European Journal of Wood and Wood Products (2025) 83:173
https://doi.org/10.1007/s00107-025-02306-z

ORIGINAL ARTICLE

Prediction of dust emissions during CNC milling of spruce and pine with machine learning

Evren Osman Çakiroğlu[1]

Received: 25 March 2025 / Accepted: 23 July 2025

Abstract

Wood dust generated by CNC machines during milling is hazardous to human health. This study aims to determine the wood dust emissions (PM2.5, PM10) according to the wood species, spindle speed (12000 rpm, 15000 rpm, and 18000 rpm), feed rate (3 m/min, 6 m/min, and 9 m/min), and cutting direction (contours, linear and spiral), and to predict them with machine learning algorithms. Oriental spruce (*Picea orientalis* L.) and Scots pine (*Pinus sylvestris* L.), known for their low and high dust emission values, respectively, were used as wood species. A blade with a diameter of 3 mm was preferred as a cutter for milling both wood species. The results of the analyses show that spindle speed, feed rate, and cutting direction parameters have a significant effect on PM. According to the PM2.5 and PM10 values, the highest wood dust emissions were measured at 121.42 µg/m³ and 173.02 µg/m³, respectively, in Scots pine wood material, with a spindle speed of 18,000 rpm, a feed rate of 3 m/min, and cutting direction being linear. The lowest wood dust emission was measured as 4.20 µg/m³ and 7.40 µg/m³ for PM2.5 and PM10 values, respectively, at a feed rate of 6 and 9 m/min, 15,000 rpm in Oriental spruce wood material under the conditions of cutting direction. However, the Cubist model performed best among the machine learning algorithms for predicting PM2.5 and PM10 levels. This study aims to provide data on wood dust emissions during CNC milling to inform the development of CNC parameter adjustments that minimize dust generation.

1 Introduction

With advancing technology, wood dust emission reduction activities, which are very important for human health, can be applied during CNC use. However, although methods to reduce the amount of wood dust are applied during production, a certain amount of dust remains exposed. Wood dust-raising processes such as cutting, milling, and carving with CNC machines significantly impact human health and production quality. These wood dusts, which mix into the air, increase the rate of respiratory problems and lung diseases among employees. As a result, it causes significant problems such as cancer. Wood dust, which impacts human health, can cause respiratory problems such as asthma and chronic bronchitis (Alapieti et al. 2020). According to the results of epidemiological studies, there is an association between exposure to wood dust from hardwoods such as oak, mahogany, beech, walnut, and birch and the possibility of developing nasal cavity cancer. With increasing rates of nasal cavity and paranasal sinus cancers in workers exposed to wood dust, the International Agency for Research on Cancer (IARC) has classified wood dust as a group 1 carcinogen to humans (Vallières et al. 2015). According to the standard set by the Occupational Safety and Health Administration (OSHA), the permissible limit of dust exposure for 8 h of work per day is 15 mg/m3. In addition, the National Institute for Occupational Safety and Health (NIOSH) has recommended that a person's permissible level of dust exposure is 1 mg/m3 (Omidianidost et al. 2015). Wood dust ranges in particle size from 1 to 500 µm and is classified into coarse, medium, or fine dust categories. These particles can spread in large quantities and significantly accumulate in the environment (Hlásková et al. 2016; Markova et al., 2018; Igaz et al. 2018; Kminiak and Dzurenda 2019). The amount of Particulate Matter (PM) produced varies depending on the properties of the wood species, the CNC machine's processing parameters, the duration of the processing, and the CNC technology applied (Palmqvist and Gustafsson 1999;

✉ Evren Osman Çakiroğlu
61evrenosman@artvin.edu.tr

[1] Artvin Vocational School, Design Department, Artvin Çoruh University, Interior Design Program, Artvin, Turkey

Published online: 08 September 2025

Ratnasingam et al. 2010; Fujimoto et al. 2011; Hlásková et al. 2015; Očkajová et al. 2016).

Wood dust-generating processes include sawing, routing, turning, and milling (Kminiak et al. 2021; Tong et al. 2018). Sanding is an essential part of the upper surface treatment (Holla et al. 2016; Kos et al. 2004). Sydor et al. (2023) investigated the importance of wood dust during the sanding process in thermally modified (160 °C and 220 °C) and non-thermally modified spruce, oak, and meranti wood materials. As in sanding, CNC processes utilizing fine-cutting tools can also create fine dust particles. In another study in which thermally modified (190 °C and 212 °C) and unmodified Scots pine (Pinus sylvestris L.) wood samples were examined, Hanincová et al. (2024) determined the amount of harmful fine wood dust.

Recently, forecasting methods with machine learning have been increasing. With machine learning, data can be obtained easily and quickly without dealing with any complex formulation. Machine learning models, trained on experimental data, can extrapolate resonance behavior beyond the experimental dataset, leading to predictions applicable to real-world conditions (Akkoyun et al. 2024). These methods provide a low-cost, time-efficient, and data-driven approach to make processes more understandable and optimised (Jawad et al. 2021). Some of the related studies include performance prediction in composite materials Kumar and Ghosh (2021), prediction of wood density Iliadis et al. (2013), prediction of wood resistance Miguel et al. (2018), prediction method based on vibration test You et al. (2022), prediction method for measuring wood volume Batista et al. (2022), prediction of wood surface colour change Mo et al. (2023).

Studies also show how estimation methods are used in CNC and wood processing. Demir et al. (2022) obtained data using an Artificial Neural Network (ANN) to determine CNC machining parameters for the best wood surface quality. Çakıroğlu et al. (2022) determined optimum cutting conditions in their study for the best wood surface quality, energy, and time-saving using artificial neural network (ANN) models. Demir et al. (2024) conducted a study to determine the optimum CNC cutting parameters by modeling the surface roughness and lacquer wettability of medium-density fibreboard (MDF) panels using ANN models on CNC machines. The study by Singer et al. (2024) aimed to develop an artificial neural network (ANN) model to predict dust emission values in the processing of some solid materials and medium-density fibreboard (MDF).

In addition to these, there are a limited number of studies on dust emissions in CNC machine production in the wood industry. Simon et al. (2022) investigated wood dust emissions while processing solid and wood-based materials. Kminiak et al. (2021) investigated dust characterization during CNC milling of MDF and solid material. Another study investigated the effect of feed rate and width of cut—transformed into kinematic average chip thickness—on the amounts of chips of respirable sizes < 10.0 µm created in the cutting zone from particleboards (PB) and medium-density fibreboards (MDF) (Pędzik et al. 2024). In another study, Rabiei and Souri (2024) investigated the parameters that affect wood dust (PM) produced in CNC milling. In another study, the effect of some CNC parameters was investigated using particleboards (PB) and medium-density fibreboards (MDF) (Juda et al. 2024). A study examining particle boards analyzed particles formed in milling with different feed rates (Koleda et al. 2023). Another study analyzed the size distribution of wood dust particles of natural wood and some wood species with low heat treatment (105, 125, and 135 °C) (Juda et al., 2023). Another study using a CNC machine aimed to determine wood dust emissions (PM2.5, PM10), according to processing parameters and wood species while performing 3D machining on wood materials (Çakıroğlu et al. 2025).

Previous studies indicate that CNC machine parameters are of great importance in generating fine, respirable wood dust for sustainable manufacturing processes in terms of human health. Therefore, most research focuses on controlling lower dust emissions by using different wood species and CNC process parameters. In this study, in addition to the spindle speed and feed rate parameters applied in the literature for dust emission measurements in CNC machines, the cutting direction parameter was added, which has not been investigated before. In line with the results and recommendations obtained from our previous study, Çakıroğlu et al. (2025) found two different wood species (Oriental spruce and Scots pine) with the lowest and highest wood dust emission.

This study aims to determine and estimate the concentration of wood dust released into the air during wood processing with a CNC machine under various CNC machine parameter settings, using machine learning algorithms. Its purpose is to reveal the danger to human health. Based on this, the Cubist Model, Support Vector Regression (SVR), Random Forest (RF), Bayesian Regularized Neural Network (BRNN), and Extreme Gradient Boosting (XGBoost) models are used, with their results being discussed and compared in detail.

