

Article

Interpretable Machine Learning for Predicting Blast-Induced Particulate Matter Emissions in Surface Mines

Yulin Zhang ¹, Chi Li ¹, Shun Yang ^{2,*}, Jian Zhou ¹ and Manoj Khandelwal ^{3,4,5,*}

¹ School of Resources and Safety Engineering, Central South University, Changsha 410083, China; yl_zhang@csu.edu.cn (Y.Z.); lccsuer@csu.edu.cn (C.L.); j.zhou@csu.edu.cn (J.Z.)

² School of Geosciences and Info-Physics, Central South University, Changsha 410083, China

³ Institute of Innovation, Science and Sustainability, Federation University Australia, Ballarat, VIC 3350, Australia

⁴ Laboratory of Sustainable Development in Natural Resources and Environment, Institute for Advanced Study in Technology, Ton Duc Thang University, Ho Chi Minh City 72915, Vietnam

⁵ Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City 72915, Vietnam

* Correspondence: yangshun52068@163.com (S.Y.); m.khandelwal@federation.edu.au or khandelwalmanoj@tdtu.ac.vn (M.K.)

Abstract

Blasting is an essential operation in surface mining, but it can generate high particulate matter concentrations within a short period. Accurate prediction of blast-induced dust concentration is useful for air quality management, worker protection, and dust control planning. In this study, an interpretable machine learning framework was developed to predict particulate matter with an aerodynamic diameter less than 10 μm (PM_{10}) and total suspended particulate matter (TSP) concentrations induced by blasting in a large surface coal mine. The dataset was derived from published field monitoring records and included blasting design parameters, monitoring distance, material-related variables, and measured dust concentrations. A total of 148 valid samples were used for model development and evaluation. Six tree-based ensemble models, including Extra Trees Regression, Random Forest, Gradient Boosting Regression Trees, extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), and categorical boosting (CatBoost), were established and compared. Bayesian optimization was used for hyperparameter tuning, and model performance was evaluated using the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). Under the adopted 80:20 hold-out validation scheme, CatBoost achieved the best test performance for PM_{10} prediction, with a test R^2 of 0.9555, RMSE of 444.87 $\mu\text{g}/\text{m}^3$, and MAE of 343.48 $\mu\text{g}/\text{m}^3$. Extra Trees Regression performed best for TSP prediction, with a test R^2 of 0.8803, RMSE of 3899.04 $\mu\text{g}/\text{m}^3$, and MAE of 3036.27 $\mu\text{g}/\text{m}^3$. Residual analysis further indicated that the optimal models had no obvious systematic bias. Shapley additive explanations (SHAP) analysis, supported by within-model feature importance rankings, showed that explosive quantity, number of blastholes, and monitoring distance were the dominant variables affecting PM_{10} and TSP predictions. The explosive quantity and number of blastholes mainly increased the predicted dust concentration, whereas the monitoring distance generally reduced it. The proposed framework may provide useful support for blast parameter optimization, monitoring point arrangement, and dust control decision-making in surface mines, but further validation using larger multi-site datasets is still needed.

Academic Editor: Tiago Filipe da Silva Miranda

Received: 21 May 2026

Revised: 28 June 2026

Accepted: 30 June 2026

Published: 21 July 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: blasting dust; PM₁₀; total suspended particulate matter; surface mine; machine learning; SHAP interpretation

1. Introduction

Several production processes, including stripping, drilling, blasting, loading, and transportation, are involved in surface coal mining [1,2]. Blasting is one of the most important operations because it fragments the rock mass and facilitates efficient excavation. However, it can also generate a large amount of dust within a short time. Blast-generated dust clouds contain fine particulate matter and total suspended particulate matter, which may have an impact on the surrounding environment, worker health, equipment operation, and air quality. For the purpose of managing mine air quality and planning dust control, precise forecasting of blast-induced dust concentration is crucial. Control measures such as water spraying, pre-wetting, local suppression, and blast parameter change can be set up more successfully if dust concentration can be calculated prior to or soon after blasting.

Numerous factors affect the particulate matter produced by blasting. Blast design parameters, explosive charge, rock and material properties, monitoring distance, and environmental conditions are some of these variables [3]. Because tiny particles can stay trapped in the air and be ingested by workers, PM₁₀ is a crucial indicator. TSP reflects the total amount of dust following blasting and represents a wider variety of airborne particles. Both PM₁₀ and TSP are helpful in assessing air pollution caused by blasts in surface mines. However, because associated physical processes—such as rock fracture, gas release, particle suspension, dispersion, and deposition—control dust formation and transport, their concentrations frequently differ significantly [1].

Prior research [4–8] has estimated the amount of dust in mining and industrial settings using neural network models, statistical regression models, and empirical formulae. Although empirical and regression-based approaches are straightforward and simple to use, they frequently make the assumption that input factors and dust concentration have linear or predetermined connections. When there are nonlinear relationships and interactions between variables in the data, these assumptions could be restricted. More complicated interactions can be described by neural networks, although sample size, model structure, and parameter choice may have an impact on how well they perform. Multiple regression and neural network models were created in the initial investigation by Roy et al. to evaluate blasting dust at a sizable surface coal mine. That study showed that data-driven approaches may be applied for PM₁₀ and TSP prediction and offered useful field monitoring data.

Machine learning and optimization methods have been increasingly applied to blasting-related environmental effects [9]. Meanwhile, ensemble learning models have been widely used for complex engineering and environmental regression problems [10–12]. Without making stringent assumptions about the distribution of the data, tree-based models like Random Forest, Extra Trees Regression, Gradient Boosting, XGBoost, LightGBM, and CatBoost can capture nonlinear connections and variable interactions. These models work particularly well with tabular data gathered from field surveillance. For engineering applications, model interpretability is just as crucial as prediction accuracy. Interpretable techniques can be used to determine which material characteristics and blasting settings are primarily responsible for dust emissions, as well as how these factors impact the anticipated amounts of TSP and PM₁₀.

Even while earlier research has offered a helpful foundation for forecasting blasting dust, a number of difficulties still require investigation. First, the nonlinear relationship

between blasting conditions and dust concentration might not be adequately captured by conventional regression models. Second, there is still a lack of discussion regarding the relative effectiveness of various contemporary ensemble learning models for blast-induced PM₁₀ and TSP prediction. Third, prior research has focused more on predicting outcomes than on model interpretation. Therefore, from a machine learning standpoint, the main factors influencing PM₁₀ and TSP emissions are still not fully understood. Lastly, more clarification is required about the practical implications for dust control planning and blast parameter optimisation.

The present study uses the field monitoring dataset reported by Roy et al. [1], which was collected from the Dudhichua Project of Northern Coalfields Limited in Singrauli, India. This dataset records blast-induced PM₁₀ and TSP concentrations under different blasting and monitoring conditions and provides a useful basis for evaluating data-driven prediction models for surface mine blasting dust.

This study develops an interpretable machine learning framework for forecasting blast-induced PM₁₀ and TSP concentrations in surface mines in order to solve these problems. The dataset was rearranged for machine learning analysis after being taken from the blasting dust monitoring records for a sizable surface coal mine that Roy et al. reported. Six ensemble models were developed and compared: Extra Trees Regression, Random Forest, Gradient Boosting Regression Trees, XGBoost, LightGBM, and CatBoost. R², RMSE, MAE, and MAPE were used to assess the model's performance. Additionally, the contribution and influence direction of important factors were explained using feature importance and SHAP analysis. Compared with the original regression and neural-network study based on the same field monitoring records, the present study advances the analysis in three aspects. It should be noted that this comparison is approximate because the source study and the present study used different data partitions, sample numbers after preprocessing, and performance metrics. Nevertheless, the source study provides a useful quantitative benchmark for evaluating the present framework. First, several modern tree-based ensemble models were systematically compared under the same preprocessing, data-splitting, and Bayesian optimization procedure. Second, PM₁₀ and TSP were modeled independently using the same set of blasting, monitoring-distance, and material-related inputs, and neither dust indicator was used as an input variable for predicting the other. Third, SHAP analysis was used to explain not only which variables were important, but also how they influenced model predictions. This provides more practical information for blast parameter optimization and dust control planning than prediction accuracy alone. The main objective of this study is to provide a reliable prediction framework and practical decision support for monitoring point arrangement, blast-parameter optimization, and blast-induced dust control in surface mining.

2. Materials and Methods

2.1. Data Source

The dataset used in this study was obtained from the blasting dust monitoring records reported by Roy et al. [1]. The original records include measured particulate matter concentrations and corresponding blasting, monitoring, and material-related parameters from a large surface coal mine. After data verification and preprocessing, 148 valid samples were retained for model development and evaluation.

According to the original field monitoring study [1], particulate matter concentrations were measured using respirable dust samplers (Instrument IPM 115BL). PM₁₀ was collected after coarse particles greater than 10 µm were separated by a cyclone, whereas TSP was collected using a glass microfiber filter. The sampling flow rate was maintained at 2.5 L min⁻¹, and the particulate matter mass was determined from the difference between the initial and final filter weights using a Sartorius balance with a sensitivity of

0.0001 g. Moisture content was determined in the Coal Analysis Laboratory of the mine using samples collected from coal and overburden benches, and it was calculated as the ratio of water weight to dry material weight. Rock density was measured using collected bench samples based on the weight-and-volume method in the laboratory. The present study did not conduct new field measurements; instead, it reorganized these published measurements for machine learning analysis.

Two dust indicators were considered as prediction targets: particulate matter with an aerodynamic diameter less than 10 μm (PM_{10}) and total suspended particulate matter (TSP). These two targets were modelled separately, and neither PM_{10} nor TSP was used as an input feature for predicting the other target. The input variables included blasting design parameters, monitoring distance, and material-related variables. Specifically, the input variables consisted of the number of blastholes, blasthole depth, blasthole diameter, burden, spacing, stemming column, explosive quantity, monitoring distance, moisture content, rock density, and rock type.

The original records were reorganized into a unified tabular dataset for machine learning analysis. The variable names, units, data types, and physical meanings are summarized in Table 1. PM_{10} and TSP were not used as input variables in the same prediction task. Instead, two independent regression tasks were established, one for PM_{10} prediction and the other for TSP prediction.

Table 1. Description of input and output variables.

Variable Type	Variable	Abbreviation	Unit	Data Type	Description
Input	Number of blastholes	NO	—	Numerical	Number of blastholes detonated in a blasting round. It reflects the blasting scale and potential dust generation source intensity.
Input	Blasthole depth	HD	m	Numerical	Depth of the blastholes used in the blast design. It is related to the blasting volume and rock-breaking range.
Input	Blasthole diameter	BD	mm	Numerical	Diameter of the blastholes. It affects the charge structure and energy release during blasting.
Input	Burden	B	m	Numerical	Distance from a blasthole to the nearest free face. It influences rock fragmentation, gas escape, and dust generation.
Input	Spacing	S	m	Numerical	Distance between adjacent blastholes in the same row. It affects the distribution of explosive energy and fragmentation characteristics.
Input	Stemming column	SC	m	Numerical	Length of stemming material placed in the upper part of the blasthole. It controls explosive gas confinement and may influence dust emission from the collar area.
Input	Explosive quantity	EQ	kg	Numerical	Total explosive charge used in a blasting round. It represents the energy input and is closely associated with dust generation intensity.
Input	Distance	D	m	Numerical	Distance between the blasting site and the dust monitoring sampler. It reflects the spatial attenuation and dispersion of particulate matter.
Input	Moisture content	MC	%	Numerical	Water content of the blasted material, determined in the original study using collected coal and overburden samples in the mine laboratory.

Input	Rock density	RD	g/cm ³	Numerical	Density of the blasted rock or material, measured in the original study using collected bench samples based on the weight-and-volume method.
Input	Rock type	RT	—	Categorical	Lithological type of the blasted material, such as coal or overburden. It represents differences in material composition and breakage behavior.
Output	Particulate matter with aerodynamic diameter less than 10 µm	PM ₁₀	µg/m ³	Numerical	Concentration of fine inhalable particulate matter generated by blasting. It is one of the prediction targets in this study.
Output	Total suspended particulate matter	TSP	µg/m ³	Numerical	Concentration of total suspended particulate matter generated by blasting. It is another prediction target in this study.

2.2. Input and Output Variables

The input variables were selected according to the original monitoring records and their relevance to blast-induced dust generation and transport. They were grouped into three categories: blasting design parameters, monitoring-distance information, and material-related variables. The blasting design parameters included the number of blastholes, blasthole depth, blasthole diameter, burden, spacing, stemming column, and explosive quantity. Monitoring distance represented the distance between the blasting source and the dust sampling point. The material-related variables included moisture content, rock density, and rock type.

The output variables were PM₁₀ and TSP concentrations. PM₁₀ represents fine inhalable particulate matter, whereas TSP represents the overall suspended particulate matter concentration after blasting. The two targets were predicted independently. The detailed definitions, abbreviations, units, data types, and descriptions of all variables are provided in Table 1.

2.3. Data Preprocessing

Before model development, the original blasting dust dataset was checked and reorganized to make it suitable for machine learning analysis. Blasting design parameters, monitoring distance, material-related variables, rock type, and two dust concentration indicators (PM₁₀ and TSP) are all included in the dataset. In order to make the variable names and units consistent with the definitions provided in Table 1, these variables were first organised in a single tabular format. In addition to preparing the dataset for later model training, this step was required to minimise potential errors brought on by inconsistent data labels. The detailed data preprocessing workflow is shown in Figure 1.

After data organization, anomalous records and missing values were examined. Records with missing key information were removed from the modeling dataset. Machine learning modelling and statistical analysis were then performed on the remaining samples. There were significant differences in a number of variables, particularly explosive quantity, monitoring distance, PM₁₀ concentration, and TSP concentration, because the dataset was gathered from field monitoring under various blasting and environmental conditions. These variations were retained because they correspond to actual field monitoring conditions rather than data-entry errors. As a result, the target variables were not artificially smoothed.

In the original dataset, rock type was a categorical variable, including coal and overburden. To ensure a consistent input format across all compared models, rock type was encoded using integer labels. The encoded values were used only as category identifiers

and were not interpreted as measured physical quantities. In the present implementation, the same encoded input matrix was used for all models, including CatBoost, so that the performance comparison was based on the same preprocessing procedure. Therefore, the CatBoost result in this study reflects the performance of the CatBoost algorithm under the unified encoded-feature setting, rather than an additional advantage from a different categorical preprocessing strategy.

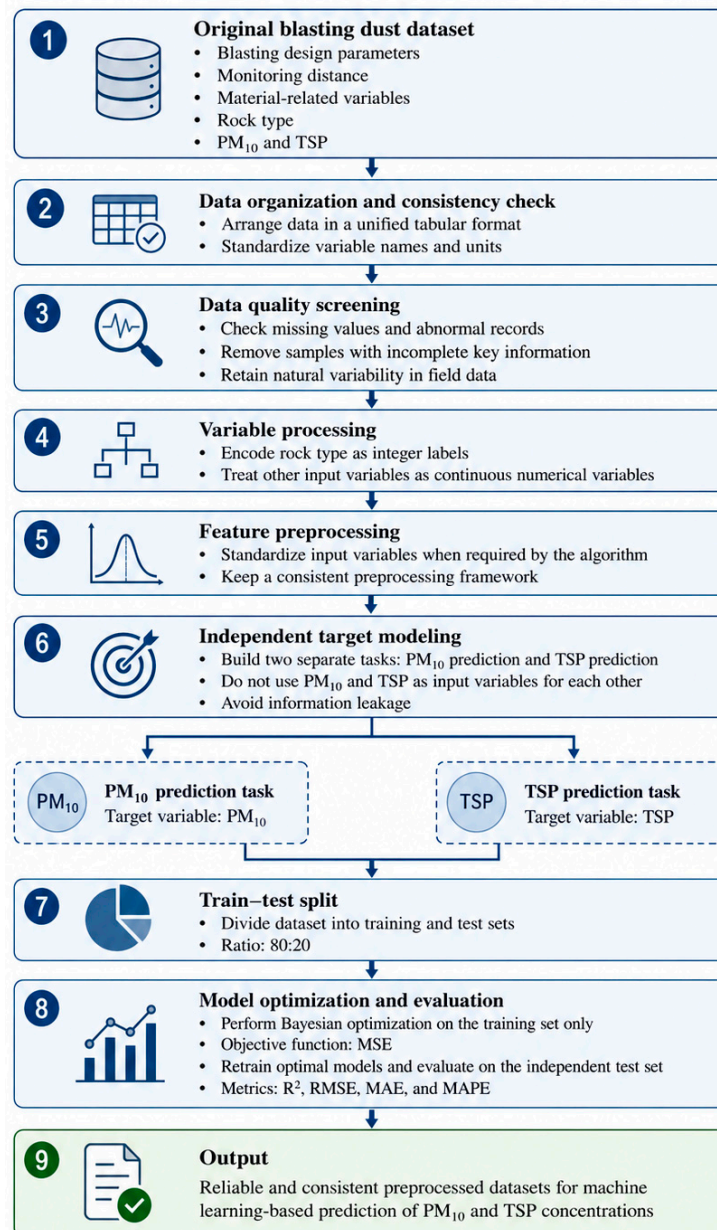


Figure 1. Flowchart of the data preprocessing procedure for PM₁₀ and TSP prediction.

