machines

Article

Comparative Analysis of Neural Networks and Decision Trees for Roughness Prediction

Mihai Banica , Andrei Osan * and Andrei Catalin Filip

Department of Engineering and Management of Technology, North University Center of Baia Mare, Technical University of Cluj-Napoca, 430083 Baia Mare, Romania; mihai.banica@imtech.utcluj.ro (M.B.); andrei.filip@campus.utcluj.ro (A.C.F.)
* Correspondence: andrei.osan@imtech.utcluj.ro

Abstract

In the current context of the manufacturing industry, optimizing the cutting parameters to achieve a controlled surface roughness involves costly and time-consuming experimental efforts. The present study addresses this challenge by developing robust machine learning-based approximate functions for predicting surface roughness (Ra) resulting from toroidal milling on a five-axis CNC. The research includes an experimental design conducted under real production conditions on C45 steel. The relatively small experimental dataset was augmented, normalized, and then scripts were written for four prediction models: two artificial neural network architectures and two models based on decision trees. Their performance was analyzed based on *MSE*, *RMSE*, R^2, and MRA metrics. The results obtained reveal significant differences between the models, highlighting solutions with high accuracy, excellent robustness, and superior generalization capacity for new data. The study highlights the high potential of prediction models in optimizing machining processes, providing an effective way to reduce costly physical experiments and increase productivity in industrial environments.

Keywords: toroidal cutter; roughness; augmentation; machine learning; neural networks; decision trees; prediction

1. Introduction

In the current conditions of machining, the focus has shifted from the question of whether a certain surface roughness can be achieved to the problem of optimizing how it can be achieved faster, safer, and cheaper, while maintaining a high degree of precision and repeatability. As there is less time and money available for experiments, materials and tools are controlled to reduce costs. This constraint makes it difficult to conduct an extensive experimental study with a large number of samples and to determine the optimal cutting regimes for different manufacturing conditions. In such a situation, the application of artificial neural networks and algorithms based on decision trees offers a solution to reduce the time and, implicitly, the costs of research.

A careful review of the recent literature shows that most studies focus on isolated applications of a single type of model (ANN or trees) using laboratory or simulated data. Few papers make a systematic comparison of the performance and robustness of multiple architectures on independent external datasets, obtained under real industrial conditions. The present research addresses precisely this gap, providing a detailed comparative evaluation between two neural network architectures and two decision tree-based models, using augmented real experimental data.

Academic Editors: Zafer Bingul and Juan Gabriel Avina-Cervantes

Received: 28 May 2026
Revised: 26 June 2026
Accepted: 14 July 2026
Published: 15 July 2026

 https://doi.org/10.3390/machines14070802

Based on this premise, this study intends to implement and evaluate artificial neural networks for predicting surface roughness obtained under different cutting conditions from those used for machine learning. Although the cutting processes have a large number of interdependent variables, the choice of the most important parameters, correlated with the experimental measurement of roughness, allows the generation of a dataset for training and testing predictive models.

Mathematical modeling of the relationships between process parameters and the roughness of the machined surface is important for understanding their impact on the results. Artificial neural networks (ANN) and decision tree-based models are modern and efficient tools for roughness analysis and surface roughness prediction due to their ability to learn complex nonlinear relationships between variables.

The study is based on a limited set of experimental data and is developed with specific augmentation techniques. Different architectures and methods for training, validating, and testing artificial neural networks and tree-based models were applied to the extended dataset, with the aim of developing a robust predictive model that can estimate surface roughness under different processing conditions, requiring only the updating of the input values, thus reducing the need for new and expensive experimental processing.

The main objectives of the research were:

1. To develop and implement a limited experimental plan under real industrial conditions;
2. To improve the experimental database to obtain a sufficient volume of information for training the models;
3. To train and test artificial neural networks and models based on decision trees;
4. To evaluate and compare the performance obtained by the models;
5. To conduct a detailed analysis of the results of the best-performing model in the learning and testing stages.
6. To conclude on the efficiency and applicability of the models developed for the prediction of the quality of the processed surfaces.

2. Experimental Methodology

The experimental research was conducted within the company S.C. Ramira S.A., situated in the city of Baia Mare, Romania, a member of the Chropynska S.A. group from the Chropyne, Czech Republic.

It is important to emphasize that all the experiments and measurements related to the study were carried out in a real industrial environment, respecting the specific conditions of production.

This chapter describes in detail the material to be processed, the geometry of the specimens, the characteristics of the cutting tool, the CNC machine used, the parameters of the cutting regime, the experimental procedure, the clamping device, and the equipment for measuring the surface roughness.

2.1. Material Used—C45 Steel

C45 steel has the following composition: carbon (0.42–0.50%), manganese (0.50–0.80%), silicon in reduced proportions, as well as controlled phosphorus and sulfur contents. Its physical properties include a density of about 7.85 g/cm^3, a modulus of elasticity of about 210 GPa, and good thermal conductivity.

The mechanical characteristics of C45 steel vary depending on the heat treatment. In the normalized state, the material has a tensile strength of 570–700 MPa and a Brinell hardness between 170 and 210 HB. By applying quenching and tempering treatments, superior hardness and strength values can be achieved. C45 steel is characterized by good

https://doi.org/10.3390/machines14070802

machinability through plastic deformation and cutting but has low corrosion resistance and moderate weldability.

Due to these properties, C45 steel is commonly used in the manufacture of shafts, axles, gears, sprockets, bolts, bushings, and other required transmission components in industrial, agricultural, and automotive equipment. The choice of this material is due to its favorable balance between mechanical strength, machinability, and low costs [1,2].

2.2. Geometric Shape of the Specimen

The shape and dimensions of the specimens were determined according to two essential requirements: reducing material consumption and ensuring adequate rigidity during the cutting process.

For this purpose, a simple parallelepiped geometry was adopted, in the form of a plate with dimensions 40 mm (length), 40 mm (width), and 20 mm (thickness).

2.3. The Tool Used for Cutting

The tool used, model JHP780160E2R400.0Z4-M64, manufactured by Seco Tools and supplied by a wholesaler in Gorinchem, Netherlands, is shown in Figure 1. It is a solid carbide monobloc milling cutter with a diameter of 16 mm, with 4 edges and a radius at the corner of 4 mm. The length of the cutting part is 32 mm, and the total length of the cutter is 100 mm. The cutter is provided with a MEGA-64 PVD coating, which provides increased resistance to wear and high temperatures.

Figure 1. Toroidal milling cutter JHP780160E2R400.0Z4-M64 from Seco Tools.

This tool is part of the manufacturer's High Performance Machining range and is designed for intensive machining with large cutting depths and widths, ensuring high stability and productivity when machining steels, stainless steels, and difficult alloys.

2.4. OKUMA MU-400VA CNC Center

The experiments were conducted on a five-axis CNC machining center, model OKUMA MU-400VA, manufactured by Okuma Corporation in Ōguchi, Aichi Prefecture, Japan.

OKUMA MU-400VA has the following technical specifications:

- Axial strokes: X ≈ 762 mm, Y ≈ 460 mm, Z ≈ 460 mm;
- Rotary table with a diameter of Ø400 mm and a maximum load capacity of 300 kg;
- Additional rotary axes: A (+20°/−110°) and C (360° continuous rotation);
- Main shaft with a maximum speed of 15,000 rpm, BT40 cone, and 22 kW power;
- Tool magazine with a capacity of 48 positions, equipped with an automatic tool changer.

With this high-precision industrial equipment, the experiments were carried out under real production conditions, ensuring rigidity, stability, and high repeatability of the machining process, as shown in Figure 2.

Figure 2. OKUMA MU-400VA CNC Center.

2.5. Process Parameters

In the experiments, the axial cutting depth (ap) was kept constant at 0.5 mm, and the radial cutting depth (a_e) was set to 0.3 mm. The machining was conducted with liquid, and the specimens were rigidly fixed in a vise.

During milling, the tool axis was kept at a constant angle of inclination to the surface being machined. The CAM programs were generated using PowerMILL v. 2015 software. For each surface, a unidirectional milling strategy was applied, with parallel tool trajectories.

The selected parameters (cutting speed, tilt angle, and feed per tooth) represent the main variables controlled in the five-axis toroidal milling process, having the greatest influence on the roughness according to the literature and practical experience in production.

The ranges of the cutting parameters were selected to cover a wider range than recommended by the tool manufacturer, allowing a complete evaluation of the behavior of the milling process:

- The cutting speed ranged from 80 to 210 m/min, exceeding the manufacturer's recommended range (140–160 m/min), allowing performance analysis at both low and high speeds;
- The feed per tooth was established around the recommended value of 0.15 mm/tooth, with a variation of ±27%, namely, 0.11 mm/tooth and 0.19 mm/tooth, to highlight the influence of this parameter on roughness;

- The angle of inclination of the tool axis was varied at three levels: 15°, 35°, and 55°. In the absence of recommendations from the manufacturer, these values cover a relevant range for simultaneous milling in five axes, allowing the analysis of the effect of the angle on the tool-material contact geometry and, implicitly, on the surface roughness.

  The parameters of the milling process are summarized in Table 1.

Table 1. Toroidal milling cutter parameters for learning.

Parameters	Values
Cutting Speed [m/min]	80, 110, 130, 150, 170, 190, 210
Tilt Angle [°]	15, 35, 55
Feed Per Tooth [mm/tooth]	0.11, 0.15, 0.19

The test parameters, which are different from the initial ones, are used to evaluate the learning models. These are presented in Table 2.

Table 2. Toroidal milling cutter parameters for testing.

Parameters	Values
Cutting Speed [m/min]	100, 120, 140, 160, 180
Tilt Angle [°]	20, 30, 40, 50, 60
Feed Per Tooth [mm/tooth]	0.12, 0.14, 0.16, 0.18, 0.20

2.6. Conducting Experiments

In the experimental study, a total of 70 surfaces were processed by milling with the toroidal mill. The analysis of the CAM programs showed that 140 trajectories, 142 trajectories for the 35° angle and 139 trajectories for the 55° angle were required for the machining of the flat surface at the 15° angle. Figure 3a shows the simulations of the toroidal milling trajectory during the generation of the CAM program, corresponding to the three values of the angle of inclination, and Figure 3b shows images of the machining process on the CNC machine.

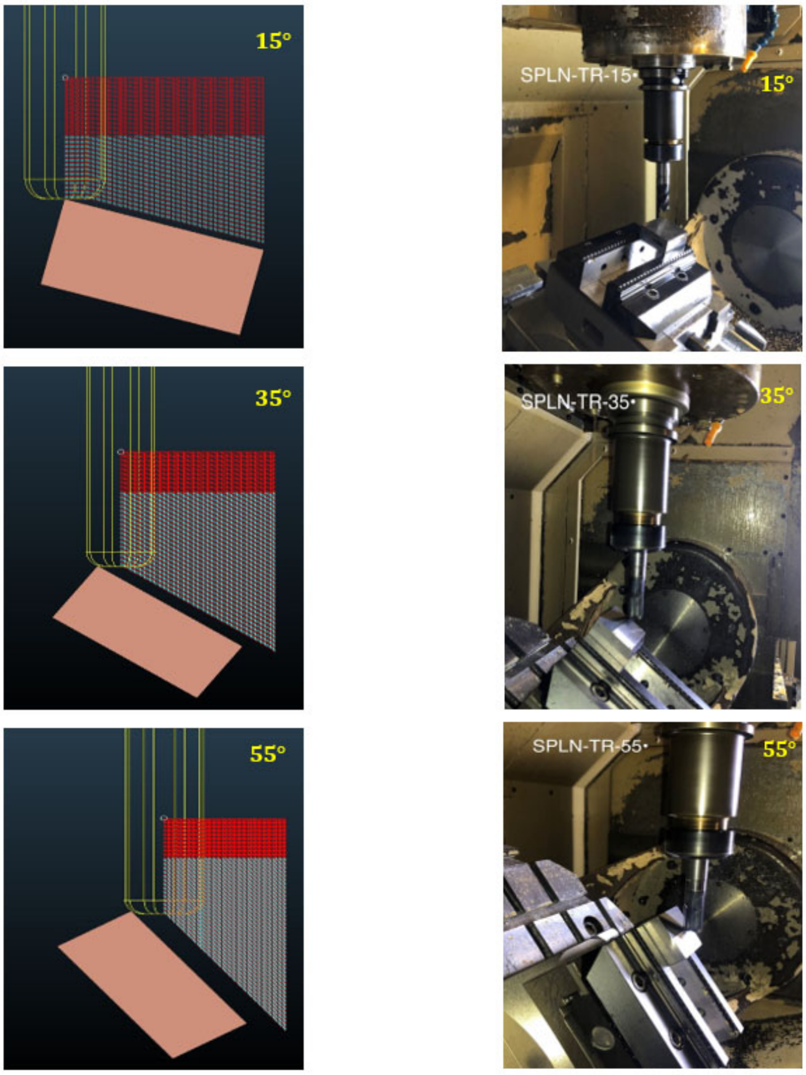
2.7. Specimen Clamping Device

The clamping of the specimens was conducted using a specialized vise for five-axis machining centers, optimized for simultaneous multi-axis machining.