2 Materials and methods

2.1 CNC machine

The samples were processed using Megatron 2128 (Bursa/Turkey), a four-axis CNC Milling Machine with 9 kW

Fig. 1 The 4-axis CNC router, model 2128, manufactured by Megatron

Table 1 CNC machine machining parameters and test set-up

4-axis CNC router, model 2128, manufactured by Megatron	
Spindle speed	12,000, 15,000, 18,000 rpm
Feet rate	3,6,9 m/min
CNC tool type	Tungsten carbide milling cutter (0450-01, Netmak Kesici Tak.Ltd.Şti., Istanbul, Turkey)
CNC tool diameter	3 mm
Milling cutter strategy	Contours, linear, spiral
Down feed	4 m/min
Wood species	Oriental spruce and Scots pine
Dimensions of workpieces	R: Ø15cm
CNC dust extraction system	Off
Depth of cut	3 mm
Dust sampling method	NIOSH 0500

spindle power and a maximum spindle speed of 24,000 rpm. The Alphacam (Ver. 2023.1, Vero Software, Cheltenham, England) was used to generate tool paths for the samples (URL-1). The contour, linear, and spiral toolpath strategies were constantly used in the CNC machining of all samples. Figure 1 shows the 4- axis CNC machine used (mod. 2128 MEGATRON, Gama Mekatronik Ltd.Şti, Bursa, Turkey). Table 1 shows the CNC machining parameters andexperimental setup for the wood samples.

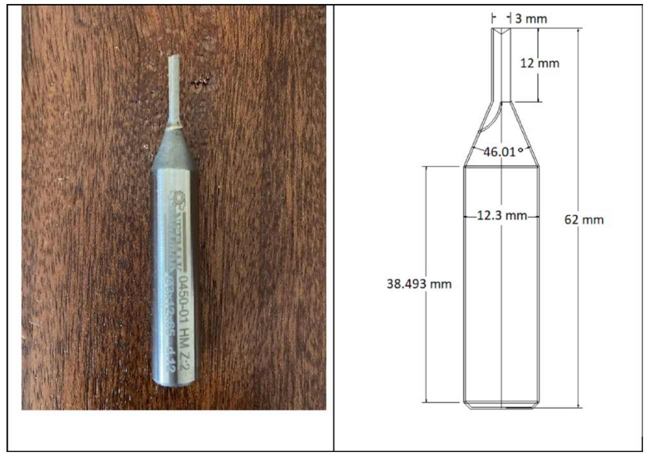

Fig. 2 Milling cutter and technical dimensions

Fig. 3 Oriental spruce and Scots pine cross-sectional images

2.2 Cutter set for CNC machine

The CNC machine has 8 different blade sections where different cutting operations are performed. Each of these cutting blades is prepared in calibrated form. The CNC tool type used in this study was the tungsten carbide milling cutter (0450-01, Netmak Kesici Tak.Ltd.Şti., Istanbul, Turkey). It is a straight cutter with a double-fluted bit. Technical specifications and drawings of the blade are shown in Fig. 2.

2.3 Wood species used

This study used Oriental spruce (*Picea orientalis* L.) and Scots pine (*Pinus sylvestris* L.) as wood species. Wood materials were obtained from the timber production facility (Artvin/Turkey) in kiln-dried form. Figure 3 shows cross-sectional images of the wood species used in the study.

2.4 Operation mode and tool path strategy

The study defined a circle to cover arcs and circular drawings, while creating a wooden model. Circular wooden samples, drawn and modeled with the Alphacam program, are prepared on a CNC machine. ‘G codes’ are created and processed with the Alphacam program. With the different CNC parameters determined, a specific dispersion characterization of wood powders emerges. One of these parameters, tool path strategy, is shown in Fig. 4.

These strategies have been determined as contours (the tool moves around the geometry and tries to conform to the shape of the geometry as far as the tool diameter allows), linear (the tool follows a linear path to extract the shape of the geometry), or spiral (the tool follows a spiral path to extract the shape of the geometry). Figure 5 shows the modeling of team strategies with the Alphacam program.

Figure 6 shows the numbering and positions of the Oriental spruce and Scots pine measurement samples, modeled with Alphacam according to CNC parameters.

Measurements are made separately for each sample of Scots pine and Oriental spruce wood species processed with a CNC machine. For both tree species, machining is carried out with an unused clean knife having a diameter of 3 mm. The boundaries of the processing area were determined as a circular region with a diameter of 15 cm. In order to make the measurements, 54 different milling operations were performed: 27 sample groups for Scots pine and 27 sample groups for Oriental spruce. According to the machining parameters, 27 “G codes” were created for tree species. These sample groups are shown in Table 2. To obtain more accurate measurement results, the samples were taken from timber cut in the same direction. For each wood sample, a measurement is taken every 10 s, and this process is repeated for a total of 54 sample groups.

Figure 7 shows the positions of Oriental spruce and Scots pine wood samples processed on the CNC machine and the distribution of wood dust generated during measurement with the PCE-MPC 10 particle measurement device.

2.5 Measurement of PMs

The PCE-MPC 10 particle measurement device (PCE Deutschland GmbH, Meschede, Deutschland) is a portable instrument developed to evaluate indoor air quality (IAQ). This device monitors the concentration of particulate matter in the air, enabling the measurement of particles in both PM2.5 and PM10 size ranges. The measurement range of the PCE-MPC 10 varies between 0 and 2000 µg/m³ for PM2.5 and PM10 particles. Additionally, the device supports the measurement of temperature and relative humidity, providing a comprehensive indoor environment analysis. When inhaled, PM2.5 and PM10 particles can penetrate the lungs, potentially leading to serious health issues. Therefore, the PCE-MPC 10 particle measurement device is a significant tool for monitoring air quality and identifying risks, particularly in occupational health and safety applications (URL-2).