To reduce the effect of different variable scales, the input variables were standardised before model training, where required by the algorithm. Standardisation was especially important for models that are sensitive to magnitude fluctuations. The same preprocessing approach was used to ensure consistency across models, despite the fact that tree-based models are less vulnerable to feature scaling. The output variables, PM₁₀ and TSP, were not included in the input feature set. Two separate regression tasks were developed,

one for PM₁₀ prediction and the other for TSP prediction. This design allowed the predictive behaviour of fine particulate matter and total suspended particulate matter to be evaluated independently under the same input-variable structure.

The dataset was divided into training and test sets at an 80:20 ratio. The training set was used for model fitting and hyperparameter optimisation, while the test set was kept apart and used just for the final performance evaluation. Each target variable had the same input variable structure, but the output variable changed according to the prediction task. The PM₁₀ and TSP models were thus developed using comparable data settings.

To find suitable hyperparameters within predefined parameter ranges for model optimisation, Bayesian optimisation was used. The optimisation process was restricted to the training set, and the goal function was the mean squared error. Once the optimal hyperparameters were discovered, each model was retrained using the training data and evaluated on a different test set. The final model performance was assessed using R², RMSE, MAE, and MAPE, as described in Section 2.5.

Preprocessing included train-test separation, feature standardisation, data verification, categorical variable encoding, and independent target modelling. Using the dataset from the first blasting dust monitoring research, these methods enabled reliable and consistent machine learning forecasts of PM₁₀ and TSP concentrations.

Considering that some input variables were correlated with each other, a correlation-aware check was conducted before model training. The correlation matrix was used only to understand the dependence structure among input variables and was not used to select the target-related variables. All retained predictors were measurable blasting design, monitoring-distance, or material-related variables available before dust concentration prediction. PM₁₀ and TSP were modeled as two independent regression targets, and the target variable of one task was not included in the input feature set of the other task.

Although several input variables showed relatively high pairwise correlations, they were not removed automatically because they represent different engineering meanings. For example, burden and spacing jointly describe the blasthole layout, while blasthole diameter, rock density, and rock type characterize different aspects of charge structure and material properties. In addition, tree-based ensemble models are less sensitive to multicollinearity than ordinary linear regression models because they perform nonlinear splitting and variable selection during tree construction.

2.4. Machine Learning Models

This study built and compared six machine learning algorithms to predict particulate matter emissions caused by blasts. Extra Trees Regression (ETR), Random Forest (RF), Gradient Boosting Regression Trees (GBRT), XGBoost, LightGBM, and CatBoost are some of these models [13–17]. They were chosen because tree-based ensemble models can characterise nonlinear relationships between input variables and target responses and are useful for tabular engineering datasets. PM₁₀ and TSP were predicted using the same input variables in this investigation, but the two target variables were modelled independently.

RF and ETR are ensemble models based on bagging. To increase prediction stability, they produce a lot of decision trees and integrate their outcomes. While ETR adds more randomisation to the tree-splitting process, RF constructs each tree using bootstrap samples and random feature selection. In particular, when the dataset includes nonlinear correlations and variable interactions, this can lower model variance and enhance generalisation.

Boosting-based ensemble models include GBRT, XGBoost, LightGBM, and CatBoost. Boosting models construct trees in a sequential manner as opposed to bagging models. Every new tree is trained to minimise the mistakes made by its predecessors. While XGBoost and LightGBM increase computing efficiency and regularisation capacity, GBRT is a traditional boosting technique. CatBoost is another GBRT technique that works well

with categorical features. For a fair comparison, all models in this investigation were trained using the same data split, and rock type was encoded as a categorical-related input variable during preprocessing.

Three important hyperparameters were chosen for optimisation for every model. This configuration was chosen to strike a compromise between model complexity and performance. Particularly for a very small field-monitoring dataset, an excessive number of hyperparameters may raise computing costs and overfitting risk. On the training set, appropriate hyperparameter combinations were found via Bayesian optimisation. Bayesian optimization was implemented using the BayesianOptimization function from the bayesian-optimization Python package 3.3.0. For each model and each prediction target, the optimization was conducted separately on the training set. The objective function was the mean squared error obtained from the training procedure, and lower MSE was considered better. The search bounds for all optimized hyperparameters are listed in Table 2. For each optimization task, 5 random initial evaluations were first performed, followed by 500 Bayesian optimization iterations. To improve reproducibility, a fixed random seed of 42 was used where applicable. Hyperparameters not listed in Table 2 were kept at their default values in the corresponding Python libraries. The test set was only retained for the final model evaluation and was not utilised during hyperparameter tuning.

Table 2. Hyperparameter search bounds and optimized values used in Bayesian optimization.

Model	Hyperparameter	Search Space	Best Value for PM ₁₀	Best Value for TSP
ETR	n_estimators	50–500	84	342
ETR	max_depth	2–30	28	8
ETR	min_samples_leaf	1–10	3	2
RF	n_estimators	50–500	187	116
RF	max_depth	2–30	13	28
RF	min_samples_leaf	1–10	2	3
GBRT	n_estimators	50–500	210	498
GBRT	learning_rate	0.005–0.300	0.0379	0.0154
GBRT	max_depth	2–8	2	2
XGBoost	n_estimators	50–800	212	763
XGBoost	learning_rate	0.005–0.300	0.0216	0.0068
XGBoost	max_depth	2–8	6	3
LightGBM	n_estimators	50–500	290	431
LightGBM	learning_rate	0.005–0.300	0.1708	0.2063
LightGBM	num_leaves	10–80	11	69
CatBoost	iterations	50–1000	821	245
CatBoost	learning_rate	0.005–0.300	0.0223	0.0706
CatBoost	depth	2–10	3	2

Note: The search bounds were used for Bayesian optimization separately for PM₁₀ and TSP prediction.

Following optimisation, the training set with the optimal hyperparameter combination was used to retrain each model. The prediction performance of the trained models was then assessed by applying them to the independent test set. The prediction accuracy of various models was compared using four metrics: R², RMSE, MAE, and MAPE. In order to properly evaluate the model's applicability for various dust concentration targets, the final comparison was carried out independently for PM₁₀ and TSP.

All computations were performed using Python 3.9.21. Data processing and numerical calculation were performed using pandas 2.2.3, NumPy 2.0.2, and SciPy 1.13.1. Extra Trees Regression, Random Forest, and Gradient Boosting Regression Trees were imple-

mented using scikit-learn 1.6.1. XGBoost, LightGBM, and CatBoost were implemented using xgboost 2.1.4, lightgbm 4.6.0, and catboost 1.2.8, respectively. Bayesian hyperparameter optimization was implemented using the bayesian-optimization package 3.3.0. SHAP analysis was performed using SHAP 0.47.1, and figure generation was completed using Matplotlib 3.9.4. The programs used in this study were written by the authors, and no external unpublished program or third-party code from the literature was directly used. A fixed random seed of 42 was used where applicable to improve reproducibility.

2.5. Model Evaluation Metrics

Four widely used regression metrics—coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE)—were employed in this study to assess the prediction performance of various machine learning models [18–20]. Because they can characterise model accuracy from many angles, these measures were chosen. While RMSE, MAE, and MAPE assess prediction errors in various ways, R^2 indicates the goodness of fit. To guarantee a fair comparison across models, the same metrics were applied to both PM_{10} and TSP prediction tasks.

The degree to which the predicted values explained the variation in the measured values was assessed using the coefficient of determination, or R^2 . Better model performance is indicated by a higher R^2 value. The model can account for the majority of the variation in the target variable when R^2 is near 1. The primary metric utilised in this study to compare the overall fitting capabilities of various models was R^2 .

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

where y_i is the measured value, $\hat{y}_i$ is the predicted value, $\bar{y}$ is the mean value of the measured data, and n is the number of samples.

The average magnitude of prediction errors was assessed using RMSE. Large errors are given more weight by RMSE since it squares the error before averaging. This measure can be used to determine whether a model generates significant deviations for certain samples. Better forecast accuracy is shown by a lower RMSE score.

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

The average absolute difference between measured and anticipated values was also calculated using MAE. MAE is less susceptible to very big errors than RMSE. As a result, it gives a clearer explanation of the typical prediction mistake. The projected values are closer to the measured values when the MAE value is lower.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

The prediction error was expressed as a percentage of the measured value using MAPE. When evaluating relative errors between several models and target variables, this statistic is useful. However, rather than being the sole assessment metric, MAPE was utilised in conjunction with RMSE and MAE because it is impacted by very small measured values. It should also be noted that the dust concentration data in this study are strongly right-skewed, with many low-concentration samples and a small number of high-concentration samples. Under this condition, R^2 and absolute-error metrics such as RMSE and MAE are more influenced by the ability of the model to fit medium- and high-concentration samples, whereas MAPE is more sensitive to relative errors in low-concentration sam-

ples. Therefore, a model may achieve a high R^2 while still showing a relatively high MAPE. In this study, MAPE was retained only as a supplementary relative-error metric, and the final model comparison was based on the combined interpretation of R^2 , RMSE, MAE, MAPE, residual analysis, and concentration-band error analysis.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (4)$$

In general, a model with a higher R^2 and lower RMSE, MAE, and MAPE was considered to have better predictive performance. For each target variable, the models were evaluated on both the training set and the independent test set. The training results were used to examine model fitting ability, while the test results were used to assess generalization performance. This evaluation strategy helped identify the most suitable model for PM_{10} and TSP prediction under the present blasting dust dataset.

2.6. Model Interpretation Method

While machine learning algorithms can increase the prediction accuracy of particulate matter emissions caused by blasts, it is frequently challenging to directly comprehend their internal decision-making process [14]. Prediction accuracy alone is insufficient for engineering applications. Determining whether blasting and material factors have a greater impact on PM_{10} and TSP concentrations is also crucial. In order to increase the transparency of the prediction findings, model interpretation was carried out following model training.

Two approaches of interpretation were employed in this study: SHAP analysis and feature importance analysis. The relative contribution of each input variable in the best prediction models was first assessed using feature importance. This approach helps determine the primary factors influencing dust emissions and offers a straightforward ranking of the variables. Higher significance variables make a greater contribution to the model's forecast. This allowed for the comparison of the effects of several blasting-related variables, such as stemming column, moisture content, monitoring distance, and explosive quantity.

Feature importance, however, simply provides a global rating of the variables. It should be noted that the feature importance values used in tree-based models are model-specific. Because the optimal PM_{10} and TSP models were different model families, namely CatBoost for PM_{10} and Extra Trees Regression for TSP, the feature importance percentages should be interpreted mainly as within-model rankings rather than as directly comparable quantities across the two prediction targets. Therefore, cross-target interpretation in this study was based primarily on SHAP results, while feature importance was used as supporting evidence within each optimal model. It is difficult to determine if a variable raises or lowers the anticipated dust concentration. SHAP analysis was established to provide additional insight into the distribution and direction of variable effects. Based on the Shapley value principle, SHAP gives each feature a contribution value for every prediction. The anticipated PM_{10} or TSP concentration is increased by a feature with a positive SHAP value, whereas the expected value is decreased by a feature with a negative SHAP value.

The correlation analysis and model performance results were utilised in conjunction with the interpretation outcomes. Because several input variables were correlated, the feature importance and SHAP results were interpreted as model-based predictive contributions rather than independent causal effects. In correlated feature groups, such as burden-spacing and blasthole diameter-rock density, the model may distribute importance among related variables. Therefore, the interpretation results should be used to identify dominant predictive patterns, but not to quantify the isolated causal effect of each individual variable. While the machine learning models identified nonlinear linkages and feature interactions, the correlation analysis offered initial statistical associations among variables. Each variable's function in the prediction process was further clarified by feature

importance and SHAP analysis. As a result, the interpretation techniques offered helpful information for engineering decision-making and helped link the data-driven models with realistic blasting dust control.

3. Results

3.1. Descriptive Statistics of the Dataset

Table 3 provides a summary of the blasting dust dataset's descriptive statistics. A total of 148 valid samples were selected for analysis and model construction following data verification and preprocessing. Blast design parameters, monitoring distance, material-related variables, and two dust concentration indicators—PM₁₀ and TSP—are all included in the dataset. In this study, these factors were rearranged for machine learning prediction based on the blasting dust monitoring records of a major surface coal mine.

Table 3. Descriptive statistics of variables.

Variable	Abbreviation	Unit	Minimum	Maximum	Mean	Standard Deviation
Number of blastholes	NO	—	4.00	94.00	19.47	15.46
Blasthole depth	HD	m	6.00	20.00	11.75	3.96
Blasthole diameter	BD	mm	100.00	269.00	185.86	53.98
Burden	B	m	3.00	10.00	7.60	1.10
Spacing	S	m	3.00	10.00	7.99	1.16
Stemming column	SC	m	4.00	9.50	6.29	1.30
Explosive quantity	EQ	kg	481.00	56,915.00	5898.34	10,639.37
Distance	D	m	15.00	160.00	72.51	41.29
Moisture content	MC	%	3.53	17.11	11.72	4.49
Rock density	RD	g/cm ³	1.54	2.96	1.79	0.26
Rock type	RT	—	1.00	2.00	1.31	0.46
PM ₁₀ concentration	PM ₁₀	µg/m ³	57.00	12,890.00	1720.78	1889.61
Total suspended particulate matter	TSP	µg/m ³	602.00	43,213.00	9636.59	9120.84

Note: RT is a categorical variable encoded numerically for model development, where 1 and 2 represent different rock types, namely coal and overburden. PM₁₀ and TSP are output variables and were modeled separately.

The input variables exhibit distinct variations in their numerical ranges, as Table 3 illustrates. The average number of blastholes is 19.47, with a range of 4 to 94. The explosive quantity varies more significantly, with a standard deviation of 10,639.37 kg and a range of 481.00 kg to 56,915.00 kg. This wide range also indicates that a small number of very large blasting rounds are included in the dataset. These cases may have an important influence on both the correlation between explosive quantity and dust concentration and the variance-based performance metric R². In the present study, these samples were retained because they are valid field observations and represent real high-energy blasting conditions that are important for dust control assessment. Therefore, they were not treated as data-entry errors or statistical outliers.

There is significant diversity in the output variables as well. With a mean of 1720.78 µg/m³ and a standard deviation of 1889.61 µg/m³, the PM₁₀ concentration falls between 57.00 µg/m³ and 12,890.00 µg/m³. With a mean of 9636.59 µg/m³ and a standard deviation of 9120.84 µg/m³, the TSP concentration ranges from 602.00 µg/m³ to 43,213.00 µg/m³. TSP has a larger distribution range and a significantly higher mean value than PM₁₀. This makes sense as TSP, which stands for total suspended particulate matter, has a larger particle-size proportion than PM₁₀.

