0.7651	1
Electrode Supply Water Flow	0.1457	4	Semi Dry Reactor Inlet Temperature	0.7935	1
Press Down Elevation-A	0.2590	3	Semi Dry Reactor Outlet Temperature	0.5817	2
Press Down Elevation-B	0.2992	3	Bag Filter Inlet Pressure	0.2683	6
Press Down Elevation-C	0.3053	3	Bag Filter Differential Pressure	0.2290	6
Power Use	0.5097	3	Induced Draft Fan Inlet Pressure	0.5280	1
Shell Cooling Water Supply Flow	0.2129	2	Induced Draft Fan Power	0.5765	2

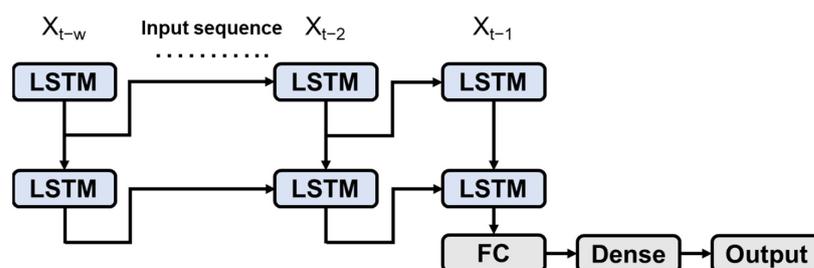


Figure 5. Schematic of the proposed NOx emissions prediction model.

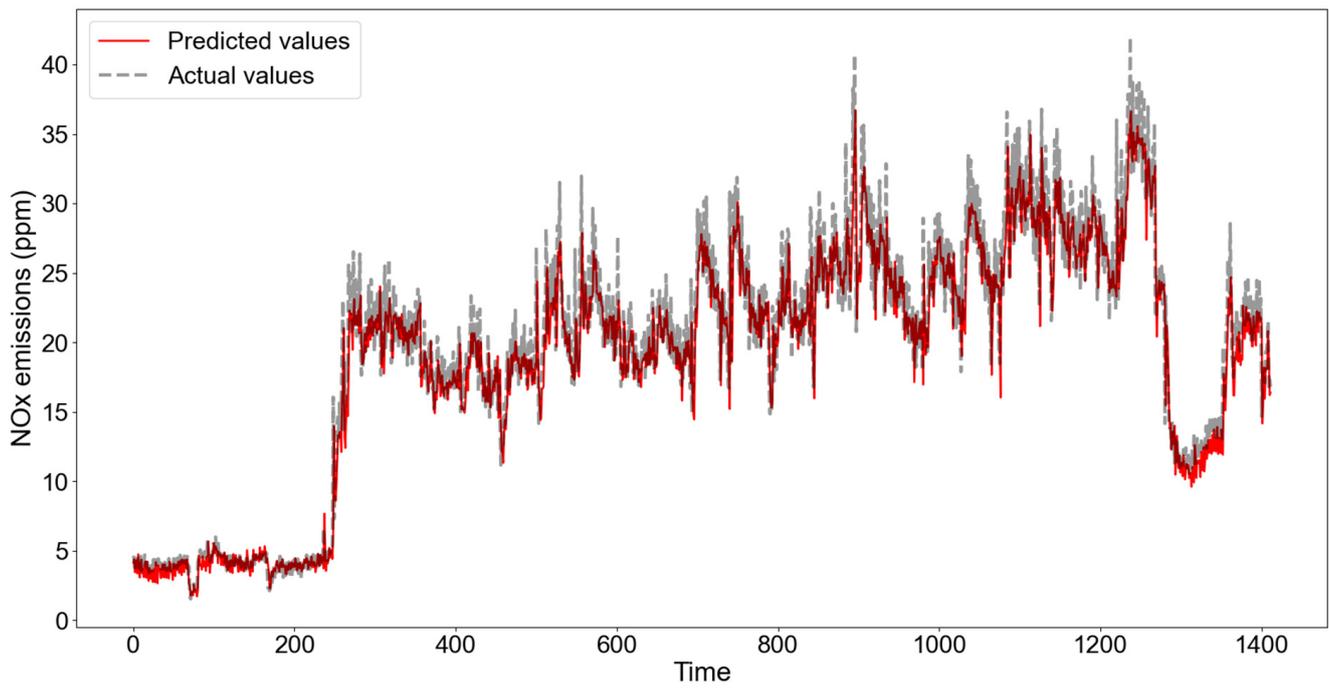


Figure 6. Comparison of predicted and actual NOx emission.

