

Article

NO_x Emission Prediction of Diesel Vehicles in Deep Underground Mines Using Ensemble Methods

Michalina Kotyla ¹, Aleksandra Banasiewicz ¹, Pavlo Krot ^{1,*}, Paweł Śliwiński ² and Radosław Zimroz ¹

¹ Faculty of Geoengineering, Mining and Geology, Wrocław University of Science and Technology, Na Grobli 15, 50-421 Wrocław, Poland; michalina.kotyla@pwr.edu.pl (M.K.); aleksandra.banasiewicz@pwr.edu.pl (A.B.); radoslaw.zimroz@pwr.edu.pl (R.Z.)

² KGHM Polska Miedź S.A., M. Skłodowskiej-Curie 48, 59-301 Lubin, Poland; pawel.sliwinski@kg hm.com

* Correspondence: pavlo.krot@pwr.edu.pl

Abstract: The mining industry faces persistent challenges related to hazardous gas emissions. Diesel engine-powered wheeled vehicles are commonly used during work shifts and are a primary source of nitrogen oxides (NO_x) in underground mines. Despite diesel engine manufacturers providing gas generation data, mining companies need to predict NO_x emissions from numerous load-haul-dumping (LHD) vehicles operating under dynamic conditions and not always equipped with gas sensors. This study focused on two ensemble methods: bootstrap aggregation (bagging) and least-square boosting (boosting) to predict NO_x emissions. These approaches combine multiple weaker statistical models to yield a robust result. The innovation of this research is in the statistical analysis and selection of LHD vehicles' working parameters, which are most suitable for NO_x emission prediction; development of the procedure of source data cleaning and processing, model building and analyzing factors, which may influence the accuracy; and the comparison of two ensemble methods and showing their advantages and limitations for this specific engineering application, which was not previously reported in the literature. For datasets obtained from the same LHD vehicle and different operators, the more efficient bagging method gave a coefficient of determination $R^2 > 0.79$ and the RMSE (root mean square error) was under 30 ppm, which is comparable with the measurement accuracy for transient regimes of physical NO_x sensors available in the market. The obtained insights can be utilized as input for mine ventilation systems, enhancing mining transport management, reducing workplace air pollution, improving work planning, and enhancing personnel safety.

Keywords: NO_x emissions; diesel engine; LHD vehicles; deep underground mine; machine learning; bagging; boosting



Citation: Kotyla, M.; Banasiewicz, A.; Krot, P.; Śliwiński, P.; Zimroz, R. NO_x Emission Prediction of Diesel Vehicles in Deep Underground Mines Using Ensemble Methods. *Electronics* **2024**, *13*, 1095. <https://doi.org/10.3390/electronics13061095>

Academic Editor: Dimitris Apostolou

Received: 23 February 2024

Revised: 14 March 2024

Accepted: 14 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 mining industry is constantly evolving and developing. This dynamic sector plays a key role in providing the raw materials needed to meet society's needs, from energy to minerals. Nevertheless, the growth of this industry brings both benefits and challenges that underground mines must resolve. Underground mines, a major segment of the mining industry, face a wide variety of natural hazards. Although different wired and wireless measurement and data collection technologies are available, which can be applied in underground mines [1], certain challenges exist related to the monitoring of underground operations including climate hazards, fire hazards, or gas hazards; the latter one is the subject of this article.

The mining enterprises continuously produce huge volumes of hazardous gases, e.g., from mining explosives containing carbon oxides and nitrogen oxides in their detonation products [2]. However, diesel engines of the powerful transportation trucks and load-haul-dumping (LHD) vehicles are a major source of nitrogen oxides (NO_x) in the underground working space. This is an issue that is of great importance for both worker safety and

environmental protection [3–5]. Unfortunately, due to the hard-working conditions of mobile machinery, most underground vehicles do not have sensors to measure parameters, including, among other things, NO_x emissions, making it difficult for mining companies to accurately monitor and predict gas emissions for appropriate ventilation planning [6]. That is why it is so important to develop modern modeling and forecasting tools based on machine learning techniques. A hybrid methodology for investigating diesel particulate matter (DPM) concentration distribution in underground mines was developed in [7] using a ventilation network solver and computational fluid dynamics (CFD). However, such models require reliable input data from every working vehicle which, when obtained from one machine, cannot be averaged over the whole fleet due to specific operational conditions and operators' steering behavior. Practically every LHD vehicle is characterized by highly dynamic regimes of motion [8]. Although hydraulic torque converters are installed in the transmission system, the V-type routes cause torsional vibrations and non-stationary torque demand during the LHD loading and unloading cycles, which cause non-optimal fuel combustion regardless of engine electronic control unit (ECU) settings.

To ensure safe working conditions, continuous operation, and reduced health risks for underground workers, onboard monitoring tools are installed, which supply data from trucks and LHD vehicles for the SYNAPSA system in the investigated copper ore mine. The study is focused on using ensemble methods to predict NO_x emissions, which is crucial for effective gas management in dynamic operating environments. Ensemble methods were earlier used for NO_x emission predictions in diesel engines [9] and coal-fired power plants [10], as well as for natural gas consumption [11] and coal-fired [12] or oil-fired boiler emissions [13] but not for diesel vehicles operated underground.

Ventilation in mines is a complex and essential aspect that significantly impacts the safety of underground workers [14,15]. Regulations governing underground mining focus on ensuring safe working conditions and protecting the health of personnel working in these specific environments. In this context, the regulations include guidelines and standards that cover a variety of aspects, such as maximum working temperatures and acceptable concentrations of harmful gases. Work in underground mines exposes personnel to risks from harmful gases such as methane, carbon monoxide, nitrogen oxide, and others. These gases can be toxic or explosive in appropriate concentrations.

One of the major concerns in underground mines is the presence of NO_x, which consist of nitrogen oxide and nitrogen dioxide. While NO_x can occur naturally, most of its volume is related to technological processes, particularly those involving diesel engines. These engines not only emit NO_x but also produce various other pollutants [16]. What makes NO_x gases particularly dangerous is their lack of odor, with NO being colorless and NO₂ being a brown color. In diesel engine exhaust, monoxide NO is predominant, constituting around 90% of the emissions. To reduce its concentration, the new approach of NO_x photocatalytic conversion was proposed in [17], which has benefits in stability, cost, efficiency, convenience in operation, and strong durability. The results of the work [18] undermined the significant effect of NO₂ on the performance of the urea-based selective catalytic reduction (SCR) system, which does not form N₂O considered as an undesired byproduct of catalytic NO_x control. Restrictions of existing sensors for NO_x measurement are discussed in [19]. The Co₃O₄-based catalysts for NO_x reduction in fast SCR processes are proposed in work [20].

The negative impact of NO_x on human health cannot be underestimated. Exposure to nitrogen oxides can result in irritation of the eyes and throat, causing discomfort and tissue irritation. Coughing is also a common result of exposure to these gases, and their presence can cause reduced lung function and difficulty breathing. However, what becomes more alarming is long-term exposure to nitrogen oxides, which can result in serious respiratory problems. Particularly dangerous is nitrogen dioxide, which is even more toxic than nitric oxide. Long-term exposure to NO₂ can trigger chronic respiratory diseases such as asthma and chronic obstructive pulmonary disease. In extreme cases, when the concentration of nitrogen oxides reaches very high levels, sudden and serious deterioration of health and

even death can occur. Concentrations of NO₂ exceeding 200 ppm are extremely dangerous to human health and can lead to rapid respiratory failure and death [14,21,22].

Diesel engines' NO_x emissions are impacted by various factors, including the vehicle's load, the experience and skill of the operator (as this impacts engine rotations and motor speed), and the type of work conducted while using the vehicle [23,24]. The significant health risks posed by NO_x emissions put mining companies under an obligation to closely monitor and control these emissions to ensure safe working conditions for underground workers. Faced with the risks associated with these toxic gases, advanced strategies are being developed that focus on minimizing NO_x exposure and protecting the health of personnel due to exhaust gases aftertreatment. The problem exists with the engine electronic control unit's (ECU) tuning because of the transient character and cyclic loading of the LHD vehicles [25].