The dispersion features of PM₁₀ and TSP concentrations are also displayed in Figure 2. Both PM₁₀ and TSP have right-skewed distributions, according to the histograms in Figure

2a,b. Only a few samples exhibit extremely high dust concentrations; the majority of samples are concentrated in the low-concentration range. PM₁₀ has a mean of 1720.78 µg/m³ and a median of 1260.50 µg/m³. The median value for TSP is 7338.50 µg/m³, and the mean value is 9636.59 µg/m³. The right-skewed nature of the dust concentration data is confirmed by the fact that the mean values for both objectives are higher than the median values.

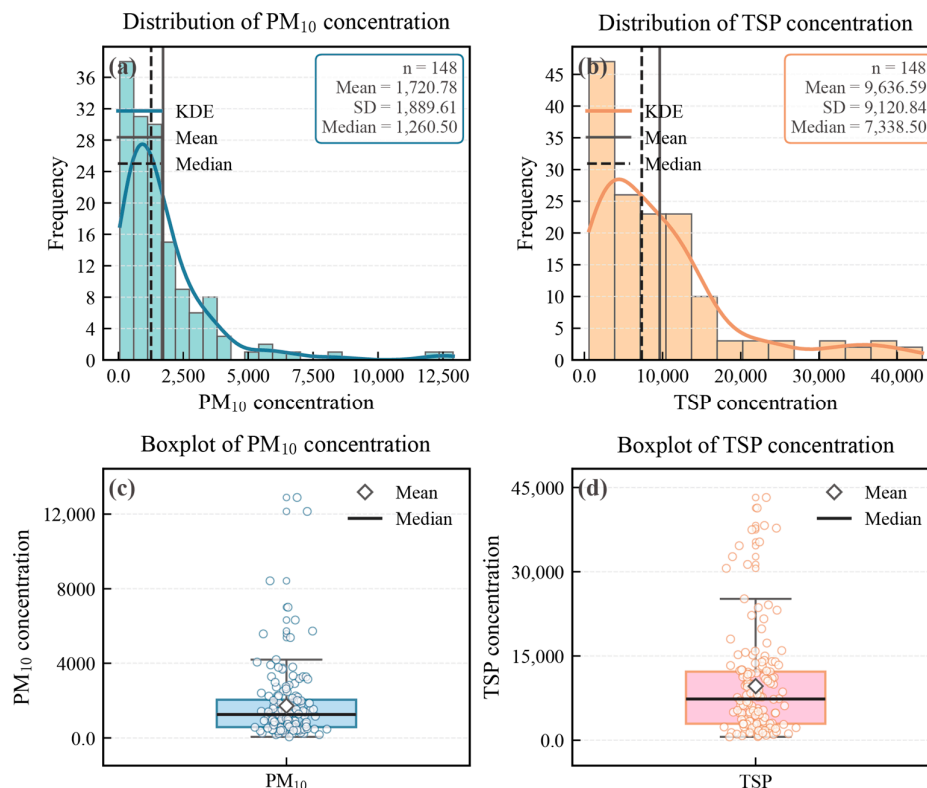


Figure 2. Distribution characteristics of PM₁₀ and TSP concentrations: (a) histogram of PM₁₀ concentration, (b) histogram of TSP concentration, (c) boxplot of PM₁₀ concentration, and (d) boxplot of TSP concentration.

Additional proof of this unequal distribution may be found in the boxplots in Figure 2c,d. For both PM₁₀ and TSP, there are a number of high-value samples, particularly in the upper concentration range. Large explosive volumes, short monitoring distances, low moisture content, or particular blasting conditions could all be associated with these high-concentration sites. Since these samples reflect actual field monitoring settings and are crucial for training algorithms to identify high-dust emission cases, they were kept in the dataset rather than eliminated.

3.2. Correlation and Preliminary Mechanism Analysis

To preliminarily examine the relationships among blasting parameters, material-related variables, and dust concentrations, Spearman rank correlation analysis was conducted, and the results are presented in Figure 3. This nonparametric method was adopted because the distribution plots in Figure 2 show that PM₁₀ and TSP exhibit noticeable right-skewed characteristics, indicating that the normality assumption required by Pearson correlation may not be fully satisfied. Compared with Pearson correlation, Spearman correlation is more suitable for evaluating monotonic associations in skewed field-monitoring data and is less sensitive to extreme values. The statistical significance of each correlation coefficient was further evaluated using a two-tailed *p*-value, and the significance levels are marked in Figure 3. It should be

noted that the correlation analysis was used only as an exploratory statistical tool before machine learning modelling and does not imply causality.

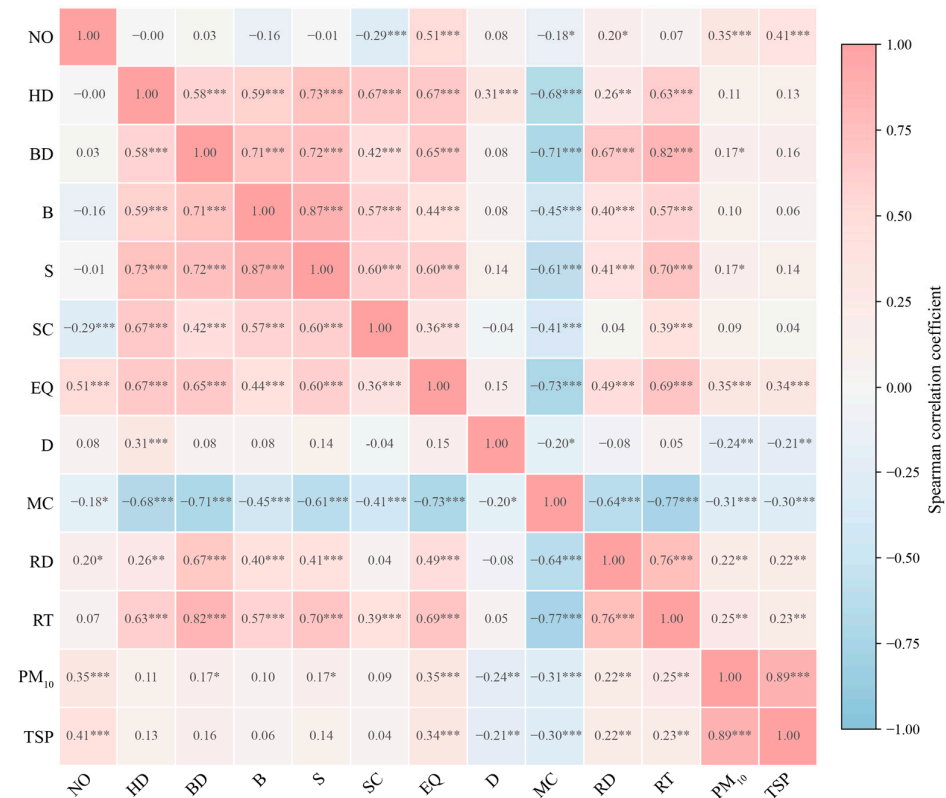


Figure 3. Spearman rank correlation heatmap of input and output variables. The values in the cells represent Spearman correlation coefficients (ρ). Asterisks indicate the statistical significance levels of the correlations: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

As shown in Figure 3, PM₁₀ and TSP exhibit a strong and statistically significant positive association, with a Spearman correlation coefficient of 0.89 ($p < 0.001$). This result indicates that high PM₁₀ concentrations are generally accompanied by high TSP concentrations under the monitored blasting conditions. This trend is reasonable because PM₁₀ represents the fine fraction of airborne particulate matter, whereas TSP reflects the overall suspended dust level. Therefore, the two dust indicators are closely associated with the same blasting-induced dust generation and dispersion process.

Among the blasting design variables, the number of blastholes and explosive quantity show significant positive associations with both PM₁₀ and TSP. The number of blastholes is positively correlated with PM₁₀ and TSP, with correlation coefficients of 0.35 ($p < 0.001$) and 0.41 ($p < 0.001$), respectively. Explosive quantity also shows significant positive correlations with PM₁₀ and TSP, with correlation coefficients of 0.35 ($p < 0.001$) and 0.34 ($p < 0.001$), respectively. These results suggest that blasting rounds with a larger scale and greater explosive charge tend to be associated with higher dust concentrations. From an engineering perspective, more blastholes and larger explosive quantities generally correspond to stronger rock breakage, higher energy release, and a larger potential dust source.

Monitoring distance shows a statistically significant negative association with both dust indicators. The correlation coefficients between distance and PM₁₀/TSP are -0.24 ($p < 0.01$) and -0.21 ($p < 0.01$), respectively. This trend indicates that dust concentration tends to decrease as the monitoring point moves farther away from the blasting source. Such a pattern is consistent with the physical processes of dust dispersion, dilution, and gravita-

tional settling during transport. Therefore, monitoring distance remains an important spatial factor for characterizing blast-induced particulate matter concentration.

Moisture content also exhibits significant negative associations with PM₁₀ and TSP, with correlation coefficients of -0.31 ($p < 0.001$) and -0.30 ($p < 0.001$), respectively. This result suggests that higher moisture content may contribute to dust suppression. A possible explanation is that water promotes particle agglomeration and increases the difficulty of fine particles becoming airborne after blasting. This negative association is consistent with practical dust control experience, although it should still be interpreted as a statistical association rather than a direct causal conclusion.

In contrast, the stemming column shows weak and statistically insignificant associations with PM₁₀ and TSP, with correlation coefficients of 0.09 and 0.04, respectively. This indicates that no clear monotonic relationship between stemming column and dust concentration can be identified from the present dataset alone. However, this does not necessarily mean that stemming has no influence on dust emissions. Its effect may be nonlinear or may depend on interactions with other blast design parameters, such as burden, spacing, blasthole diameter, and explosive quantity. Therefore, stemming column was still retained as an input variable in the subsequent machine learning models.

The heatmap also reveals clear correlations among several input variables. For example, burden and spacing show a strong positive association, with a correlation coefficient of 0.87 ($p < 0.001$). Blasthole diameter is strongly associated with rock type and rock density, with correlation coefficients of 0.82 ($p < 0.001$) and 0.67 ($p < 0.001$), respectively. Moisture content is negatively correlated with several blasting and material-related variables, including explosive quantity, blasthole diameter, and rock type. These relationships indicate that the input variables in the field dataset are not fully independent. This is common in practical blasting operations because blast design parameters are often adjusted jointly according to bench geometry, production requirements, and material conditions.

The presence of these correlations indicates that some predictors contain partly overlapping information, but it does not indicate target leakage. All correlated variables were input-side variables measured before or during blast design and field monitoring, whereas PM₁₀ and TSP were used only as prediction targets. Therefore, the high correlations among predictors mainly reflect the coupled nature of field blasting design rather than the inclusion of output information in the model. These correlations were considered when interpreting feature importance and SHAP results.

The scatter plots in Figure 4 provide a visual comparison between selected variables and PM₁₀/TSP concentrations. Explosive quantity shows an overall increasing trend with both PM₁₀ and TSP, indicating that larger explosive charges are generally associated with higher dust concentrations. Distance and moisture content show generally decreasing trends, which is consistent with the attenuation of dust concentration during transport and the dust-suppression effect of material moisture. In contrast, stemming column does not show a clear monotonic trend with either PM₁₀ or TSP. These visual patterns are generally consistent with the Spearman correlation analysis shown in Figure 3. However, the scatter plots should be interpreted as exploratory evidence only, because the field data are right-skewed and several blasting parameters are correlated with each other. Therefore, the subsequent machine learning and SHAP analyses were used to further examine the nonlinear and multivariable relationships between blasting conditions and dust concentration.

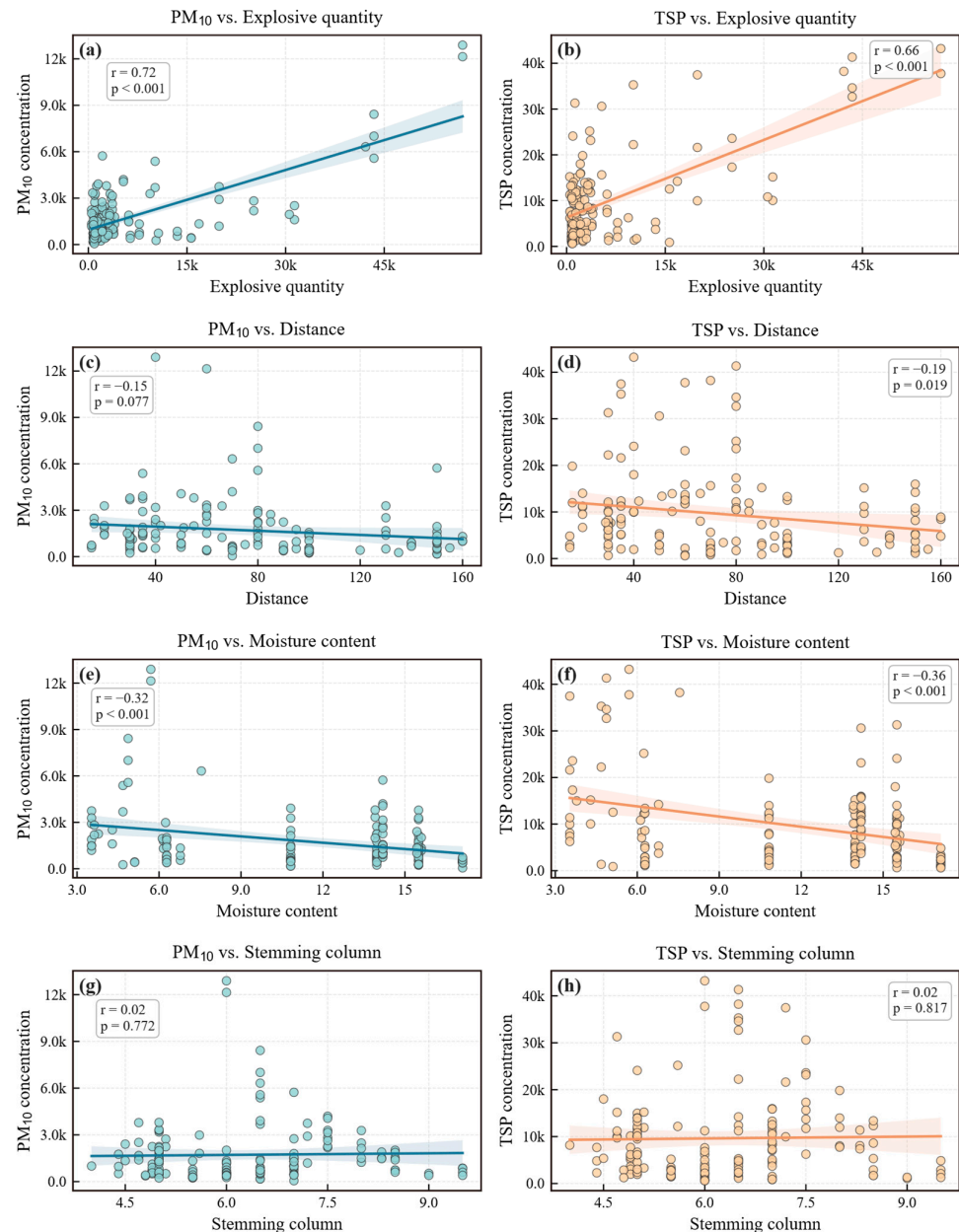


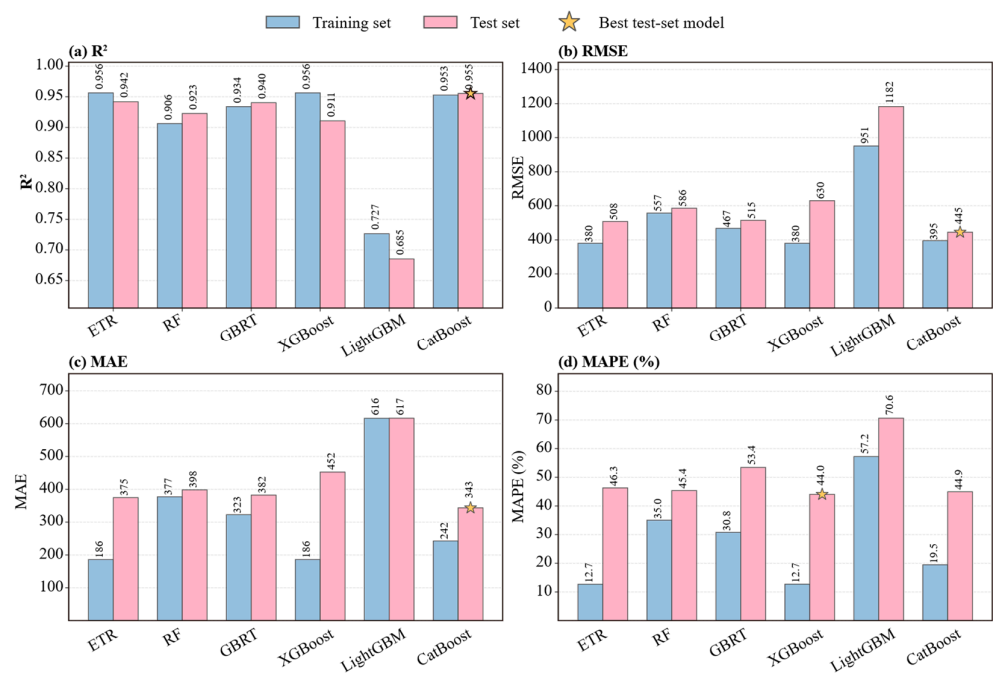
Figure 4. Scatter plots between key variables and PM₁₀/TSP concentrations: (a) PM₁₀ versus explosive quantity; (b) TSP versus explosive quantity; (c) PM₁₀ versus distance; (d) TSP versus distance; (e) PM₁₀ versus moisture content; (f) TSP versus moisture content; (g) PM₁₀ versus stemming column; and (h) TSP versus stemming column. The solid lines represent fitted linear trends, and the shaded regions indicate 95% confidence intervals.