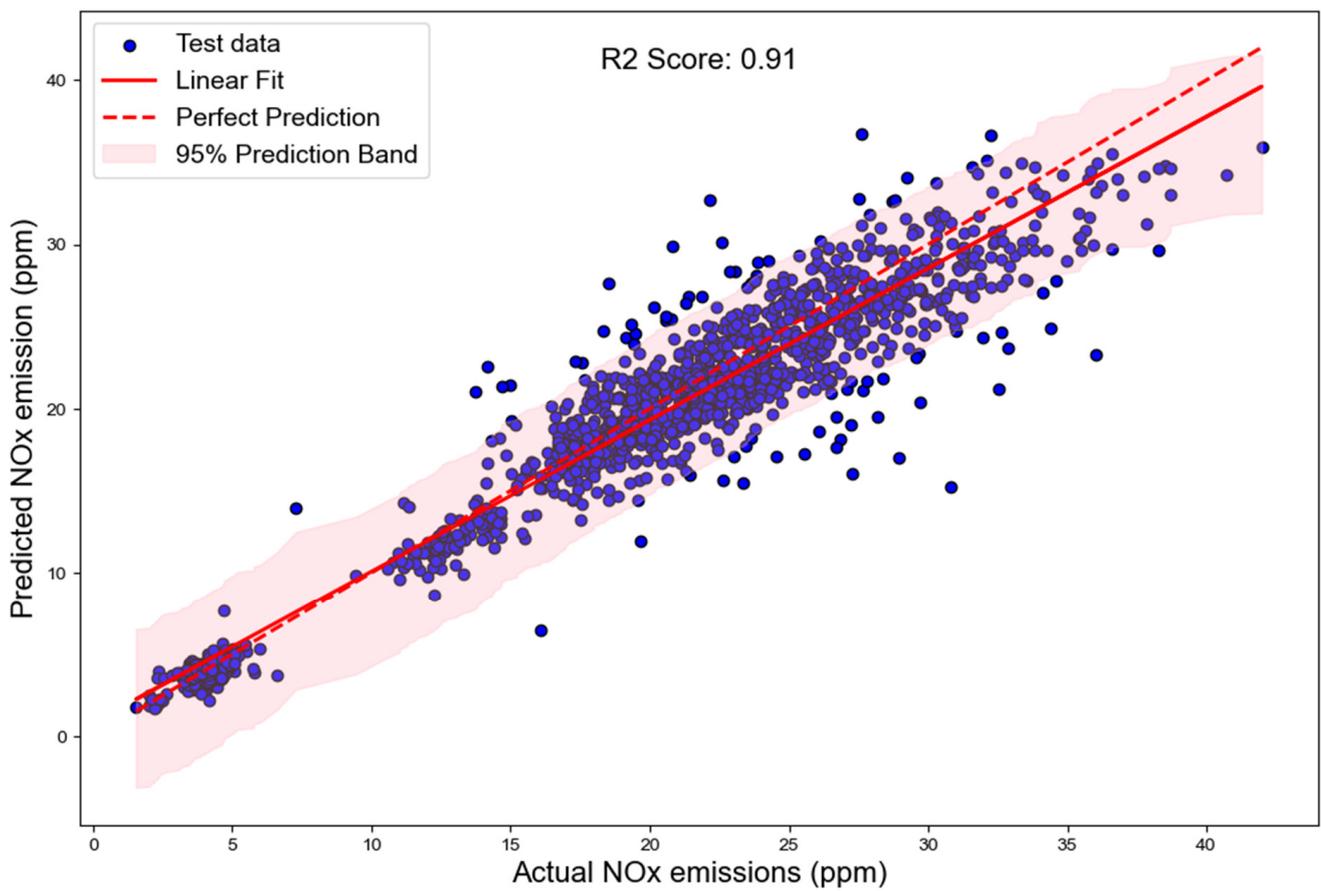


Figure 7. Scatter plots of the prediction model on the test set.

Table 3. Impacts of incorporating temporal factors or previous time-step NOx emissions on the performance (bold: indicates the best model).

Model	MAPE	R ²	MAE	MSE
Proposed Model (NOx)	9.4506	0.9145	1.7823	6.4525
Model without NOx	42.6532	0.5859	4.8642	31.2548
Model with only NOx	12.7176	0.5655	2.3918	10.1512
Linear Regression	14.5906	0.8841	2.2690	8.7988
Deep Neural Network (DNN)	14.1423	0.8751	2.3234	9.4811
Gradient Boosting Regression	19.4197	0.7903	3.1028	15.9172
Random Forest Regression	16.0896	0.8777	2.3040	9.2833

4.3. Interpretation of the NOx Emission Prediction

The SHAP algorithm was applied to the constructed model to calculate the importance of each variable over time. Specifically, SHAP assigns an importance value to each feature for each prediction, based on additive feature attribution methods that comply with a set of variables. In the test dataset, 1000 data points were randomly selected to derive SHAP values. The results of SHAP analysis provide information on how variables influence the model's predictions but do not directly indicate causality. Therefore, it is important to be aware of this limitation when interpreting the results obtained from SHAP analyses. The average absolute SHAP values for each variable were calculated and plotted to visually represent the impact of these variables on the NOx prediction at different time points (Figure 8). The purpose of the various colors is to distinguish between variables, and therefore, colors are unrelated to whether something is worse or better. Based on the SHAP analysis, the temperature measured in the dust duct and SDR before passing through the SDR device and the NOx emissions at the previous time-steps contributed to the predictions. In the semi-dry reactor (SDR), there is a process where liquid is sprayed into the exhaust gas to lower its temperature. Indeed, a noticeable difference in the area between the SDR inlet temperature and the SDR outlet temperature can be observed. This suggests that the contributions of the semi-dry reactor inlet temperature and the dust duct temperature, which are related to the temperature of the exhaust gas before passing through the semi-dry reactor, may be linked to the actual NOx emissions.

Summary and bar plots were employed to illustrate how the input features contributed to the predicted output values (Figures 9 and 10). Summary plots allow us to understand the global trend of the SHAP values of a feature. Specifically, the summary plots show the distribution of SHAP values for each feature. Each point represents the SHAP value of the feature for an individual prediction. Points moving to the right indicate a positive impact on the model output, whereas points to the left indicate a negative impact. Red points represent "high" NOx emissions, whereas blue points represent "low" NOx emissions. The bar plots represent the importance of the features; their importance decreases from top to bottom. Figure 8 depicts how different variables affect NOx predictions across time. Figure 9 examines the variables' impact on lower NOx emission data points, whereas Figure 10 focuses on higher emission points. Thus, while Figure 8 offers a global view of variable impacts over time, Figures 9 and 10 provide more local insights into their effects at particular emission levels. Figures 9 and 10 show the average SHAP values for each feature in the bar graph. The bar lengths indicate the importance of the features, with longer bars indicating more important features. To derive the SHAP values for both low and high NOx emission levels, we selected 200 data points for each category from the test dataset. The first 200 data points were designated to represent the low NOx emission level, while data points from the 1000th to the 1200th position were chosen to represent the high NOx emission level. Figure 9a shows the summary plots, and Figure 9b shows the bar plots when the NOx emissions are low. Features such as the induced draft fan power, induced draft fan inlet pressure, and bag filter differential pressure were identified as important features when NOx emissions were low. Figure 9b illustrates the variables with high contributions at points of low NOx emissions. The exhaust gas facilities maintain pressure to discharge