Continuous research and development of technologies to predict NO_x emissions are crucial to improving safety in mining conditions. With advanced monitoring and data analysis systems, it is possible to accurately track NO_x levels in various areas of a mine, identify higher-risk areas, and take appropriate countermeasures [26,27]. The authors of the work [28] intended to quantify gas concentrations inside an urban tunnel based on available traffic intensity data. They did not address the working parameters of every individual vehicle. Hence, the result of the prediction is a summarized pollutant emission in certain places (urban tunnels). Instead, this paper deals with individual vehicles' monitoring parameters and prediction of their emissions in different places on the route.

Innovative approaches, such as numerical modeling, make it possible to predict NO_x emissions according to various factors, such as the type of mining operation, rock type, and atmospheric conditions. This makes it possible to predict high-risk areas, which in turn makes it possible to take preventive measures to reduce workers' exposure to dangerous concentrations of NO_x.

There are known approaches to forecasting the NO_x emissions levels in underground diesel vehicles. In the previous works, neural networks [29] and polynomial regression [30] approaches were used. Those publications showed that it is possible to predict hazardous gas levels with satisfying results that increase safety in deep underground mines, allow better management of the workload, and also have a positive impact on the effectiveness of the mine.

In this paper, a model aggregation approach, known as the ensemble approach, was used with a focus on two ensemble methods—bootstrap aggregation (bagging) and least-square boosting (boosting). Both of these types of algorithms are based on the concept of combining multiple weak models to produce a single strong result. A comparison of model aggregation methods for regression is given in [31]. Starting from works [32,33], these methods have gained popularity in different domains [34,35].

The conducted research was focused on two powerful ensemble methods: bootstrap aggregation and boosting, both relying on random forests as the base models. Ensemble methods are well suited for enhancing prediction accuracy and reducing overfitting by combining multiple models to produce a robust and generalized result of gas emission [36].

By using random forests as the base of the models, the ensemble approach benefits from the ability to handle complex relationships in the data, leading to improved NO_x emission predictions. The application of random forest with ensemble methods ensures more reliable and consistent results.

This study has great significance for improving safety in the mining industry by focusing on predicting and monitoring NO_x emissions from LHD vehicles. The implementation of advanced methods brings multidimensional benefits that are crucial to the efficient and reliable operation of mining machines. The proposed methods can significantly help mining companies create safer and healthier working conditions for their personnel, as well as ensure the continuity of work underground.

Continuing investigation in this direction will contribute to the creation of thoughtful practices that will enhance worker safety in the mining industry. In the long-term per-

spective, it is presumed that this research will influence the creation of better standards and regulations for the protection of the working environment, in line with the drive for sustainable and responsible management of mining operations. Ultimately, through the application of these advanced methods, mining companies will be able to ensure the continuity of underground operations while minimizing risks to personnel and environmental protection.

2. Materials and Methods

2.1. Ensemble Methods

The goal of ensemble methods is to combine multiple models to improve the prediction results. The main advantage of the usage of the ensemble methods is increasing accuracy and robustness according to single models. Ensemble learning could be defined as in [37].

Ensemble learning is a process that uses a set of models, each of them obtained by applying a learning process to a given problem. This set of models (ensemble) is integrated in some way to obtain the final prediction.

The process can be divided into 3 phases:

1. Generation—creation of a set of models to be used further, usually redundant;
2. Pruning—selection and training models;
3. Integration—the output of the model generates the final result.

For this research purpose, two of the ensemble methods were chosen—bootstrap aggregation and least-square boosting.

The bootstrap aggregation (bagging) method was first introduced in 1996 by Breiman [32]. The method has three main steps:

- Bootstrapping—at this stage, the algorithm creates diverse samples. From the training set different subsets are generated by random selection with replacement. This allows the possibility that within the sample the same instance could be selected multiple times.
- Training—bootstrap samples are trained independently and in parallel using weak learners.
- Aggregation—the last step, where an average of all predicted outputs is taken, known as soft voting.

Least-square boosting (boosting) is a supervised learning method. Similar to the bagging method, random subsets of data are created and weak learners are created. In opposition, the learners are trained sequentially, not in parallel. This means that in every iteration, the prediction is compared to the expected result, residuals are calculated and the learner is updated to be stronger. Those types of algorithms (see Figure 1) combine multiple weak models into a single, strong result.

Bagging reduces variance and improves model accuracy. It combines multiple base models trained on different subsets of training data. In the chosen approach, several decision trees are created and trained, where each is trained on a different random subset of the original training data.

Usually, prediction $\hat{f}(x)$ at input x is obtained by fitting a model to training data $Z = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$. For solving regression problems such as NOx emission prediction, bootstrap aggregation was chosen. This method averages this prediction over a bootstrap sample set, which implies reducing its variance. For each bootstrap sample Z^{*b} , $b = 1, 2, \dots, B$, model is fitted giving prediction $\hat{f}^{*b}(x)$. Bagging estimate definition [35]:

$$\hat{f}_{bag}(x) = \frac{1}{B} \sum_{b=1}^B \hat{f}^{*b}(x) \quad (1)$$

where x —input, $\hat{f}_{bag}(x)$ —bagging estimate, B —number of bootstrap samples, b —bootstrap sample number, and $\hat{f}^{*b}(x)$ —prediction of bootstrap sample.

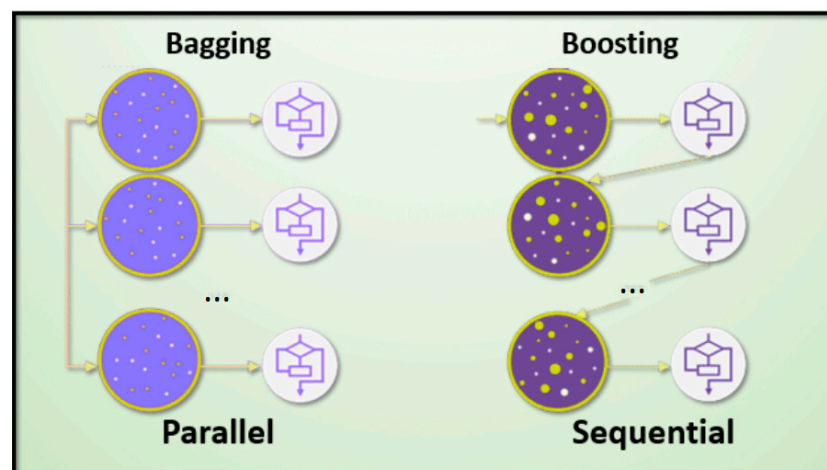


Figure 1. Schematic diagrams of bagging and boosting methods (points—intermediate solutions; arrows—trees; schematic algorithms—selection of best solution).

Boosting is another ensemble learning technique that combines weak learners into a strong model. Differently than bagging, boosting sequentially trains multiple models, where each subsequent model focuses more on the errors made by the previous model. The goal of boosting is to give more weight to the misclassified data points in each iteration, thereby improving the model's performance with a focus on hard-to-classify instances. The steps of the algorithm are as follows:

- Creating data samples—similar to bagging, multiple bootstrap samples are created from the original dataset.
- Build trees—each decision tree is trained on each of the bootstrap samples using randomly chosen features.
- Boosted training—a series of boosting iterations is conducted, where each iteration builds a new model to correct the errors of the previous model.
- Prediction—the final prediction for a new data point is obtained by adding the predictions from each model in the ensemble, considering their respective weights.

For the NO_x emission prediction problem, the least-square boosting method was chosen. Similar to the bagging algorithm, to receive the final prediction $\hat{f}(x)$, multiple models need to be trained on the created bootstrap sample set. In each iteration, a model is improved based on the loss function, which is defined as:

$$L(y, \hat{f}(x)) = (y - \hat{f}(x))^2 \quad (2)$$

where $\hat{f}(x)$ is actually trained model and y is current cumulative output:

$$y_i = \beta_0 + \sum_{j=1}^{i-1} \beta_j h_j + \beta_i h_i = y_{i-1} + \beta_i h_i, \quad (3)$$

where h is a parameterized function of input variables x , and β is a hyperparameter to control divergence.

Training error is computed as follows:

$$E = \sum_{t=1}^N [\beta_t h_t - L(y, \hat{f}^t(x))] \quad (4)$$

where $\hat{f}^t(x)$ is a residual error and β_i is calculated from $\frac{\partial E}{\partial \beta_i} = 0$. According to this, we can define LS-boosting as [31,34]:

$$\hat{f}_{boost}(x) = \arg_{\min} E_{x,y}[L(y, \hat{f}(x))] \quad (5)$$

where x —input, $\hat{f}_{boost}(x)$ —boosting estimate, $E_{x,y}$ —training error, b —bootstrap sample number, $\hat{f}(x)$ —prediction of bootstrap sample, and $L(y, \hat{f}(x))$ —loss function.