3.3. Model Performance Comparison for PM₁₀ Prediction

Table 4 and Figure 5 show the six machine learning models' PM₁₀ concentration prediction ability. Each model's fitting ability and generalisation performance were compared using four evaluation metrics: R², RMSE, MAE, and MAPE. Model fitting was examined using the training set, and prediction accuracy on unseen samples was assessed using the test set. This contrast is crucial since the dataset's PM₁₀ concentrations exhibit a right-skewed distribution and include a number of high-value samples, which could make accurate prediction more challenging.

Table 4. Prediction performance of different models for PM₁₀.

Model	Dataset	R ²	RMSE (µg/m ³)	MAE (µg/m ³)	MAPE (%)
Extra Trees Regression	Training set	0.9564	379.83	185.88	12.68
Extra Trees Regression	Test set	0.9419	507.87	374.57	46.29
Random Forest	Training set	0.9062	556.99	377.11	35.03
Random Forest	Test set	0.9227	585.84	398.24	45.37
Gradient Boosting	Training set	0.9340	467.31	322.58	30.79
Gradient Boosting	Test set	0.9404	514.73	382.29	53.43
XGBoost	Training set	0.9564	379.83	186.03	12.70
XGBoost	Test set	0.9107	629.93	452.48	44.04
LightGBM	Training set	0.7266	950.96	616.27	57.22
LightGBM	Test set	0.6852	1182.49	616.58	70.61
CatBoost	Training set	0.9528	395.04	242.49	19.46
CatBoost	Test set	0.9555	444.87	343.48	44.94

**Figure 5.** Comparison of model performance for PM₁₀ prediction: (a) R²; (b) RMSE; (c) MAE; and (d) MAPE. The blue bars represent the training set, the pink bars represent the test set, and the star indicates the best test-set model.

The majority of models had acceptable PM₁₀ prediction performance, as Table 4 illustrates. CatBoost achieved the best test-set performance out of the six models. With RMSE, MAE, and MAPE values of 444.87 µg/m³, 343.48 µg/m³, and 44.94%, respectively, its test R² was 0.9555. This suggests that, given the current dataset, CatBoost has the best generalisation ability for PM₁₀ prediction. With an RMSE of 507.87 µg/m³ and a test R² of 0.9419, ETR also did well. Although GBRT's test R² of 0.9404 was comparable to CatBoost's, its MAE and MAPE were marginally higher. Because the dataset contains only 148 samples and the test set contains 30 samples, these rankings should be interpreted as performance under the adopted hold-out split rather than as a definitive universal ranking. This is especially relevant for PM₁₀, where CatBoost, ETR, and GBRT showed relatively close test R² values.

The benefits of CatBoost are further supported by the results shown in Figure 5. CatBoost maintained a slight discrepancy between training and test performance and displayed the greatest R² value on the test set. This implies that there was no clear overfitting and that the model accurately represented the primary nonlinear connection between the

input variables and PM₁₀ concentration. Although XGBoost and ETR had good training R² values, their test-set outcomes differed. While XGBoost showed a noticeable decline in test R², ETR continued to perform well. This suggests that although XGBoost had poorer generalisation in this instance, it fit the training data quite well.

The error-based metrics show a similar trend. CatBoost produced the lowest test RMSE and MAE, indicating that its predicted PM₁₀ values were closest to the measured values on average. Although XGBoost achieved a relatively low training error, its test RMSE increased to 629.93 µg/m³, and its MAE increased to 452.48 µg/m³. RF showed moderate performance, with a test R² of 0.9227 and RMSE of 585.84 µg/m³. In contrast, LightGBM showed the weakest PM₁₀ prediction ability, with the lowest test R² of 0.6852 and the highest test RMSE of 1182.49 µg/m³. This suggests that LightGBM was less suitable for this small field-monitoring dataset under the current parameter setting.

The model's performance before and after Bayesian optimisation is contrasted in Figure 6. The findings demonstrate that most models' prediction performance was enhanced by hyperparameter optimisation, however the degree of increase varied depending on the algorithm. Several models' test performance was improved using R² Bayesian Optimisation, particularly for tree-based ensemble approaches that are sensitive to model complexity. The optimised models typically displayed lower RMSE, MAE, or MAPE values for error metrics, suggesting that Bayesian optimisation contributed to a decrease in prediction error. The degree of improvement was not consistent, though, indicating that the impact of optimisation is dependent on both data quality and model structure.

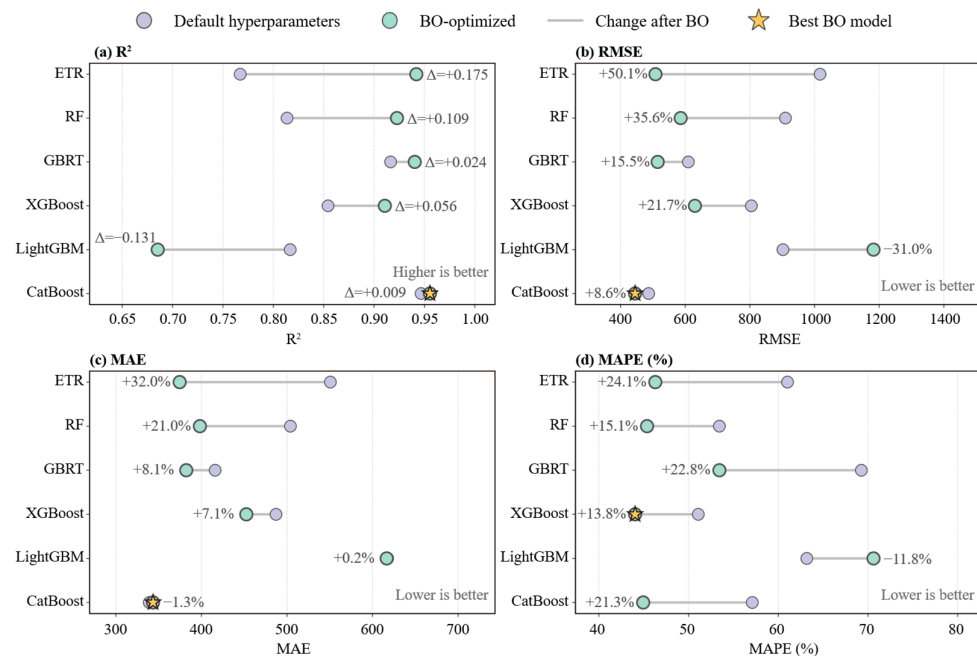


Figure 6. Before-and-after comparison of PM₁₀ prediction performance using Bayesian Optimization: (a) R²; (b) RMSE; (c) MAE; and (d) MAPE. The purple circles indicate the model performance with default hyperparameters, the green circles indicate the BO-optimized results, the grey horizontal lines show the change after BO, and the star marks the best BO model. For R², higher values indicate better performance, whereas for RMSE, MAE, and MAPE, lower values indicate better performance.

3.4. Model Performance Comparison for TSP Prediction

Table 5 and Figure 7 show the prediction performance of the six machine learning models for TSP concentration. Model performance was assessed using R², RMSE, MAE, and MAPE on both training and test sets, as in the PM₁₀ prediction task. TSP has a greater

standard deviation and a wider concentration range than PM₁₀, making precise prediction more challenging. In order to find a reliable prediction technique for total suspended particulate matter emissions, it is crucial to compare various models.

Table 5. Prediction performance of different models for TSP.

Model	Dataset	R ²	RMSE (µg/m ³)	MAE (µg/m ³)	MAPE (%)
Extra Trees Regression	Training set	0.9257	2301.65	1213.01	20.83
Extra Trees Regression	Test set	0.8803	3899.04	3036.27	51.87
Random Forest	Training set	0.8646	3107.45	2202.02	44.12
Random Forest	Test set	0.8113	4894.55	3861.15	83.73
Gradient Boosting	Training set	0.8989	2685.16	1908.76	36.43
Gradient Boosting	Test set	0.8311	4630.55	3692.53	82.95
XGBoost	Training set	0.9257	2301.65	1213.24	20.84
XGBoost	Test set	0.8596	4222.80	3027.54	47.73
LightGBM	Training set	0.7387	4316.87	3190.96	64.47
LightGBM	Test set	0.8028	5004.56	3865.41	91.58
CatBoost	Training set	0.9196	2394.47	1562.20	28.48
CatBoost	Test set	0.8306	4638.52	3678.35	68.47

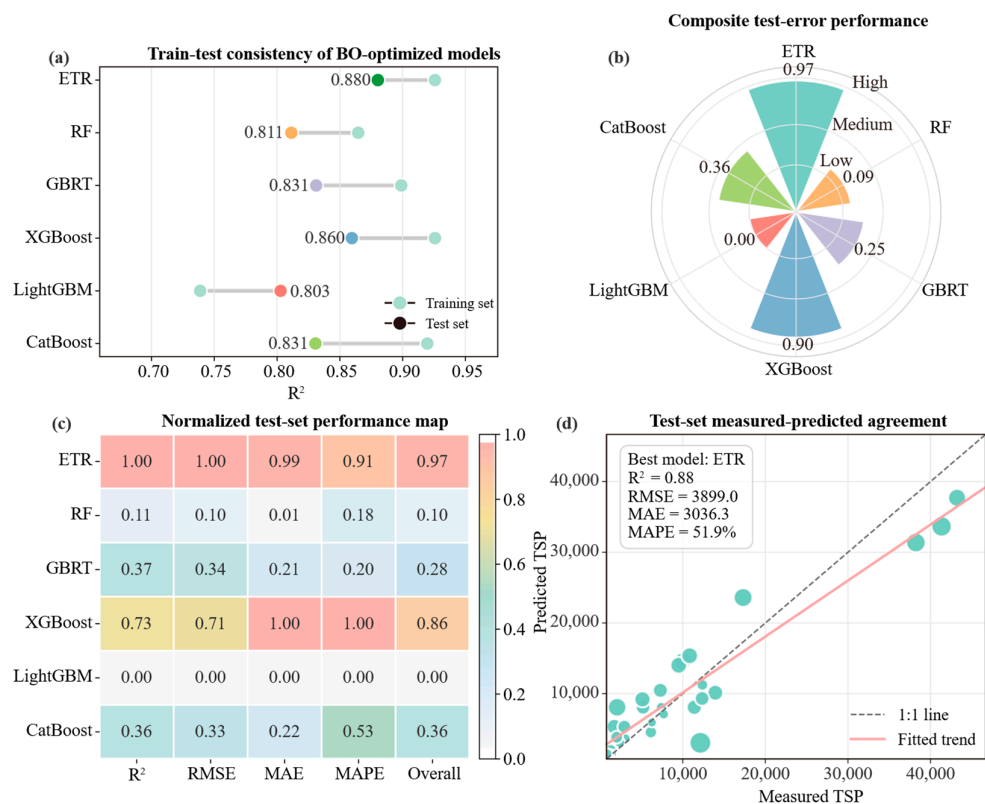


Figure 7. BO-optimized model performance for TSP prediction.

ETR produced the strongest overall test performance for TSP prediction, as Table 5 illustrates. The matching RMSE and MAE values were 3899.04 µg/m³ and 3036.27 µg/m³, respectively, while the test R² of ETR was 0.8803. According to these findings, ETR outperformed the other six models in terms of generalisation for the TSP prediction test. With a test R² of 0.8596, RMSE of 4222.80 µg/m³, and the lowest test MAE of 3027.54 µg/m³, XGBoost also demonstrated strong prediction accuracy. This implies that XGBoost was successful in lowering the average absolute prediction error as well.

The results in Figure 7 further show the performance differences among the BO-optimized models. ETR and XGBoost achieved relatively high R^2 values on both the training and test sets, indicating that these two models captured the main nonlinear relationship between blasting variables and TSP concentration. CatBoost showed a high training R^2 of 0.9196, but its test R^2 decreased to 0.8306. This indicates that CatBoost fitted the training data well but had weaker generalization for TSP than for PM_{10} . RF and GBRT showed moderate test performance, with test R^2 values of 0.8113 and 0.8311, respectively.

The error-based metrics show a similar pattern. ETR produced the lowest test RMSE, which means that it reduced large prediction errors more effectively than the other models. XGBoost achieved the lowest MAE and MAPE on the test set, with values of $3027.54 \mu\text{g}/\text{m}^3$ and 47.73%, respectively. This indicates that XGBoost had an advantage in controlling the average relative error. In contrast, LightGBM showed the weakest overall performance for TSP prediction. Although its test R^2 was 0.8028, its RMSE, MAE, and MAPE were higher than those of most other models, suggesting that it was less stable for this dataset.

The TSP prediction performance before and after Bayesian optimisation is contrasted in Figure 8. The findings demonstrate that while hyperparameter optimisation enhanced a number of models, the benefit was not consistent. Bayesian optimisation improved model fitting and generalisation for some models by raising R^2 and lowering prediction errors. The improvement was minimal for other models, particularly when a portion of the data structure had already been covered by the default model. This finding implies that Bayesian optimisation can enhance model performance, however the ultimate outcome is still contingent upon the model type and TSP dataset complexity. The Bayesian optimization convergence curves are shown in Figure 9. For both PM_{10} and TSP prediction, the best MSE decreased rapidly during the early iterations and then gradually approached a stable value, indicating that the optimization process effectively searched the hyperparameter space. The convergence trends also show that the improvement differed among models, which is consistent with the before-and-after performance comparison in Figures 6 and 8. These results indicate that Bayesian optimization contributed to model tuning, but the final prediction performance was still affected by the model structure and the complexity of the PM_{10} and TSP datasets.