exhaust gas outside the chimney. Observing that variables such as induced draft fan power, induced draft fan inlet pressure, and bag filter differential pressure have high contributions, it is apparent that at points of low NOx emissions, the internal pressure of the exhaust gas facilities has a greater influence than temperature or operational variables. Figure 10a shows the summary plots, and Figure 10b shows the bar plots when the NOx emissions are high. Features such as exhaust NOx, bag filter differential pressure, and bag filter inlet pressure were identified as important when the NOx emissions were high. As can be seen in Figure 10b, the NOx emissions from previous points have a very high impact. Therefore, it can be inferred that there is some inertia effect with the emissions at a particular level. Figure 11 represents the actual temperature of the dust duct at data points where NOx levels are low and high. In Figure 11, when comparing the temperature in the dust duct, which collects gases emitted from the electric arc furnace (EAF), across two segments, it is observed that there is about a twofold difference. Considering both Figures 8 and 11, they suggest a possible correlation between NOx emissions and the temperature in the exhaust gas system. However, the direct comparison of the SHAP value between low and high emission levels may not entirely reflect an equal analysis due to the dataset containing a higher number of samples at high NOx emission levels. Despite the dataset’s imbalance, the figures reveal the relationships between features and the target, offering insights into the variables’ impacts on low or high NOx emissions.

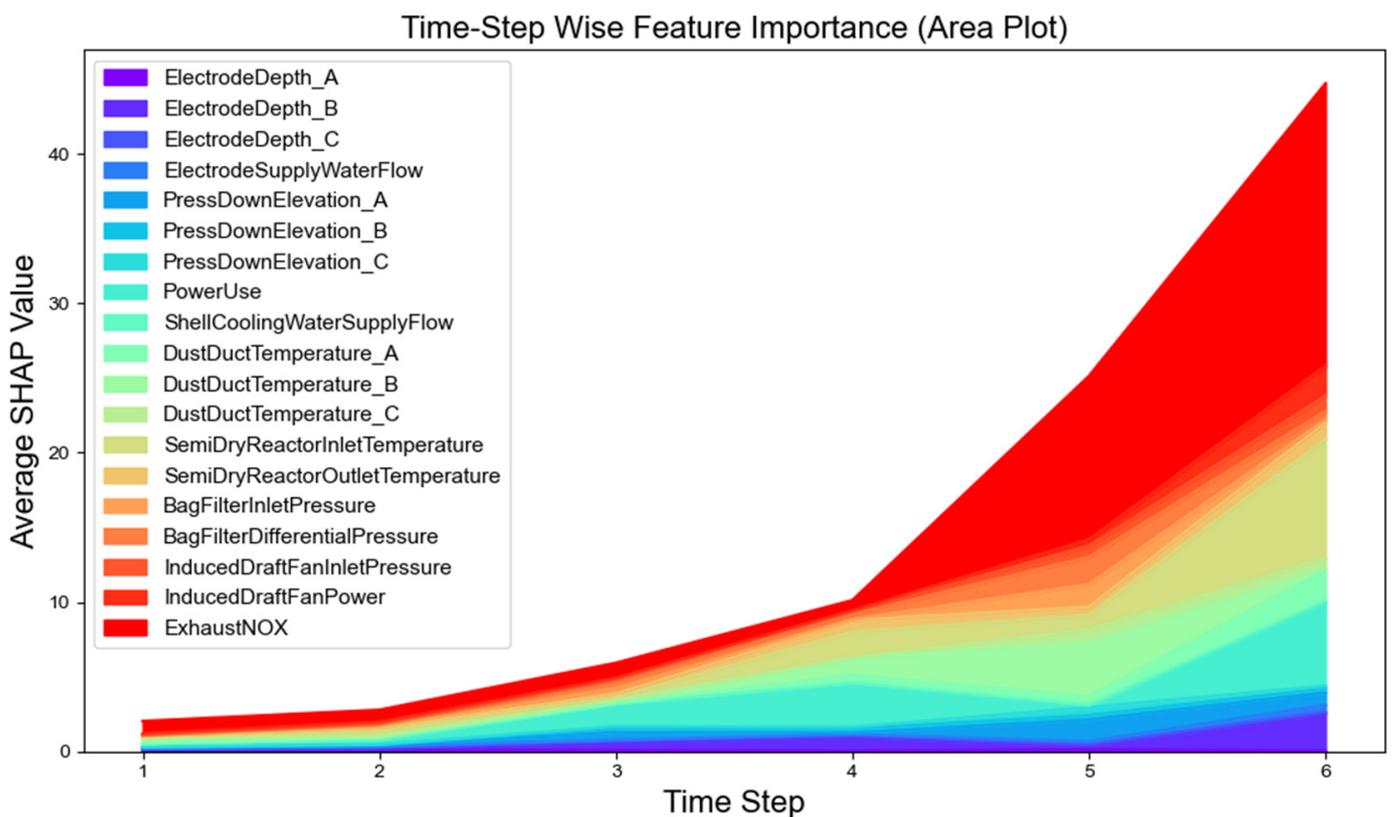
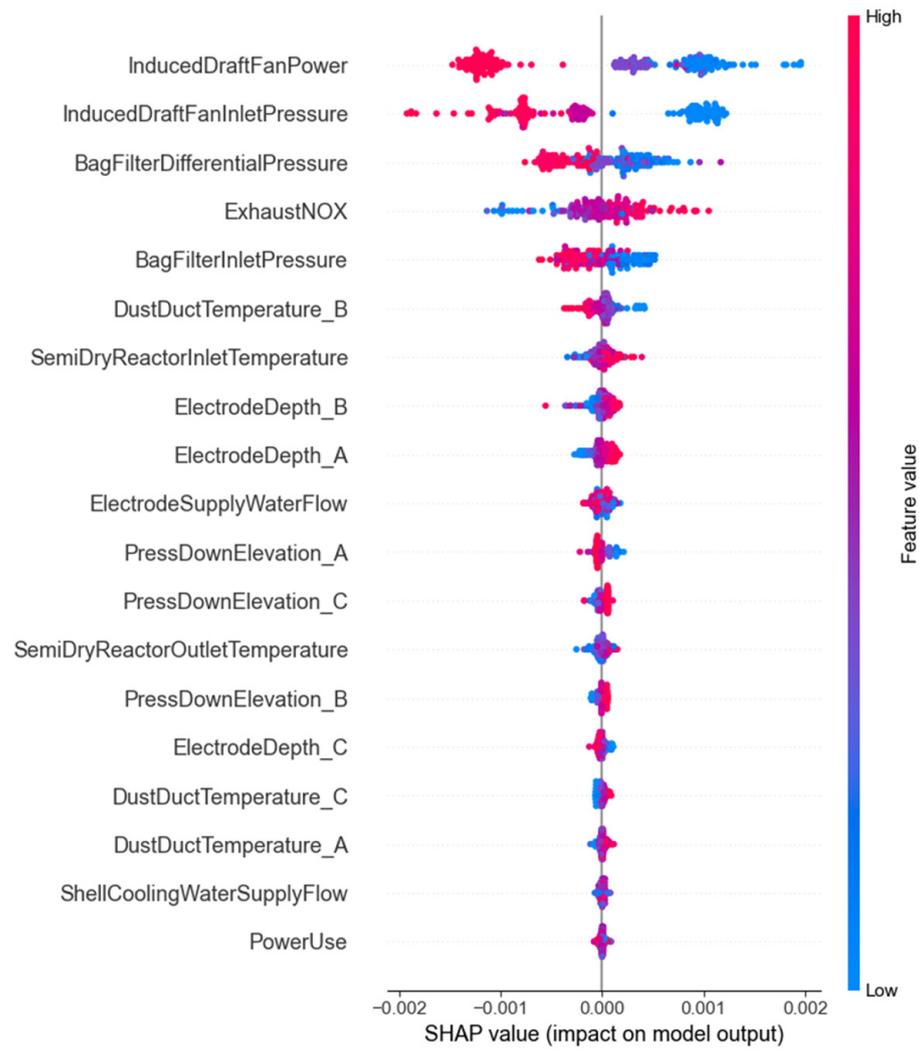
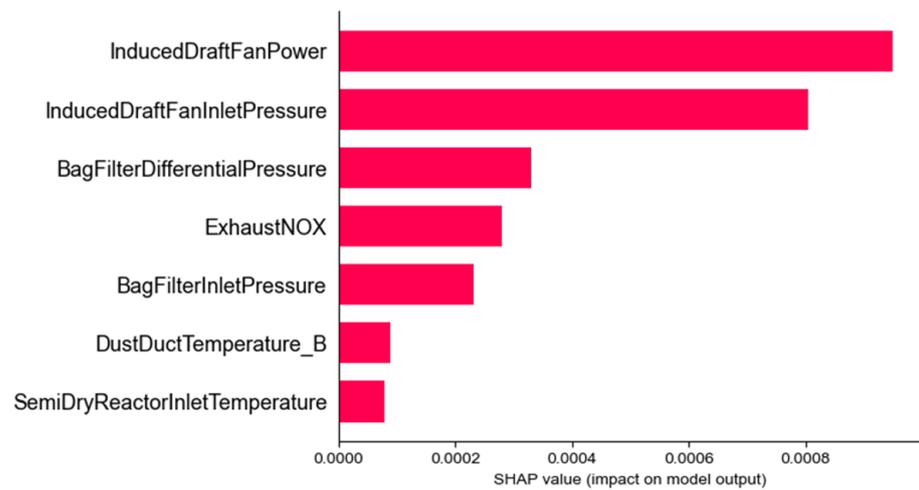


