machines

MDPI

Article

Data-Driven Tool Condition Monitoring Platform for Grooving Operations with Chip Control Issues

Iñigo Llanos *, Anne Sanz, Irantzu Ermina, Ainhoa Robles and Jokin Munoa *

Ideko Research Centre, Arriaga Industrialdea 2, E-20870 Elgoibar, Spain
* Correspondence: illanos@ideko.es (I.L.); jmunoa@ideko.es (J.M.); Tel.: +34-943748000 (I.L. & J.M.)

Abstract

Accurate Tool Condition Monitoring (TCM) is key to maximizing tool life and minimizing downtime in machining. However, real-time tool wear assessment remains challenging, particularly in turning, due to non-periodic signals and chip entanglement issues. This study frames tool wear prediction in grooving operations as a supervised regression problem using machining monitoring data as inputs. Multi-source sensor data were acquired and processed using an edge platform during groove machining to train six machine learning regression models. Also, a feature and sensor importance analysis was performed to identify opportunities to simplify instrumentation and reduce industrial implementation. The results demonstrate that Random Forests and XGBoost consistently achieve high accuracy ($R^2 > 0.9$), effectively handling the non-linear, noisy nature of turning data. Based on these results, a final ensemble model combining both approaches was defined and deployed at the industrial edge as part of a final condition tool monitoring system. To this end, the ensemble model was encapsulated within a Docker-based container and was executed on an edge computing platform, enabling reliable, portable, and scalable real-time deployment for TCM.

Keywords: tool condition monitoring; tool wear prediction; grooving operations; machine learning in manufacturing; edge computing; chip entanglement effects

1. Introduction

The identification of the tool state is key to enabling optimized tool management during machining operations. Such tool control can allow the most to be got out of cutting tools while avoiding unexpected stops and part scraping due to uncontrolled tool failure, thus minimizing manufacturing costs and maximizing productivity. Nevertheless, such strategies are scarcely applied in industry due to the difficulty in accurately identifying the actual state of the cutting tools. The typical strategy applied is the replacement of the cutting tools after a given time or number of machined parts. This strategy usually leads to lightly used tools and, more importantly, it still does not avoid the appearance of part quality issues or even tool breakage.

Tool Condition Monitoring (TCM) has been a recurrent research topic in bibliography since the 1980s [1]. In these kinds of systems, the primary source of information regarding tool condition consists of process variables associated with tool wear, commonly referred to as tool wear indicators. Among these, the most frequently utilized indicators include those derived from the cutting process, such as power, torque, vibrations, and Acoustic Emissions (AE), which are typically acquired through different sensors [2].

Check for updates

Academic Editors: Danko Brezak, Tomislav Staroveški and Miho Klaić

Received: 30 May 2026
Revised: 14 July 2026
Accepted: 15 July 2026
Published: 17 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Most TCM research focuses on rotary machining processes such as milling and drilling [3,4]. For these cases, the harmonic characteristics of the acquired signals facilitate the extraction of predictive features for assessing tool condition. In contrast, turning processes present unique challenges: the signals often lack repetitive patterns, and continuous chip formation can lead to chip entanglement. These factors introduce significant noise into the acquired signals, negatively impacting the accuracy and robustness of TCM system outputs.

Recent studies have explored advanced methods for TCM in turning operations. For example, Sadhukhan et al. [5] used an unscented Kalman filter (UKF) to estimate flank wear during Inconel 718 turning, demonstrating superior accuracy compared to the extended Kalman filter (EKF). Similarly, Çakir et al. [6] investigated tool wear in ceramic and carbide inserts used in grooving operations on Inconel 718 and Ti6Al4V alloys, respectively. Their comparison of traditional wear measurement techniques with the electromechanical impedance (EMI) method highlighted EMI's effectiveness in detecting wear levels and estimating tool life, emphasizing the critical importance of sensor placement. Bhuiyan et al. [7] analyzed AE and vibration signals to examine the relationships among tool wear, chip formation, and surface roughness. Their findings indicated that AE is highly sensitive to tool wear while vibration components correlate with surface roughness and chip formation. In another study, Denkena et al. [8] investigated the effectiveness of an online process monitoring system, using feature selection techniques to determine the sensitivity of various indicators to wear and achieving successful wear detection in seven out of eight experimental scenarios.