The vise is equipped with a direct mechanical clamping system without intermediate force transmission, operated manually. The construction includes high clamping jaws, an upper shaft positioned close to the clamping area for optimal force transmission, as well as long jaw guides in a compact construction with SKB claw jaws. The device is produced by ROHM GmbH, Sontheim an der Brenz, Heidenheim, Germany (Figure 4).

2.8. Roughness Measurement with the TR200 Roughness Meter

The surface roughness measurement was conducted according to the ISO 4287:2010 [3] standard. On each processed surface, three measurements were made parallel to the feed direction, using a TR-200 roughness meter made în Beijing, China, with an uncertainty value of ±0.01 µm. The measurement parameters (cut-off length of 0.8 mm and evaluation length of 4 mm) were selected according to ISO 4288:1996. The positioning of the roughness meter was performed on a stand with a granite column and base, according to Figure 5.



(**a**) (**b**)

Figure 3. Experimental performance on OKUMA MU-400VA: (**a**) simulation of the CAM program; (**b**) CNC machining.

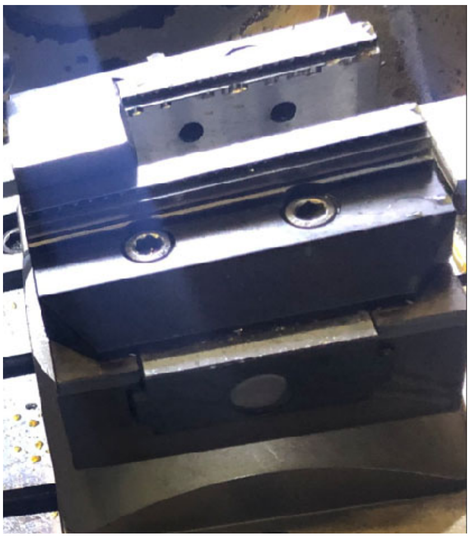


Figure 4. Gripping device with impression.

Figure 5. Roughness measurement.

The TR-200 roughness meter can measure several parameters, namely Ra, Rz, Ry, Rq, Rp, Rm, Rt, R3z, Rmax, Sk, S, Sm, and tp. The results and graphs are displayed on an LSD 128 × 64 screen, and the device can also be connected to a TA-220 printer to print all parameters and graphs. The TR-200 has high inductance accuracy, which makes it compatible with four standards, namely ISO, DIN, ANSI, and JIS.

3. Supervised Learning

Supervised learning is a method within machine learning in which a model is trained on a dataset for which the inputs and outputs are known. The goal is for the model to learn the relationship between inputs and outputs so that it can make accurate predictions on new, previously unseen data.

This category includes both artificial neural networks, such as those trained with the MATLAB v.2025.b trainlm (Levenberg–Marquardt) and trainbr (Bayesian Regularization) functions, as well as decision trees (fitrtree) and tree assemblies (fitrensemble). All these techniques are considered supervised learning methods because they rely on known data to solve regression problems.

The choice of the four models (two neural network architectures—trainbr and trainlm—and two tree-based models—fitrtree and fitrensemble) was motivated by the lack of complementary comparison of performance and explainability. Neural networks offer a high capacity to approximate complex nonlinear relationships (black-box models), while tree-based models provide transparency and interpretability (white-box/gray-box models), allowing the analysis of the importance of predictors. This hybrid approach allows for a balanced assessment of the trade-off between accuracy and explainability in the context of machining processes.

3.1. Artificial Neural Networks

In the context of modern industrial processes, artificial neural networks play an essential role in predicting and optimizing the performance of complex systems from the early stages of the technological process. Through their ability to analyze large volumes of data and identify hidden relationships between variables, they help make accurate and effective predictions. The use of artificial neural networks allows the detection of complex interactions within processes, facilitating rapid interventions and improving production efficiency [4,5]. Thus, this type of approach becomes an essential tool for increasing product quality and strengthening competitiveness in modern industries [6–9].

https://doi.org/10.3390/machines14070802

Currently, artificial intelligence is experiencing accelerated development, being increasingly used due to the significant increase in its performance. It has found applications in a wide range of fields, such as health, education, natural language processing, economics, trade, and industrial production [10–14].

Artificial neural networks are computational models inspired by biological neural networks, with the ability to approximate complex nonlinear relationships without requiring explicit mathematical formulation. In this research, a multilayer architecture comprising an input layer, one or more fully interconnected hidden layers, and an output layer was used.

The implementation was conducted in the MATLAB v.2025b environment, using dedicated scripts for the trainbr and trainlm algorithms. MATLAB enables rapid prototyping, efficient training, and robust model evaluation [15–17].

The use of a multilayer architecture with feedforward propagation offers advantages in modeling and analyzing processing operations. One of the main benefits lies in the ability to model non-linearly, as machining processes involve complex relationships between technological parameters such as cutting speed, feed per tooth, and tool tilt angle, and the value of roughness resulting from measurements. The network can learn these relationships without requiring explicit mathematical formulations [18].

Another important advantage is the ability to generalize. After the training stage, the neural network can make predictions for new and unknown data, making it suitable for industrial applications where processing conditions can vary significantly. The network architecture also offers a high degree of flexibility and scalability by adjusting the number of neurons, hidden layers, and activation functions, depending on the complexity of the data and the desired level of accuracy.

Artificial neural networks show robustness in the presence of industrial experimental noise, being able to identify relevant patterns even when the data contain natural variations or errors specific to real processing conditions [19].

3.2. Algorithms Based on Decision Trees

Algorithms based on decision trees show extensive applicability in a variety of fields, due to their high flexibility and ability to efficiently organize and analyze complex data. They can replace or complement traditional statistical methods, being used to identify patterns in datasets, extract relevant information from texts, handle missing values, and optimize search and classification processes. In addition, these algorithms have significant applications in fields such as medicine, industry, and economics, where they help to substantiate the decision-making process.

In the context of machine learning, tree-based methods are considered among the most widely used regression and prediction techniques due to their ability to build interpretable and easy-to-visualize models. They can effectively manage both numerical and categorical variables while providing a logical structure that allows the relationships between data to be understood.

At the same time, these methods are used to develop predictive models capable of estimating behaviors, results, or trends based on the available data. By continuously comparing the predicted values with the actual ones, the relationships between the variables can be identified, which contributes to the progressive improvement of the accuracy of the model.

The main purpose of using tree-based algorithms in such applications is to obtain robust and interpretable predictive models, which facilitate the evaluation and optimization of the analyzed processes. Due to their hierarchical structure and decision-making rules, these algorithms allow a clear understanding of the influence of each variable on the final result, thus supporting informed and effective decision-making.

https://doi.org/10.3390/machines14070802

Several tree-based algorithms are implemented within the MATLAB environment, of which the most suitable ones for modeling the parameters of the cutting process are fitrtree and fitrensemble. The fitrtree algorithm builds an individual regression tree, while fitrensemble combines multiple trees through bagging or boosting methods to improve predictive performance [20–23].

3.3. Output Dataset Augmentation and Normalization

Gaussian Data Augmentation contributes significantly to improving the accuracy and robustness of predictions made by machine learning models trained on limited datasets [24,25]. In the case of a small number of experimental samples, there is a risk of the phenomenon of overfitting, in which case the model excessively learns the particularities of the training data and loses its ability to generalize.

To reduce this effect and increase the performance of the model on new data, the randn (Gaussian) function was applied in the study, which generates additional data by introducing controlled statistical variations into the existing data [26,27]. This approach helps to improve the generalizability of the model and increase the stability of predictions made by machine learning algorithms.

The experimental set comprises 210 observations resulting from different toroid milling regimes, obtained by the controlled variation of the cutting speed (Vc), the angle of inclination of the milling axis, and the feed per tooth (fz). The output variable is the surface roughness Ra (μm), measured with the roughness meter.

In order to increase the volume of data and improve the robustness and generalization of machine learning models, an augmentation procedure was applied that extended the training set by generating 6 augmented samples for each original observation, using the following methodology:

- The inputs (Vc, angle of inclination, fz) have been preserved;
- Gaussian noise was added exclusively to the output variable Ra, with the relation: noiseY = randn(size(output)) * (yNoiseLevel * yStd), where randn(size(output)) creates random values with zero mean and standard deviation 1. These values are then multiplied item by item by the product of the noise level (0.1) and the standard deviation of the original data;
- The augmented values were limited to the minimum–maximum range of the original experimental data;
- Reproducibility was ensured by setting the random number generator.

The controlled addition of Gaussian noise exclusively on the variable Ra reflects the natural measurement variations specific to industrial conditions, without altering the causal relationships between the input parameters and the physical surface processing process.

The procedure has been implemented in MATLAB, and the test set has remained unchanged to allow for an objective and unbiased assessment of generalization performance. Noise augmentation controlled solely on the output variable is suitable for machining processes, as the measurement of surface roughness under industrial conditions is affected by significant noise. This technique allows the realistic simulation of variations encountered in the industrial environment while maintaining the physical relationships between the input and output parameters. The augmentation was achieved before the separation of the drive, validation, and test sets. The external test set remained completely untouched by augmentation, thus ensuring an objective assessment of generalization without the risk of data leakage. Following the augmentation, from the initial 210 experimental data points, 1260 synthetic data points were generated, thus resulting in a final set of 1470 examples.

Normalization is a technique by which the values of a dataset are adjusted on a common scale or around the average. This is used to make the data comparable to each other and to avoid large variables disproportionately influencing the calculations. In this way, numerical analysis and algorithms such as machine learning become more stable and efficient. In the research, normalization was applied after the augmentation stage on the final dataset. It contains three continuous numerical predictors: cutting speed, tool axis tilt angle, and feed per tooth, as well as the output variable represented by surface roughness.

It is essential to emphasize that although augmentation was applied to the initial set of 210 observations to increase the volume of training data, the external test set remained completely untouched, allowing an objective assessment of generalization on real non-augmented experimental data. The controlled addition of Gaussian noise exclusively on the output range Ra (with a level of 0.1 $\times$ standard deviation) faithfully reflects the measurement variations inherent in industrial conditions, without distorting the inputs (Vc, tilt angle, fz) or the causal physical relationships between the process parameters and roughness.

This strategy avoids the introduction of arbitrary synthetic artifacts and is supported by the literature dedicated to augmentation in manufacturing applications, where measurement noise is a real component of the data. The results obtained on the validation and external test sets confirm that the models (especially fitrensemble) maintain solid and robust performance, demonstrating that controlled augmentation improves training stability without compromising the credibility of predictions on real data.

3.4. Experimental Strategy

Based on the complete dataset, four supervised machine learning models were developed and compared: two artificial neural networks (trainlm and trainbr) and two decision tree-based models (fitrtree and fitrensemble).

The experimental strategy was structured in two main stages:

- Determining the architecture and structure of the models, which involves identifying the best architectures of neural networks and the optimal structures of the decision trees;
- Establishing the best-performing model.

3.5. Parameters for Determining the Performance of Learning Algorithms

The performance of the models was quantified using four complementary metrics:

a. Mean squared error (*MSE*):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2 \tag{1}$$

where n is the number of observations, x_i is the observed value, and y_i is the estimated value.

b. Root mean squared error (*RMSE*):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \tag{2}$$

RMSE provides a measure of absolute error in the same unit as the output variable (µm).

c. Coefficient of determination (R^2):

$$R^2 = 1 - \frac{\sum(x_i - y_i)^2}{\sum(x_i - \overline{x})^2} \tag{3}$$

where $\overline{x}$ is the mean of the observed data.