2.2. Data Aggregation

The measurement data that formed the basis for the study were obtained from the SYNAPSA system. This system plays a key role in collecting and analyzing data related to the operation of various mining vehicles. In this case, the data were obtained from a specific LHD vehicle (see Figure 2), which was equipped with a turbocharged diesel engine DEUTZ TCD 12.0 V6 and a selective catalytic reduction (SCR) system, as shown in Figure 3.



Figure 2. Load-haul-dumping (LHD) vehicle for underground operations [38].

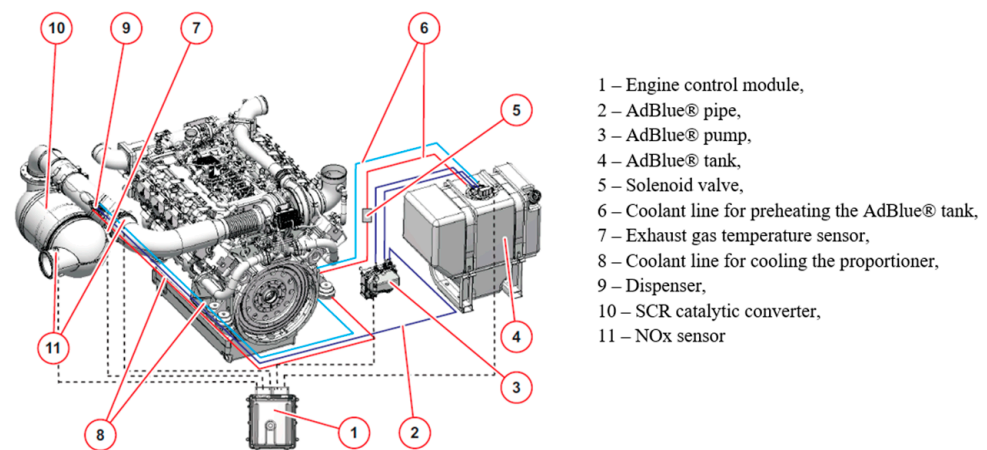


Figure 3. The DEUTZ TCD 12.0 V6 diesel engine with its SCR system [39].

The engine parameters that are relevant to the analysis are detailed in Table 1. The best point consumption values, as defined in the manual [39], are based on the use of diesel fuel with a specified density of 0.835 kg/dm³ at 15 °C.

Table 1. Parameters of diesel engine DEUTZ TCD 12.0 V6 [39].

Parameter		Value	Units
Power output (ISO 14396) [40]		390	kW
	at speed	2100	rpm
Max. torque		2130	Nm
	at speed	1400	rpm
Min. idling speed		600	rpm
Specific fuel consumption		194	g/kWh

The power and torque characteristics that are associated with the operation of this engine are presented in Figure 4.

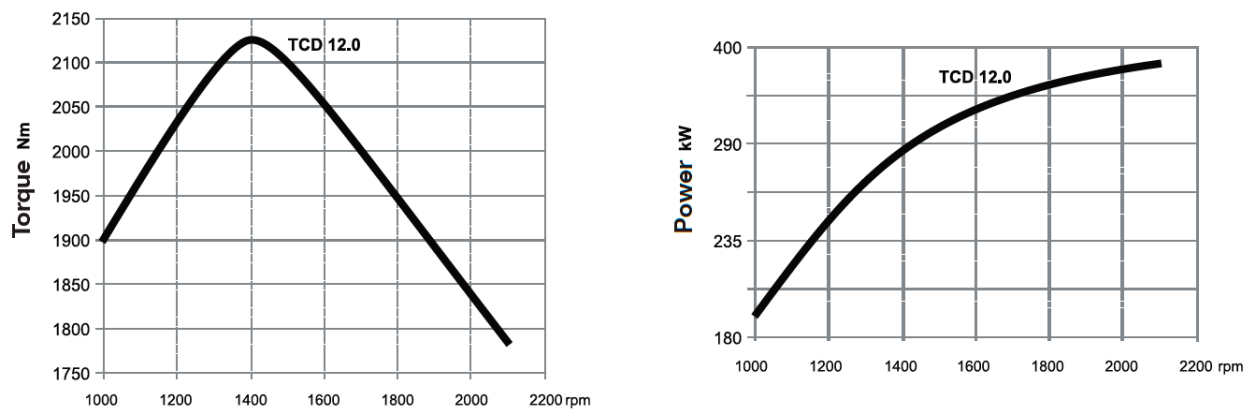


Figure 4. Power and torque functions of diesel engine DEUTZ TCD 12.0 V6 [39].

These graphs show the relationship between power, torque, and engine rotations. These characteristics are important because they allow us to understand in which range of engine speeds the highest performance is achieved and how these parameters vary with speed. The manual [39] provides accurate information on best point consumption, which relates to optimal engine performance. This information is extremely important in the context of evaluating engine performance and efficiency, especially regarding economical operation. On the other hand, the power and torque characteristics in Figure 4 allow a visual analysis of engine dynamics, which is crucial for understanding engine emissions under varying conditions.

Parameters of the typical NO_x sensor are presented in Table 2. Current regulations require an accuracy of $\pm 20\%$ or ± 0.1 g/bhp-h [19] of NO_x mass measurements, which could not be achieved under transient load conditions. Many sensors are characterized by noise levels around 10 ppm, which is caused by the presence in the exhaust system of residual amounts of NO_x.

Table 2. NO_x sensor parameters.

Parameter	Value	Units
Measuring range (NO _x)	0–1500	ppm
Accuracy	± 10 (20)	%
Operating temperature	−40... + 105	°C
Exhaust gas temperature	<800	°C

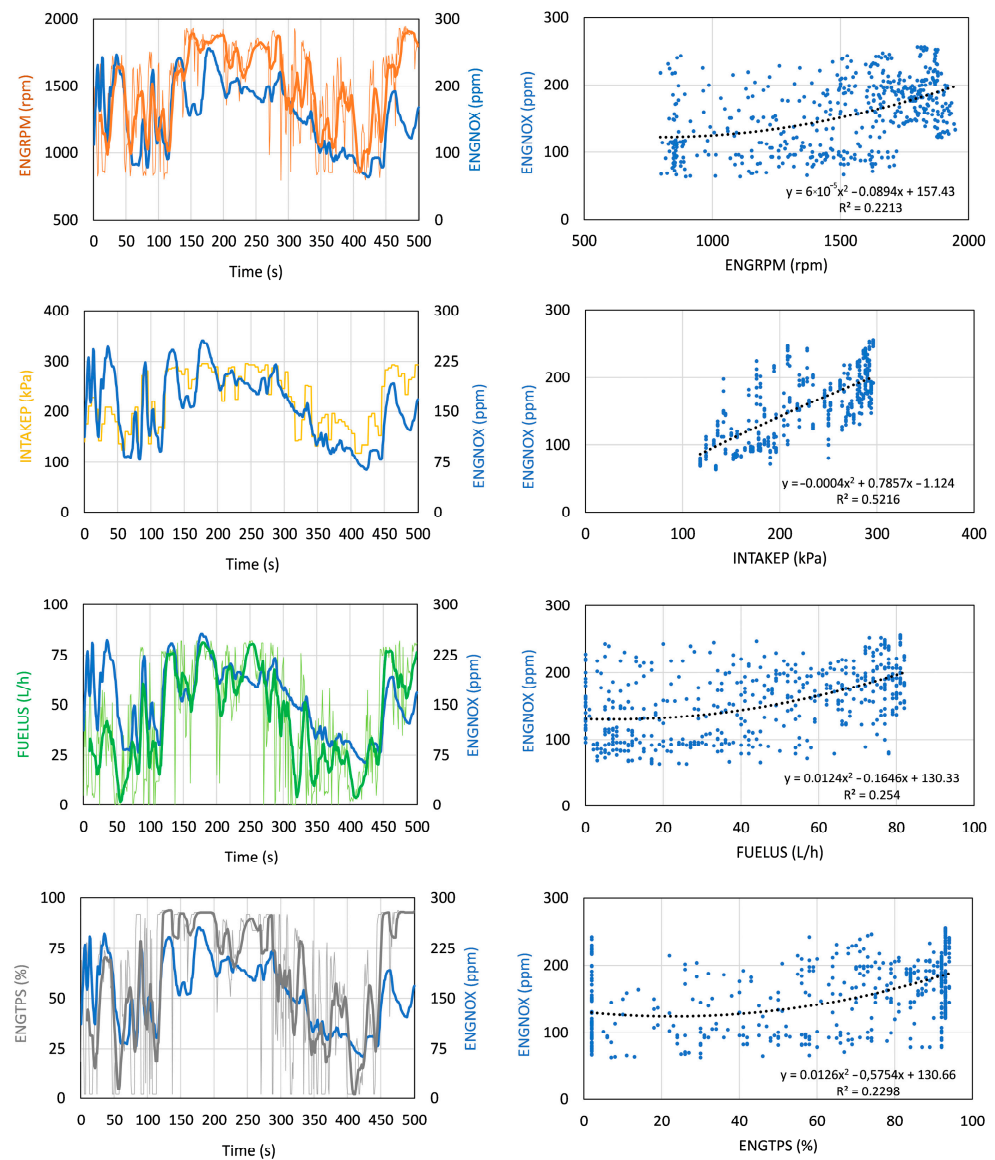
The SYNAPSA system collects parameters of diesel engine and LHD vehicle operation via CAN bus. The time interval of the sampled parameters is 1 s. This system includes not only the on-board part of hardware collecting signals from different sensors and vehicle internal units. It is also the wireless and cable facilities for data transferring. The stored local data once per 6 h are wirelessly uploaded from LHDs to the enterprise server with database and software for data processing and reports generation for different departments.