In recent decades, a significant amount of research has explored the use of Machine Learning (ML) algorithms for tool wear monitoring, as they enable effective processing of sensor-derived features and the modeling of relationships with tool wear. ANN-based approaches have demonstrated accurate wear prediction with uncertainty estimation [9], while deep learning methods using low-cost smartphone sensors have improved prediction reliability under varying machining conditions [10]. Multi-source data fusion combined with Gaussian Process Regression has enabled highly accurate real-time monitoring, outperforming state-of-the-art models such as LSTM and XGBoost [11]. Physics-Informed Machine Learning frameworks have further reduced experimental requirements while maintaining high predictive accuracy [12]. Additionally, sensor fusion strategies based on force and vibration measurements, coupled with Random Forest and SVM models, have achieved excellent performance for tool condition monitoring in smart manufacturing environments [13].

When dealing with the implementation of a TCM system, both regression and classification approaches are proposed in the literature. However, since tool wear evolves as a continuous physical process and the flank wear width is measured as a continuous variable, a regression approach is generally more appropriate than classification. Regression models preserve the full information contained in wear measurements and allow the actual wear level to be estimated, whereas classification requires discretizing the wear into predefined categories, leading to information loss [14]. Thus, tool wear prediction is undertaken in the present work as a supervised regression problem based on monitoring input from machine signals and external sensors.

From a methodological perspective, existing TCM approaches for turning can be grouped into model-based methods, sensor-based monitoring techniques, and machine-learning approaches. Model-based and sensor-based solutions can provide accurate wear estimation under stable operating conditions but are often sensitive to modeling assumptions, sensor placement, and process disturbances [5–8]. More recently, ML-based approaches have achieved high prediction accuracy through advanced learning frameworks

and sensor fusion strategies [9–13]. However, despite their methodological differences, these studies are generally validated under controlled laboratory conditions and primarily focus on predictive performance. To the best of the authors' knowledge, the influence of severe chip-entanglement phenomena during grooving under different operating conditions has not been systematically investigated. Furthermore, industrial deployment and post-deployment validation of TCM models receive limited attention in the literature. These gaps are particularly relevant because machine-learning models are susceptible to concept drift when operating conditions differ from those observed during training [15].

The present work differs from previous approaches by focusing on supervised regression models trained with data acquired under industrial grooving conditions characterized by recurrent chip-entanglement disturbances. Furthermore, the selected model is deployed on an edge device and subsequently validated, allowing the assessment of robustness under realistic industrial variability. Accordingly, the main contributions of this work are threefold: (i) the comparative assessment of six supervised regression algorithms for continuous tool-wear estimation under industrial grooving conditions; (ii) the evaluation of model robustness and generalization capability under different operating conditions; and (iii) the deployment of the selected ensemble model on an edge-computing platform, addressing practical challenges associated with concept drift and demonstrating its applicability in production-oriented TCM environments.

To achieve these objectives, experimental grooving tests were performed to acquire monitoring signals under diverse cutting conditions. A dataset was generated through a dedicated data-curation process and used to train six supervised regression models: Linear Regression (LR), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbor Regressor (KNR), Gaussian Process Regressor (GPR), and Extreme Gradient Boosting (XGBoost) Regressor. The obtained results identify RF and XGBoost as the most suitable algorithms for the considered task. Finally, an ensemble combining both models is containerized using Docker and deployed on an edge device, providing the basis for an industrialized TCM platform.