Values close to 1 indicate a very good fit.

d. Relative amplitude (RA)

$$RA = \frac{abs(x_i - y_i)}{mean(x_i, y_i)} \times 100 \tag{4}$$

It is a metric proposed in the study, which provides a scale-independent measure of relative error, being useful for comparing performance over different roughness ranges.

e. Mean relative amplitude (MRA)

$$MRA = \frac{1}{n} \sum \frac{abs(x_i - y_i)}{mean\ (x_i, y_i)} \times 100 \tag{5}$$

4. Model Benchmarking

4.1. Artificial Neural Network Architectures

The trainbr algorithm implements Bayesian Regularization, which minimizes a combination of error and network complexity, effectively reducing the risk of overlearning. The trainlm algorithm uses the Levenberg–Marquardt method, a fast optimizer based on the approximation of the Hessian matrix, suitable for nonlinear regression problems.

In order to identify the best performing architectures, scripts have been written for the trainbr and trainlm algorithms so that each architecture would undergo ten trainings, thereby reducing the influence of random initialization of weights and obtaining more stable and reproducible results. In the process of training neural networks, an essential parameter that influences both the accuracy of the model and the calculation time is the number of epochs. An epoch represents a complete traversal of the entire set of training data [28,29].

The architectures evaluated in the study are defined as follows:

- Input layer: 10 neurons
- Output layer: 1 neuron
- Hidden layers: two layers with the architectures [10, 5], [10, 10], [15, 10], [15, 15], [20, 10], [20, 15], and [20, 20], where the first value represents the number of neurons in the first hidden layer, and the second value represents the number of neurons in the second hidden layer
- Function fitting network: fitnet
- Training algorithm: trainbr and trainlm
- Transfer functions: tansig for hidden layers and purelin for the output layer
- Data splitting: trainbr uses 100% of the data for training, whereas trainlm uses 80% of the data for training and 20% for internal validation
- Number of epochs: 3000 for trainbr and 1000 for trainlm, respectively

Considering the peculiarities of the two algorithms, the choice of the optimal architecture was based on different criteria: in the case of the trainlm algorithm, the average of the *RMSE* values resulting from the internal validation was used because MATLAB uses the concept of early stopping, with the training stopping when the error on validation starts to increase. In the case of trainbr, the average of the *RMSE* values obtained in the training process was considered because Bayesian regularization automatically reduces overfitting.

4.1.1. The Best Architecture for the Trainbr Model

The best architecture is the one with 10 neurons hidden in the first layer and 10 neurons hidden in the second layer, with $RMSE$ = 0.062206341 and R^2 = 0.938851954. The choice of the best architecture was made based on the $RMSE$ average of ten runs of the script, and the results, sorted in ascending order, are presented in Table 3.

Table 3. Centralization results in choosing the best training architecture.

Architecture	Mean of *RMSE* (Train)	Mean R^2 (Train)
[10, 10]	0.062206341	0.938851954
[20, 20]	0.062206730	0.938851188
[15, 10]	0.062206731	0.938851187
[20, 15]	0.062207042	0.938850575
[15, 15]	0.062207083	0.938850494
[10, 5]	0.062207856	0.938848974
[20, 10]	0.062208359	0.938847985

4.1.2. Best Architecture for Trainlm

The best architecture is the one with 20 neurons hidden in the first layer and 20 neurons hidden in the second layer, with *RMSE* = 0.065520729 and R^2 = 0.933255833. The choice of the best architecture was made based on the *RMSE* average of ten runs of the script, and the results, sorted in ascending order, are presented in Table 4.

Table 4. Centralization of results for the purpose of choosing the best trainlm architecture.

Architecture	Mean of *RMSE* (Validation)	Mean of R^2 (Validation)
[20, 20]	0.065520729	0.933255833
[15, 15]	0.065953881	0.929871116
[10, 10]	0.066016937	0.927038526
[20, 15]	0.066098198	0.929639556
[20, 10]	0.066260614	0.930017742
[10, 5]	0.066581975	0.932336651
[15, 10]	0.066865269	0.929365313

4.2. Structure of Decision Trees

Fitrtree builds a single regression tree, while fitrensemble implements tree assemblies by boosting (LSBoost) or bagging methods, improving robustness and reducing prediction variance.

The optimal structure is the one that minimizes *RMSE*. In the case of tree-based algorithms such as fitrtree and fitrensemble, the selection and optimization of the model structure are automatically determined by integrated mechanisms for searching for the best structures.

4.2.1. The Best Structure for the Fitrtree Model

The algorithm allows the optimization of the hyperparameters used in learning: MinLeafSize = 10 (the minimum number of data needed by a final node) and MaxNumSplits = 135 (the maximum number of splits by a single tree). These values directly influence how complex the decision tree is, which is reflected in *RMSE* and R^2 by balancing detailed learning and generalization. A very complex tree (MaxNumSplits actually large, MinLeafSize relatively small) tends to reduce *RMSE* on training data and increase R^2 because it fits very well with data, but on new data, it can increase *RMSE* and decrease R^2 due to overlearning. In contrast, a simpler shaft has higher *RMSE* and lower R^2 per workout but can have more stable performance on the test data, avoiding overlearning and generalizing better.

4.2.2. The Best Structure for the Fitrensemble Model

In the case of the fitrensemble model, optimal hyperparameters are: LSBoost algorithm, NumLearningCycles = 497, LearnRate = 0.993800, MinLeafSize = 19, MaxNumSplits = 1102. This configuration indicates a powerful boosting model, in which shafts are built sequentially to correct previous errors, resulting in the progressive reduction of the *RMSE* and the increase of R^2. The relatively high learning rate makes each tree contribute significantly to the final model, accelerating adaptation to the data, allowing the trees to be highly detailed and capture complex relationships from the data. Overall, this combination resulted in high performance with low errors and good ability to explain data variation, representing a balance between complexity and accuracy.

4.3. Performance of Models for Learning Data

In order to reduce the variability introduced by the random initialization of the weights and the stochastic process of tree construction, the models with the selected architectures/structures were run 15 times, and the best models were saved.

4.3.1. The Trainbr Model

Because the trainbr algorithm incorporates validation during training, the model with the best validation performance was saved as net_trainbr_best (for reference, the best model was the second one). This model achieved the lowest $RMSE$ = 0.0622050455378155 (Figure 6).

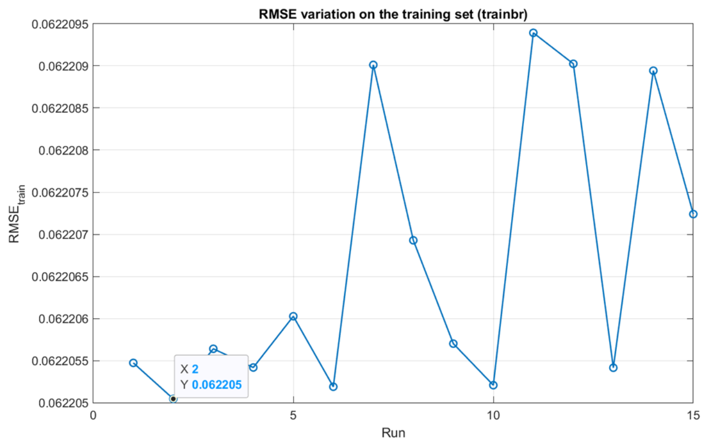
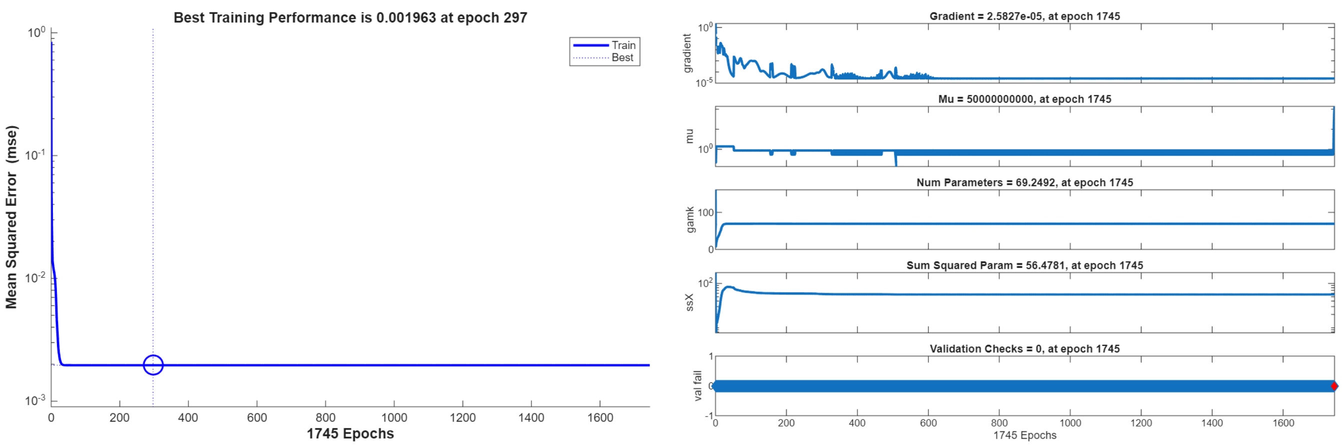


Figure 6. *RMSE* variation for training data.

For the best model, the best performance was achieved at epoch 297, with a value of MSE = 0.001963 (Figure 7a). According to Figure 7b, the training was completed at the time of 1745, at which time the gradient decreased to 2.5827×10^{-5}.



(**a**) (**b**)

Figure 7. (**a**) Best training performance; (**b**) completion report.

The error histogram for the training set (Figure 8) reveals a nearly normal error distribution centered around zero in the range of [−0.1799, +0.169] µm.

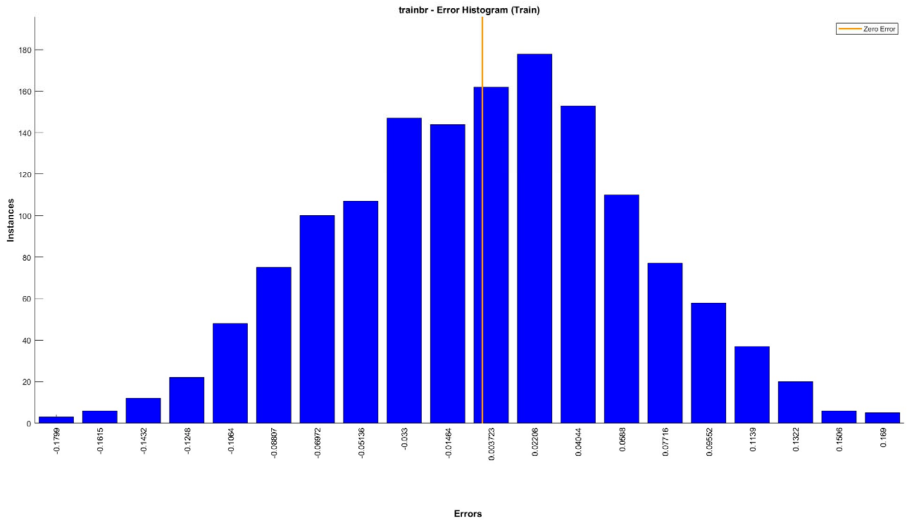


Figure 8. Error histogram for training data.

The regression graph confirms the quality of the model on the training set with R^2 = 0.9389 (Figure 9).

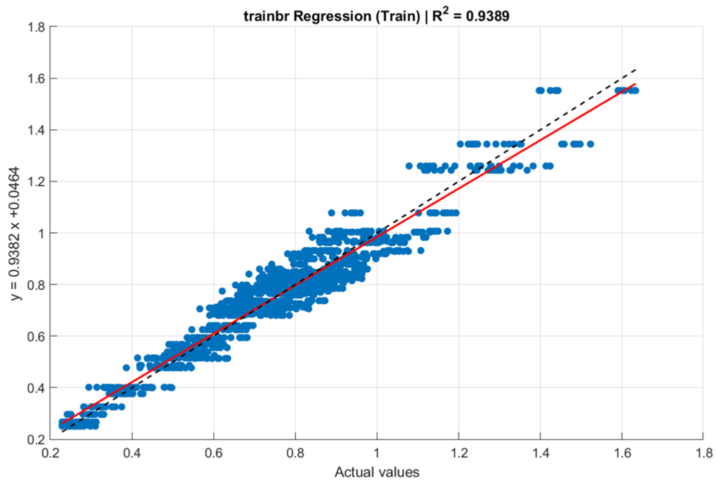


Figure 9. Regression for training data.