2.3. Data Preparation

The full list of monitored parameters contains 46 variables. For further examination, after deep analysis, only those ones chosen were the ones which affect the value of NO_x concentrations (see Table 3). Graphs of the NO_x emission relationships with some of the selected parameters are shown in Figure 5. The correlations are not enough strong to use each of these parameters alone for NO_x prediction.

Table 3. Monitoring system parameters chosen for analysis.

Nr	Parameter	Description	Units
0	ENGNOX	Engine NOx emissions	ppm
1	ENGCOOLT	Engine coolant temperature	°C
2	ENGOILP	Engine oil pressure	kPa
3	ENGRPM	Engine rotations	rpm
4	ENGTPS	Engine acceleration	%
5	FUELUS	Fuel consumption	L/h
6	GROILP	Gear oil pressure	kPa
7	GROILT	Gear oil temperature	°C
8	HYDOILP	Hydraulic oil pressure	MPa
9	INTAKEP	Intake air pressure	kPa
10	INTAKET	Intake air temperature	°C
11	SPEED	Vehicle speed	km/h

**Figure 5.** Correlation between NOx emissions and some parameters of the LHD vehicle (dark lines show smoothed original data).

The cycles somehow repeated but, when looking at them in detail, they are certainly not the same, and the NOx prediction task is not trivial. The cycles seem similar, but their duration and signals' patterns are quite different.

The collected data are sequential; however, depending on the operator and workload of the vehicle, some short-term outliers can be observed. To obtain the best results from collected data, those short peaks were smoothed with the moving average method for better visibility of dependencies. When the load is higher, more power is needed, so fuel consumption increases which results in higher NOx emissions.

Input and output variables were characterized by non-linear dependencies. The original and smoothed data of NOx emission from LHD vehicles are presented in Figure 6. The time series of parameters chosen as input signals are presented in Figure 7. To model the non-linear dependencies between input and output variables, ensemble models were chosen. Combining multiple models into predictions gives better results than a single model.

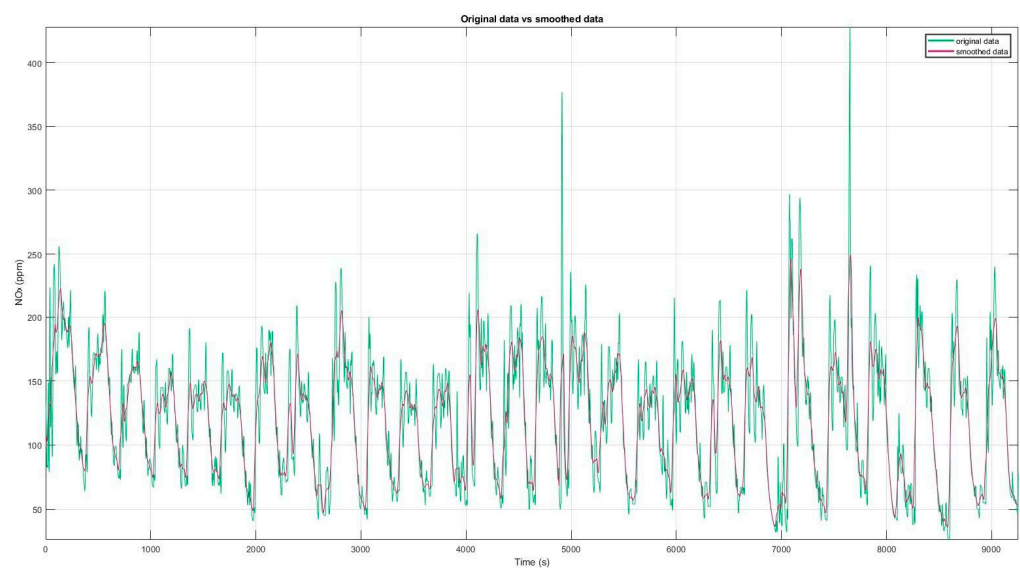


Figure 6. NOx emission levels during work shift cycles of LHD operation.

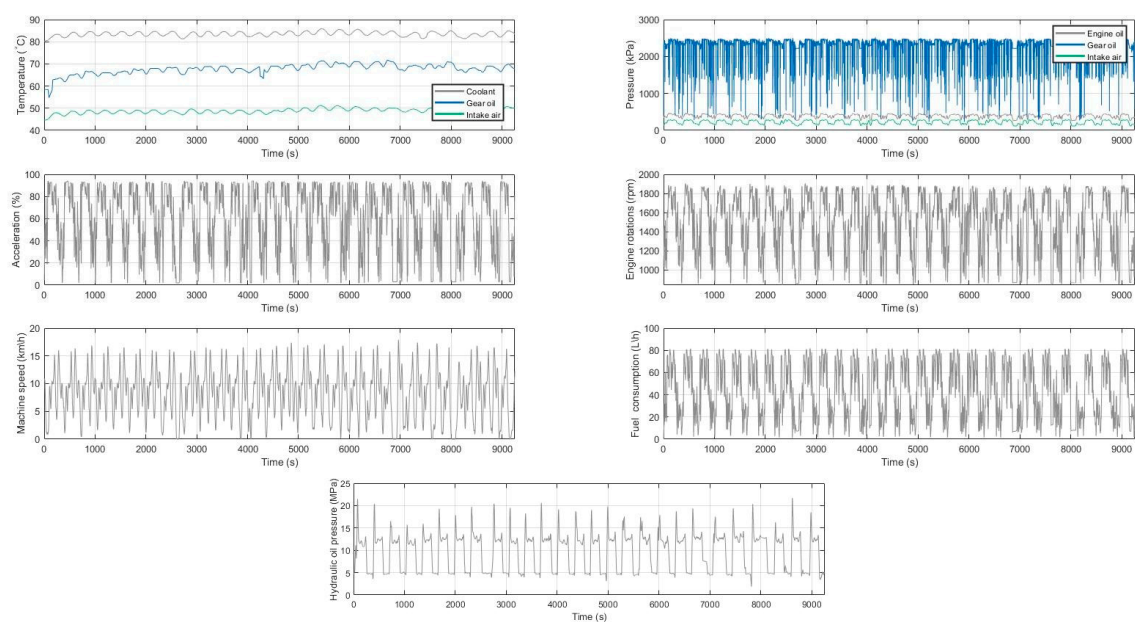


Figure 7. Smoothed input variables time series.

The dataset covers 27 cycles of the machine working, collected during one working shift in the underground mine (around 9000 samples). The dataset was divided into two parts: 90% of the training data and 10% of the test data. Both methods, bagging and boosting, used 700 decision trees to perform prediction. The number of learners was chosen to be optimal, not causing overfitting. The performed tests included examination of MAE, RMSE, and R^2 metrics for each algorithm, respectively, to the number of learning trees (Figure 8).