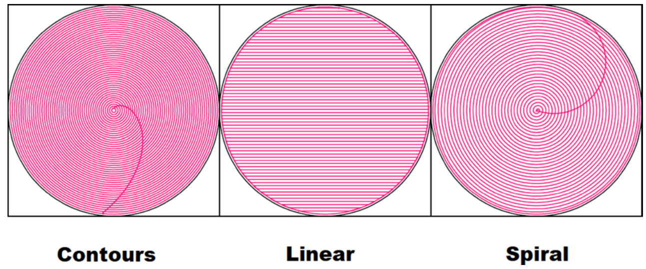

Fig. 4 Tool path strategy created with Alphacam programme

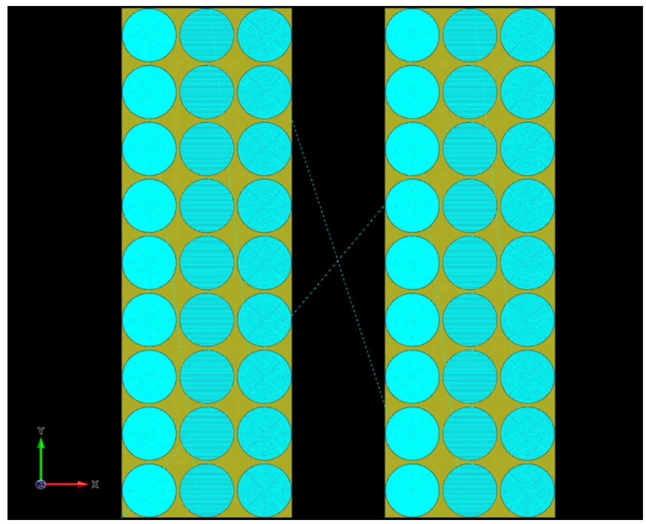

Fig. 5 Model created in alphacam programme

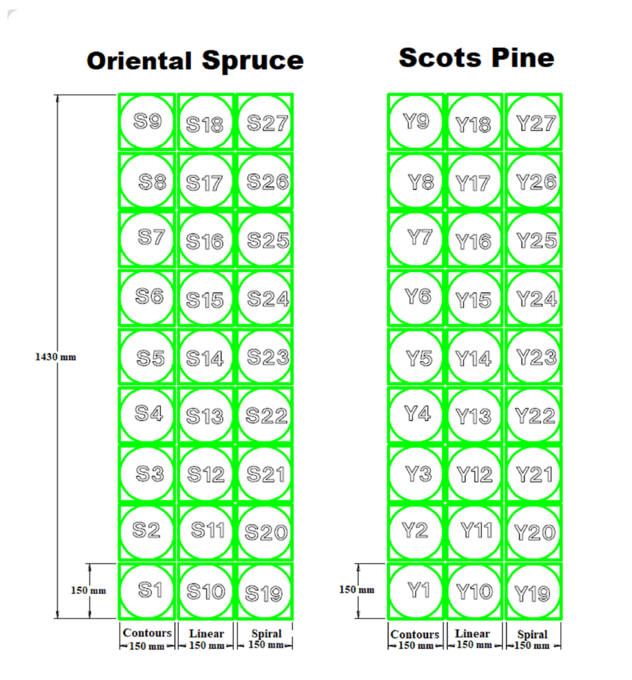

Fig. 6 Oriental spruce (S) and Scots pine (Y) measurement samples modelled with Alphacam Programme

Table 2 Sample groups formed according to different parameters for oriental spruce and scots pine

Wood	Tool Path Strategy			Spindle Speed (rpm)	Feed Rate (m/min)
	Contours	Linear	Spiral		
Oriental spruce	S1	S10	S19	12,000	3
	S2	S11	S20	12,000	6
	S3	S12	S21	12,000	9
	S4	S13	S22	15,000	3
	S5	S14	S23	15,000	6
	S6	S15	S24	15,000	9
	S7	S16	S25	18,000	3
	S8	S17	S26	18,000	6
	S9	S18	S27	18,000	9
Scots Pine	Y1	Y10	Y19	12,000	3
	Y2	Y11	Y20	12,000	6
	Y3	Y12	Y21	12,000	9
	Y4	Y13	Y22	15,000	3
	Y5	Y14	Y23	15,000	6
	Y6	Y15	Y24	15,000	9
	Y7	Y16	Y25	18,000	3
	Y8	Y17	Y26	18,000	6
	Y9	Y18	Y27	18,000	9

Fig. 7 Wood samples treated with CNC machine and particle measurement device

Before starting the measurements, witness data were taken on Monday when the workshop had yet to commence operations. The purpose of taking witness measurements on Monday morning is to ensure that particulate matter in the air settles over the weekend, as the workshop is closed on Saturdays and Sundays. Using this method, the PM levels in the ambient air were measured at the lowest point before the workshop began and were accepted as reference levels.

2.6 Machine learning algorithms

2.6.1 Cubist model

It is an explainable regression model based on decision trees. Cubist, which is an extension of Quinlan's M5 model tree (Quinlan 1992), uses linear regression equations at the leaves of the decision tree, and makes flexible predictions by transforming the branches in the tree into 'if-then' rule sets (Kuhn and Quinlan 2021b). This approach produces more flexible and interpretable results using decision trees with a rule-based system (Amrani et al. 2024; Bayram et al. 2023).

2.6.2 Support vector regression (SVR)

SVR is a powerful regression technique that projects data into high-dimensional space using the kernel trick to model non-linear relationships (Drucker et al. 1997). This method captures non-linear patterns by mapping the data in the input space to a higher-dimensional feature space, thus learning the relationships between inputs and outputs (Smola and Schölkopf 2004; Yeşilkanat and Akkoyun 2023). Thanks to kernel functions, data that cannot be decomposed linearly in the original space are decomposable in high-dimensional space (Awad and Khanna 2015; Hastie et al. 2009).

2.6.3 Random forest (RF)

Random Forest is an ensemble learning method. It produces a single result by combining the output of multiple decision trees and thus reduces the generalization error compared to a single tree (Breiman 2001; Neudecker et al. 2020). When multiple decision trees work together, higher accuracy and generalization performance are achieved, especially if the correlation between trees is low. This method also reduces the risk of overfitting by averaging the tree results (Yeşilkanat 2020).

2.6.4 Bayesian regularized neural network (BRNN)

BRNN focuses on penalizing model complexity and preventing overfitting using Bayesian regularisation in artificial neural networks (Mackay 1992). These networks resist overlearning because Bayesian regularisation effectively 'trims' unnecessary weights during training, it reduces the number of effective weights by turning off unnecessary parameters and creates simpler, more generalizable models (Burden and Winkler 2009). BRNNs are more robust than classical backpropagation networks and can learn without overfitting, even on small data sets (Akkoyun et al. 2024).

2.6.5 Extreme gradient boosting (XGBoost)

XGBoost is an algorithm based on the tree boosting learning technique where multiple weak models are transformed into strong predictors by successive training. XGBoost (Extreme Gradient Boosting), developed by Chen and Guestrin, is an optimized gradient boosting library with regularization added (Chen and Guestrin 2016a). This model is characterized by its fast and high performance, which avoids overfitting through regularisation and support parallel processing (Zhang et al. 2021, 2022). As a result, successful prediction performance have been obtained in many different applications in the literature with this algorithm, which has a high error reduction capacity (Acıslı-Celik and Yeşilkanat 2023; Bentéjac et al. 2021; Chen et al. 2021).

2.7 Performance metrics

The model performances used in the study were evaluated using statistical metrics. The metrics used in this study are MAE (Mean Absolute Error; Eq. (1)), RMSE (Root Mean Squared Error; Eq. (2)), R^2 (Coefficient of Determination; Eq. (3)), and ρc (Lin's Correlation Coefficient; Eq. (4)). MAE (Mean Absolute Error) is a straightforward metric that measures the error magnitude of the estimates. MAE has the same unit as the original data. It represents the average deviation across all samples, independent of other errors, therefore giving equal weight to the error of each sample in the data and being less sensitive to outlier observations. The RMSE (Root Mean Squared Error) is the standard deviation of the model's prediction errors. The smaller the RMSE value, the less error the model has on the data, meaning that its predictions are closer to reality. Its increased sensitivity is attributed to the squaring of significant errors. R^2 (Coefficient of Determination) is a statistical metric showing the model's explanatory power. The R^2 value is expressed as a percentage of how much of the total variance in the dependent variable is explained by the model. This metric takes a value between 0 and 1, and the closer the value is to 1 (i.e., 100%), the better the model explains the variability in the data. Lin's Concordance correlation coefficient (ρ_c) is a measure that simultaneously assesses the accuracy and precision of model predictions. This coefficient indicates how close the predicted and actual values are to the 45° line of perfect agreement (identity line). ρc takes a value between −1 and +1: +1 means perfect agreement (all points are on the 45° line), 0 means randomness, and −1 means opposite agreement. In the scope of this study, The metric is used to summarise both the systematic error and the random error of the models in a single value.

$$\mathrm{MAE}=\frac{1}{n}\sum{}_{i=1}^{n}|y_i-\widehat{y_i}| \tag{1}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum{}_{i=1}^{n}(y_i-\widehat{y_i})^2} \tag{2}$$

$$R^2=1-\frac{\sum{}_{i=1}^{n}(y_i-\widehat{y_i})^2}{\sum{}_{i=1}^{n}\left(y_i-\overline{y}\right)^2} \tag{3}$$

$$\rho_c=\frac{2r\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2+\left(\mu_x-\mu_y\right)^2} \tag{4}$$

Here *n* represents the number of data points, y_i. refers to the actual value, $\widehat{y_i}$ indicates the prescribed value, $\overline{y}$ the average of all true values, *r* is the Pearson correlation coefficient between predicted and actual values, μ_x ve μ_x. are the averages of the predicted and actual values, respectively and σ_x ve σ_x are their respective standard deviations.