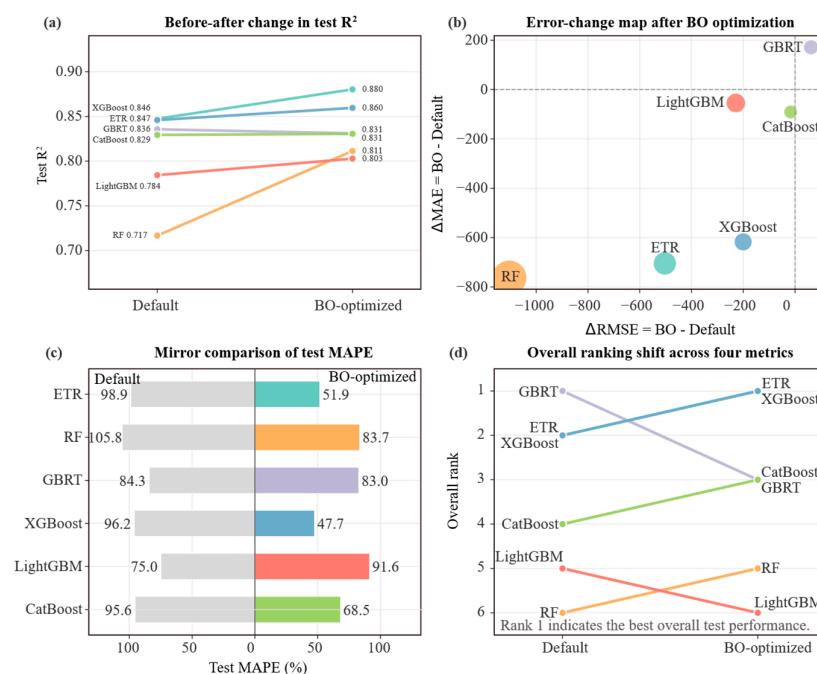


Figure 8. Comparison of TSP prediction performance before and after Bayesian optimization.

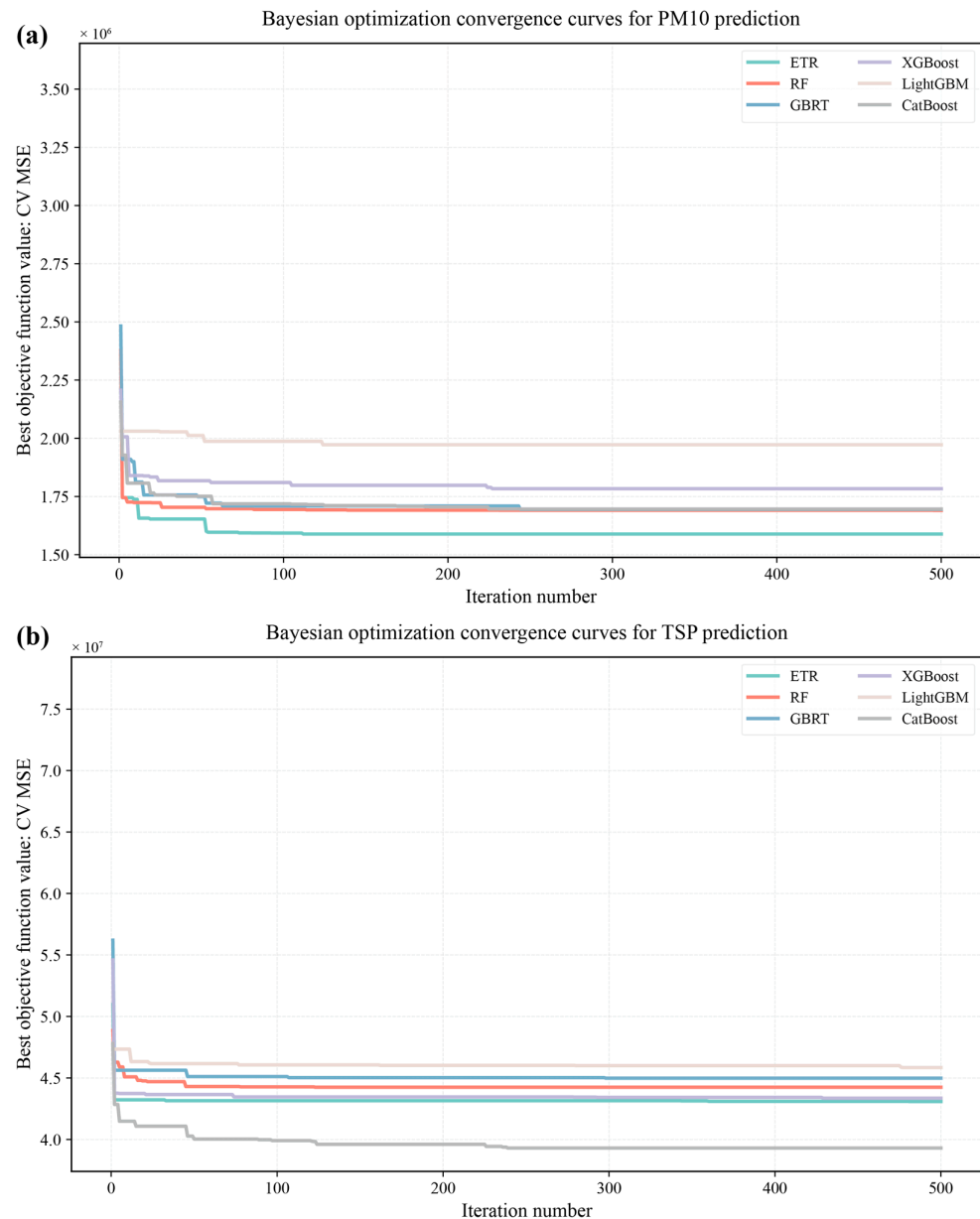
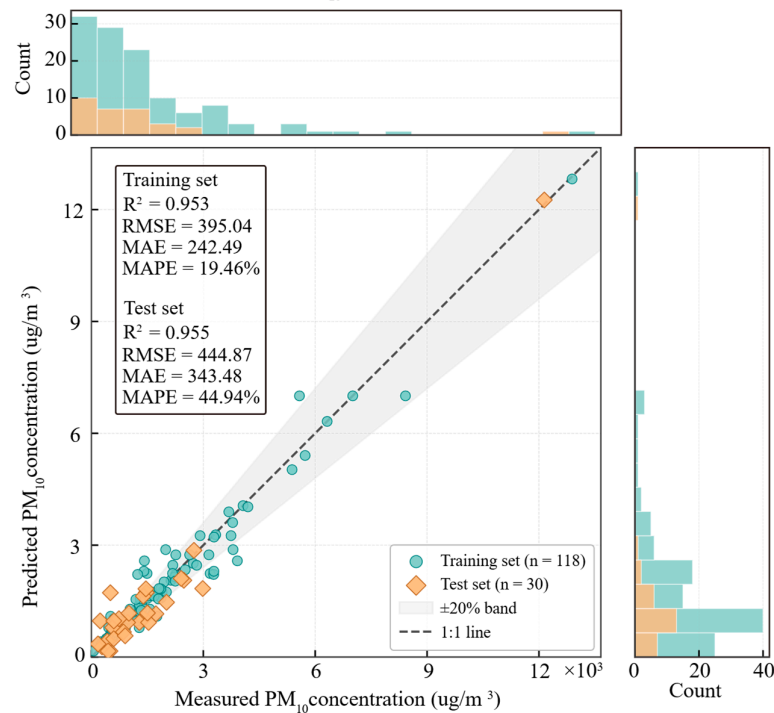
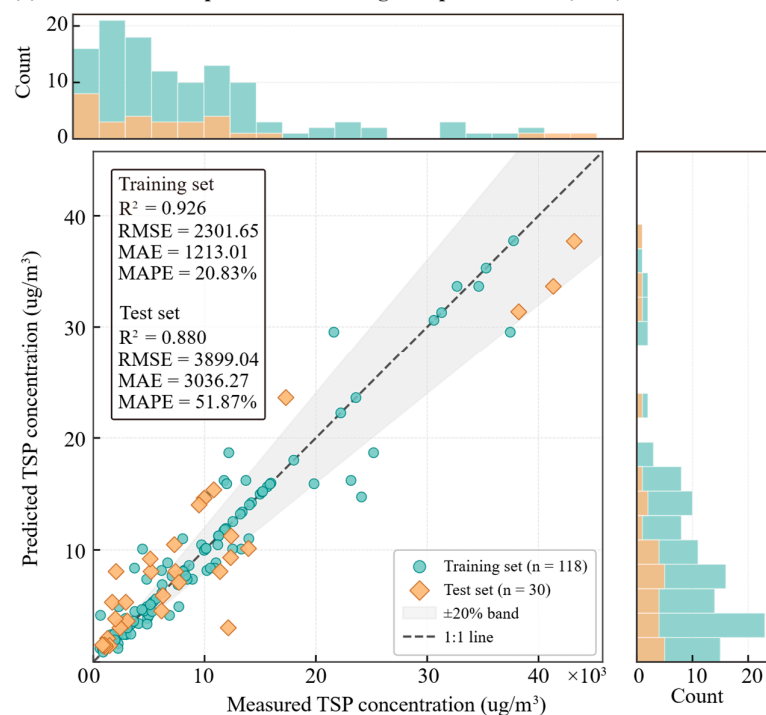


Figure 9. Bayesian optimization convergence curves for PM₁₀ and TSP prediction: (a) PM₁₀ prediction and (b) TSP prediction.

3.5. Prediction Performance of the Optimal Model

The best model for PM₁₀ prediction was CatBoost, and the best model for TSP prediction was ETR, according to the model comparison results. The measured and anticipated PM₁₀ and TSP values for the training and test sets are shown in Figure 10. The ideal prediction condition was represented by the 1:1 reference line, and the deviation range between measured and anticipated values was assessed by adding the $\pm 20\%$ band. The prediction accuracy and generalisation capacity of the best models are directly evaluated visually in this graphic.

(a) Measured versus predicted PM_{10} using the optimal model (CatBoost)**(b) Measured versus predicted TSP using the optimal model (ETR)****Figure 10.** Measured versus predicted PM_{10} and TSP using the optimal models.

The CatBoost model demonstrated good agreement between measured and projected values for PM_{10} . The model was able to capture the primary variable pattern of PM_{10} concentration since the majority of the training and test data were distributed along the 1:1 line. With RMSE, MAE, and MAPE values of 395.04 $\mu g/m^3$, 242.49 $\mu g/m^3$, and 19.46%, respectively, the training set attained an R^2 of 0.953. With RMSE, MAE, and MAPE values of 444.87 $\mu g/m^3$, 343.48 $\mu g/m^3$, and 44.94%, respectively, the model achieved an R^2 of 0.955

on the test set. The close training and test R^2 values suggest that the CatBoost model did not show obvious overfitting under the adopted hold-out validation scheme.

The majority of PM_{10} samples are clustered in the low-concentration region, as further demonstrated by the marginal histograms in Figure 10. The descriptive statistics displayed in Section 3.1 are in line with this. The CatBoost model still makes accurate predictions for the majority of samples, despite the existence of a few high-concentration samples. Due to the small number of extreme dust episodes in the dataset, some variations can be seen in the high-value range. Such high-concentration cases in field monitoring data are frequently influenced by a number of interrelated parameters, such as local blasting circumstances, moisture content, explosive quantity, and distance.

The ETR model also shows strong predictive ability for TSP prediction. With RMSE, MAE, and MAPE values of 2301.65 $\mu\text{g}/\text{m}^3$, 1213.01 $\mu\text{g}/\text{m}^3$, and 20.83%, respectively, the training set obtained an R^2 of 0.926. The test set's R^2 score was 0.880, and its RMSE, MAE, and MAPE values were 3899.04 $\mu\text{g}/\text{m}^3$, 3036.27 $\mu\text{g}/\text{m}^3$, and 51.87%, respectively. TSP had a higher test-set error than PM_{10} . This makes sense because TSP has a larger concentration range and more variability, making prediction more challenging. The relationship between RMSE and the standard deviation of the target variable was also considered to assess whether the low RMSE values might indicate overfitting. A RMSE value lower than the standard deviation of the measured target does not by itself indicate overfitting. Instead, it means that the model performs better than a simple mean-value predictor. For PM_{10} , the standard deviation of the measured concentration was 1889.61 $\mu\text{g}/\text{m}^3$, while the CatBoost model achieved RMSE values of 395.04 $\mu\text{g}/\text{m}^3$ on the training set and 444.87 $\mu\text{g}/\text{m}^3$ on the test set. The test RMSE was therefore about 23.5% of the overall PM_{10} standard deviation. For TSP, the standard deviation was 9120.84 $\mu\text{g}/\text{m}^3$, while the ETR model achieved RMSE values of 2301.65 $\mu\text{g}/\text{m}^3$ on the training set and 3899.04 $\mu\text{g}/\text{m}^3$ on the test set. The test RMSE was about 42.8% of the overall TSP standard deviation. These results indicate that the models captured meaningful variance in the dust concentration data rather than simply reproducing the mean response.

The TSP panel in Figure 10 illustrates that many samples fall inside or close to the $\pm 20\%$ region, and the majority of projected values match the overall trend of the measured values. Some test samples, particularly those in the medium and high concentration ranges, do, however, depart from the 1:1 line. This suggests that while the model can explain the general relationship between TSP concentration and blasting circumstances, its accuracy declines in certain intricate high-dust scenarios. Air humidity, wind direction, wind speed, and short-term atmospheric disturbance following blasting are examples of variables not included in the dataset that could be the cause of these differences.

3.6. Residual Analysis

To further assess the prediction accuracy of the best models, residual analysis was carried out. The ETR model was examined for TSP, and the CatBoost model was examined for PM_{10} . The residual distributions and the correlation between residuals and predicted values are shown in Figure 11. The ideal condition is represented by the zero-residual line, and the dispersion of prediction errors was examined using the ± 1 SD and ± 2 SD bands. After fitting the measured dust concentration data, this technique aids in determining if the models exhibit a clear systematic bias.

The CatBoost model's residuals for PM_{10} prediction were mostly dispersed around the zero-residual line. The residual distribution was nearly symmetric, with a mean residual of 11.61 $\mu\text{g}/\text{m}^3$ and a skewness of 0.12. Only a few samples exceeded the broader error range, with the majority falling within the ± 1 SD band. This implies that there was no discernible tendency for the CatBoost model to overestimate or underestimate PM_{10} levels. In line with the fact that the majority of PM_{10} samples were in the low-concentration in-

interval, the residual scatter plot also demonstrates that the majority of mistakes were clustered in the low-prediction range.

Compared to PM_{10} , the residual distribution for TSP prediction was more spread. The skewness was 0.55, the standard deviation was $2711.78 \mu\text{g}/\text{m}^3$, and the mean residual was $-35.68 \mu\text{g}/\text{m}^3$. In relation to the TSP concentration scale, the mean value was nearly zero, suggesting that there was no significant general bias in the ETR model. The higher standard deviation, however, indicates that TSP's prediction error was more erratic. This outcome is in line with the model performance comparison, in which PM_{10} had lower RMSE and MAE values than TSP. Additionally, the residual scatter plot shows that certain samples, particularly in the higher expected concentration range, had considerable positive or negative residuals.