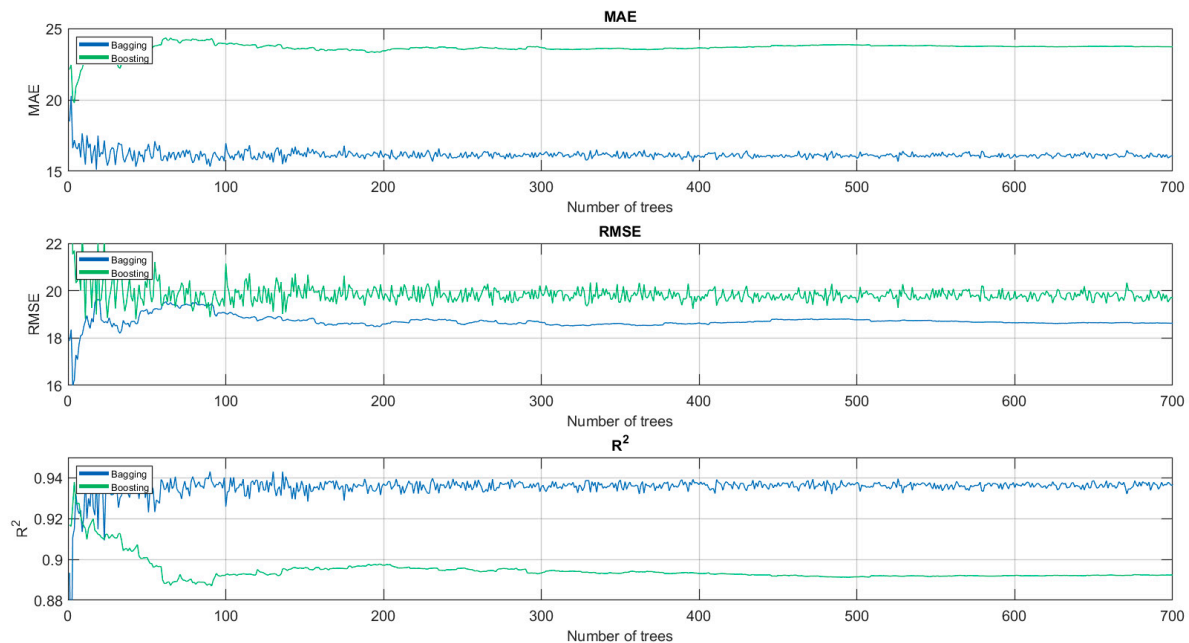


Figure 8. Tests of methods performance depending on a number of trees.

3. Results and Discussion

Both the bagging and boosting methods gave satisfying results, making them suitable candidates for application as ‘soft-sensors’ of NO_x emissions from diesel vehicles in underground mines. The original data smoothing in the preprocessing stage enhanced the quality of the predictions, providing valuable insights into NO_x concentration levels. Data cleanup was needed as the sensors were working in very hard conditions and the measurements were fraught with distortions.

3.1. Results for One Working Shift Time Period

At the initial stage of the prediction model verification, the one working shift time period was taken for learning and testing. Results of the calculation showed that the bagging algorithm, being more robust against overfitting, offered a safer choice in terms of prediction reliability (see Figure 9).

The application of boosting requires more caution to avoid overfitting issues. However, both methods delivered accurate predictions for NO_x emissions. Relationships between measured and predicted values of NO_x obtained by two methods are shown in Figure 10. For bagging, the coefficient of determination R^2 value exceeded 0.90 and the lower RMSE (root mean square error) was under 20 ppm, which is comparable with the accuracy for transient regimes of physical NO_x sensors available in the market. The achieved results prove the accuracy and reliability of estimating NO_x emissions by this method. With this information, it becomes feasible to manage more LHD vehicles in one place while respecting harmful gas concentrations and ensuring the safety of underground mine workers.

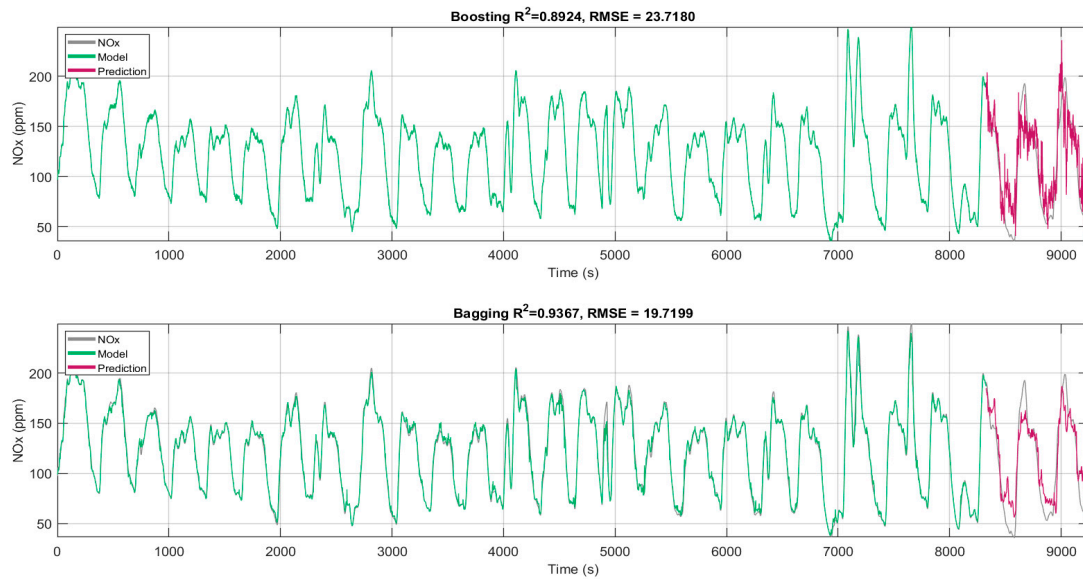


Figure 9. Original data, trained model, and predicted data for both methods for the first dataset.

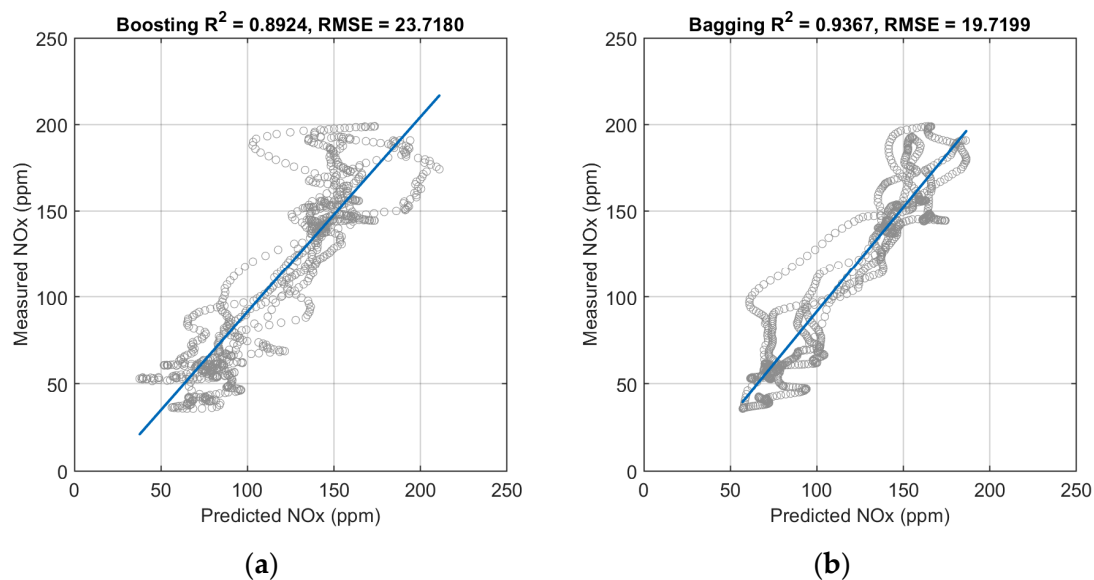


Figure 10. Relationships between measured and predicted values of NOx (ppm) by two methods: (a) boosting (linear trend (blue line) $y = 1.13x - 21.60$); and (b) bagging (linear trend (blue line) $y = 1.21x - 29.59$) for the first dataset.

The presented above results were achieved from the data aggregated during a single work shift. Collection data from multiple shifts and updating information about each operator could provide better predictions. These could have significant meaning for the further improvement of ventilation system management, as each operator's work style impacts emissions and creates a specific pattern.