Figure 8. The absolute SHAP value of each variable for time-wise steps.

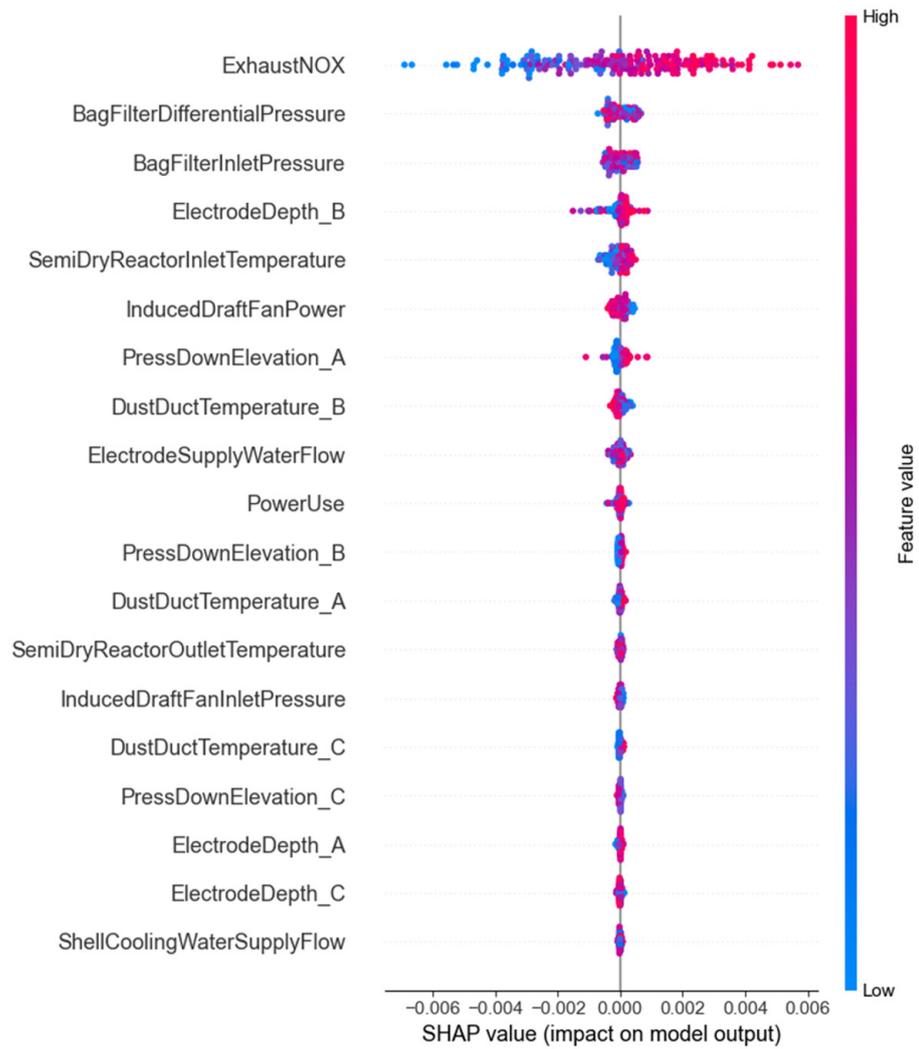


(a) Summary Plot (Low NOx)

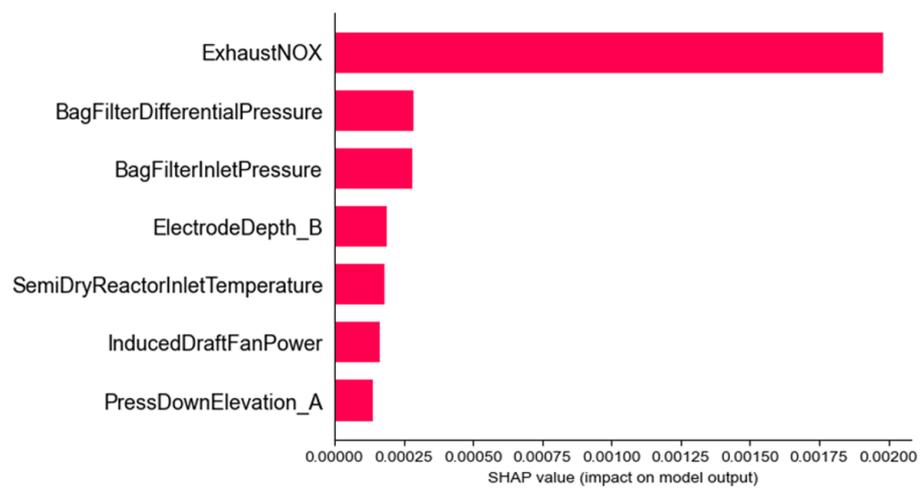


(b) Bar Plot (Low NOx)

Figure 9. Summary and bar plots with low NOx emissions.



(a) Summary Plot (High NOx)



(b) Bar Plot (High NOx)

Figure 10. Summary and bar plots with high NOx emissions.