2.8 Used programmes

All model development and evaluation stages were carried out using the R programming language. The caret package, which contains many statistical learning algorithms and simplifies the training process, was used for the model's training and testing processes (Kuhn 2019; Kuhn et al. 2020). This package provides a unified framework for processes such as data splitting, preprocessing, model building, and hyperparameter setting. Each machine-learning algorithm was implemented through the relevant packages in R. The ranger package (Wright and Ziegler 2017) was used for the random forest model; the xgboost package (Chen and Guestrin 2016b) for the XGBoost model; the kernlab package (Karatzoglou et al. 2004) for the SVR model; the Cubist package (Kuhn and Quinlan 2021a) for the Cubist model; and the BRNN package (Rodriguez and Gianola 2022) for the BRNN model. In addition, dplyr (Wickham et al. 2023) and open-air (Carslaw and Ropkins 2012) packages were used for data processing and visualisation.

3 Results and discussion

3.1 Data set and structure

The data set used in the study includes variables and measurements obtained from experiments performed on CNC machines under different conditions. For each sample observation, the dataset contains PM2.5 and PM10 acceptable dust emission values resulting from process parameters such as wood species (Oriental spruce or Scots pine), processing time (in seconds), rotational speed per minute (RPM), cutting speed (m/min), cutting direction (contour, linear, spiral). This dataset, totaling 273 samples, covers different wood species and process combinations that produce low and high dust emissions. In this manner, the models were trained to learn dust emission under various treatment conditions.

3.2 Training process

The data set used in the modeling studies was initially divided into two subsets: training and testing. According to the ratio given in the literature, 80% is used for model training, while the remaining 20% is reserved as test data to test the model's success (Gholamy et al. 2018). This separation was randomized, and care was taken to ensure that both subsets represented the original data distribution. During the examination of the data set, the right-skewed distribution observed in PM2.5 and PM10 measurements showed that the data have a log-normal structure. Therefore, a logarithmic transformation was consistently applied to both the training and test subsets to improve the model's prediction performance and reduce the effect of extreme outliers. This transformation contributes to a more balanced scale, especially for the PM (Particulate Matter) values that spread over wide ranges, and thus helps estimate the model parameters in a statistically robust manner.

In Fig. 8, the dataset distributions in raw data (top part) and after logarithmic transformation (bottom part) are presented as box plots for the training and test subsets. In the box plots for the raw data, the effect of high right skewness and some extreme outliers is noticeable. In contrast, the skewness is reduced in the logarithmically transformed data distribution, and data appears more symmetric. This approach allowed the error terms to better fit the assumed statistical

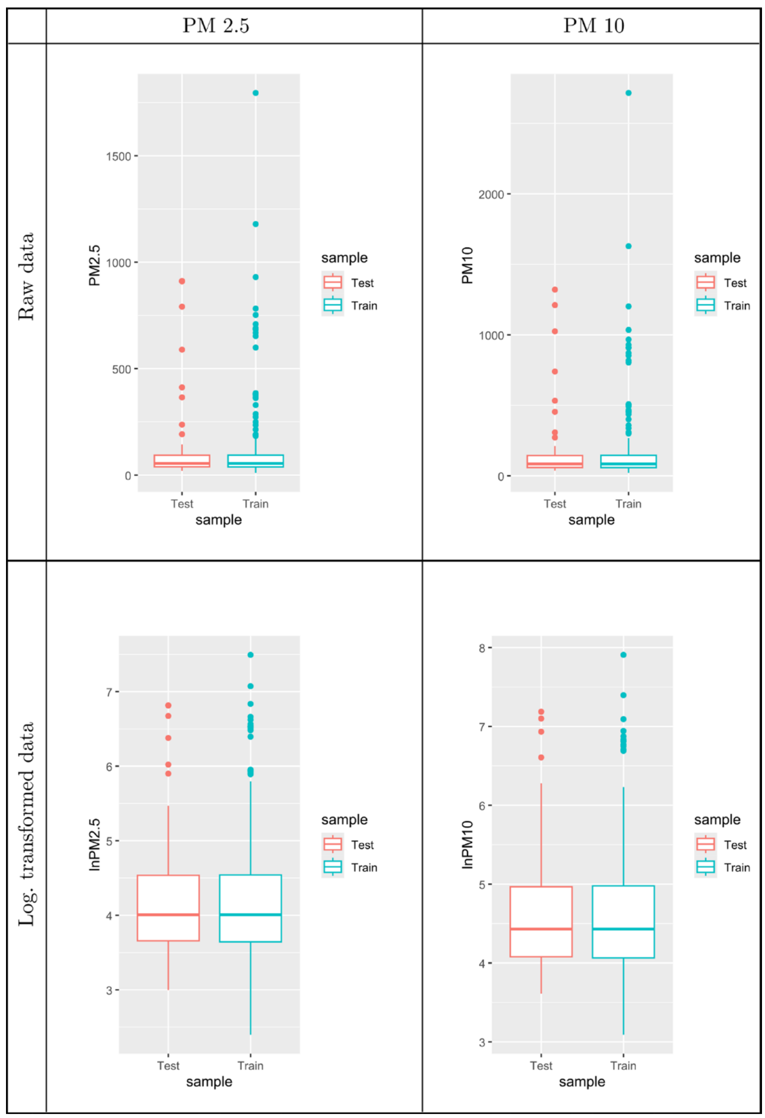

Fig. 8 Raw and log. Transformed Comparative distribution of training and test data

Table 3 Average validation performance metrics for the training process with 10-fold cross-validation

Particulate Matter (PM)	RMSE	Rsquared	MAE	ML methods
PM 2.5	**0.354**	**0.833**	**0.211**	**Cubist**
	0.361	0.828	0.239	SVR
	0.367	0.825	0.259	BRNN
	0.380	0.829	0.256	RF
	0.478	0.727	0.324	XGBoost
PM 10	**0.343**	**0.826**	**0.196**	**Cubist**
	0.336	0.830	0.216	SVR
	0.331	0.846	0.230	BRNN
	0.399	0.780	0.274	RF
	0.433	0.739	0.299	XGBoost

properties in both the training and testing phases and thus positively impacted the model's overall performance.

In the training phase, a 10-fold cross-validation method was applied to increase the robustness of the models and to measure their prediction performance more reliably. In this method, the training data were divided into 10 equal parts, each time, 9 parts were used for training the model and 1 part for validation, and the process was repeated 10 times. Thus, 10 different validation results were obtained for each model and averaged. Cross-validation allows the detection of overfitting by testing the consistency of the model on different data subsets, especially when the amount of data is relatively limited. In this study, the average training performance of each model with 10-fold cross-validation was calculated and presented in Table 3.

When the results are analyzed, Cubist and BRNN models are found to produce lower RMSE and MAE values and higher R^2 values compared to other algorithms. While the average RMSE value of the Cubist model for PM2.5 was 0.354, however, this value increased up to 0.478 in the XGBoost model. Similarly, while Cubist's R^2 value was 83.3%, XGBoost gave the lowest value of 72.7%. A similar picture is observed in the PM10 variable: The BRNN model shows the best performance with the RMSE = 0.331 and R^2 = 0.846, while XGBoost is behind the other methods with RMSE = 0.433 and R^2 = 0.739. SVR, Random Forest, and Cubist models were close to each other in terms of performance, with average errors in the range of 0.33–0.38, and R^2 values in the 82–85% band. These results show that the rule-based (Cubist) and Bayesian regularised artificial neural network (BRNN) approaches on the training data are effective. The XGBoost model is successful in capturing complex relationships; however, it has a slightly higher error rate than the other methods. This could be because the parameter settings of XGBoost are not fully optimized, or the dataset is relatively small. Nevertheless, all models showed acceptable learning performance.