3.2. Results for the Different Time Period

To check the robustness of the developed model, the different working shift time period was taken for testing at the next stage of the NOx prediction model verification. Results of the calculation showed that both methods also had satisfactory results for the data collected during the work of another operator on the same LHD vehicle. The original data (see Figure 11) were not as repetitive as the data from the first operator dataset (Figure 6). Time series also required smoothing as in the first case and spikes were truncated. The

truncation level (250 ppm) is presented by a line in the below figure. It was assigned based on data analysis and iterative testing of the prediction model to satisfy the requirement of minimal manual intrusion into a dataset and maximum model accuracy. In the future, this level may be changed after historical data processing for many operators and LHD vehicles.

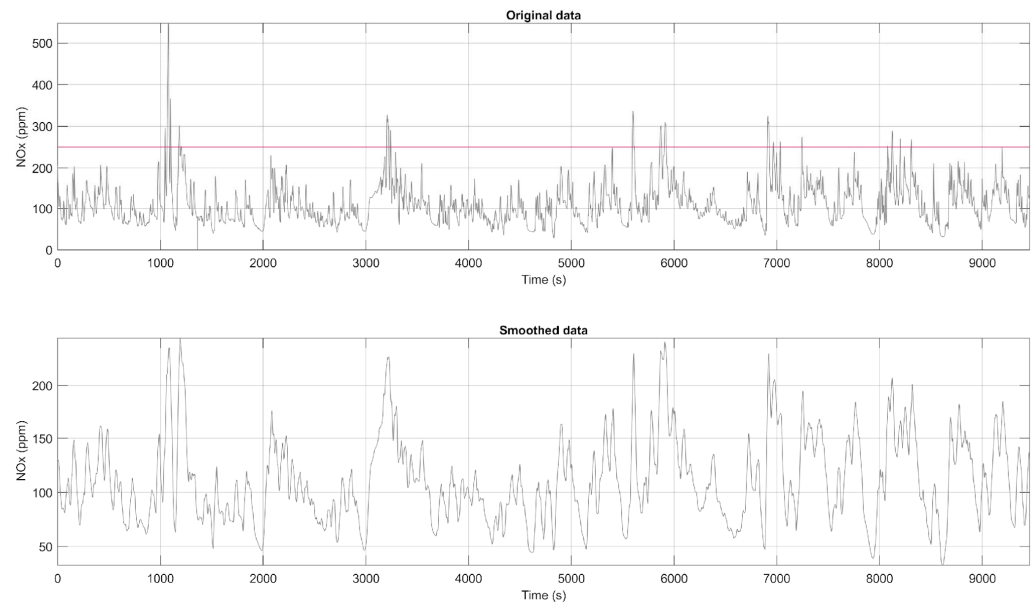


Figure 11. Original data with the cut-off level (horizontal pink line) and smoothed data.

The cutting of spikes above the level of 250 ppm resulted in manipulation in less than 2% of all the samples. The cumulative NOx removed for the simulation purposes (but was present in the underground mine) was 10,129 ppm which was less than 1% percent (0.96%) and had no significant impact on the generated by the model NOx level. The graphs in Figure 12 represent cumulative NOx after cutting and removing the chosen samples.

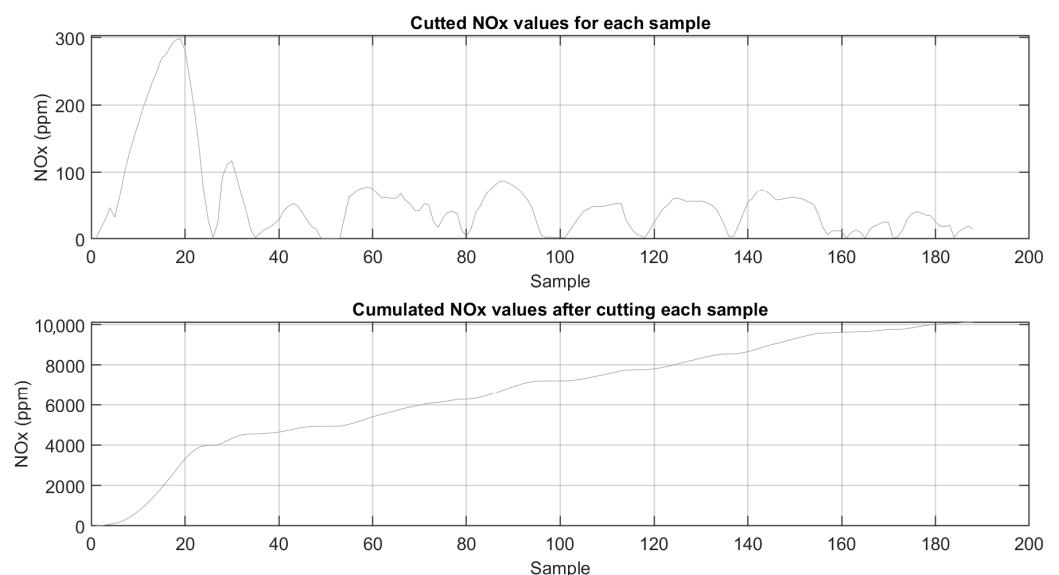


Figure 12. An analysis of samples limited by the amplitude.

The final results provide satisfying NOx prediction, shown below in Figures 13 and 14. Although the data were not such explicitly periodical as in the first working shift (the second operator had a more dynamic manner of driving), for the same learning parameters, the model still gave a good result. For bagging, the R^2 value was about 0.79 and the lower

RMSE (root mean square error) was under 29 ppm, which is comparable with the accuracy for transient regimes of physical NOx sensors available on the market.

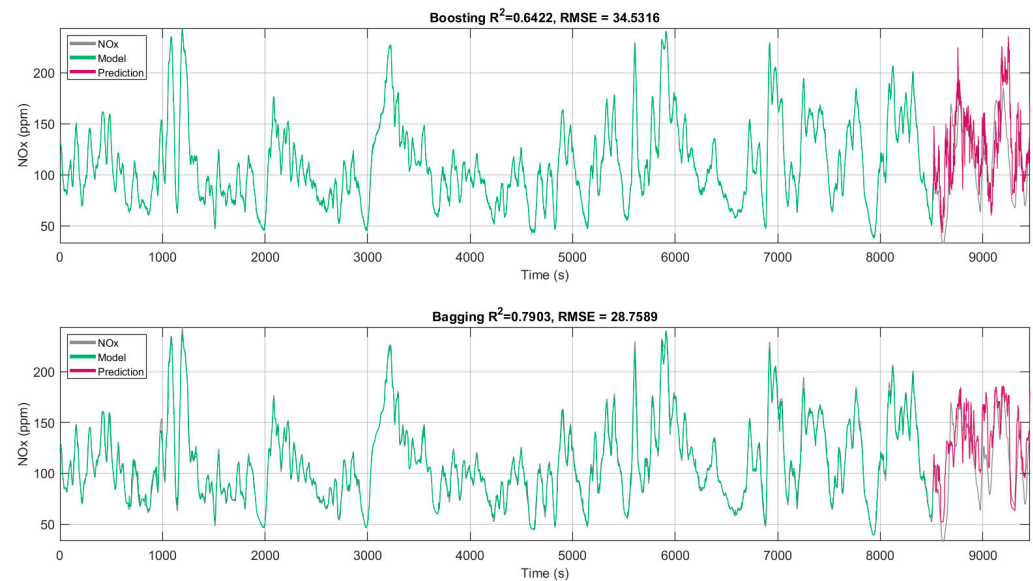


Figure 13. Original data, trained model, and predicted data for both methods for the second dataset.