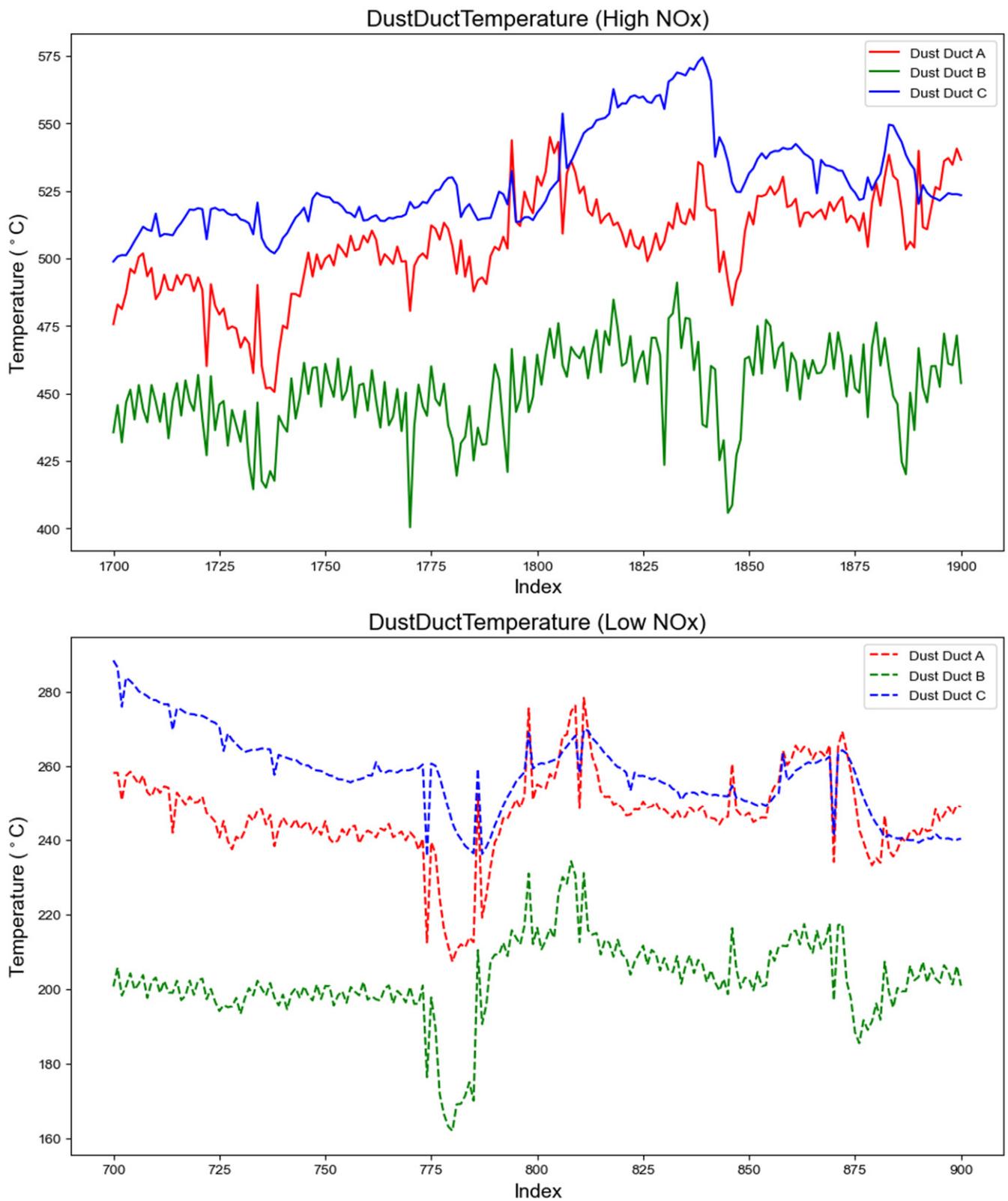


Figure 11. Comparison of dust duct temperature at different NOx emission levels.

5. Discussion and Conclusions

This study proposes a model for predicting NO_x emissions suitable for the EAFs of ferroalloy production sites. A Kalman-filter-based smoothing algorithm was used to denoise the NO_x emission data from the EAFs and construct the training data. The study presented an interpretable model using variables collectable from EAFs at ferroalloy production sites and was able to identify key influencing variables in prediction through the utilization of explainable AI. The NO_x emission prediction model employs real-time data collected from the EAFs of the ferroalloy production workplace, thereby offering insights for practitioners aiming to establish a real-time prediction system with data collection and NO_x prediction capability. With increasing environmental regulations, practitioners involved in related industries need to prepare for these changes, which can serve as a basis for proactive adaptation in ferroalloy production.

This study developed an interpretable model for predicting NO_x emissions in EAFs by adopting LSTM and identified the variables with a significant impact on NO_x emission predictions from the collected data through explainable AI methods. Owing to this research, it is possible to provide guidance for building a NO_x prediction system in EAFs, and it hints at ways to reduce NO_x emissions at ferroalloy production sites through NO_x prediction. For practical applications, NO_x prediction can be implemented in real-world settings, with potential expansion to both chimney and internal exhaust gas emissions. However, the key to effective NO_x emission prediction lies in the ability to collect data. Real-time data transmission from manufacturing and exhaust gas facilities to systems capable of immediate data management and collection is essential. From the perspective of building an NO_x emission prediction system, this study can be helpful in establishing a system for the prediction of EAFs at ferroalloy production sites, where a data collection system has not yet been implemented. This study outlines the collected data, key variables, and data collection locations, offering guidance for workplaces looking to initiate data collection and management for NO_x prediction. Many EAFs in ferroalloy production face challenges in establishing a data collection system for real-time historical processes and observational data from chimneys. Moreover, identifying the specific data required for accurate real-time NO_x emissions prediction from the collectable data is necessary to these facilities. Owing to the limited prior research on predicting NO_x emissions from EAFs, it is necessary to identify data that can be collected and that are essential for the prediction of EAFs at ferroalloy production.

Regarding potential impacts, this research can assist ferroalloy plant operators planning to reduce NO_x emissions. NO_x prediction can significantly contribute to NO_x reduction efforts, both pre- and post-management. For pre-management, by identifying key operating variables during the NO_x prediction process, it is possible to apply them to the operating systems of production facilities, attempting to adjust variables for the reduction in NO_x emissions. In this study, through SHAP analysis, the operational variables were determined when the NO_x emission levels were high and low. However, the variables with high importance values were measurements, whereas the actual operational variables, such as the depth of the electrode bars and power usage, showed low importance. If future research develops a high-performance predictive model based on operational variables, it will be possible to identify combinations of operational variables to reduce NO_x emissions using an interpretable method. In post-management techniques, NO_x prediction can contribute to exhaust systems using selective catalytic reduction (SCR) facilities. Denitrification facilities (e.g., SCR) remove NO_x emissions from exhaust gases through chemical reactions, and the rate of NO_x removal varies depending on the amount of ammonia used as a reducing agent. The excessive injection of ammonia can cause ammonia slip, leading to potential equipment failure and reduced dust collection efficiency, whereas too little ammonia reduces the NO_x reduction. Therefore, a system that can adjust the amount of ammonia injection by predicting NO_x emissions in real time is required.

Despite these contributions, further studies are required. First, the study could be applied to various EAF environments as the types, variables, and specifications of EAFs

can vary, and broader application in diverse settings could enhance the generalizability of this research. By expanding our collection of operational variables and enhancing the depth of interpretable AI analysis in future research, NO_x prediction can transform into a proactive management tool. This progress is expected to facilitate the implementation of operational strategies specifically targeted at reducing NO_x emissions, thereby advancing toward active environmental management. Second, as this study was conducted between May and July, there is a need to use data over a more extended period. With significant seasonal temperature variations in Korea, collecting more data to consider seasonality could improve the applicability of this study. Continued research in this area could lead to broader and more universal applications of this study for various EAFs at ferroalloy production sites. Third, a systematic approach to assign hyperparameters of the Kalman Filter smoothing algorithm is required. In this study, they are designed based on trial-and-error methods. An enhanced and systematic approach can improve this study. Finally, the intrinsic limitation of SHAP should be acknowledged. Because SHAP is a model-agnostic method utilizing the average value of the local one, the derived values are inherently influenced by the specific samples used. This challenge is not exclusive to this particular issue but is relevant to the interpretation of deep learning models at large. Further research into the interpretation of NO_x prediction is necessary to deepen our understanding of NO_x generation from EAFs.