The trainbr model demonstrated a stable evolution and a good ability to adapt to the data analyzed, without obvious signs of instability in the training process. The compact error distribution and consistency of the results obtained in independent runs indicate a high level of precision and a good generalization capacity.

4.3.2. The Trainlm Model

The best (twelfth) model resulting from the validation, saved under the name net_trainlm_best, is the one that achieved the lowest $RMSE$ = 0.0634539101458346 (Figure 10).

For the best model, the best validation performance was achieved at epoch 9, with a value of MSE = 0.0020426 (Figure 11a). The training stopped at epoch 11 (Figure 11b), with a very small gradient (1.0981×10^{-10}) and two validation errors. This behavior indicates a rapid convergence, but also a possible overlearning effect.

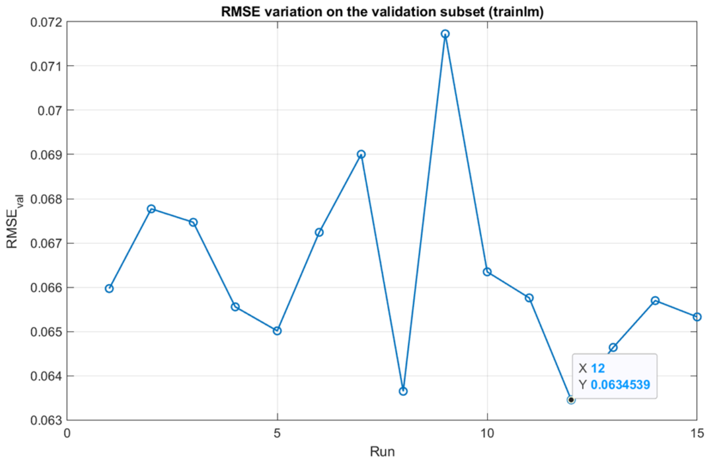
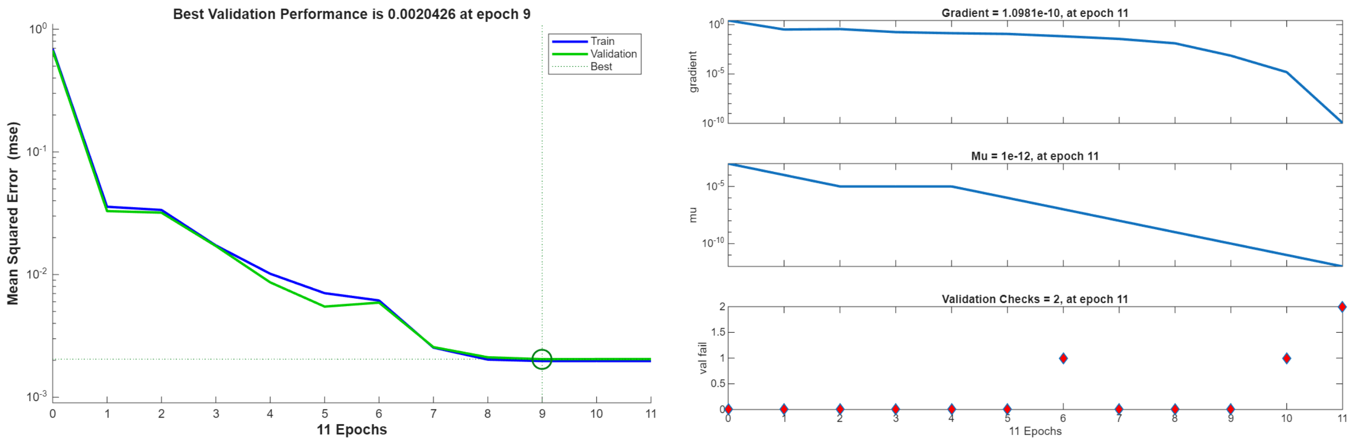


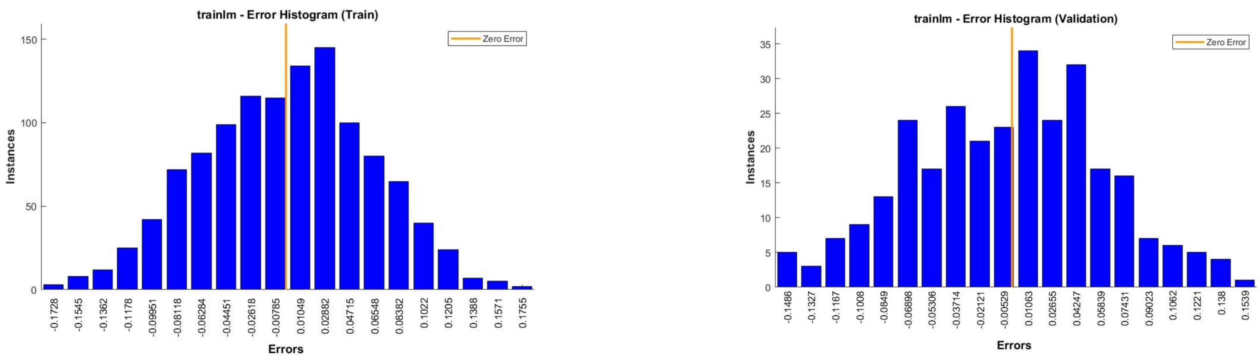
Figure 10. *RMSE* variation for validation data.



(**a**) (**b**)

Figure 11. (**a**) Best training performance; (**b**) completion report.

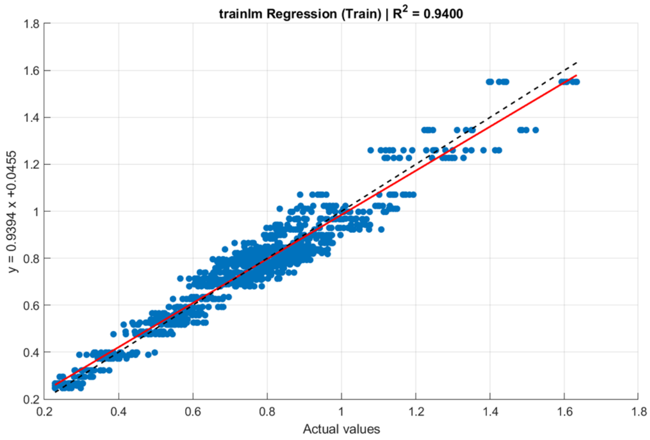
The errors on the histograms on the training set (Figure 12a) and validation (Figure 12b) show distributions concentrated around zero.



(**a**) (**b**)

Figure 12. Histogram of errors: (**a**) for training data; (**b**) for validation data.

The regression plots show a good fit on the training set, R^2 = 0.9400 (Figure 13a), and on the validation set, R^2 = 0.9297 (Figure 13b).



(**a**) (**b**)

Figure 13. Regressions: (**a**) for training data; (**b**) for validation data.

The results obtained highlight a good performance of the trainlm model for both training and validation data, characterized by rapid convergence, stability, and a high generalization capacity. Error distribution indicates accurate predictions close to the actual values, and independent runs confirm model consistency.

4.3.3. The Fitrtree Model

The best model (tenth) resulting from the validation, saved as model_fitrtree_best, is the one that achieved the lowest *RMSE* = 0.0614338195414177 (Figure 14).

The histogram of errors on the training set (Figure 15a) and the validation histogram (Figure 15b) confirm a symmetrical and well-centered distribution around zero.

The analysis of the importance of the predictors performed on the validation set (Figure 16) reveals their ranking, although the absolute values are not important. In the case of the fitrtree model, the cutting speed takes first place, followed by the angle of inclination and the feed per tooth.

The regression graphs confirm the excellent performance of the model: R^2 = 0.9390 on the training set (Figure 17a) and R^2 = 0.9326 on the validation set (Figure 17b).

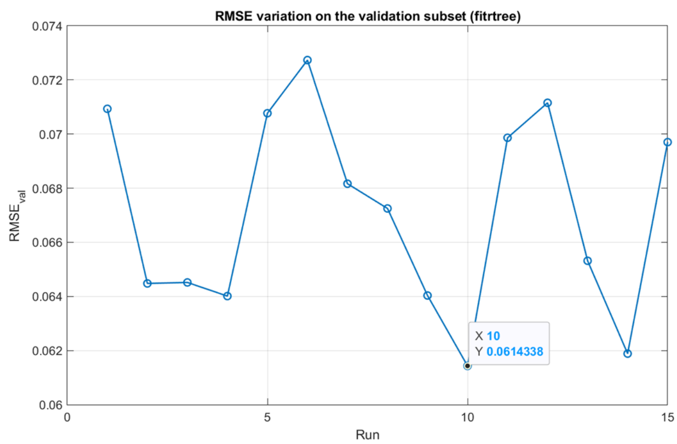
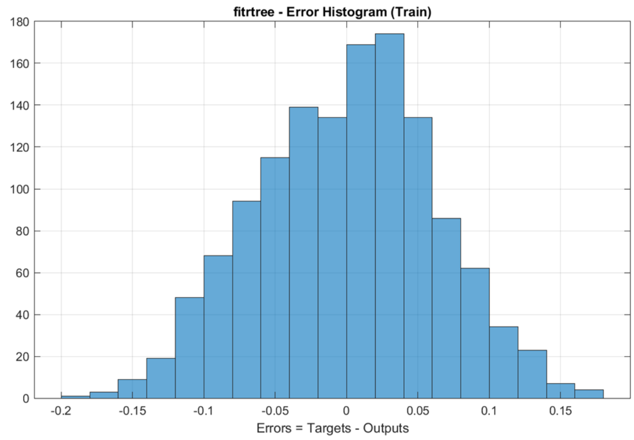
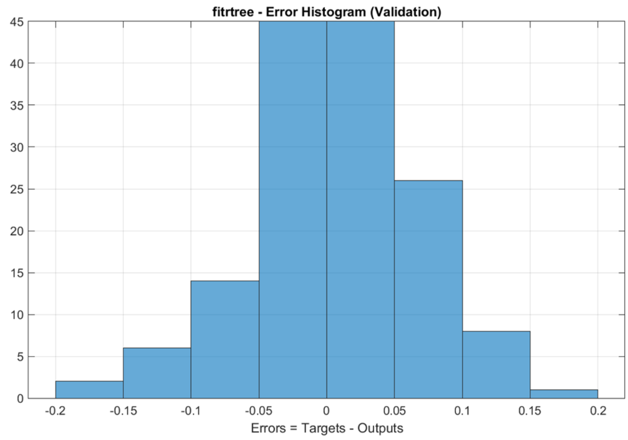


Figure 14. *RMSE* variation for validation data.

The fitrtree model shows stable behavior and good predictability, supported by balanced error distribution and low variations between independent runs. The results obtained indicate an adequate fit of the model for both training and validation data, which confirms a good generalization capacity. The analyses carried out show that the cutting speed has the most important influence on the result, while the angle of inclination of the axis and the feed per tooth contribute to a lesser extent. Overall, the model provides accurate and consistent results for the data analyzed.



(**a**)



(**b**)

Figure 15. Histogram of errors: (**a**) for training data; (**b**) for validation data.