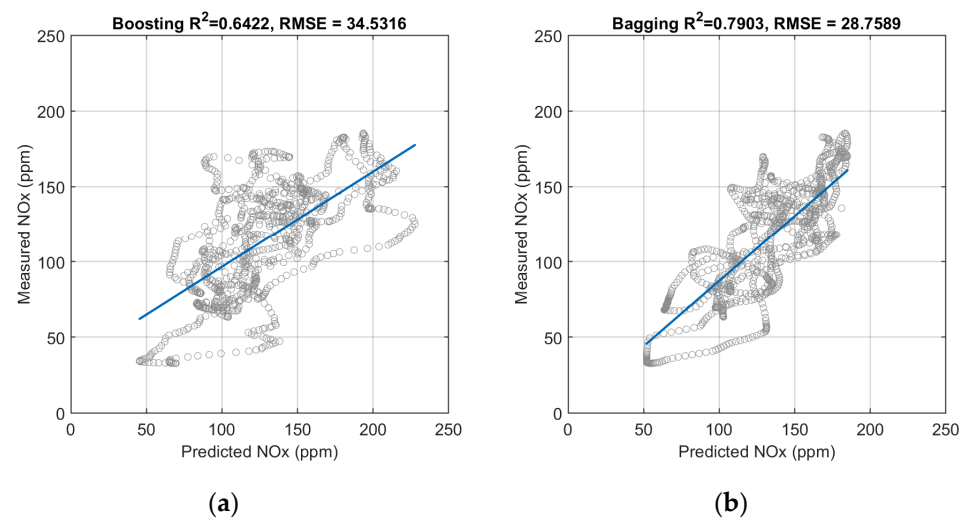


Figure 14. Relations between measured and predicted values of NOx (ppm) by two methods: (a) boosting (linear trend (blue line) $y = 0.87x - 0.72$); and (b) bagging (linear trend (blue line) $y = 0.63x - 33.24$) for the second dataset.

The obtained RMSE value was less than sensor accuracy but of the same order. The matter is that the model learning process addresses rather certain patterns of the signal than statistical metrics. Sensor recordings are needed to input such patterns into the model. Therefore, prediction models under certain conditions may have statistical metrics comparable with a sensor. Moreover, the sensor has a certain time delay due to gas transportation in the exhaust system but the model deals with the main original parameters causing the volumes of NOx emission (acceleration, engine rotations, etc.), which was accounted for during model learning.

The implementation of software is usually a less costly solution than sensor installation on the vehicle, or its integration into onboard ECU and monitoring system with subsequent sensors maintenance or replacement because of frequent failures due to high temperatures in the underground mines.

In general, ensemble methods are not widely represented in the literature, especially for NOx emission prediction of LHD diesel vehicles in deep underground mines. In our previous studies of NOx emissions from heavy-duty diesel engines [29,30], we used other methods for NOx value prediction, namely, ANN-based regression and multivariate polynomial model, respectively. However, the current paper represents better prediction accuracy with fewer input variables and for bigger datasets. This is an outstanding result for the highly dynamic working conditions of LHD vehicles.

Overall, the proposed approach of using bagging and boosting methods as ‘soft-sensors’ for NOx emissions showcases the potential for significantly improving safety, planning, and operational efficiency in underground mining. The accurate prediction of harmful gas concentrations empowers mining companies to make informed decisions, implement better ventilation management, and create a safer working environment for their personnel.

The NOx emissions depend on the engine model, fuel composition, and design of the vehicle equipped with a specific transmission. However, we developed an approach and procedure, that can be applied to any diesel LHD vehicle in the deep underground mines but the model needs rebuilding and learning again on a sufficient dataset from other environments and for vehicles of interest.

We investigated how the operator’s manner of driving influences NOx emission. This factor affects even more than possible working parameters’ deviations during the considered period of data recording.

Regarding the fuel influence, all diesel fuel supplied to the mine is certified and checked for chemical composition and its deviation. Hence, this factor can be neglected; but, of course, for other mines, vehicles, engines, and fuels, e.g., biodiesel, our model should be updated and learned on new datasets.

We did not investigate the influence of vehicle and engine technical conditions on the accuracy (bias) in NOx prediction by the developed method since the maintenance process is well organized in this mine and engines are regularly checked on-site. Probably, models need adaptation parameters, which will account for the current stage of the vehicle and engine deterioration during exploitation, but this is a long-term perspective for our next research.

4. Conclusions

The innovation of the presented research is in the following:

- (1) The conducted statistical analysis and selection made of diesel engine LHD vehicles’ working parameters, which are most suitable for NOx emission prediction in deep underground mines.
- (2) The development of the procedure of source data cleaning and processing, NOx prediction model building, and analyzing factors which may influence the accuracy of prediction.
- (3) The comparison of two ensemble methods and showing their advantages and limitations for this specific engineering application, which was not previously reported in the literature.

The implementation of the proposed models improves knowledge about the working environment and harmful gases in the atmosphere within underground mines in scenarios involving LHD vehicles equipped with diesel engines. The absence of NOx sensors in the majority of these vehicles poses a considerable challenge. However, the innovative approach put forth in this study holds the potential to markedly improve working conditions by accurately predicting the concentration of harmful gases.

Conducted research recognized that achieving a better fit to the training data does not universally translate to superior predictive capabilities. The bagging algorithm is more robust in handling changing input signals, making it a reasonable choice, especially considering the diverse experiences and operational styles of different vehicle operators, which can impact NOx emission levels.

The achieved accuracy differs for two working shifts when the same LHD vehicle was operated by two operators. The first operator had a more explicit cyclic manner of driving, and the bagging method showed the $R^2 = 0.9367$ (0.7903), and $RMSE = 19.7199$ (28.7589) parameters for the first (second) dataset, respectively. These results reflect specific patterns in the NOx data that will be accounted for during the implementation of this model in practice.

Furthermore, the significance of the acquired information cannot be overstated. Integrating this predictive data with the ventilation system of the underground mines offers an adaptive approach to enhance working conditions. By optimizing ventilation strategies based on accurate NOx emission predictions, mining companies can achieve two-fold benefits: ensuring a safer environment for underground workers and simultaneously reducing air pollution.

In conclusion, the proposed models offer a promising pathway to unraveling the complexities of NOx emissions in underground mining. By addressing the challenges, such as the lack of NOx sensors and leveraging ensemble techniques, the study not only enhances understanding of the working atmosphere but also provides actionable insights for safer and more efficient mining operations. The integration of this predictive approach into ventilation systems management could pave the way for enhanced safety, improved planning, and reduced environmental impact in underground mining.

Author Contributions: Conceptualization, A.B. and R.Z.; methodology, M.K.; software, M.K.; validation, P.Ś. and P.K.; formal analysis, P.K., P.Ś., and R.Z.; investigation, A.B. and P.Ś.; resources, R.Z.; data curation, P.Ś. and P.K.; writing—original draft preparation, A.B., M.K., and P.K.; writing—review and editing, M.K. and P.K.; visualization, M.K.; supervision, R.Z.; project administration, A.B.; funding acquisition, R.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the European Institute of Innovation and Technology (EIT), a body of the European Union, under the Horizon Europe, the EU Framework Programme for Research and Innovation. EIT RawMaterials GmbH under Framework Partnership Agreement No. 21119 (VOT3D—Ventilation Optimizing Technology based on 3D-scanning).

Data Availability Statement: Data are unavailable due to privacy restrictions.

Conflicts of Interest: Author Paweł Śliwiński was employed by the company KGHM Polska Miedź S.A. The remaining 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.