This study was implemented using the Python 3.8.18 programming language, alongside TensorFlow 2.10.0 for deep learning model development and SHAP 0.42.1 for interpretability analysis.

Author Contributions: Conceptualization, Y.S., S.L., J.L., C.-W.K., H.S.B., Y.B., and J.Y.; methodology, Y.S., S.L., J.L., and J.Y.; software, Y.S., S.L., C.-W.K., and J.Y.; validation, J.L., C.-W.K., and J.Y.; formal analysis, S.L.; investigation, Y.S., S.L., J.L., C.-W.K., J.Y., H.S.B., Y.B., and J.Y.; resources, C.-W.K. and J.Y.; data curation, Y.S., H.S.B., and Y.B.; writing—original draft preparation, Y.S., S.L., J.L., and J.Y.; writing—review and editing, C.-W.K., H.S.B., Y.B., and J.Y.; visualization, Y.S., S.L., and H.S.B.; supervision, J.Y.; project administration, C.-W.K., J.Y., H.S.B., Y.B., and J.Y.; funding acquisition, Y.B. and J.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Korea Environment Industry & Technology Institute (KEITI) through the R&D Project for Intelligent Optimum Reduction and Management of Industrial Fine Dust funded by the Korea Ministry of Environment (MOE) (2022003580004), the Human Resources Program in Energy Technology of the Korea Institute of Energy Technology Evaluation and Planning (KETEP), and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No. 20204010600220).

Data Availability Statement: The datasets presented in this article are not readily available because the data were provided by the ferroalloy plants in Gwangyang. Requests to access the datasets should be directed to RIST.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Agency, I.E. *Iron and Steel Technology Roadmap: Towards More Sustainable Steelmaking*; OECD Publishing: Berlin, Germany, 2020.
2. Jonidi Jafari, A.; Charkhloo, E.; Pasalari, H. Urban air pollution control policies and strategies: A systematic review. *J. Environ. Health Sci. Eng.* **2021**, *19*, 1911–1940. [[CrossRef](#)]
3. Trnka, D. *Policies, Regulatory Framework and Enforcement for Air Quality Management: The Case of Korea*; OECD Publishing: Berlin, Germany, 2020.
4. Fichte, R. *Ferroalloys*. *Ullmann's Encyclopedia of Industrial Chemistry*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2000.
5. Kirschen, M.; Voj, L.; Pfeifer, H. NO_x emission from electric arc furnace in steel industry: Contribution from electric arc and co-combustion reactions. *Clean Technol. Environ. Policy* **2005**, *7*, 236–244. [[CrossRef](#)]
6. Weschler, C.J. Ozone's impact on public health: Contributions from indoor exposures to ozone and products of ozone-initiated chemistry. *Environ. Health Perspect.* **2006**, *114*, 1489–1496. [[CrossRef](#)]
7. Yang, G.T.; Wang, Y.N.; Li, X.L. Prediction of the NO emissions from thermal power plant using long-short term memory neural network. *Energy* **2020**, *192*, 116597. [[CrossRef](#)]