Table 4 Performance metrics for testing data

Particulate Matter (PM)	Models	RMSE	Rsquared	MAE
PM 2.5	**Cubist**	**0.210**	**0.962**	**0.150**
	RF	0.227	0.943	0.171
	SVR	0.267	0.936	0.169
	BRNN	0.289	0.913	0.217
	XGBoost	0.355	0.870	0.265
PM 10	**Cubist**	**0.191**	**0.968**	**0.128**
	BRNN	0.234	0.935	0.166
	SVR	0.239	0.942	0.136
	RF	0.279	0.916	0.201
	XGBoost	0.332	0.862	0.250

3.3 Testing process

The models' potential to predict real-world data was determined by their predictive performance on the test dataset. The results of the models on the allocated 20% test data are summarized in Table 4. According to the test results, the Cubist was the most successful model; the lowest error and highest accuracy values for both PM2.5 and PM10 were determined in it. While Cubist's RMSE value for PM2.5 predictions was only 0.210, it was 0.227 for its closest competitor, Random Forest, and 0.355 for XGBoost, which showed the lowest prediction performance. Again, the Cubist model explained the measurements at a high level with an R^2 of 96.2% for PM2.5. PM10 test results similarly showed the superiority of the Cubist model (RMSE = 0.191, R^2 = 0.968). The XGBoost model was also the poorest-performing algorithm in the test phase, its R^2 for PM2.5 was the lowest at 87%, while its error was the highest at RMSE = 0.355. The BRNN and SVR models performed moderately well. Random Forest achieved a high accuracy, although with a slightly higher error than Cubist. This comparison reveals that the Cubist model is more consistent and successful than other methods, especially in the test data, and that the model's generalization ability is high. The main reason is that Cubist's rule-based structure may have provided an advantage in capturing the relationships in the data, for extrapolation. On the other hand, the poor test performance of XGBoost, a tree-based boosting method, may be due to the overly complex structure of the model, which is not well adapted to this dataset.

In addition to the above finding, a validation diagram for the test data was created to visually analyze the models' prediction performance. This diagram, shown in Fig. 9, compares the prediction results obtained by the models with the actual values for the test data. In this graph, the proximity of the points to the 45° identity line indicates the accuracy of the model predictions. In Fig. 9, the Cubist model (for PM2.5 $\rho c = 0.977$, for PM10 $\rho c = 0.979$) shows that the points are primarily centered around the diagonal, thus providing both high accuracy and high agreement. BRNN (for

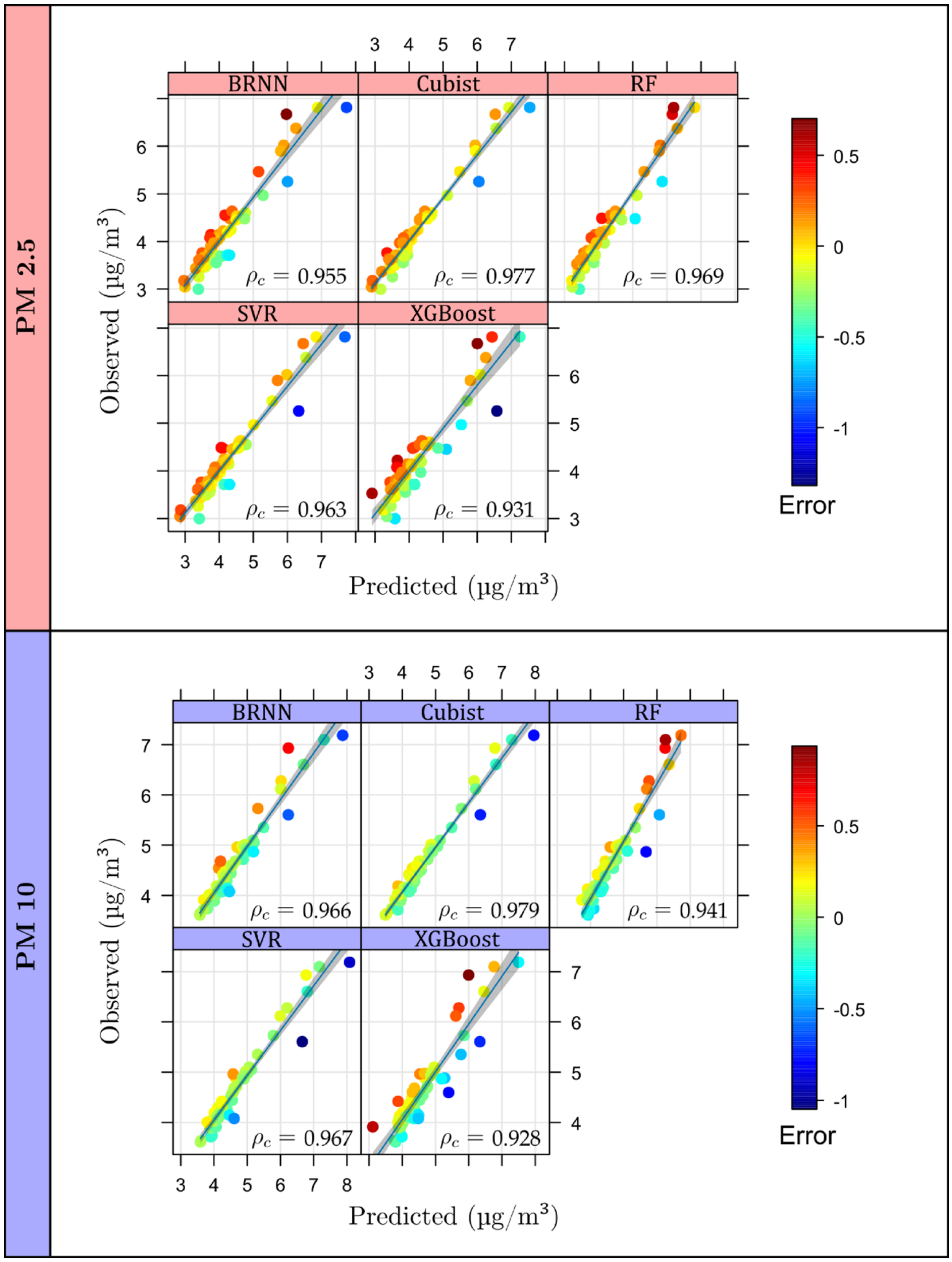

Fig. 9 Cross-validation diagrams for the test data. (The axes in the diagrams are shown in a logarithm-transformed scale to better visualize local variations)

PM2.5 ρc = 0.955, for PM10 ρc = 0.966), SVR (for PM2.5 ρc = 0.963, for PM10 ρc = 0.967) and Random Forest (for PM2.5 ρc = 0.969, for PM10 ρc = 0.941) models also show a relatively good distribution.

In contrast, some deviations from the actual values were noted in the XGBoost model with the lowest performance level (for PM2.5 ρc = 0.931; for PM10 ρc = 0.928). However, the overall trend for all models shows that the predicted values are close to the actual measurements and that there is a good correlation, with varying degrees of correlation by model. Another conclusion from the cross-validation figure is that the prediction correlations between models are also high, indicating that multiple methods consistently capture patterns in the data set. In particular, the success of the Cubist model in cross-validation supports its superior performance in the test data. It emphasizes that the model can learn the general trend without overfitting.

As a result of the research, the average dust emission values for two different wood species, measured against all parameters, are given in Table 5. In addition, in order to make Table 5 more understandable, Table 2 shows the sample groups created according to different parameters (tool path strategy, spindle speed and feed rate) for Oriental spruce sample groups and Scots pine samples.