References

1. Zhang, H.; Li, B.; Karimi, M.; Saydam, S.; Hassan, M. Recent Advancements in IoT Implementation for Environmental, Safety, and Production Monitoring in Underground Mines. *IEEE Internet Things J.* **2023**, *10*, 14507–14526. [\[CrossRef\]](#)
2. Zawadzka-Malota, I. Testing of mining explosives with regard to the content of carbon oxides and nitrogen oxides in their detonation products. *J. Sustain. Min.* **2015**, *14*, 173–178. [\[CrossRef\]](#)
3. Acuna, E.I.; Lowndes, I.S. A review of primary mine ventilation system optimization. *Interfaces* **2014**, *44*, 163–175. [\[CrossRef\]](#)
4. Oluwoye, I.; Dlugogorski, B.Z.; Gore, J.; Oskierski, H.C.; Altarawneh, M. Atmospheric emission of NOx from mining explosives: A critical review. *Atmos. Environ.* **2017**, *67*, 81–96. [\[CrossRef\]](#)
5. Trzop, K.; Palka, D.; Szurgacz, D.; Sobik, L. Preventive measures to minimize natural hazards in the Polish underground mining industry. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *684*, 012009. [\[CrossRef\]](#)
6. Cao, Y.; Ji, H.; Pei, F. Research on vertical change trend of environmental indicators of mining face and its assessment in gold mine. *Min. Sci.* **2017**, *24*, 99–115. [\[CrossRef\]](#)
7. Zhang, H.; Fava, L.; Cai, M.; Vayenas, N.; Acuña, E. A hybrid methodology for investigating DPM concentration distribution in underground mines. *Tunn. Undergr. Space Technol.* **2021**, *115*, 104042. [\[CrossRef\]](#)
8. Krot, P.; Zimroz, R.; Sliwinski, P.; Gomolla, N. Safe Operation of Underground Mining Vehicles Based on Cyclic Fatigue Monitoring of Powertrains. In *Structural Integrity and Fatigue Failure Analysis*; VCMF 2020. Structural Integrity; Lesiuk, G., Szata, M., Blazejewski, W., Jesus, A.M.d., Correia, J.A., Eds.; Springer: Cham, Switzerland, 2022; Volume 25, pp. 283–292. [\[CrossRef\]](#)
9. Shin, S.; Lee, Y.; Kim, M.; Park, J.; Lee, S.; Min, K. Deep neural network model with Bayesian hyperparameter optimization for prediction of NOx at transient conditions in a diesel engine. *Eng. Appl. Artif. Intell.* **2022**, *94*, 103761. [\[CrossRef\]](#)
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. Meira, E.; Oliveira, F.L.C.; de Menezes, L.M. Forecasting natural gas consumption using Bagging and modified regularization techniques. *Energy Econ.* **2022**, *106*, 105760. [CrossRef]
12. Lv, Y.; Liu, J.; Yang, T.; Zeng, D. A novel least squares support vector machine ensemble model for NOx emission prediction of a coal-fired boiler. *Energy* **2013**, *55*, 319–329. [CrossRef]
13. Park, M.H.; Hur, J.J.; Lee, W.J. Prediction of oil-fired boiler emissions with ensemble methods considering variable combustion air conditions. *J. Clean. Prod.* **2022**, *375*, 134094. [CrossRef]
14. Andersen, Z.J.; Hvidberg, M.; Jensen, S.S.; Ketzel, M.; Loft, S.; Sørensen, M.; Tjønneland, A.; Overvad, K.; Raaschou-Nielsen, O. Chronic obstructive pulmonary disease and long-term exposure to traffic-related air pollution: A cohort study. *Am. J. Respir. Crit. Care Med.* **2011**, *183*, 455–461. [CrossRef]
15. Shaw, S.; Van Heyst, B. An evaluation of risk ratios on physical and mental health correlations due to increases in ambient nitrogen oxide (NOx) concentrations. *Atmosphere* **2022**, *13*, 967. [CrossRef]
16. Subramaniam, N.; Yusof, N. Modelling of CO₂ emission prediction for dynamic vehicle travel behavior using ensemble machine learning technique. In Proceedings of the 2021 IEEE 19th Student Conference on Research and Development (SCoReD), Kota Kinabalu, Malaysia, 23–25 November 2021; pp. 383–387. [CrossRef]
17. Abdelsalam, E.M.; Mohamed, Y.M.A.; Abdelkhalik, S.; Nazer, H.A.E.; Attia, Y.A. Photocatalytic oxidation of nitrogen oxides (NOx) using Ag- and Pt-doped TiO₂ nanoparticles under visible light irradiation. *Environ. Sci. Pollut. Res.* **2020**, *27*, 35828–35836. [CrossRef]
18. Acharya, R.; Alam, M.; Boehman, A. *Fuel and System Interaction Effects on Urea-SCR Control of NOx in Diesel Exhaust Aftertreatment*; SAE Technical Paper; SAE: Warrendale, PA, USA, 2006; 2006-01-0638. [CrossRef]
19. Funk, S. Real world NOx sensor accuracy assessment and implications for REAL NOx tracking. *SAE Int. J. Adv. and Curr. Prac. Mobil.* **2021**, *3*, 2761–2769. [CrossRef]
20. Irfan, M.F.; Goo, J.H.; Kim, S.D. Co₃O₄ based catalysts for NO oxidation and NOx reduction in fast SCR process. *Appl. Catal. B Environ.* **2008**, *78*, 267–274. [CrossRef]
21. Halim, A. Ventilation requirements for diesel equipment in underground mines—Are we using the correct values. In *Society for Mining, Metallurgy and Exploration 2017, Proceedings of the 16th North American Mine Ventilation Symposium, Golden, CO, USA, 17–22 June 2017*; Brune, J.F., Ed.; Volume 27, pp. 1–7. Available online: <http://www.diva-portal.org/smash/get/diva2:1134809/ATTACHMENT01.pdf> (accessed on 1 February 2024).
22. Mabahwi, N.A.B.; Leh, O.L.H.; Omar, D. Human health and wellbeing: Human health effect of air pollution. *Procedia-Soc. Behav. Sci.* **2014**, *153*, 221–229. [CrossRef]
23. Agarwal, A.K.; Mustafi, N.N. Real-world automotive emissions: Monitoring methodologies, and control measures. *Renew. Sustain. Energy Rev.* **2021**, *137*, 110624. [CrossRef]
24. Duan, K.; Li, Q.; Liu, H.; Zhou, F.; Lu, J.; Zhang, X.; Zhang, Z.; Wang, S.; Ma, Y.; Wang, X. Dynamic NOx emission factors for main engines of bulk carriers. *Transp. Res. Part D Transp. Environ.* **2022**, *107*, 103290. [CrossRef]
25. Krot, P.; Sliwinski, P.; Zimroz, R.; Gomolla, N. The identification of operational cycles in the monitoring systems of underground vehicles. *Measurement* **2020**, *151*, 107111. [CrossRef]
26. Arrègle, J.; López, J.J.; Guardiola, C.; Monin, C. On board NOx prediction in diesel engines: A physical approach. In *Automotive Model Predictive Control; Lecture Notes in Control and Information Sciences*; Del Re, L., Allgöwer, F., Glielmo, L., Guardiola, C., Kolmanovsky, I., Eds.; Springer: London, UK, 2010; Volume 402, pp. 25–36. [CrossRef]
27. Sadeghi, S.; Soltanmohammadlou, N.; Nasirzadeh, F. Applications of wireless sensor networks to improve occupational safety and health in underground mines. *J. Saf. Res.* **2022**, *83*, 8–25. [CrossRef]
28. López-de Abajo, L.; Alberti, M.G.; Gálvez, J.C. Pollutant concentration prediction from traffic data analysis for concrete durability studies in Madrid Calle 30 urban tunnels. *Tunn. Undergr. Space Technol.* **2024**, *144*, 105477. [CrossRef]
29. Banasiewicz, A.; Moosavi, F.; Kotyla, M.; Śliwiński, P.; Krot, P.; Wodecki, J.; Zimroz, R. Forecasting of NOx Emissions of Diesel LHD Vehicles in Underground Mines—An ANN-Based Regression Approach. *Appl. Sci.* **2023**, *13*, 9965. [CrossRef]
30. Banasiewicz, A.; Śliwiński, P.; Krot, P.; Wodecki, J.; Zimroz, R. Prediction of NOx Emission Based on Data of LHD On-Board Monitoring System in a Deep Underground Mine. *Energies* **2023**, *16*, 2149. [CrossRef]
31. Barutçuoğlu, Z.; Alpaydin, E. A comparison of model aggregation methods for regression. In *International Conference on Artificial Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 76–83. [CrossRef]
32. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
33. Freund, Y. Boosting a Weak Learning Algorithm by Majority. *Inf. Comput.* **1995**, *121*, 256–285. [CrossRef]
34. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
35. Hastie, T.; Tibshirani, R.; Friedman, J.H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: New York, NY, USA, 2009; Volume 2. [CrossRef]
36. Sharma, M.; Maity, T. Review on machine learning-based underground coal mines gas hazard identification and estimation techniques. *Arch. Comput. Methods Eng.* **2023**, *31*, 371–388. [CrossRef]
37. Mendes-Moreira, J.; Soares, C.; Jorge, A.M.; Sousa, J.F.D. Ensemble approaches for regression: A survey. *ACM Comput. Surv.* **2012**, *45*, 1–40. [CrossRef]
38. KGHM ZANAM LHD LKP-1701. Available online: https://www.kghmzanam.com/wp-content/uploads/2020/12/LKP_1701_PL.pdf (accessed on 16 December 2023).

39. DEUTZ TCD 12.0 and TCD 16.0 Diesel Engines. Available online: <https://www.deutzsupport.com/product-details/tcd-12-0-and-tcd-16-0/> (accessed on 16 December 2023).
40. *ISO 14396:2002*; Reciprocating internal combustion engines Determination and method for the measurement of engine power. ISO: Geneva, Switzerland, 2002.

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.