8. Tang, Z.H.; Wang, S.K.; Chai, X.Y.; Cao, S.X.; Ouyang, T.H.; Li, Y. Auto-encoder-extreme learning machine model for boiler NO emission concentration prediction. *Energy* **2022**, *256*, 124552. [[CrossRef](#)]
9. Wang, F.; Ma, S.; Wang, H.; Li, Y.; Zhang, J. Prediction of NOx emission for coal-fired boilers based on deep belief network. *Control Eng. Pract.* **2018**, *80*, 26–35. [[CrossRef](#)]
10. Yuan, Z.; Meng, L.; Gu, X.; Bai, Y.; Cui, H.; Jiang, C. Prediction of NOx emissions for coal-fired power plants with stacked-generalization ensemble method. *Fuel* **2021**, *289*, 119748. [[CrossRef](#)]
11. Korpela, T.; Kumpulainen, P.; Majanne, Y.; Häyrynen, A.; Lautala, P. Indirect NOx emission monitoring in natural gas fired boilers. *Control Eng. Pract.* **2017**, *65*, 11–25. [[CrossRef](#)]
12. Wang, Z.; Peng, X.; Cao, S.; Zhou, H.; Fan, S.; Li, K.; Huang, W. NOx emission prediction using a lightweight convolutional neural network for cleaner production in a down-fired boiler. *J. Clean. Prod.* **2023**, *389*, 136060. [[CrossRef](#)]
13. Wang, H.; Zhang, C.; Liu, X. Heat transfer calculation methods in three-dimensional CFD model for pulverized coal-fired boilers. *Appl. Therm. Eng.* **2020**, *166*, 114633. [[CrossRef](#)]
14. Belošević, S.; Tomanović, I.; Beljanski, V.; Tucaković, D.; Živanović, T. Numerical prediction of processes for clean and efficient combustion of pulverized coal in power plants. *Appl. Therm. Eng.* **2015**, *74*, 102–110. [[CrossRef](#)]
15. Chan, E.; Riley, M.; MJ, T.; EJ, E. Nitrogen oxides (NOx) formation and control in an electric arc furnace (EAF): Analysis with measurements and computational fluid dynamics (CFD) modeling. *ISIJ Int.* **2004**, *44*, 429–438. [[CrossRef](#)]
16. Zhou, H.-C.; Lou, C.; Cheng, Q.; Jiang, Z.; He, J.; Huang, B.; Pei, Z.; Lu, C. Experimental investigations on visualization of three-dimensional temperature distributions in a large-scale pulverized-coal-fired boiler furnace. *Proc. Combust. Inst.* **2005**, *30*, 1699–1706. [[CrossRef](#)]
17. Zhang, Y.; Wang, W.; Shao, S.; Duan, S.; Hou, H. ANN-GA approach for predictive modelling and optimization of NOx emissions in a cement precalcining kiln. *Int. J. Environ. Stud.* **2017**, *74*, 253–261. [[CrossRef](#)]
18. Ding, X.; Feng, C.; Yu, P.; Li, K.; Chen, X. Gradient boosting decision tree in the prediction of NOx emission of waste incineration. *Energy* **2023**, *264*, 126174. [[CrossRef](#)]
19. Fleurialt, C.; Grogan, J.; White, J. Electric arc smelting. *JOM* **2019**, *71*, 321–322. [[CrossRef](#)]
20. Singh, R. *Applied Welding Engineering: Processes, Codes, and Standards*; Butterworth-Heinemann: Oxford, UK, 2020.
21. Kim, J.; Lee, G.; Lee, S.; Lee, C. Towards expert-machine collaborations for technology valuation: An interpretable machine learning approach. *Technol. Forecast. Soc. Chang.* **2022**, *183*, 121940. [[CrossRef](#)]
22. Faravelli, T.; Bua, L.; Frassoldati, A.; Antifora, A.; Tognotti, L.; Ranzi, E. A new procedure for predicting NOx emissions from furnaces. In *Computer Aided Chemical Engineering*; Elsevier: Amsterdam, The Netherlands, 2000; Volume 8, pp. 859–864.
23. Lv, M.; Zhao, J.; Cao, S.; Shen, T. Prediction of the 3D Distribution of NOx in a Furnace via CFD Data Based on ELM. *Front. Energy Res.* **2022**, *10*, 848209. [[CrossRef](#)]
24. Safdarnejad, S.M.; Tuttle, J.F.; Powell, K.M. Dynamic modeling and optimization of a coal-fired utility boiler to forecast and minimize NOx and CO emissions simultaneously. *Comput. Chem. Eng.* **2019**, *124*, 62–79. [[CrossRef](#)]
25. Shen, Q.; Wang, G.; Wang, Y.; Zeng, B.; Yu, X.; He, S. Prediction Model for Transient NOx Emission of Diesel Engine Based on CNN-LSTM Network. *Energies* **2023**, *16*, 5347. [[CrossRef](#)]
26. Li, N.; Lv, Y.; Hu, Y. Prediction of NOx Emissions from a Coal-Fired Boiler Based on Convolutional Neural Networks with a Channel Attention Mechanism. *Energies* **2022**, *16*, 76. [[CrossRef](#)]
27. Adadi, A.; Berrada, M. Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access* **2018**, *6*, 52138–52160. [[CrossRef](#)]
28. Doshi-Velez, F.; Kim, B. Towards a rigorous science of interpretable machine learning. *arXiv* **2017**, arXiv:1702.08608.
29. Molnar, C. *Interpretable Machine Learning*; Lulu.Com: Raleigh, NC, USA, 2020.
30. Das, A.; Rad, P. Opportunities and challenges in explainable artificial intelligence (xai): A survey. *arXiv* **2020**, arXiv:2006.11371.
31. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [[CrossRef](#)] [[PubMed](#)]
32. Jiang, P.; Liu, Z.; Abedin, M.Z.; Wang, J.; Yang, W.; Dong, Q. Profit-driven weighted classifier with interpretable ability for customer churn prediction. *Omega* **2024**, *125*, 103034. [[CrossRef](#)]
33. Liu, Z.; Jiang, P.; Wang, J.; Du, Z.; Niu, X.; Zhang, L. Hospitality order cancellation prediction from a profit-driven perspective. *Int. J. Contemp. Hosp. Manag.* **2023**, *35*, 2084–2112. [[CrossRef](#)]
34. Rabby, M.F.; Tu, Y.; Hossen, M.I.; Lee, I.; Maida, A.S.; Hei, X. Stacked LSTM based deep recurrent neural network with kalman smoothing for blood glucose prediction. *BMC Med. Inform. Decis. Mak.* **2021**, *21*, 101. [[CrossRef](#)]
35. Xue, G.; Qi, C.; Li, H.; Kong, X.; Song, J. Heating load prediction based on attention long short term memory: A case study of Xingtai. *Energy* **2020**, *203*, 117846. [[CrossRef](#)]
36. Staal, O.M.; Sælid, S.; Fougner, A.; Stavdahl, Ø. Kalman smoothing for objective and automatic preprocessing of glucose data. *IEEE J. Biomed. Health Inform.* **2018**, *23*, 218–226. [[CrossRef](#)]
37. Song, M.; Xue, J.; Gao, S.; Cheng, G.; Chen, J.; Lu, H.; Dong, Z. Prediction of NOx concentration at SCR inlet based on BMIFS-LSTM. *Atmosphere* **2022**, *13*, 686. [[CrossRef](#)]
38. Wen, X.; Li, K.; Wang, J. NOx emission predicting for coal-fired boilers based on ensemble learning methods and optimized base learners. *Energy* **2023**, *264*, 126171. [[CrossRef](#)]

39. Bostani, H.; Sheikhan, M. Hybrid of binary gravitational search algorithm and mutual information for feature selection in intrusion detection systems. *Soft Comput.* **2017**, *21*, 2307–2324. [[CrossRef](#)]
40. Kraskov, A.; Stögbauer, H.; Grassberger, P. Estimating mutual information. *Phys. Rev. E* **2004**, *69*, 066138. [[CrossRef](#)] [[PubMed](#)]
41. Shapley, L.S. *Additive and Non-Additive Set Functions*; Princeton University: Princeton, NJ, USA, 1953.
42. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 4765–4774.
43. Chen, H.; Lundberg, S.; Lee, S.-I. Explaining models by propagating Shapley values of local components. In *Explainable AI in Healthcare and Medicine: Building a Culture of Transparency and Accountability*; Springer: Cham, Switzerland, 2021; pp. 261–270.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.