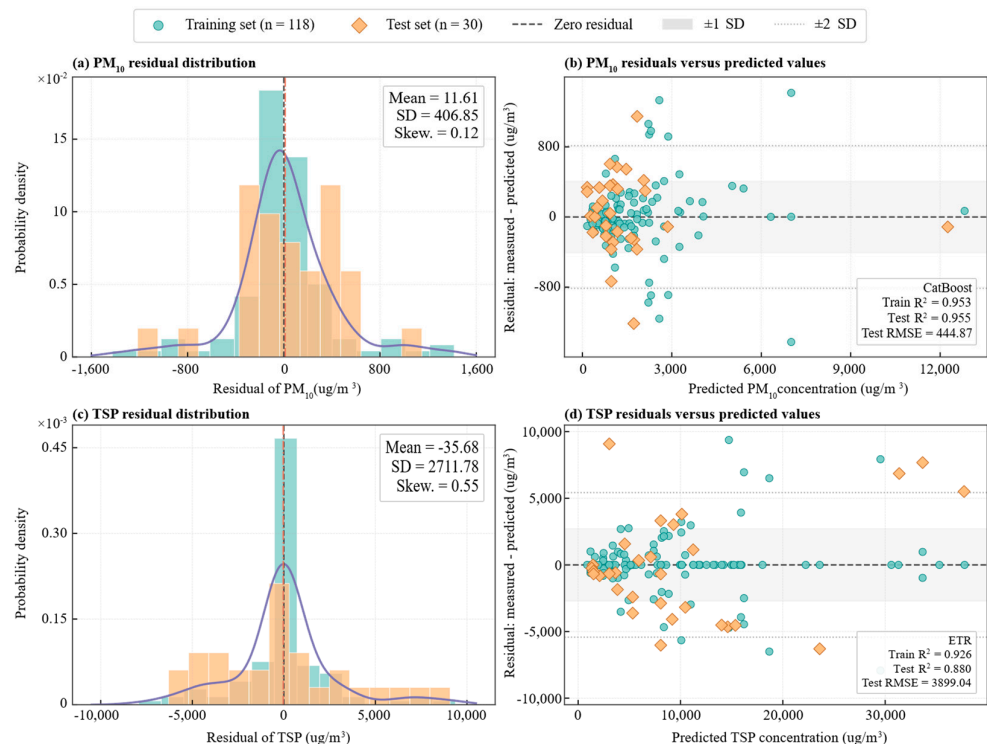


Figure 11. Residual analysis of the optimal model.

The distinct features of the two target variables can account for the discrepancy between PM_{10} and TSP residuals. In the dataset, PM_{10} exhibited a comparatively lower concentration range and represented a smaller particle proportion. TSP, on the other hand, has greater variability and covers a wider variety of suspended particles. As a result, complicated field elements such as dust cloud movement, particle settling, wind disturbance, and local monitoring circumstances have a greater impact on TSP prediction. The greater residual dispersion of the TSP model could be explained by the incomplete inclusion of these elements in the available input variables.

The high-concentration interval error analysis is shown in Figure 12 to further investigate model performance in high-dust situations. The 75th percentile of the measured concentration was used to define the high-concentration interval in this figure. Eight test samples fell within the high-concentration range for PM_{10} , which had a threshold of $1524.25 \mu\text{g}/\text{m}^3$. The CatBoost model maintained adequate prediction accuracy for high PM_{10} values, as evidenced by its high-concentration MAPE of 21.54%. This finding implies that CatBoost can reliably identify situations with significant PM_{10} emissions.

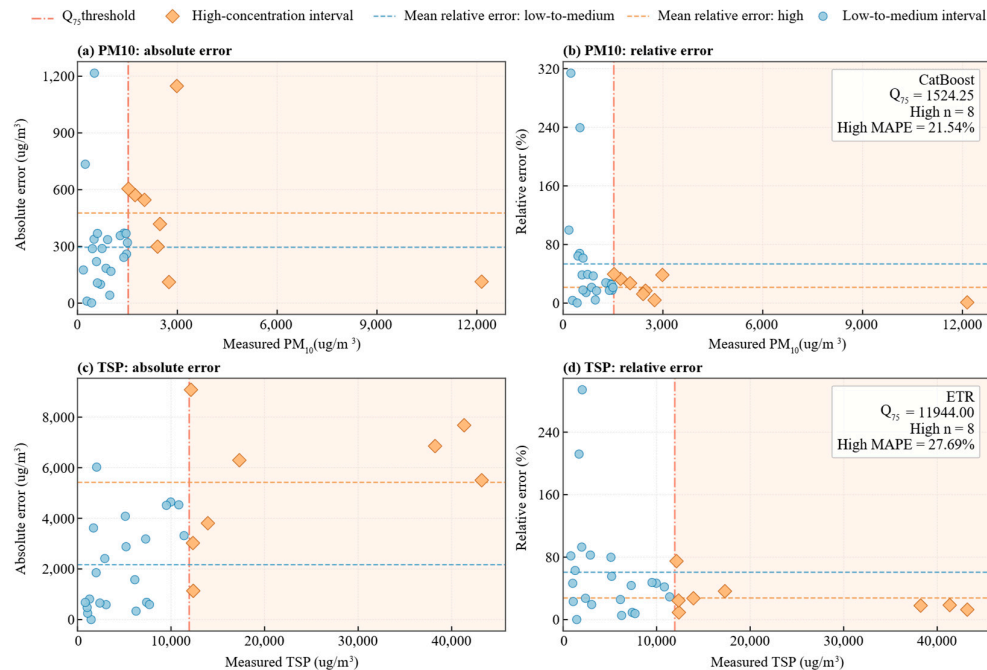


Figure 12. High-concentration interval error analysis of the optimal models.

Eight test samples fell within the high-concentration threshold of 11,944.00 $\mu\text{g}/\text{m}^3$ for TSP. The ETR model's high-concentration MAPE was 27.69%. This score was greater than PM_{10} , indicating that it was more challenging to predict high TSP. Larger discrepancies were seen in a number of high-concentration TSP samples, particularly when the measured concentration was quite high. This could be connected to the intricate process of total suspended particle production and movement following blasting. Short-term airflow, falling trajectory, and sample position have a greater impact on large suspended particles, potentially increasing forecast uncertainty. The concentration-band error analysis also helps explain why the optimal models achieved high R^2 values but still showed relatively high test MAPE values. The test R^2 values of the optimal models were 0.9555 for PM_{10} and 0.8803 for TSP, indicating that the models captured the main variance of the measured concentrations. However, the corresponding test MAPE values were 44.94% and 51.87%, respectively. This apparent discrepancy is mainly related to the right-skewed distribution of PM_{10} and TSP. A small number of high-concentration samples contribute strongly to the total variance and therefore have a large influence on R^2 . In contrast, MAPE gives greater relative weight to low-concentration samples, where even a moderate absolute error can produce a large percentage error. The high-concentration interval analysis showed that the MAPE values of the optimal models were 21.54% for PM_{10} and 27.69% for TSP, which were lower than the overall test MAPE values. This indicates that the high overall MAPE was largely affected by relative errors in low- and medium-concentration samples rather than by poor prediction of high-concentration dust events. Therefore, R^2 , RMSE, MAE, and MAPE should be interpreted together instead of relying on a single metric. It should be further noted that the high test R^2 values may be partly affected by the presence of several high-concentration or high-explosive-quantity cases, because R^2 is sensitive to the total variance of the measured target values. Therefore, the reported R^2 values should not be interpreted as evidence that the models are equally accurate for all blasting scales. Instead, they indicate that the models captured the main variance pattern of the present dataset. The relatively high MAPE values and the residual dispersion, especially for TSP, show that prediction uncertainty remains for low- and medium-concentration cases. Thus, the model performance should be interpreted together

with the residual analysis, concentration-band error analysis, and the limitations of the available field dataset.

3.7. Feature Importance Analysis

Feature importance analysis was used to identify the main predictors within each optimal model. The PM₁₀ model was based on CatBoost, whereas the TSP model was based on Extra Trees Regression. Therefore, the importance percentages in Figure 13 should be interpreted as model-specific rankings within each target rather than as directly comparable values between PM₁₀ and TSP. In this section, the feature importance results are used to describe the dominant variables for each model separately. Cross-target comparison is made mainly in the subsequent SHAP analysis, where the direction and sample-level contribution of each variable can be examined in a more consistent way. In addition, the input variables were not fully independent, as shown by the correlation heatmap. Therefore, the importance assigned to variables within correlated groups may be shared or redistributed by the model. For example, burden and spacing jointly describe the blasthole layout, whereas blasthole diameter, rock density, and rock type partly reflect material and charge-structure conditions. Thus, Figure 13 should be interpreted as a model-based ranking of predictive contribution rather than as a strict measure of the independent physical effect of each variable.

With an importance value of 33.1%, explosive quantity ranked first in the PM₁₀ model. This finding suggests that the model forecast was most strongly influenced by the quantity of explosive utilised during a blasting round. Higher energy release, stronger rock fracture, and more severe dust formation are typically associated with larger explosive quantities. This result is also in line with the correlation analysis, which revealed a strong positive association between explosive quantity and PM₁₀ concentration.

With an importance value of 19.7%, the quantity of blastholes was the second most significant factor for PM₁₀. The size of a blasting operation is reflected in this variable. A larger blasting area and more disruption to the rock mass are often indicated by more blastholes. As a result, the quantity of blastholes may be an indirect indicator of the strength of a possible dust source. Distance came in third place with a 13.1% significance value. This demonstrates that PM₁₀ prediction was clearly impacted by the monitoring position. Air dilution, particle dispersion, and deposition all have an impact on PM₁₀ concentration as one gets farther away from the blasting source.

With a importance value of 25.7%, explosive quantity was also the most significant factor for TSP prediction. This demonstrates that both fine particulate matter and total suspended particulate matter are significantly influenced by explosive energy input. Within the TSP model, explosive quantity was also ranked as the most important variable. This indicates that explosive energy input was a dominant predictor for TSP under the ETR model. However, because the PM₁₀ and TSP importance values were obtained from different model families, the percentage values should not be directly compared between the two targets.

With a importance value of 18.5%, distance came in second for TSP prediction. Within the TSP model, distance ranked second in feature importance, indicating that spatial attenuation was an important predictor for TSP concentration. This interpretation is also consistent with the SHAP results and the physical process of particle dilution and settling during transport. Particle settling during transport may have a greater impact on observed concentration since TSP contains bigger suspended particles. With a significance score of 16.1%, the quantity of blastholes came in third. Like PM₁₀, this variable contributes to the quantity of dust created during blasting and represents the blasting scale.

Although their significance values were lower, other variables also played a role in model prediction. Stemming column, blasthole diameter, burden, moisture content, rock density, and rock type all contributed moderately or weakly to PM₁₀. These factors also

have secondary effects on TSP. They may nevertheless have an impact on dust emissions through nonlinear interactions, even though their individual relevance values were not as high as those of explosive quantity, distance, and number of blastholes. For instance, while moisture content can encourage particle agglomeration and decrease dust suspension, a stemming column can affect collar dust release and explosive gas confinement.

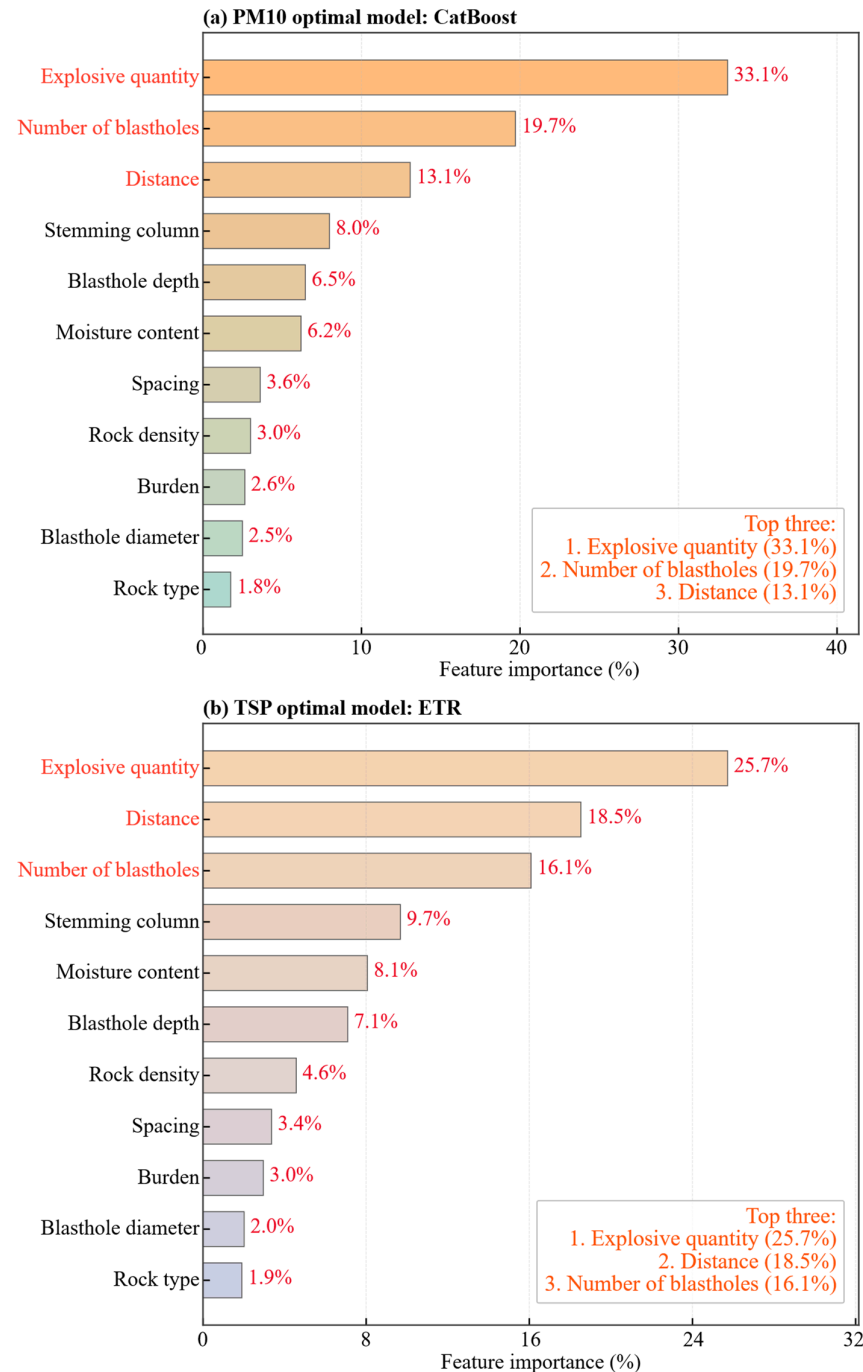


Figure 13. Feature importance ranking for PM₁₀ and TSP.

3.8. SHAP Interpretation

The prediction behaviour of the optimal models was further explained using SHAP analysis [14]. Recent mining and rock-engineering studies have also demonstrated the

usefulness of SHAP for interpreting nonlinear ensemble models [21]. In contrast to feature importance analysis, SHAP can display the direction of each variable's influence on the anticipated dust concentration in addition to its relative importance. SHAP values for the best PM₁₀ model based on CatBoost and the best TSP model based on ETR were computed in this study. Figures 14 and 15 display the outcomes. Because the optimal PM₁₀ and TSP models were different model families, the cross-target comparison of variable effects was based primarily on the SHAP results rather than on the raw feature importance percentages. In the present study, SHAP was used to compare the relative ranking, influence direction, and sample-level contribution patterns of key variables for PM₁₀ and TSP. The absolute SHAP magnitudes were not directly compared between the two targets because PM₁₀ and TSP have different concentration ranges. Moreover, because some predictors were strongly correlated, SHAP values may allocate contribution among correlated variables. Therefore, the SHAP plots were interpreted together with the correlation heatmap and engineering knowledge, rather than as fully independent causal attributions for each feature.