Table 5 Dust emission values according to different parameters

		PM2.5 (µg/m³)	PM10 (µg/m³)		PM2.5 (µg/m³)	PM10 (µg/m³)		PM2.5 (µg/m³)	PM10 (µg/m³)
Orİental spruce Spruce	S1	8.57	13.88	S10	10.87	17.04	S19	6.50	10.75
	S2	9.37	14.91	S11	10.15	15.95	S20	5.27	8.53
	S3	*12.04*	*18.64*	S12	9.25	15.15	S21	4.53	8.27
	S4	11.11	17.74	S13	10.55	16.25	S22	*12.73*	*18.80*
	S5	**4.20**	**7.67**	S14	8.97	14.43	S23	11.56	16.88
	S6	**4.28**	**7.40**	S15	8.30	13.05	S24	11.60	16.25
	S7	4.92	8.33	S16	6.89	11.78	S25	10.87	16.11
	S8	7.50	12.50	S17	6.92	11.52	S26	10.44	15.48
	S9	7.73	12.83	S18	6.80	11.00	S27	10.93	16.13
	Ort.	7.75	12.66	Ort	8.74	14.02	Ort.	9.38	14.13
Scots pİne Spruce	Y1	**8.18**	**13.22**	Y10	64.63	85.80	Y19	65.48	84.18
	Y2	10.27	16.33	Y11	10.24	16.40	Y20	12.16	17.92
	Y3	10.04	15.36	Y12	11.33	17.47	Y21	12.00	17.87
	Y4	22.38	31.73	Y13	12.62	18.47	Y22	71.83	93.55
	Y5	9.46	13.46	Y14	**8.16**	**12.48**	Y23	10.20	15.60
	Y6	9.95	14.15	Y15	11.50	16.70	Y24	9.20	14.47
	Y7	28.28	37.97	Y16	*121.42*	*173.02*	Y25	12.08	17.78
	Y8	12.91	20.49	Y17	37.92	51.16	Y26	9.84	15.08
	Y9	*95.30*	*118.65*	Y18	46.80	63.25	Y27	10.87	16.13
	Ort.	22.97	31.26	Ort.	36.07	50.53	Ort.	23.74	32.51

As shown in Table 5, the dust emission values of Oriental spruce sample groups were considerably lower than those of Scots pine samples. The highest emission values of Oriental spruce samples were measured in sample group S22 (12.73 µg/m³ for PM2.5 and 18.80 µg/m³ for PM10) and S3 (12.04 µg/m³ for PM2.5 and 18.64 µg/m³ for PM10). The highest emission values of Scots pine samples were measured in sample group Y16 (121.42 µg/m³ for PM2.5 and 173.02 µg/m³ for PM10) and Y9 (95.30 µg/m³ for PM2.5 and 118.65 µg/m³ for PM10).

The lowest emission values of Oriental spruce samples were measured in the sample groups S5 (4.20 µg/m³ for PM2.5 and 7.67 µg/m³ for PM10) and S6 (4.28 µg/m³ for PM2.5 and 7.40 µg/m³ for PM10). The lowest emission values of Scots pine samples, Y9 (95.30 µg/m³ for PM2.5 and 118.65 µg/m³ for PM10) and Y16 (121.42 for PM2.5 and 173.02 for PM10), were measured in the sample group.

This difference in dust emission values between Oriental spruce and Scots pine samples was also reported in a study; it was stated that this difference between the two coniferous woods may be due to the microscopic and anatomical properties of the wood (Çakıroğlu et al. 2025). In another study conducted by Sydor et al. (2023) on the ratios of Oriental spruce, oak, and Meranti wood sanding powders in the total dust mass according to particle sizes, it was reported that the ratios of Oriental spruce wood dust particles < 2.5, < 4, and < 10 µm were lower than those of other wood species, i.e., the average particle size of Oriental spruce wood dust was larger than that of other wood species. Accordingly, the fact that the dust emission of spruce samples in this study is lower than that of Scots pine samples coincides with the literature. However, in a study conducted by Koleda et al. (2023), the effect of feeding rates was analyzed. Accordingly, it was determined that the smallest particles occurred at 4 m/min, and the largest amount of wood dust occurred at 6 m/min feeding speed. Therefore, low feed rates that cause the formation of tiny wood dust that endanger people and the environment inside the wood plant during milling are not recommended. A similar result was obtained using birch, European beech, and alder wood powders and it was reported that milling with a low feed rate (2 m/min.) produced finer dust than milling with high feed rates (4 and 6 m/min.) (Juda et al. 2023). In our study, the highest fine dust emission was observed at the lowest feed rates (3 m/min) for Oriental spruce (S22) and Scots pine (Y16) wood species.

In the study 4-axis CNC machine used (mod. 2128 MEGATRON, Gama Mekatronik Ltd.Şti, Bursa, Turkey). In addition, a single CNC tool type (Tungsten carbide milling cutter (0450-01, Netmak Kesici Tak.Ltd.Şti., Istanbul, Turkey) was used. In future studies, different CNC machines and different milling cutters can be used and their effects can be investigated. However, considering that the CNC machine used in this study and similar 4-axis machines are used in the wood industry, the importance of the results of this study in practical applications increases. As a result, the parameters determined in CNC machining of Scots pine and Oriental spruce wood species can be exemplary. In addition, the Cubist model can be taken into account during the processing of wood species with similar specific gravity and anatomical structure, enabling CNC operators to produce accurate and less wood dust emissions. For this, with a software integrated into the CNC machine, PM 2.5 and PM 10 levels can be accurately and successfully predicted with the Cubist model. CNC operators can realize their production without wasting time.

NIOSH (National Institute for Occupational Safety and Health) sets the wood dust exposure level at 1 mg/m³ as TWA (Time-Weighted Average). In addition, OSHA (Occupational Safety and Health Administration) gives two types of TWA limits. These are 15 mg/m³ (total dust) including all particle fractions and 5 mg/m³ (respirable fraction) particles that can directly reach the lungs (URL-3). PM2.5 and PM10 levels fall within the respirable fraction. According to the PM2.5 and PM10 values, the highest wood dust emission s were measured at 121.42 µg/m³ and 173.02 µg/m³, respectively, in Scots pine wood material, with a spindle speed of 18,000 rpm, a feed rate of 3 m/min, and cutting direction being linear. Although this situation caused more dust emission for the "Y16" group than the other groups, it was below the limit values according to NIOSH and OSHA. As it is known, NIOSH and OSHA set the occupational exposure limit (OEL) to wood dust as 8 h. However, although the measurements in this study were 5 min each, if the same process was applied for 8 h, the average limit values could have reached these levels.

4 Conclusion

The results obtained from the study show that spindle speed, feed rate, and cutting direction parameters significantly affect PM. According to PM2.5 and PM10 values, the highest wood dust emissions were measured at 121.42 µg/m³ and 173.02 µg/m³ in Scots pine wood material, with a spindle speed of 18,000 rpm, a feed rate of 3 m/min, and a linear cutting direction. The lowest wood dust emission was measured as 4.20 µg/m³ and 7.40 µg/m³ at 15,000 rpm, feed rate 6 and 9 m/min, and cutting direction contours in Oriental spruce wood material according to PM2.5 and PM10 values, respectively. However, the processing methods of the wood also affect the amount of wood dust produced. Within the scope of the study, 3 different cutting directions (Contours, Linear, and Spiral) were applied, and the highest average dust emission values in Oriental spruce samples were measured in the spiral cutting direction. The highest average dust emission values were obtained for pine samples in the linear cutting direction. Using appropriate air filtration and emission control units is recommended, especially for wood species with high emission concentrations, as the environmental impact limits vary during CNC milling. In order to minimize wood dust emissions without disrupting the workflow for CNC operators, the processing parameters of Scots pine wood material are recommended in the range of spindle speed 12,000–15,000 rpm and feed rate 3–6 m/min, while for Oriental spruce wood material, spindle speed is recommended within the limits of 15,000 and feed rate 6–9 m/min.

In conclusion, after extensive comparisons and evaluations, the Cubist model best predicted PM2.5 and PM10 levels. Both the low error rates obtained by cross-validation on the training data and the high accuracy on the unknown test data showed that the Cubist approach is more suitable for this problem than other machine learning techniques. By combining the explainability of decision trees with the precision of linear models, this model successfully predicted dust emissions during CNC wood processing. Although other methods were also successful to some extent, Cubist's rule-based flexible structure and overall performance made it stand out. This finding indicates that rule-based machine learning approaches can effectively model and optimize wood dust emissions. The study results demonstrate the value of data-driven approaches in predicting and controlling wood dust exposure, which is especially critical for occupational health. While the aim is to minimize chip removal, the results may not directly relate the optimized parameters to this aspect. In this context, the model developed to determine the combinations that will minimize wood dust emission by optimizing CNC machining parameters was successful. It is also recommended that these findings be tested with industrial environment data and real-time applications in future studies and investigate the effect of different wood species and different types of processing (e.g. grooving, drilling, profiling) on dust emission, more detailed particle size distribution analysis under various parameters and to develop a user-friendly emission estimation tool for CNC operators.