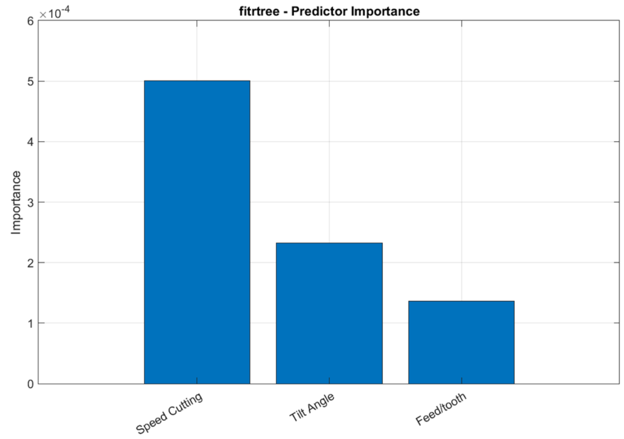
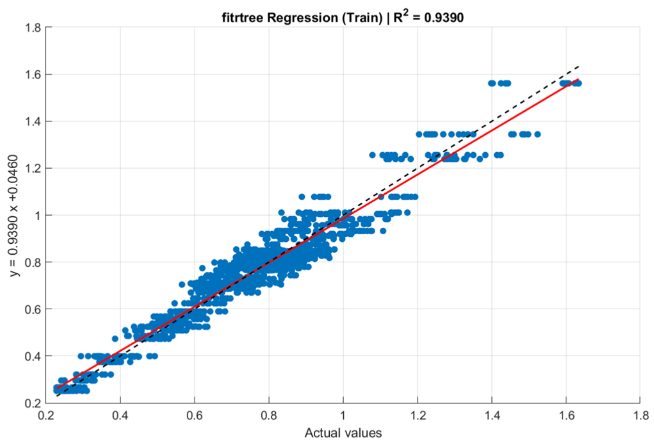
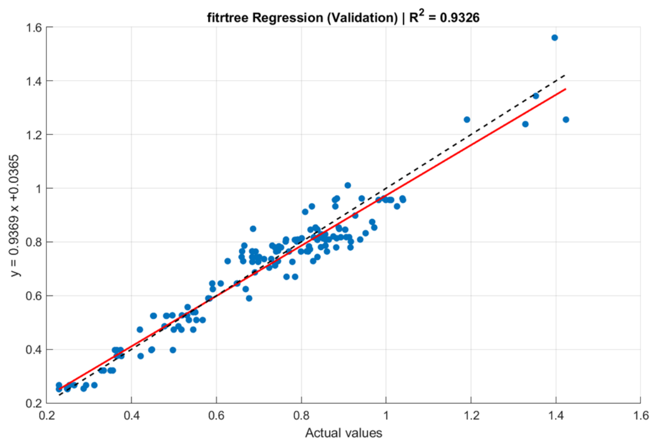


Figure 16. Analyzing the importance of predictors.



(**a**)



(**b**)

Figure 17. Regressions: (**a**) for training data; (**b**) for validation data.

4.3.4. The Fitrensemble Model

The best model (tenth) resulting from the validation, saved under the name model_fitrensemble_best, is the one that achieved the lowest *RMSE* = 0.0614338195356529 (Figure 18).

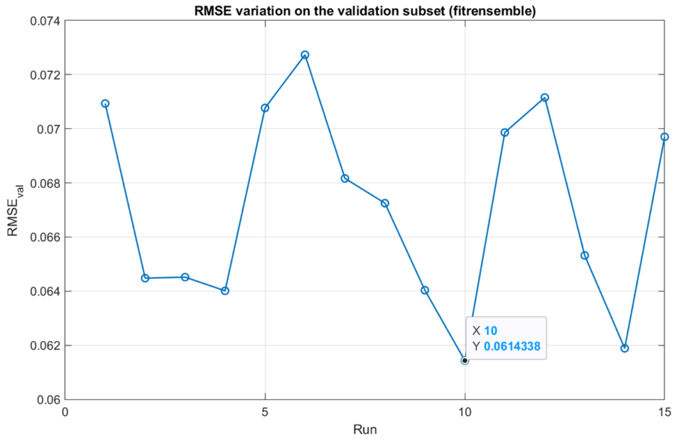
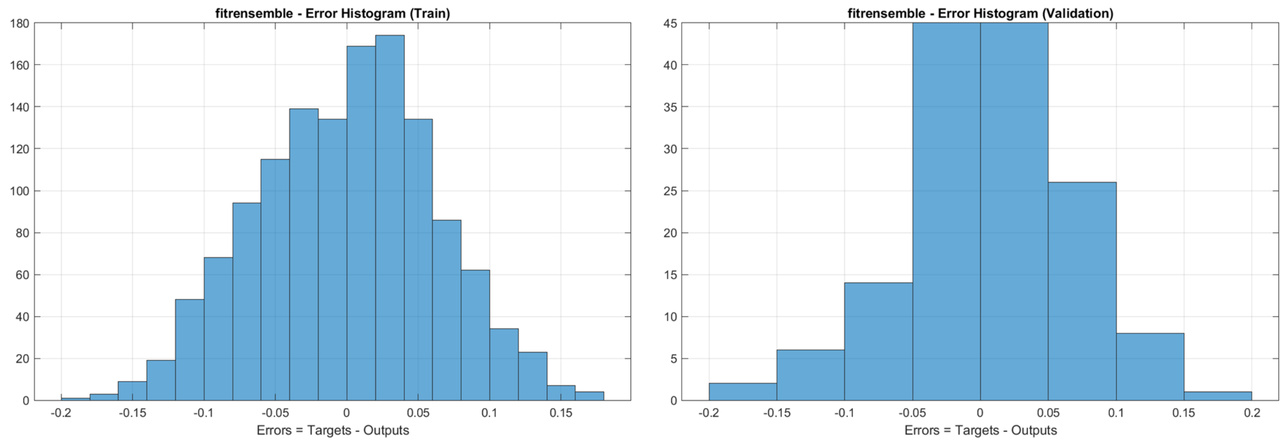


Figure 18. *RMSE* variation for validation data.

The error histograms on the training set (Figure 19a) and on the validation set (Figure 19b) confirm a symmetrical and well-centered distribution around zero, similar to that of the fitrtree model.



(**a**) (**b**)

Figure 19. Histogram of errors: (**a**) for training data; (**b**) for validation data.

The analysis of the importance of predictors performed on the validation set (Figure 20) reveals that the cutting speed is the most important predictor, followed by the angle of inclination and the feed per tooth, similar to the fitrtree model.

The regression graphs confirm the very good performance of the model: R^2 = 0.9390 on the training set (Figure 21a) and R^2 = 0.9326 on the validation set (Figure 21b).

The fitrensemble model presents stable and precise results, evidenced by the balanced distribution of errors and the low variations between independent runs. The analysis of predictors shows that the cutting speed has the most important contribution to the result, while the angle of inclination of the axis and the feed per tooth have a secondary influence. Overall, the model stands out for its very good fit and robust behavior in the modeling process.

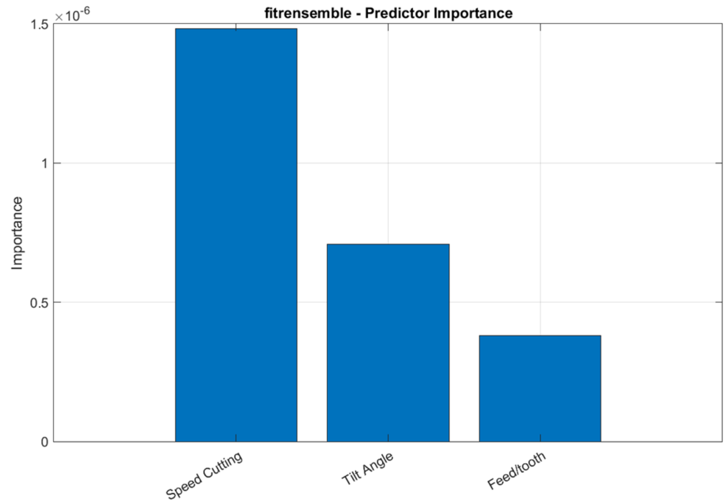


Figure 20. Analyzing the importance of predictors.

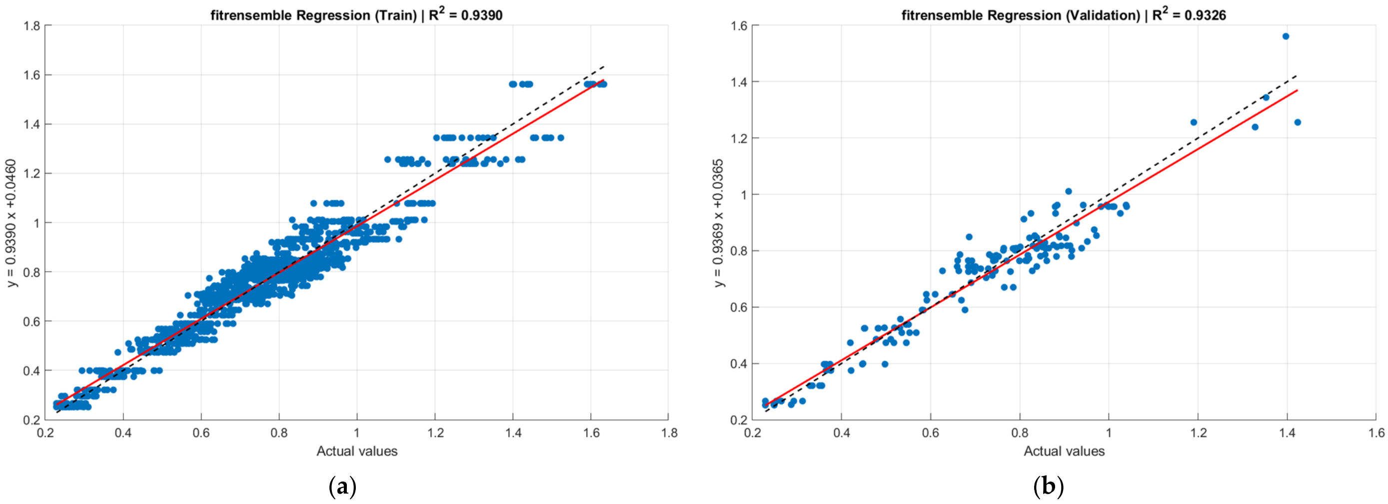


Figure 21. Regressions: (**a**) for training data; (**b**) for validation data.

4.4. Training and Validation Results for the Four Models

In the training and validation stage, very good behavior of all models is transmitted on the training data, with *RMSE*_train in the range [0.0622, 0.0624] according to Table 5.

The classification of the models was based on the *RMSE* value on the validation set. Models based on decision trees occupied the first two positions, showing almost identical values in both training and validation. These results indicate the good generalization capacity and high robustness of these models.

The trainbr model ranked third, using the *RMSE* value on the training set in the absence of a separate validation set. Although it shows comparable performance on training data, the absence of independent validation limits the possibility of evaluating generalization. The trainlm model ranked last, recording the highest *RMSE* on the validation set and a slight degradation of performance from training to validation, suggesting the beginning of overlearning even in the internal training stage.

From the perspective of the coefficient of determination R^2, all models recorded high values, learning over 93.8% of the variability of roughness. On the validation set, fitrtree and fitrensemble got R^2 = 0.9326, while trainlm dropped to 0.9297, highlighting a slight superiority of tree models in generalization.

Table 5. Performance of models on the training and validation sets.

Rank	Model	*RMSE*_Train	*RMSE*_Validation	R^2_Train	R^2_Validation
1	fitrensemble	0.0624388488594664	0.0614338195356529	0.939027409654278	0.932634123148549
2	fitrtree	0.0624388488594664	0.0614338195414177	0.939027409654278	0.932634123135906
3	trainbr	0.0622050455378155	-	0.938854500010402	-
4	trainlm	0.0623326308518884	0.0634539101458346	0.940009692431846	0.92966894671263

5. Model Results for Test Data

The performance of the models was evaluated for external test data using the best run saved from the training and validation stage, selected according to the *RMSE* criterion. The classification presented in Table 6 was made in ascending order of the value of *RMSE*_test.

The comparative analysis shows that the trainbr model recorded the best performance, obtaining the lowest error values and a good generalization capacity.

The tree-based models showed very close results, with high values of the coefficient of determination R^2 and low relative errors. In contrast, the trainlm model demonstrated significantly lower performance, confirming a high degree of overlearning and a limited ability to generalize to new data.

Table 6. Performance of saved models for external test data.

Rank	Model	*RMSE*_Test	R^2_Test	*MSE*_Test	*MRA*_Test [%]	Max*RA*_Test [%]
1	trainbr	0.0364442129999644	0.986962170926124	0.00132818066118677	4.650	7.846
2	fitrensemble	0.0443317576982868	0.994900577826597	0.00196530474061961	5.912	7.568
3	fitrtree	0.0447601289798864	0.995642863476776	0.00200346914629607	5.978	7.871
4	trainlm	0.1552993847370080	0.462250428415499	0.02411789889969320	20.552	50.211

The regression graph of the trainbr model in Figure 22 shows a very good alignment between the actual and predicted values, with most of the points being located close to the ideal regression line. With a coefficient of determination R^2 = 0.987, the model explains 98.7% of the variance in the external test data, demonstrating excellent generalization and accurate prediction.