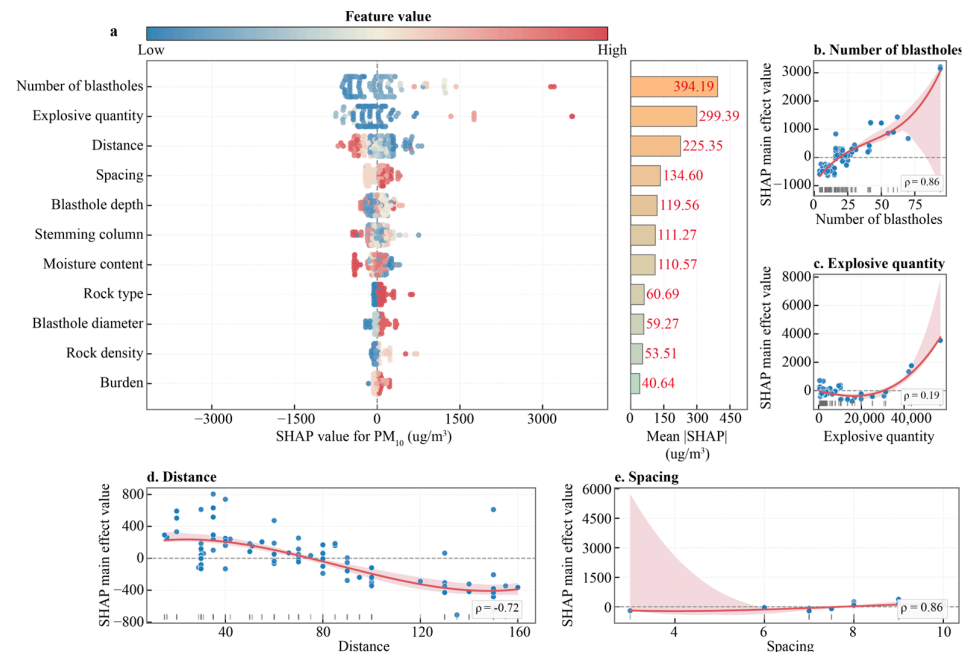


Figure 14. SHAP summary plot for PM₁₀ using the optimal model (CatBoost).

The SHAP summary graphic in Figure 14 demonstrates that the most important factors for PM₁₀ prediction were explosive quantity, number of blastholes, and distance. In general, this outcome agrees with the ranking of feature importance. Positive SHAP values were mostly generated by high explosive quantity values, suggesting that higher explosive charges raised the anticipated PM₁₀ concentration. Higher explosive energy can increase rock fragmentation and produce more fine particles during blasting, which makes this trend understandable. Additionally, the dependence plot exhibits a distinct upward trend, indicating that explosive quantity significantly improves PM₁₀ prediction.

Additionally, the quantity of blastholes demonstrated a significant impact on PM₁₀ prediction. The anticipated PM₁₀ concentration often increased with the number of blastholes. This variable is directly correlated with the total volume of disturbed rock and represents the size of the blasting round. As a result, it can depict the possible source intensity of dust production. The impact of blasthole number, however, may interact with

other factors, including explosive quantity, hole diameter, burden, and spacing, because the SHAP distribution also exhibits certain sample-level variations.

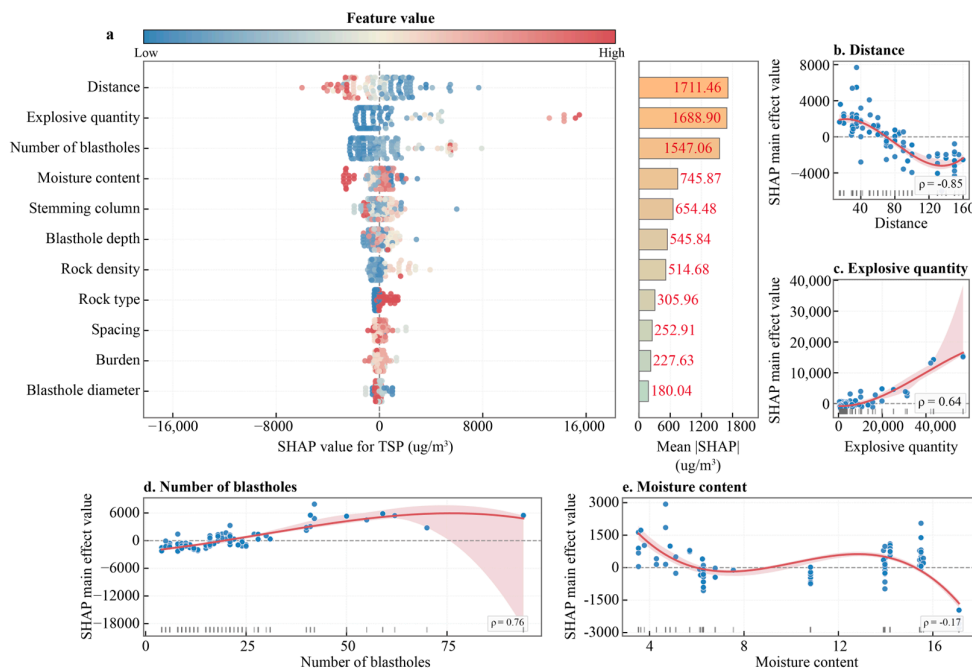


Figure 15. SHAP summary plot for TSP using the optimal model (ETR).

The PM₁₀ prediction was mostly negatively impacted by distance. Longer monitoring distances decreased the anticipated PM₁₀ concentration, as evidenced by the correlation between higher distance values and negative SHAP values. This is consistent with the physical process of dust deposition and dispersion following blasting. Particulate matter is diluted, and some particles settle gradually as the dust cloud moves away from the blasting source. As a result, distance plays a significant role in characterising the spatial attenuation of PM₁₀ concentration.

The SHAP summary figure in Figure 15 displays a comparable but distinct pattern for TSP prediction. The number of blastholes, explosive quantity, and distance were also important factors. The anticipated TSP concentration was often higher for high explosive quantity values, indicating that energy input plays a significant role in the creation of total suspended particulate matter. Larger blasting scales may result in stronger dust emissions and higher concentrations of suspended particles, according to the positive SHAP trend.

TSP prediction was significantly impacted negatively by distance. Greater monitoring distances tended to lower the expected TSP concentration, according to the SHAP dependence plot. According to the feature importance results, this effect was more noticeable for TSP than for PM₁₀. TSP is more impacted by transport distance and gravity settling because it contains particles with a larger size range, including larger suspended particles. As a result, as one gets farther away from the blasting source, the observed TSP concentration may drop more noticeably.

Additionally, the SHAP results show that the model's output was influenced by a few secondary factors. Although they had lower average SHAP values, variables like moisture content, stemming column, blasthole diameter, burden, spacing, rock density, and rock type still had an impact on individual forecasts. For instance, the correlation analysis showed that moisture content generally had a dust-suppression tendency, but the SHAP results indicate that its influence might change depending on the blasting settings. This implies that rather

than functioning as a straightforward linear element, moisture content may interact with explosive quantity, rock type, and particle transport pathways.

As a whole, the SHAP interpretation shows that the main factors influencing blast-induced PM_{10} and TSP predictions are explosive quantity, number of blastholes, and distance. The anticipated dust concentration is mostly increased by explosive quantity and number of blastholes, but it is typically decreased by distance. These results align with the results of the feature importance analysis and the preliminary correlation analysis. More significantly, SHAP enhances the machine learning models' interpretability and promotes their application in engineering dust control decision-making by offering sample-level evidence for the direction and amount of variable impacts.

4. Discussion

4.1. Comparison of Model Performance

The findings of the model comparison demonstrate that tree-based ensemble models were appropriate for forecasting PM_{10} and TSP concentrations caused by blasts. With the highest R^2 and the lowest RMSE and MAE, CatBoost had the best test performance for PM_{10} prediction. ETR performed the best overall for TSP prediction, particularly when it came to test R^2 and RMSE. These findings show that blasting parameters and dust concentration have a nonlinear connection that cannot be adequately explained by straightforward linear hypotheses [22]. A quantitative comparison with the source study by Roy et al. [1] was conducted to better clarify the relative performance of the proposed framework. In the source study, the best multiple regression models reported R^2 values of 0.774 for PM_{10} and 0.758 for TSP. In comparison, the present optimal ensemble models achieved test R^2 values of 0.9555 for PM_{10} and 0.8803 for TSP, corresponding to absolute increases of 0.1815 and 0.1223, respectively. In relative terms, the R^2 values increased by approximately 23.4% for PM_{10} and 16.1% for TSP compared with the reported multiple regression benchmarks. Roy et al. [1] also reported ANN correlation coefficients of 0.97 for PM_{10} and 0.96 for TSP, which correspond approximately to R^2 values of 0.941 and 0.922 if squared correlation is used as a rough reference. Compared with these ANN results, the present CatBoost model showed comparable or slightly higher performance for PM_{10} , whereas the present ETR model showed a lower R^2 for TSP. Therefore, the improvement of the present framework should not be interpreted as a uniform accuracy increase over all models in the source study. Rather, the main advantage lies in the systematic comparison of modern ensemble models, multi-metric test-set evaluation, residual and concentration-band error analysis, and SHAP-based interpretation of variable effects. The model ranking should be interpreted with caution because the dataset is small and the final test set contained only 30 samples. A k-fold or repeated k-fold validation strategy would provide a more comprehensive estimate of metric variability. However, adding a post hoc non-nested k-fold validation only for the two selected optimal models may not fully resolve the uncertainty in model ranking, because model selection and hyperparameter tuning were originally performed under the fixed hold-out protocol. A rigorous cross-validation-based comparison would require repeating the full optimization and model-selection process for all candidate models within each fold, preferably using nested or repeated cross-validation. Given the limited sample size and the strongly right-skewed concentration distribution, such an expanded validation design was not included in the present revision. Therefore, the present results are interpreted as hold-out test performance under the current dataset, and close model rankings, especially among CatBoost, ETR, and GBRT for PM_{10} , should be regarded as indicative rather than definitive. It should be emphasized that a RMSE value lower than the standard deviation of the target variable is not direct evidence of overfitting. The standard deviation represents the dispersion of the measured concentrations around their mean, whereas RMSE represents the average prediction error. If a model captures useful relationships between the input variables and the target concen-

tration, its RMSE is expected to be lower than the standard deviation. Overfitting is more appropriately assessed by the gap between training and test performance, residual behavior, and validation design. In this study, the PM₁₀ model showed very close training and test performance, indicating no obvious overfitting. For TSP, the test error was higher than the training error, suggesting weaker generalization than PM₁₀; therefore, the TSP results were interpreted more cautiously and the need for further validation was acknowledged.

Attention should also be paid to the distinction between PM₁₀ and TSP prediction. TSP was not as well predicted as PM₁₀. This could be due to the fact that TSP encompasses a larger range of particle sizes and is more significantly impacted by local sampling conditions, air disturbance, and particle settling. PM₁₀, on the other hand, may have a more consistent relationship with blasting size and monitoring distance and indicates finer particles. Consequently, given the current dataset, the higher prediction error for TSP makes sense. A log-transform of the target variables and a relative-error-based loss were also considered during the methodological design. However, they were not adopted in the main analysis for three reasons. First, PM₁₀ and TSP concentrations need to be interpreted in their original physical units ($\mu\text{g}/\text{m}^3$) for dust control decision-making. Second, the original-scale MSE objective, together with RMSE and MAE, gives sufficient attention to high-concentration dust events, which are usually more important for risk identification and engineering control. Third, with the limited sample size of the present dataset, back-transformation from a logarithmic model may introduce additional bias and uncertainty. Nevertheless, log-transformed targets and relative-error-based optimization may help reduce relative errors in low-concentration samples, and they should be further examined in future studies using larger datasets.

Several models' efficiency was enhanced using Bayesian optimisation, albeit not all algorithms saw the same increase. This implies that while hyperparameter adjustment is helpful, the final prediction ability is still determined by the model structure and data properties. All things considered, CatBoost and ETR provide a decent trade-off between prediction accuracy and generalizability. As a result, they were chosen as the best models for additional analysis. Compared with previous studies that used empirical regression, multiple linear regression, or neural-network models for mine dust prediction, the present results show that tree-based ensemble models can provide competitive prediction performance for blast-induced PM₁₀ and TSP under nonlinear field conditions. Roy et al. [1] demonstrated the feasibility of regression and neural-network models for blasting dust assessment, but their discussion mainly focused on predictive accuracy. In contrast, the present study combines ensemble learning with feature importance and SHAP interpretation, which allows both prediction performance and variable effects to be analysed. Similar to recent open-pit dust prediction studies [22,23], the results confirm the usefulness of machine learning for dust concentration estimation. However, this study focuses specifically on short-duration blast-induced particulate matter and further distinguishes the predictive behaviour of PM₁₀ and TSP. The present study focused on tree-based ensemble models because they are well suited for small tabular engineering datasets and can capture nonlinear relationships without strict assumptions about data distribution. In addition, these models provide direct access to feature importance analysis and are compatible with SHAP interpretation. However, this model selection also has limitations. Non-tree-based algorithms, such as support vector regression, Gaussian process regression, or neural networks, and hybrid physics-informed models that incorporate dust-dispersion mechanisms may provide different advantages. These approaches were not fully compared in the present work because the main objective was to evaluate interpretable ensemble learning models using the available field monitoring dataset. Future studies with larger datasets should compare tree-based models with non-tree-based and physics-informed hybrid approaches to better identify their relative strengths and weaknesses.

4.2. Influence Mechanism of Key Variables

Based mainly on SHAP interpretation, and supported by correlation analysis and within-model feature importance rankings, explosive quantity, number of blastholes, and monitoring distance were identified as the primary factors influencing PM_{10} and TSP predictions. It should also be noted that some blasting design and material-related variables were correlated with each other. This reflects the actual coupling of field blast design rather than data leakage. For example, burden and spacing are usually adjusted together according to bench geometry, while blasthole diameter, rock type, and rock density are related to material and charge-structure conditions. However, the correlations among predictors may affect the allocation of feature importance and SHAP values. For example, burden and spacing describe related aspects of the blasthole layout, while blasthole diameter, rock density, and rock type are also partly coupled in the present field dataset. Under such multicollinearity, tree ensembles can still provide useful predictions, but the attribution assigned to one variable may partly reflect the information shared with its correlated variables. Therefore, the interpretation results in Figures 13–15 should be understood as model-based predictive contributions rather than independent causal effects. In both SHAP analyses, explosive quantity showed a dominant positive contribution pattern, while the within-model feature importance rankings also placed explosive quantity among the leading predictors. More energy is typically released, rock breakage intensity is increased, and more fine particles are produced with a bigger explosive quantity [22]. Consequently, it makes sense that this variable raised the anticipated dust concentration. These findings are generally consistent with earlier blasting dust studies, which reported that blast scale, explosive quantity, moisture condition, and monitoring distance are important factors affecting particulate matter concentration. The similarity indicates that the dominant predictors identified by the machine learning models are physically reasonable. However, the present results also show some differences from conventional correlation-based analyses. For example, stemming column showed weak pairwise correlation with PM_{10} and TSP, but it still contributed to individual predictions through nonlinear interactions in the ensemble models. This suggests that a variable with weak marginal correlation may still contain useful predictive information when combined with other blasting parameters. Therefore, the machine learning interpretation provides a more comprehensive view than single-factor correlation analysis.

Another significant component, particularly for PM_{10} prediction, was the quantity of blastholes. This variable represents the volume of disturbed rock and the size of the blasting round. A larger dust source region and a higher potential for particle release are typically associated with more blastholes. However, other design factors like hole diameter, weight, spacing, and stemming length may also have an impact. This explains why a straightforward single-factor analysis is not as appropriate as the machine learning approach.

Dust concentration was clearly negatively impacted by monitoring distance, particularly for TSP. Dust particles are spread, diluted, and deposited as one gets farther away from the blasting source. In TSP, larger suspended particles are more susceptible to gravitational settling and transit distance. The link between moisture content and dust concentration was weaker but generally negative. This implies that increased moisture content could promote particle agglomeration and lessen dust suspension. Although these findings are in line with real-world blasting dust behaviour, they should be viewed as data-driven correlations rather than rigid causal conclusions.

4.3. Engineering Implications

The engineering implications of this study should be understood based on the combined evidence from model performance, residual analysis, feature importance, SHAP interpretation, and the comparison with previous studies. The high test performance of the optimal models indicates that blast-induced PM₁₀ and TSP concentrations can be reasonably estimated from blasting design parameters, monitoring distance, and material-related variables. However, the prediction results should not be regarded only as a numerical forecasting exercise. More importantly, the interpretation results provide a basis for identifying which blasting conditions are more likely to cause high dust emissions and why these conditions require additional control measures.