Author contributions E.O.C. Collection of data for the whole article, making measurements, modelling, literature searches, article writing.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

References

Acıslı-Celik S, Yeşilkanat CM (2023) Predicting science achievement scores with machine learning algorithms: a case study of OECD PISA 2015–2018 data. Neural Comput Appl 35:21201–21228. https://doi.org/10.1007/s00521-023-08901-6

Akkoyun S, Yeşilkanat CM, Bayram T (2024) Machine learning predictions for cross-sections of 43,44Sc radioisotope production by alpha-induced reactions on Ca target. Nuclear Instruments

Methods Phys Res Sect B: Beam Interact Mater Atoms 549(February). https://doi.org/10.1016/j.nimb.2024.165293

Alapieti T, Mikkola R, Pasanen P, Salonen H (2020) The influence of wooden interior materials on indoor environment: a review. Eur J Wood Prod 78:617–634

Amrani N, Yeşilkanat CM, Akkoyun S (2024) Neutron-Alpha reaction cross section determination by machine learning approaches. J Fusion Energy 43:1–9. https://doi.org/10.1007/s10894-024-00461-4

Awad M, Khanna R (2015) Support Vector Regression. In efficient learning machines (Issue Vc, pp. 67–80). https://doi.org/10.1007/978-1-4302-5990-9_4

Batista TS, Teodoro LPR, De Azevedo GB, Azevedo GTDOS, Poersch NL, Borges MVV, Teodoro PE (2022) Artificial neural networks and non-linear regression for quantifying the wood volume in Eucalyptus species. Southern Forests: a Journal of Forest Science, 84(1), 01–07. https://doi.org/10.2989/20702620.2021.1976604

Bayram T, Yeşilkanat CM, Akkoyun S (2023) Applications of different machine learning methods on nuclear charge radius estimations. Phys Scr 98:125310. http://iopscience.iop.org/article/10.1088/1402-4896/ad0434

Bentéjac C, Csörgő A, Martínez-Muñoz G (2021) A comparative analysis of gradient boosting algorithms. Artific Intell Rev. https://doi.org/10.1007/s10462-020-09896-5

Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Burden F, Winkler D (2009) Bayesian regularization of neural networks. Artific Neural Netw Methods Appl. https://doi.org/10.1007/978-1-60327-101-1_3

Çakıroğlu EO, Demir A, Aydın İ, Büyüksarı Ü (2022) Prediction of optimum CNC cutting conditions using artificial neural network models for the best wood surface quality, low energy consumption, and time savings. BioResources 17(2):2501. https://doi.org/10.15376/biores.17.2.*2501-2524*

Çakıroğlu EO, Demirarslan KO, Taşdemir T (2025) Characterization of wood dust emission according to some wood species in 3D machining applied with CNC machine. Eur J Wood Prod 83(2):59. https://doi.org/10.1007/s00107-025-02222-2

Carslaw DC, Ropkins K (2012) Openair- an r package for air quality data analysis. Environ Modell Software 27–28, 52–61. https://doi.org/10.1016/j.envsoft.2011.09.*008*

Chen T, Guestrin C (2016a) XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. https://doi.org/10.1145/2939672.2939785

Chen T, Guestrin C (2016b) Extreme gradient boosting [R package Xgboost version 1.5.0.2]. Comprehensive R archive network (CRAN). https://doi.org/10.1145/2939672.2939785

Chen Z, Luo H, Xu L (2021) Machine learning models of ischemia/hemorrhage in Moyamoya disease and analysis of its risk factors. Clin Neurol Neurosurg 209:106919. https://doi.org/10.1016/j.clineuro.2021.106919

Demir A, Cakiroglu EO, Aydin I (2022) Determination of CNC processing parameters for the best wood surface quality via artificial neural network. Wood Mater Sci Eng 17(6):685–692. https://doi.org/10.1080/17480272.2021.1929466

Demir A, Birinci AU, Ilhan O, Ozturk H (2024) Identification of optimal cutting parameters regarding lacquer wettability of CNC processed MDF panels by artificial neural network. Wood Mater Sci Eng. https://doi.org/10.1080/17480272.2024.2426780

Drucker H, Surges CJC, Kaufman L, Smola A, Vapnik V (1997) Support vector regression machines. Adv Neural Inf Process Syst 1:155–161

Fujimoto K, Takano T, Okumura S (2011) Difference Inmass concentration of airborne dust during circular sawing offive wood-based materials. J Wood Sci 57(2):149–154. https://doi.org/10.1007/s10086-010-1145-y

Gholamy A, Kreinovich V, Kosheleva O (2018) Why 70 / 30 or 80 / 20 relation between training and testing sets: a pedagogical explanation. J Intell Technol Appl Stat 11(2):105–111. https://doi.org/10.6148/IJITAS.201806_11(2).0003

Hanincová L, Pędzik M, Majka J, Sydor M, Rogoziński T (2024) Influence of thermal modification and sanding parameters on finest particle content in Pinewood dust. Wood Mater Sci Eng 19(4):887–895. https://doi.org/10.1080/17480272.2023.2289612

Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning. New York: Springer. https://doi.org/10.1007/978-0-387-84858-7

Hlásková L, Rogozınski T, Dolny S, Kopecký Z, Jedinák M (2015) Content of respirable and inhalable fractions in dust created while sawing Beech wood and its modifications. Drewno 58:135–146. https://doi.org/10.12841/wood.1644-3985.096.11

Hlásková L, Rogoziński T, Kopecký Z (2016) Influence of feed speed on the content of fine dust during cutting of two-side-laminated particleboards. Wood Industry/Drvna Industrija. https://doi.org/10.5552/drind.2016.1417

Holla K, Ristvej J, Moricova V, Novak L (2016) Results of a survey among SEVESO establishments in the Slovak Republic. J Chem Health Saf 23(2):9–17. https://doi.org/10.1007/s00107-021-01764-5

Igaz R, Kminiak R, Krišťák Ľ, Němec M, Gergeľ T (2018) Methodology of temperature monitoring in the process of CNC machining of solid wood. Sustainability 11:95. https://doi.org/10.3390/su11010095

Iliadis L, Mansfield SD, Avramidis S, El-Kassaby YA (2013) Predicting Douglas-fir wood density by artificial neural networks (ANN) based on progeny testing information. Holzforschung 67(7):771–777. https://doi.org/10.1515/hf-2012-0132

Jawad J, Hawari AH, Zaidi SJ (2021) Artificial neural network modeling of wastewater treatment and desalination using membrane processes: a review. Chem Eng J 419:129540. https://doi.org/10.1016/j.cej.2021.129540

Juda M, RogoziÅ T, Kminiak R (2024) The effect of the feed rate and revolutions of the cutting tool on the amounts of created chips in dust and respirable sizes from milling particleboards, and medium-density fibreboards. Global For J 2(02):147–160. https://doi.org/10.32734/gfj.v2i02.16735

Júda M, Sydor M, Rogoziński T, Kučerka M, Pędzik M, Kminiak R (2023) Effect of low-thermal treatment on the particle size distribution in wood dust after milling. Polymers 15(4):1059. https://doi.org/10.3390/polym15041059

Karatzoglou A, Smola A, Hornik K, Zeileis A (2004) kernlab - an S4 package for kernel methods in R. J Stat Softw 11(9):1–20. http://www.jstatsoft.org/v11/i09/

Kminiak R, Dzurenda L (2019) Impact of sycamore maple thermal treatment on a granulometric composition of chips obtained due to processing on a CNC machining mentre. Sustainability 11(3), 718. https://doi.org/10.3390/su11030718