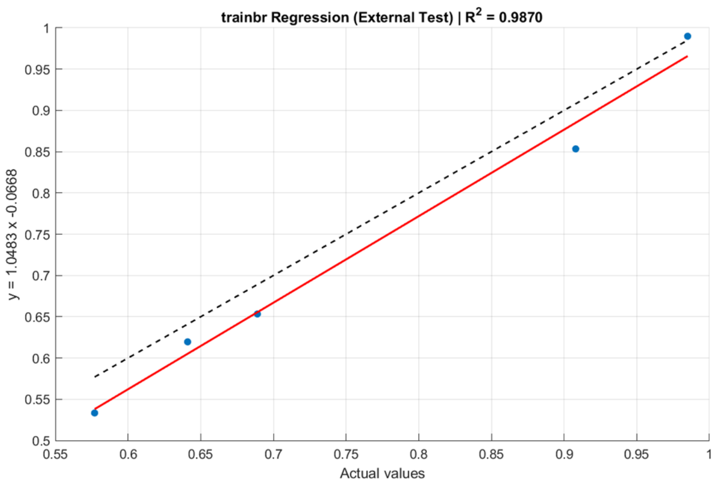


Figure 22. Regression for the trainbr model test data.

The regression graph of the trainlm model in Figure 23 shows a significant dispersion of the points from the regression line, indicating a weak correlation between the actual and predicted values. The reduced value of R^2 = 0.4623 shows that the model explains only 46.23% of the variance in the data, confirming a high degree of overlearning and poor performance on the external test set.

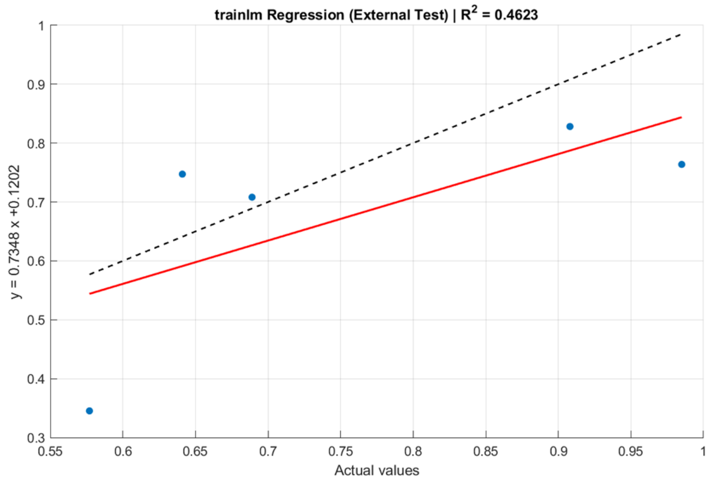


Figure 23. Regression for the trainlm model test data.

The regression graph of the fitrtree model in Figure 24 shows a very good alignment of the points on the regression line, with a minimum dispersion and R^2 = 0.9956, indicating that the model explains 99.56% of the variance of the test data, demonstrating very high accuracy and superior generalization.

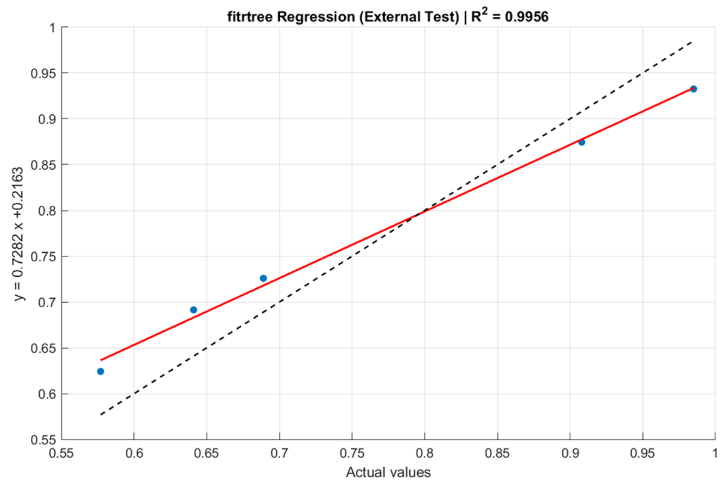


Figure 24. Regression for the fitrtree model test data.

The regression graph of the fitrensemble model in Figure 25 shows a very good correlation between the actual and predicted values, with the points being tightly grouped around the regression line. With R^2 = 0.9949, the model explains 99.49% of the data variance, confirming very good performance and robustness of the shaft assembly.

The *RA* analysis on the observations from the test set highlights a better performance of the trainbr model (Figure 26), with the *RA* values reduced between 0.5% and 7.9%, which indicates high accuracy and good consistency in predictions.

In contrast, the trainlm model (Figure 27) shows significantly greater variations in *RA*, with values between 3% and 50%, reflecting lower stability and more uneven accuracy in individual observations in the test set.

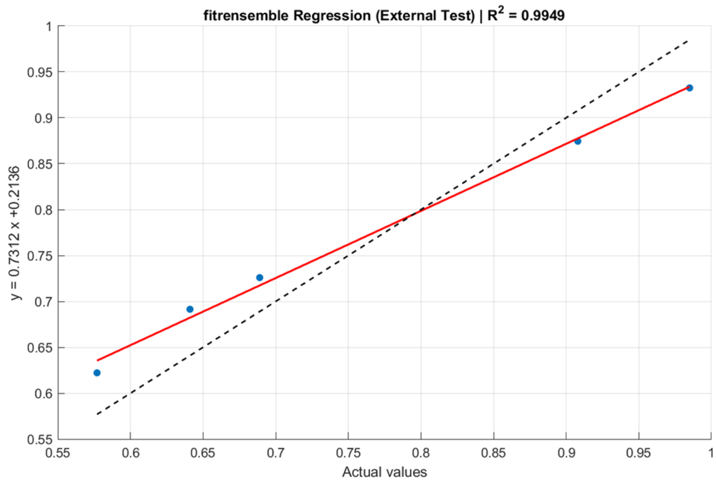


Figure 25. Regression for the fitrensemble model test data.

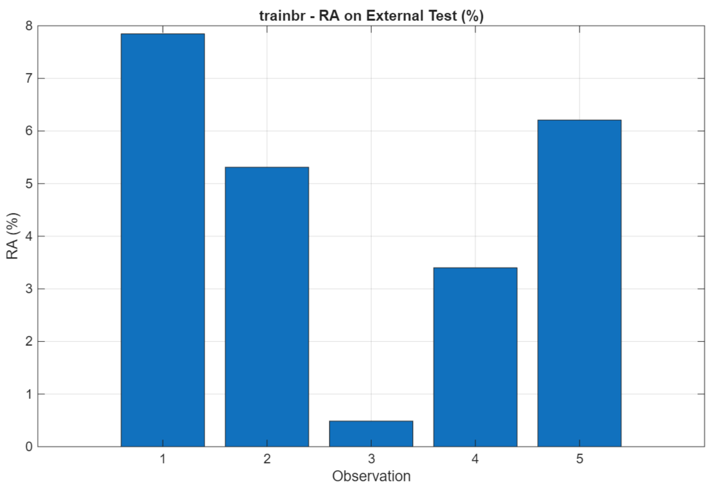


Figure 26. *RA* analysis for the trainbr model.

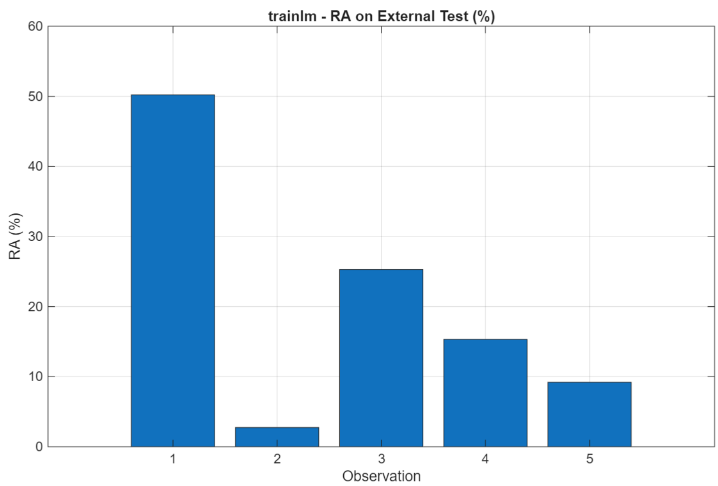


Figure 27. *RA* analysis for the trainlm model.

The *RA* analysis on the individual observations from the test set shows that the fitrtree model (Figure 28) has values between 3.8% and 7.9%, indicating good accuracy and controlled variation of predictions.

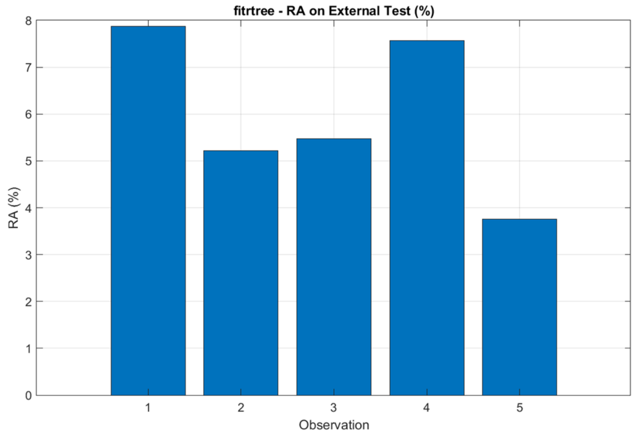


Figure 28. *RA* analysis for the fitrtree model.

Similarly, the fitrensemble model (Figure 29) records the same ranges, between 3.8% and 7.9%, which confirms consistent behavior and comparable performance between the two models on the individual observations in the test set.

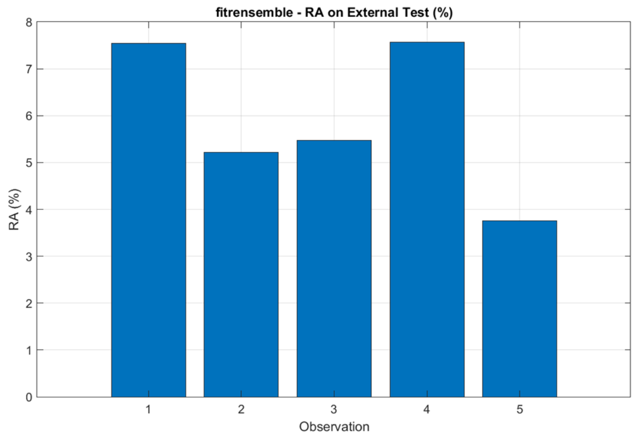


Figure 29. *RA* analysis for the fitrensemble model.

The sensitive predictor analysis performed for the test data for the trainbr model shown in Figure 30 shows a dominant influence of the cutting speed (*Vc*) (~56%), followed by the angle of inclination of the axis (~30%) and the feed per tooth (*fz*) (~13%). This highlights the predominant influence of the cutting speed, while the other parameters have a lower contribution.

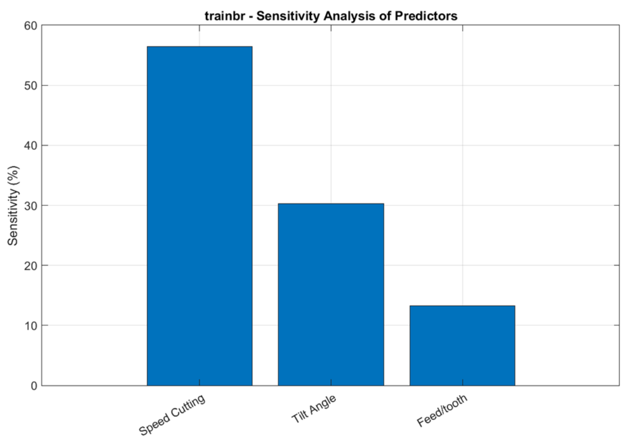


Figure 30. Sensitive predictor analysis for the trainbr model.

The sensitive predictor analysis performed for the test data for the trainlm model shown in Figure 31 indicates a dominant influence of axis tilt angle (~44%), followed by cutting speed (~31%) and feed per tooth (~25%).