Explosive quantity was consistently identified as the most influential predictor for both PM₁₀ and TSP. This result was supported by three types of evidence. First, the correlation analysis showed a significant positive association between explosive quantity and the two dust indicators. Second, within each optimal model, feature importance analysis ranked explosive quantity among the leading predictors, although the percentages were not directly compared across PM₁₀ and TSP because different model families were used. Third, SHAP analysis consistently showed that higher explosive quantity generally produced positive SHAP values, indicating an increase in predicted dust concentration. This finding is consistent with the physical mechanism of blasting, because larger explosive charges release more energy, disturb a larger rock volume, and increase the probability of fine-particle generation. It is also consistent with previous studies that emphasized the role of blast scale and explosive energy in dust generation. Therefore, for blasting operations with high charge quantities, dust control measures such as pre-wetting, water spraying, and local suppression should be planned in advance. This implication is not only a general recommendation, but is directly derived from the model evidence that explosive quantity strongly increases the predicted dust concentration.

The number of blastholes was another important predictor, especially for PM₁₀. This variable represents the scale of the blasting round and is closely related to the disturbed rock volume. The SHAP results showed that a larger number of blastholes generally increased the predicted PM₁₀ concentration, while feature importance analysis also ranked this variable among the top predictors. This indicates that dust control planning should not focus only on the total explosive charge, but should also consider the spatial scale of the blast. Compared with simple regression-based conclusions, this result suggests that blast layout and blast scale may jointly influence dust emissions through nonlinear interactions with burden, spacing, blasthole diameter, and stemming length. Therefore, practical dust reduction should be linked with the optimization of the overall blast design rather than the adjustment of a single parameter.

Monitoring distance showed a clear negative effect on dust concentration, particularly for TSP. This result was supported by the negative correlation coefficients and the SHAP distributions, where larger distances generally reduced the predicted dust concentration. The stronger role of distance in TSP prediction is reasonable because TSP includes a broader particle-size range and is more affected by dilution, gravitational settling, and transport distance. This result is consistent with the general understanding of dust dispersion in surface mines, but the present model further quantifies its predictive contribution. From an engineering perspective, this finding indicates that monitoring point arrangement should cover both near-field and far-field locations. Near-field monitoring is useful for capturing high-concentration dust episodes, whereas far-field monitoring is necessary for evaluating the attenuation range and potential environmental influence of the dust cloud.

Moisture content showed a generally negative association with PM₁₀ and TSP, although its importance was lower than that of explosive quantity, number of blastholes, and monitoring distance. This indicates that moisture condition is a secondary but still meaningful

factor in blast-induced dust control. The negative effect is physically reasonable because water can promote particle agglomeration and reduce the suspension of fine particles. This result supports the practical use of pre-wetting or water-based suppression in dry blasting conditions. However, the SHAP results also indicate that the effect of moisture content may vary among samples, suggesting that its dust-suppression effect can interact with explosive quantity, rock type, and transport conditions. Therefore, moisture control should be regarded as a supporting measure rather than an independent solution.

Compared with previous studies, the main contribution of the present engineering discussion is that the suggested control strategies are supported by interpretable machine learning evidence rather than only by empirical judgement. Earlier studies have shown that blasting dust is affected by explosive quantity, moisture content, monitoring distance, and mine operating conditions. The present study reaches a similar general understanding, but further shows how these variables contribute to PM₁₀ and TSP prediction in different ways. PM₁₀ prediction was more stable and was more strongly associated with blast scale variables, whereas TSP prediction showed larger uncertainty and stronger sensitivity to transport-related effects. This difference suggests that dust control planning should distinguish between fine particulate matter and total suspended particulate matter. In practical applications, the proposed framework can be used as a decision-support tool to identify high-risk blasting scenarios, optimize monitoring locations, and select appropriate dust control measures. Nevertheless, these implications should be applied together with field monitoring and site-specific engineering judgement, because meteorological factors such as wind speed, wind direction, humidity, and atmospheric stability were not included in the current dataset. The practical value of interpretable machine learning has also been demonstrated in blasting-engineering applications, including tunnel overbreak prediction and open-pit flyrock forecasting [24,25]. These studies show that interpretation methods can connect model outputs with engineering parameters and support engineering decision-making. In the present study, the positive SHAP effects of explosive quantity and number of blastholes indicate that high-charge and large-scale blasting rounds are more likely to increase PM₁₀ and TSP predictions. In contrast, the negative SHAP effect of monitoring distance reflects the general attenuation of dust concentration away from the blasting source. Therefore, SHAP helps translate the fitted models into actionable evidence for identifying high-risk blasts, arranging near- and far-field monitoring points, and selecting appropriate pre-wetting or local suppression measures. In contrast, the negative SHAP effect of monitoring distance explains why dust concentration generally decreases as the sampling point moves away from the blasting source. Therefore, SHAP helps connect model prediction with actionable dust control decisions, such as identifying high-risk blasts, arranging near-field and far-field monitoring points, and selecting appropriate pre-wetting or local suppression measures.

4.4. Limitations and Future Work

Although the proposed machine learning framework achieved good prediction performance for PM₁₀ and TSP in the present dataset, several limitations should be acknowledged. First, the dataset was obtained from published field monitoring records from one large surface coal mine. After data verification and preprocessing, only 148 valid samples were available for model development and evaluation. This sample size is relatively small for training machine learning models, especially when nonlinear ensemble models are used. Therefore, although the training and test results showed good performance, the possibility of overfitting cannot be completely excluded. This is especially relevant for TSP, where the training–test error gap was larger than that of PM₁₀. The results should be interpreted as site-specific prediction performance under the present data conditions rather than as a universally validated model for all surface mines. Another limitation is that the

present study did not perform an exclusion-based validation by removing very large blasting rounds from the test set. Because the dataset is small and the high-explosive-quantity samples are valid field observations, removing them would further reduce the already limited number of test samples and would change the original field-data distribution. Nevertheless, these samples may partly influence the EQ–dust association and the high R^2 values. Future studies should collect larger datasets with more balanced distributions of blasting scale and should explicitly evaluate model robustness on non-extreme, medium-scale, and high-energy blasting subsets.

Second, the current dataset mainly includes blasting design parameters, monitoring distance, and material-related variables. Crucial meteorological factors, including wind speed, wind direction, temperature, humidity, and atmospheric stability, were not included. This limitation may affect the engineering application of the proposed framework, especially for TSP prediction and far-field dust dispersion assessment. Dust deposition, dilution, and diffusion following blasting can be significantly influenced by these factors. Therefore, the higher prediction errors for TSP, particularly in high-concentration samples, could be partially explained by the absence of meteorological information.

Third, the model validation was based on an 80:20 train–test split, which allowed an independent test set to be retained for final performance evaluation. However, for a small dataset, the estimated performance may still be sensitive to the specific data partition. Although k-fold cross-validation can provide mean and standard-deviation estimates of model performance, a rigorous use of cross-validation for model ranking would require the full model-selection and hyperparameter-optimization process to be repeated within each fold. In the present study, such nested or repeated cross-validation was not performed. Therefore, the reported model ranking should be interpreted as the result under the adopted hold-out validation scheme. Future studies should use repeated k-fold cross-validation, nested cross-validation, or bootstrap resampling to obtain more reliable estimates of model generalization performance and metric uncertainty.

Fourth, the present study did not include external validation using data from other mines. Differences in mine geometry, lithology, blasting practice, explosive type, climate, and dust control measures may affect the relationship between input variables and dust concentration. Therefore, future work should collect larger datasets from multiple surface mines and different seasons, incorporate meteorological and operational variables, and evaluate the model using independent external datasets. Such efforts would improve the external validity and practical applicability of the proposed framework.

Finally, the interpretation results obtained from feature importance and SHAP analysis should be understood as model-based predictive associations rather than strict causal relationships. Future research can combine machine learning with physical dust-dispersion models, field sensor networks, uncertainty analysis, and real-time monitoring data. This would support the development of more reliable early-warning and decision-support systems for blast-induced dust control in surface mines. In addition, future studies should compare the present tree-based ensemble framework with non-tree-based algorithms and physics-informed hybrid models that incorporate dust-generation and dust-dispersion mechanisms.

5. Conclusions

A machine learning-based framework was developed to predict blast-induced PM_{10} and TSP concentrations using a published field monitoring dataset from a large surface coal mine. The information, which comprised blasting parameters, monitoring distance, material conditions, and dust concentration indicators, was gathered from publicly available blasting dust monitoring records. In order to enhance model performance, six tree-

based ensemble models were compared using Bayesian optimisation. The following are the primary conclusions:

- (1) PM₁₀ and TSP concentrations had considerable variability and right-skewed distributions, according to descriptive statistics and correlation analysis. There was a significant positive correlation between PM₁₀ and TSP. While monitoring distance and moisture content generally exhibited negative relationships, explosive quantity clearly indicated a positive correlation with dust concentration.
- (2) The model comparison indicated that ensemble tree-based models achieved good prediction performance under the adopted hold-out validation scheme. CatBoost showed the highest test R² for PM₁₀ prediction, whereas ETR showed the highest test R² for TSP prediction. However, because only 148 samples were available and the test set contained 30 samples, the model ranking should be interpreted with caution and further verified using repeated or nested cross-validation in future studies. These findings imply that the nonlinear relationship between blasting settings and dust emissions can be effectively described by ensemble learning models. Compared with the source study, the proposed models clearly improved upon the reported multiple regression benchmarks, with R² increases of approximately 23.4% for PM₁₀ and 16.1% for TSP. Compared with the reported ANN results, the present PM₁₀ model showed comparable or slightly higher performance, whereas the TSP model showed a lower R². Therefore, the main contribution of the proposed framework lies not only in prediction accuracy, but also in systematic ensemble-model comparison, multi-metric evaluation, and interpretable variable-effect analysis.
- (3) The best models demonstrated strong prediction stability and no discernible systematic bias, according to residual analysis and measured-versus-predicted plots. TSP displayed greater residual variation, whereas PM₁₀ was predicted more precisely. This discrepancy could be explained by TSP's larger dispersion uncertainty and broader particle-size range.
- (4) Based mainly on SHAP interpretation and supported by within-model feature importance rankings, explosive quantity, number of blastholes, and monitoring distance were consistently identified as the main factors influencing PM₁₀ and TSP predictions. While monitoring distance generally decreased the projected dust concentration, explosive quantity and number of blastholes primarily raised it. The optimisation of blast parameters, dust monitoring arrangement, and dust control planning in surface mines can all benefit from these findings.

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

Funding: This research is partially supported by the National Major Science and Technology Project of China (No. 2025ZD1010703) and the Scientific Research Foundation of Hunan Provincial Education Department (25A0011).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The cleaned dataset and Python scripts run under Python 3.9.21 used for data preprocessing, model training, Bayesian optimization, model evaluation, SHAP analysis, and figure generation are available from the corresponding author upon reasonable request. The

original field monitoring records were obtained from the published dataset reported by Roy et al. [1].

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

References

- Roy, S.; Adhikari, G.R.; Renaldy, T.A.; Jha, A.K. Development of multiple regression and neural network models for assessment of blasting dust at a large surface coal mine. *J. Environ. Sci. Technol.* **2011**, *4*, 284–301.
- Xie, Z.; Huang, C.; Zhao, Z.; Xiao, Y.; Zhao, Q.; Lin, J. Review and prospect the development of dust suppression technology and influencing factors for blasting construction. *Tunn. Undergr. Space Technol.* **2022**, *125*, 104532.
- Shi, J.; Zhang, W.; Guo, S.; An, H. Numerical modelling of blasting dust concentration and particle size distribution during tunnel construction by drilling and blasting. *Metals* **2022**, *12*, 547.
- Hosseini, S.; Pourmirzaee, R. Green policy for managing blasting induced dust dispersion in open-pit mines using probability-based deep learning algorithm. *Expert Syst. Appl.* **2024**, *240*, 122469.
- Bui, X.N.; Lee, C.W.; Nguyen, H.; Bui, H.B.; Long, N.Q.; Le, Q.T.; Moayedi, H. Estimating PM₁₀ concentration from drilling operations in open-pit mines using an assembly of SVR and PSO. *Appl. Sci.* **2019**, *9*, 2806.
- Li, L.; Zhang, R.; Sun, J.; He, Q.; Kong, L.; Liu, X. Monitoring and prediction of dust concentration in an open-pit mine using a deep-learning algorithm. *J. Environ. Health Sci. Eng.* **2021**, *19*, 401–414.
- Lal, B.; Tripathy, S.S. Prediction of dust concentration in open cast coal mine using artificial neural network. *Atmos. Pollut. Res.* **2012**, *3*, 211–218.
- Luan, B.; Zhou, W.; Jiskani, I.M.; Wang, Z. An improved machine learning approach for optimizing dust concentration estimation in open-pit mines. *Int. J. Environ. Res. Public Health* **2023**, *20*, 1353.
- Zhou, J.; Zhang, Y.; Qiu, Y. State-of-the-art review of machine learning and optimization algorithms applications in environmental effects of blasting. *Artif. Intell. Rev.* **2024**, *57*, 5.
- Rane, N.; Choudhary, S.P.; Rane, J. Ensemble deep learning and machine learning: Applications, opportunities, challenges, and future directions. *Stud. Med. Health Sci.* **2024**, *1*, 18–41.
- Shahhosseini, M.; Hu, G.; Pham, H. Optimizing ensemble weights and hyperparameters of machine learning models for regression problems. *Mach. Learn. Appl.* **2022**, *7*, 100251.
- Ren, Y.; Zhang, L.; Suganthan, P.N. Ensemble classification and regression-recent developments, applications and future directions. *IEEE Comput. Intell. Mag.* **2016**, *11*, 41–53.
- Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. *arXiv* **2018**, arXiv:1706.09516.
- Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. *arXiv* **2017**, arXiv:1705.07874.
- Rigatti, S.J. Random forest. *J. Insur. Med.* **2017**, *47*, 31–39.
- Ahmad, M.W.; Reynolds, J.; Rezgui, Y. Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees. *J. Clean. Prod.* **2018**, *203*, 810–821.
- Prettenhofer, P.; Louppe, G. Gradient boosted regression trees in scikit-learn. In Proceedings of the PyData 2014, London, UK, 21–23 February 2014.
- Ozer, D.J. Correlation and the coefficient of determination. *Psychol. Bull.* **1985**, *97*, 307.
- Hodson, T.O. Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geosci. Model Dev. Discuss.* **2022**, *15*, 5481–5487.
- De Myttenaere, A.; Golden, B.; Le Grand, B.; Rossi, F. Mean absolute percentage error for regression models. *Neurocomputing* **2016**, *192*, 38–48.
- Huang, S.; Yong, W.; Wang, L.; Zhou, J. From measurement while drilling data to rock strength: A robust LMWOA-XGBoost model for predicting uniaxial compressive strength with SHAP interpretability. *Expert Syst. Appl.* **2026**, *325*, 132535.
- Wang, G.; Zhou, W.; Wang, Z.; Lu, X.; Zhang, Y. Characterization and concentration prediction of dust pollution in open-pit coal mines. *Atmosphere* **2024**, *15*, 1408.
- Yang, Y.; Zhou, W.; Wang, Z.; Jiskani, I.M.; Yang, Y. Accurate long-term dust concentration prediction in open-pit mines: A novel machine learning approach integrating meteorological conditions and mine production intensity. *J. Clean. Prod.* **2024**, *436*, 140411.

24. Zhang, Y.; Zhou, J.; Li, J.; He, B.; Armaghani, D.J.; Huang, S. Advancing overbreak prediction in drilling and blasting tunnel using MVO, SSA and HHO-based SVM models with interpretability analysis. *Geomech. Geophys. Geo-Energy Geo-Resour.* **2025**, *11*, 53.
25. Zhang, Y.; Qiu, Y.; Du, K.; Nguyen, H.; Armaghani, D.J.; Zhou, J. Optimizing flyrock forecasting in open-pit blasting using hybrid machine learning models. *Rock Mech. Rock Eng.* **2025**, *58*, 12523–12550.

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.