Kminiak R, Kučerka M, Kristak L, Reh R, Antov P, Očkajová A, Pędzik M (2021) Granulometric characterization of wood dust emission from CNC machining of natural wood and medium density fiberboard. Forests 12(8):1039. https://doi.org/10.3390/f12081039

Koleda P, Koleda P, Hrčková M, Júda M, Hortobágyi Á (2023) Experimental granulometric characterization of wood particles from CNC machining of chipboard. Appl Sci 13(9):5484. https://doi.org/10.3390/app13095484

Kos A, Beljo-Lučić R, Šega K, Rapp AO (2004) Influence of woodworking machine cutting parameters on the surrounding air dustiness. Eur J Wood Prod 62(3):169–176. https://doi.org/10.1007/s00107-004-0473-2

Kuhn M (2019) The caret Package. [if gte mso 9]> Normal 0 21 false false false TR X-NONE X-NONE. https://topepo.github.io/caret/index.html

Kuhn M, Quinlan R (2021a) Cubist: Rule- And Instance-Based Regression Modeling, (R package version 0.4.1). https://cran.r-project.org/Web/Packages/Cubist/Index.Html. Accessed on 02 Feb 2023

Kuhn M, Quinlan R (2021b) Cubist: Rule- and instance-based regression modeling

Kuhn M, Wing J, Weston S et al (2020) caret: classification and regression training [R package caret version 6.0–86]. https://github.com/topepo/caret/

Kumar S, Ghosh SK (2021) Statistical and artificial neural network technique for prediction of performance in AlSi10Mg-MWCNT based composite materials. Mater Chem Phys 273:125136. https://doi.org/10.1016/j.matchemphys.2021.125136

Mackay D (1992) A practical bayesian framework for backpropagation networks. Neural Comput. https://doi.org/10.1162/neco.1992.4.3.448

Marková I, Ladomerský J, Hroncová E, Mračková E (2018) Thermal parameters of Beech wood dust. BioResources 13(2):3098–3109. https://doi.org/10.15376/biores.13.2.3098-3109

Miguel EP, Melo RRD, Junior S, L., Menezzi CHSD (2018) Using artificial neural networks in estimating wood resistance. Maderas Ciencia Y Tecnología 20(4):531–543. https://doi.org/10.4067/S0718-221X2018005004101

Mo J, Tamboli D, Haviarova E (2023) Prediction of the color change of surface thermally treated wood by artificial neural network. Eur J Wood Prod 81(5):1135–1146. https://doi.org/10.1007/s00107-023-01969-w

Neudecker D, Grosskopf M, Herman M, Haeck W, Grechanuk P, Vander Wiel S, Rising ME, Kahler AC, Sly N, Talou P (2020) Enhancing nuclear data validation analysis by using machine learning. Nuclear Data Sheets 167:36–60. https://doi.org/10.1016/j.nds.2020.07.*002*

Očkajová A, Kučerka M, Banski A, Rogoziński T (2016) Factors affecting the granularity of wood dust particles. Chip Chipless Woodwork Process 10:137–144

Omidianidost A, Ghasemkhani M, Azari MR, Golbabaei F (2015) Assessment of occupational exposure to dust and crystalline silica in foundries. Tanaffos 14(3):208

Palmqvist J, Gustafsson SI (1999) Emission of dust in planing and milling of wood. Holz Roh- Werkst 57(3):164–170. https://doi.org/10.1007/s001070050035

Pędzik M, Júda M, Kminiak R, Czerniejewska-Wolska H, Rogoziński T (2024) The effect of average chip thickness on the potentially respirable dust from CNC finish milling of wood-based materials. Drewno Prace Naukowe Doniesienia Komunikaty. https://doi.org/10.53502/wood-191936

Quinlan JR (1992) Learning with continuous classes. Australian Joint Conf Artific Intell 92:343–348

Rabiei F, Souri S (2024) Investigation of computer numerical control wood milling parameters for occupationally safer by the minimization of produced wood dust. Wood Mater Sci Eng 19(3):674–682. https://doi.org/10.1080/17480272.2023.2286624

Ratnasingam J, Scholz F, Natthondan V (2010) Particle size distribution of wood dust in rubberwood (Hevea brasiliensis) furniture manufacturing. Eur J Wood Prod 68:241–242. https://doi.org/10.1007/s00107-009-0369-2

Rodriguez PP, Gianola D (2022) brnn: Bayesian regularization for feed-forward neural networks. R package version 0.9.2. https://CRAN.R-project.org/package=brnn

Simon C, Donnot A, Méausoone PJ (2022) Indoor wood dust analysis during machining with an original experimental device. Eur J Wood Prod. https://doi.org/10.1007/s00107-021-01764-5

Singer H, Ilçe AC, Şenel YE, Burdurlu E (2024) Artificial neural network–based prediction model to minimize dust emission in the machining process. Saf Health Work 15(3):317–326. https://doi.org/10.1016/j.shaw.2024.06.006

Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. Statistics Comput. 14(3):199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Sydor M, Majka J, Hanincová L, Kučerka M, Kminiak R, Kristak L, Rogoziński T (2023) Fine dust after sanding untreated and thermally modified spruce, oak, and meranti wood. Eur J Wood Prod 81(6):1455–1464. https://doi.org/10.1007/s00107-023-01971-2

Tong R, Cheng M, Zhang L, Liu M, Yang X, Li X, Yin W (2018) The construction dust-induced occupational health risk using Monte-Carlo simulation.J Cleaner Production. 184:598–608. https://doi.org/10.1016/j.jclepro.2018.02.*286*

URL-1/Hexagon program for woodworking cad-cam-software. Access Add. https://hexagon.com/products/product-groups/computer-aided-manufacturing-cad-cam-software/alphacam

URL-2 PCE Instruments (t.y.). PCE-MPC 10 Particle Measuring Instrument Technical Catalogue. Access add. https://www.pce-instruments.com/turkish/api/getartfile?_fnr=1095665&_dsp=inline

URL-3 NIOSH pocket guide to chemical hazards https://www.cdc.gov/niosh/npg/npgd0667.html

Vallières E, Pintos J, Parent ME, Siemiatycki J (2015) Occupational exposure to wood dust and risk of lung cancer in two population-based case–control studies in Montreal. Can Environ Health 14:1–9. https://doi.org/10.1186/1476-069X-14-1

Wickham H, François R, Henry L, Müller K, Vaughan D (2023) dplyr: a grammar of data manipulation

Wright MN, Ziegler A (2017) Ranger: A fast implementation of random forests for high dimensional data in C++and R. J Stat Softw 77(1 SE–Articles):1–17. https://doi.org/10.18637/jss.v077.i01

Yeşilkanat CM (2020) Spatio-temporal Estimation of the daily cases of COVID-19 in worldwide using random forest machine learning algorithm. Chaos Solitons Fractals. https://doi.org/10.1016/j.chaos.2020.110210

Yeşilkanat CM, Akkoyun S (2023) Estimation of fission barrier heights for even-even superheavy nuclei using machine learning approaches. J Phys G Nuclear Particle Phys. 50(5):55101. https://doi.org/10.1088/1361-6471/*acbaaf*

You G, Wang B, Li J, Chen A, Sun J (2022) The prediction of MOE of bamboo-wood composites by ANN models based on the nondestructive vibration testing. J Build Eng 59:105078. https://doi.org/10.1016/j.jobe.2022.105078

Zhang W, Wu C, Zhong H, Li Y, Wang L (2021) Prediction of undrained shear strength using extreme gradient boosting and random forest based on bayesian optimization. Geosci Front 12(1):469–477. https://doi.org/10.1016/j.gsf.2020.03.007

Zhang Y, Shi X, Zhang S, Abraham A (2022) A XGBoost-based lane change prediction on time series data using feature engineering. IEEE Trans Intell Transp Syst 23(10):19187–19200. https://doi.org/10.1109/TITS.2022.3170628