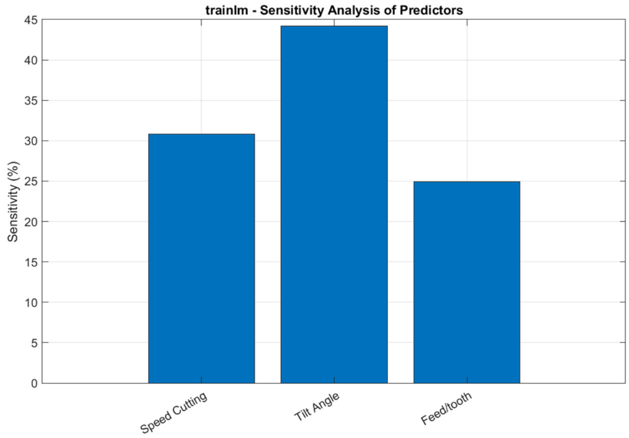


Figure 31. Sensitive predictor analysis for the trainlm model.

The sensitive predictor analysis performed for the test data for the fitrtree model shown in Figure 32 indicates a dominant influence of the cutting speed (Vc) (~48%), followed by the axis tilt angle (~37%) and feed per tooth (fz) (~15%).

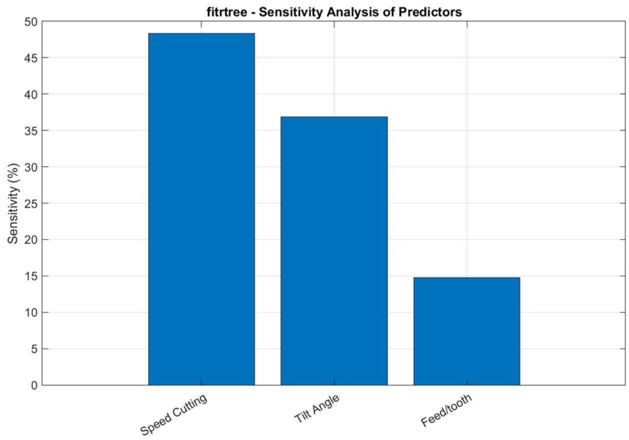


Figure 32. Sensitive predictor analysis for the fitrtree model.

The sensitive predictor analysis performed for the test data for the fitrensemble model shown in Figure 33 indicates the same hierarchy as for the fitrtree model: the cutting speed (Vc) has the dominant influence (~48%), followed by the axis tilt angle (~37%) and feed per tooth (fz) (~15%).

For an overall assessment of the performance of the models, an overall score was calculated based on the sum of the ranks obtained in the main performance metrics on the training, validation, and external test datasets from the perspective of the best run. According to Table 7, the overall score obtained by summing the ranks indicates a superior overall performance of the fitrensemble model (score 16), followed by trainbr (score 19) and fitrtree (score 21). Although trainbr dominates on test data, the fitrensemble model offers certainty on the validation side and consistency in the other metrics. In addition to the overall score, the constancy, generalization, and results obtained in training and validation confirm that the fitrensemble model offers additional safety because the test data are variable, and it is based on learning data.

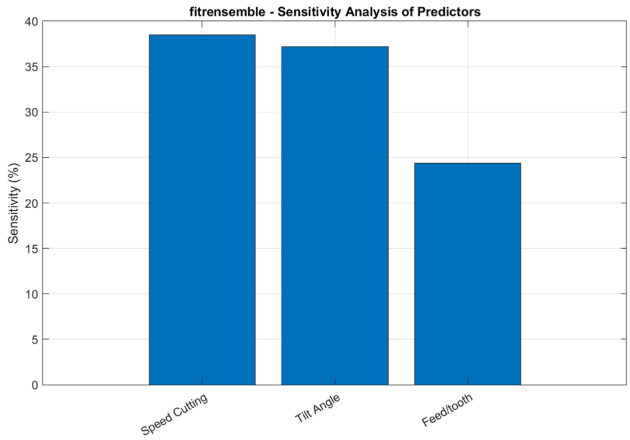


Figure 33. Sensitive predictor analysis for the fitrensemble model.

Table 7. The global score of the four models for the best run.

Model	Fitrensemble	Trainbr	Fitrtree	Trainlm
*RMSE*_test Rank	2	1	3	4
*RMSE*_train Rank	2	1	2	4
*RMSE*_validation Rank	1	3	2	4
R^2_test Rank	2	3	1	4
R^2_train Rank	2	4	2	1
R^2_validation Rank	1	4	2	3
*MRA*_test Rank	2	1	3	4
*MSE*_test Rank	2	1	3	4
*MaxRA*_test Rank	2	1	3	4
Global Score	16	19	21	32

Model Predictions for Test Data

The graph in Figure 34 for the trainbr model shows a very good match between the actual (blue) and predicted (red) values. The two lines are almost overlapping over the entire interval, with minimal deviations. The value of R^2 = 0.9870 confirms the very good accuracy of predictions on the external test set.

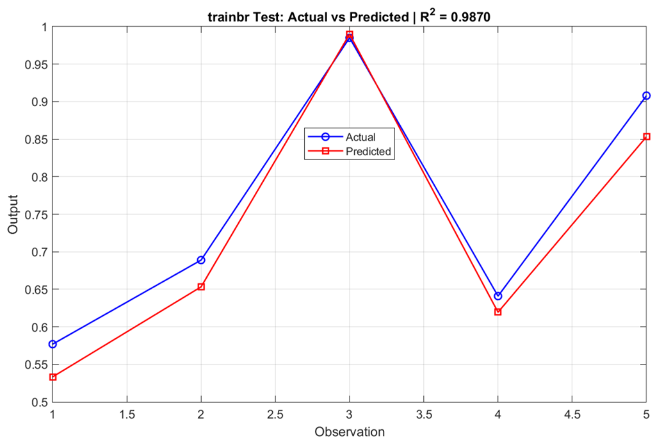


Figure 34. Actual vs. predicted trainbr model.

The graph in Figure 35 of the trainlm model shows significant deviations between the actual and predicted values, especially in observations 1, 3, and 5. The red line (predicted) moves significantly away from the blue (actual) one. R^2 = 0.4623 indicates a very weak correspondence and confirms the limited generalization ability of the model.

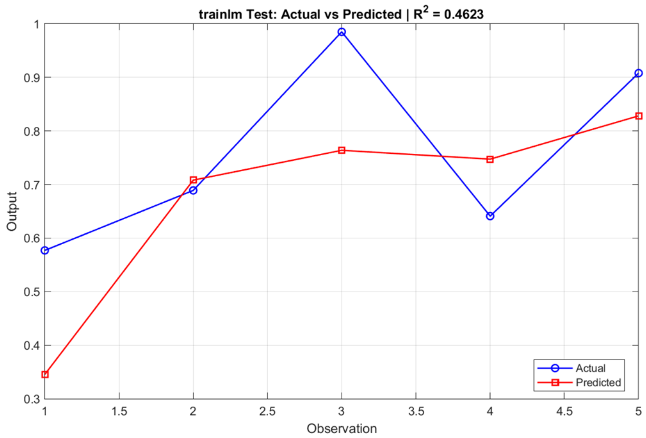


Figure 35. Actual vs. predicted trainlm model.

The graph in Figure 36 of the fitrtree model shows a very good alignment between the actual and predicted values, with the two lines being close. Deviations are minimal on all five observations. With an R^2 = 0.9956, the model demonstrates very good accuracy and very good generalization.

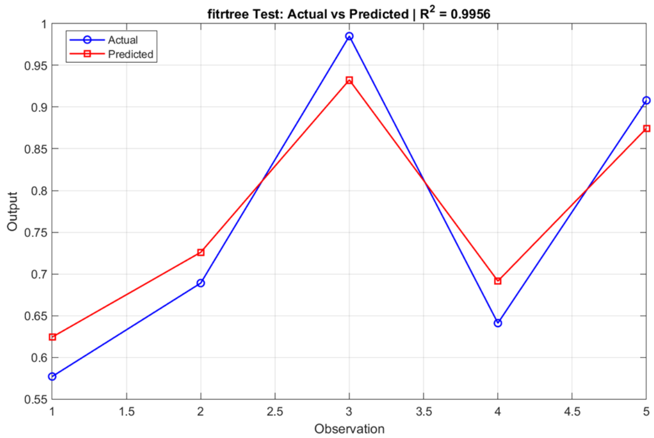


Figure 36. Actual vs. predicted fitrtree model.

The graph in Figure 37 of the fitrensemble model shows a very good overlap between the actual and predicted values. The red line closely follows the blue line throughout. R^2 = 0.9949 reflects a very good performance, close to that of the fitrtree model, confirming the robustness of the shaft assembly.

To assess the actual predictive capacity of the models, Table 8 shows the actual measured roughness values and the values predicted by the best models.

Analysis of the predicted values on the external test set highlights the superior performance of decision tree-based models. Both fitrensemble and fitrtree generate highly accurate predictions, with relative errors ranging from 3.76% to 7.87%, demonstrating high consistency across the entire dataset. The trainbr model gives very good results on most observations but shows greater variations in relative errors. In contrast, the trainlm model registers significantly higher relative errors (up to 50.21%), confirming once again the pronounced trend of overlearning and the limited ability to generalize.

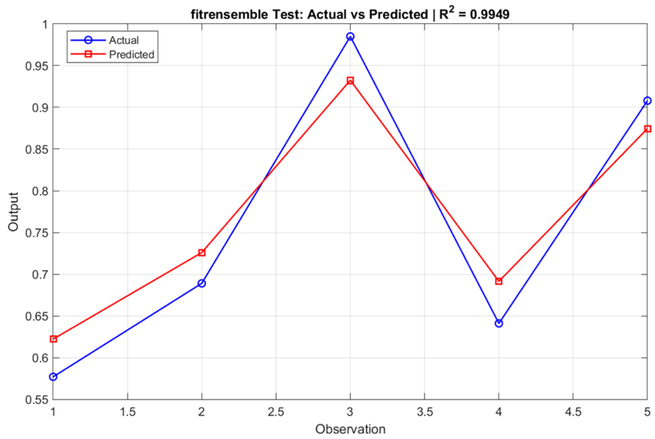


Figure 37. Actual vs. predicted fitrensemble model.

Table 8. Model predictions for test data.

Cutting Speed [m/min]	Tilt Angle [°]	Feed per Tooth [mm/tooth]	Ra Real [µm]	Trainbr		Trainlm		Fitrtree		Fitrensemble	
				Prediction [µm]	*RA* [%]	Prediction [µm]	*RA* [%]	Prediction [µm]	*RA* [%]	Prediction [µm]	*RA* [%]
100	20	0.12	0.577	0.533	7.846	0.345	50.211	0.624	7.871	0.622	7.541
120	30	0.14	0.689	0.653	5.311	0.708	2.742	0.725	5.217	0.725	5.217
140	40	0.16	0.985	0.989	0.486	0.763	25.290	0.932	5.474	0.932	5.474
160	50	0.18	0.859	0.619	3.401	0.747	15.323	0.691	7.569	0.691	7.569
180	60	0.20	0.908	0.853	6.208	0.828	9.196	0.874	3.758	0.874	3.758
MRA					4.65%		20.55%		6.03%		5.91%

6. Discussion

The analysis is based on the comparison of the four models of supervised learning, two artificial neural network architectures (trainbr and trainlm) and two models based on decision trees (fitrtree and fitrensemble). This analysis reveals important insights into their suitability for predicting surface roughness when milling with the toroidal mill on a five-axis CNC of C45 steel under real industrial conditions. The dominant influence of cutting speed aligns with known physical mechanisms: increasing speed reduces local thermal forces and improves chip formation, leading to lower roughness in the investigated domain.

All models demonstrated strong performance in the training and internal validation stages, with *RMSE* values closely concentrated between 0.0614 and 0.0635 µm and R^2 coefficients exceeding 0.929. This confirms that the augmented dataset (1470 observations derived from 210 experimental points) provided enough information to learn the complex, nonlinear relationships between the input parameters (cutting speed, tilt angle, and feed per tooth) and the output variable Ra.

A key finding is the superior behavior of the fitrensemble model and the firtree model on the training and validation sets. The fitrensemble model (LSBoost with optimized hyperparameters: NumLearningCycles = 497, LearnRate = 0.9938, Min-LeafSize = 19, MaxNumSplits = 1102) achieved the lowest *RMSE* on the validation set (0.06143 µm), along with an R^2 of 0.9326. The fitrtree model produced almost identical results, highlighting the robustness of tree-based approaches to this regression task.

These models benefit from their ability to manage mixed interactions with variables, to effectively manage noise through division mechanisms, and to provide interpretable rankings of the importance of features. In both shaft designs, the cutting speed became the dominant predictor, followed by the angle of inclination and feed per tooth, which correlates well with the machining theory, where the cutting speed strongly influences thermal effects, chip formation, and surface roughness in toroidal milling.

The trainbr model ranked second, close to the fitrtree model. Overall, it showed a very good generalization on the external test set ($RMSE$_test = 0.0364 μm, R^2 = 0.987, MRA = 4.65%). In contrast, the trainlm model showed clear signs of overlearning, with a visible degradation from validation ($RMSE$ = 0.0635 μm) to testing ($RMSE$ = 0.1553 μm, R^2 decreasing to 0.462), confirming the limitations of this algorithm when working with relatively small experimental datasets, even after Gaussian augmentation.

Considering the global results at all stages (training, validation, and testing), the fitrensemble model is considered the most reliable and robust choice. Although trainbr achieved superior accuracy on the test dataset, external test data are variable and may not fully represent future operating conditions. Validation performance remains the strongest indicator of true generalization ability, as it reflects the model's behavior on unseen data taken from the same distribution as the training set.

The fitrensemble's lowest validation error, combined with its high and consistent performance on metrics, provides greater confidence that it will maintain good predictive accuracy under varied industrial conditions. This stability is particularly valuable in production environments where process parameters and batches of materials can fluctuate.

Sensory analysis supports this conclusion: both tree-based models consistently identified the same hierarchical parameter influence observed in experimental trends, while trainbr and especially trainlm showed less stable sensitivity rankings. The Gaussian augmentation strategy proved highly effective in improving the robustness of the model without introducing unrealistic data.

Although the accuracy obtained is high, it must be interpreted in the context of the augmented dataset and the limited range of parameter variation. The models demonstrated good generalization on the external test set, but performance on data outside the experimental range remains a subject for further investigation.

From an industrial perspective, these models offer an easy way to reduce costly practical experiments. Once implemented, the fitrensemble model can serve as an essential digital component for real-time roughness prediction and parameter optimization, and can be integrated into CAM software (PowerMILL v.2015) or CNC controllers.

Although the study focuses on toroidal milling of C45 steel, the proposed comparative methodology, including controlled augmentation, multi-metric evaluation and predictor sensitivity analysis, is generalizable to a wide range of machining processes. The results demonstrate the potential of ensemble models for predicting surface quality in variable industrial environments, helping to optimize digital manufacturing processes and reduce experimental costs in the context of Industry 4.0.

Limitations of the study include the relatively small range of axial/radial cutting depths (kept constant) and the concentration on a single material (C45).

Future research could consider expanding the dataset to include additional materials, tool wear effects, and complex cutting regimes. Additionally, hybrid ANN-tree architectures may be explored to further enhance interpretability and accuracy.

7. Conclusions

This study successfully developed and compared four supervised learning models for predicting surface roughness when milling surfaces with a toroidal mill on a five-axis CNC of C45 steel. The research demonstrates that both artificial neural networks and decision tree-based assemblies can achieve high predictive performance, providing a viable alternative to extensive experimentation in industrial environments.

The proposed objectives have been fully met:

- A limited but relevant experimental plan was developed and implemented on the Okuma MU-400VA CNC;
- The database has been expanded by Gaussian augmentation, ensuring sufficient volume for training;
- Four machine learning models (trainbr, trainlm, fitrtree, and fitrensemble) were trained, validated, and tested;
- The models were comparatively evaluated using metrics (*MSE*, *RMSE*, R^2, *MRA*) and stability analysis based on 15 independent runs;
- A detailed analysis of the model with the best performance was carried out.

Main conclusions:

1. All four models effectively learned the basic relationships, with R^2 values above 0.93 for the training/validation data;
2. The fitrensemble model is considered the best. It achieved the lowest *RMSE* in the validation set and maintained good consistency across the training, validation, and testing phases. Its superior validation performance provides the highest confidence in generalizing new conditions, even though the trainbr model showed slightly better accuracy on the external test set used in this study. The choice of the optimal model depends on the priorities of the application: trainbr for maximum accuracy on similar new data, and fitrensemble for consistency and partial explainability under variable industrial conditions;
3. Models based on decision trees (fitrensemble and fitrtree) demonstrated greater robustness and stability compared to the trainlm neural network, which suffered from overlearning;
4. The cutting speed was consistently identified as the most influential parameter, followed by the angle of inclination and feed per tooth, confirming the physical mechanisms of surface generation with the toroidal cutter;
5. The combination of real industrial experiments, data augmentation, data training, and validation has enabled the development of accurate and implementable predictive models capable of supporting process optimization and reducing experimental costs;

The proposed methodology and fitrensemble model provide a practical, high-precision tool for manufacturers who want to control and optimize surface quality in five-axis machining operations. Future work will focus on expanding parameters, integrating tool wear monitoring, and developing adaptive online versions of the model for intelligent manufacturing systems.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data supporting the findings of this study are included within the article.

Acknowledgments: During the preparation of this manuscript, the authors used QuillBot (Web version) Online for English language corrections (especially grammar). After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication. The authors gratefully acknowledge the administrative and technical support provided by the Technical University of Cluj-Napoca during the development of this research.

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

Abbreviations

The following abbreviations are used in this manuscript:

a_e	Radial depth of cut
a_p	Axial depth of cut
ANN	Artificial neural network
CAM	Computer-aided manufacturing
CNC	Computer numerical control
MAE	Absolute average error
MLP	Multilayer perceptron
MRA	Mean of the relative amplitude
MSE	Mean squared error
RA	Relative amplitude
R^2	Coefficient of determination
Ra	Arithmetic average roughness
RMSE	Root mean squared error

References

1. Vu, T.-H.; Hsu, C.-H. Experimental Investigation and Predictive Modeling of Surface Roughness in Dry Turning of AISI 1045 Steel Using Power-Law and Response Surface Approaches. *Appl. Sci.* **2026**, *16*, 1392. [CrossRef]
2. Wang, J.; Gu, S.; Ma, H.; Yu, H.; Yang, C.; Zhao, J.; Zhang, X. HDA Coating on AISI 1045 Steel with Enhanced Corrosion and Wear Performance. *Coatings* **2026**, *16*, 95. [CrossRef]
3. *ISO 4287:1997/Amd 1:2010*; Geometrical Product Specifications (GPS)—Surface Texture: Profile Method—Terms, Definitions and Surface Texture Parameters. ISO: Geneva, Switzerland, 2010.
4. Weichert, D.; Link, P.; Stoll, A.; Rüping, S.; Ihlenfeldt, S.; Wrobel, S. A review of machine learning for the optimization of production processes. *Int. J. Adv. Manuf. Technol.* **2019**, *104*, 1889–1902. [CrossRef]
5. Burawat, P. Improvement of Productivity by Using Means of Lean Manufacturing, Just in Time, and Production Technology in Thai SMEs Manufacturing Industry. *Des. Constr. Maint.* **2024**, *4*, 102–108. [CrossRef]
6. El-Baz, M.A.; Abd-Elwahed, M.S. Enhancing Product Quality Using Artificial Neural Networks and Genetic Algorithms. *J. Manuf. Mater. Process.* **2025**, *9*, 322. [CrossRef]
7. Sunanta, A.; Suranuntchai, S. Optimization of Process Parameters for Advanced High-Strength Steel JSC980Y Automotive Part Using Finite Element Simulation and Deep Neural Network. *J. Manuf. Mater. Process.* **2025**, *9*, 197. [CrossRef]
8. Moor, J. The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. *AI Mag.* **2006**, *27*, 87.
9. Michalski, R.S.; Carbonell, J.G.; Mitchell, T.M. *Machine Learning: An Artificial Intelligence Approach*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
10. Haenlein, M.; Kaplan, A. A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *Calif. Manag. Rev.* **2019**, *61*, 5–14. [CrossRef]
11. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
12. Albaşkara, M.; Gerçekcioğlu, E. A Taguchi-Based and Data-Driven Assessment of Surface Roughness and Wettability in FDM-Printed Polymers. *Micromachines* **2026**, *17*, 322. [CrossRef] [PubMed]
13. Mtsweni, S.; Bakare, B.F.; Rathilal, S. Artificial Neural Network-Based Prediction of Clogging Duration to Support Backwashing Requirement in a Horizontal Roughing Filter: Enhancing Maintenance Efficiency. *Water* **2025**, *17*, 2319. [CrossRef]
14. Plati, C.; Armeni, A.; Kyriakou, C.; Asoniti, D. AI for Predicting Pavement Roughness in Road Monitoring and Maintenance. *Infrastructures* **2025**, *10*, 157. [CrossRef]
15. Haykin, S. *Neural Networks and Learning Machines*, 3rd ed.; Pearson: Upper Saddle River, NJ, USA, 2009.
16. Balonji, S.; Tartibu, L.K.; Okokpujie, I.P. Surface Roughness Prediction in CNC Milling of Al6061 Using Soft Computing Techniques. *Appl. Sci.* **2023**, *13*, 4147. [CrossRef]
17. El-Mounayri, H.; Briceno, J.F.; Gadallah, M. A New Artificial Neural Network Approach to Modeling Ball-End Milling. *Int. J. Adv. Manuf. Technol.* **2010**, *47*, 527–534.

18. Freiesleben, T.; Grote, T. Beyond Generalization: A Theory of Robustness in Machine Learning. *Synthese* **2023**, *202*, 109–134. [CrossRef]
19. Fischer, M.M. Neural Networks: A General Framework for Non-Linear Function Approximation. *Trans. GIS* **2006**, *10*, 521–533. [CrossRef]
20. Li, B.; Kang, Y.; Hao, P.; Zhu, W.; He, P.; Zhen, H.; Xu, D.; Bai, K.; Liu, Y.; Wang, Y.; et al. Prediction of Three-Dimensional In Situ Stress in Deep Coal Rocks Considering Heterogeneity and Physical Information. *Processes* **2026**, *14*, 1535. [CrossRef]
21. Zhang, Y.; Xiong, W.; Liu, J.; Yuan, X.; Lu, Z.; Zheng, F. A Data-Driven Method for Constructing Planning Evaluation Indicators for Emerging Distribution Networks. *Energies* **2026**, *19*, 2310. [CrossRef]
22. Wu, L.; Meng, X.; Ren, Y.; Ding, Y.; Sun, L.; Li, S.; Feng, H. Research on Bidirectional Prediction Model Between Drying Process Parameters and Quality of *Fritillaria ussuriensis* Maxim Based on WOA-PSO-RF. *Appl. Sci.* **2026**, *16*, 4773. [CrossRef]
23. Matzavela, V.; Alepis, E. Decision tree learning through a predictive model for student academic performance in intelligent m-learning environments. *Comput. Educ. Artif. Intell.* **2021**, *2*, 100035. [CrossRef]
24. Balasuadhakar, A.; Thirumalai Kumaran, S.; Uthayakumar, M. Machine Learning Prediction of Surface Roughness in Sustainable Machining of AISI H11 Tool Steel. *Smart Mater. Manuf.* **2025**, *3*, 100075. [CrossRef]
25. Avuçlu, E. A New Data Augmentation Method to Use in Machine Learning Algorithms Using Statistical Measurements. *Measurement* **2021**, *180*, 109577. [CrossRef]
26. Mahdi, M.; Matthias, R.; Bill, L.; Christos, L. Stable Diffusion-based Data Augmentation for Federated Learning with Non-IID Data. *arXiv* **2024**, arXiv:2405.07925.
27. Pawanr, S.; Gupta, K. Analysis of Surface Roughness and Machine Learning-Based Modeling in Dry Turning of Super Duplex Stainless Steel Using Textured Tools. *Technologies* **2025**, *13*, 243. [CrossRef]
28. Repiso-López, I.; Martín-Béjar, S.; Trujillo-Vilches, F.J.; Sevilla-Hurtado, L. Optimization of ANN Used to Predict the Surface Roughness in Dry Turning of UNS A97075 Aeronautical Aluminum Alloy. *Acta Tech. Napoca.* **2023**, *66*, 317–326.
29. Vardhan, M.V.; Sankaraiah, G.; Yohan, M. Prediction of Surface Roughness and Material Removal Rate for Machining of P20 Steel in CNC Milling Using Artificial Neural Networks. *Mater. Today Proc.* **2018**, *5*, 18376–18382. [CrossRef]