2. Materials and Methods

2.1. Experimental Setup

Experimental grooving tests were performed on rolled rings of ST-52 low-carbon steel with 120 HB hardness with an original size of Ø900 × Ø600 × 200 mm. This material was selected due to the high ductility related to its ferritic microstructure, which makes it difficult for chip breakage to occur and thus produces long continuous chips during grooving operations. The cutting tools employed were N123H2-0400-0004-GF 1125 inserts coated with a PVD TiAlN coating and C2R-QS25-LH25CB toolholder, both from Sandvik Coromant (Sandviken Sweden). The selected insert had 4 mm width and lateral nose radiuses of 0.4 mm, a local rake angle of 6° and a clearance angle of 7°. The grooving tests were performed on a Soraluce (Bergara, Spain) FMT-4000 multitasking machine with a Heidenhain (Traunreut, Germany) TNC 640 numerical control. The test rings were clamped through 4 jaws on the vertical turning table, which provided the cutting speed, and the tools were clamped to the machine head, which provided the feed motion to the tool in the X machine axis. No cutting fluid was employed during the grooving tests. The outline for the test set-up can be seen in Figure 1.

To perform the grooving tests, preliminary cleaning of the inner surface of the rings was performed by cylinder turning. Then, internal grooving tests were performed at different axial locations on the part, leaving a 3 mm wall between adjacent grooves. After the entire inner surface of the rings was machined, it was cleaned again by cylindrical turning and subsequent grooving tests were performed on the new inner surface. This way,

the grooving tests covered by the present work were performed on different axial (Z) and radial (X) locations for the part. After preliminary tests, the cutting conditions were selected following the toolmaker's recommendations, trying to keep the cutting tool life between 10 and 30 min. Accordingly, a series of four tests (T1–T4) was established. The values selected for the cutting speed and the feed per turn can be seen in Table 1. Two repetitions were executed for each condition shown in Table 1 to include experimental variability for the results.

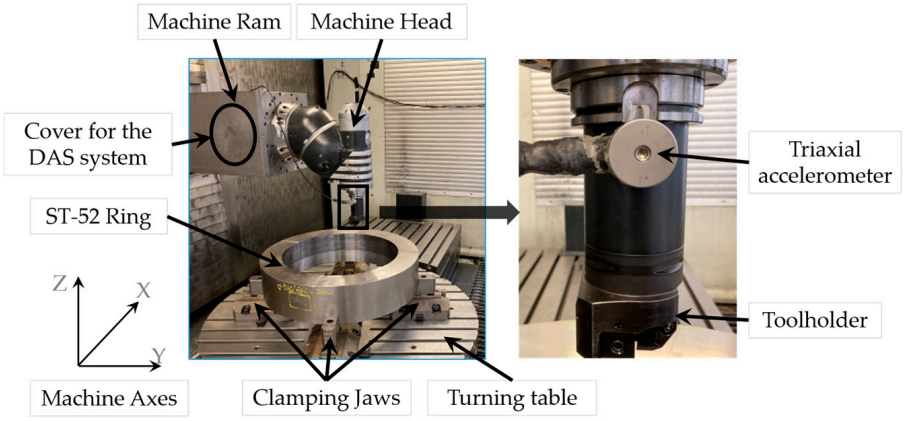


Figure 1. Outline of the test set-up and details of the toolholder and accelerometer location.

Table 1. Cutting conditions employed in each test.

Test	Cutting Speed—m/min	Feed per Turn—mm/rev
T1	350	0.05
T2	350	0.075
T3	300	0.05
T4	300	0.075

2.2. System Architecture

The proposed TCM framework combines a Heidenhain TNC640 NC control, an INGESYS IC3 controller from Ingeteam (Zamudio, Spain) for sensor integration, an industrial edge device for data acquisition and processing, and an external cloud-based PC for offline data analysis and artificial intelligence (AI) model development, supported by a fog platform for edge management and orchestration [16], as illustrated in Figure 2. This distributed setup enables the integrated acquisition, synchronization, and exploitation of high-frequency internal Computerized Numerical Control (CNC) signals and external sensor data generated during machining experiments. Data are captured and preprocessed at the edge level, further processed on the external server for feature extraction and model training, and subsequently returned to the edge for on-site inference, following edge architectures reported in the literature [17].

During the execution of the tests described in Section 2.1, monitoring signals were continuously acquired. On the one hand, several internal signals were obtained directly from the Heidenhain control using internal TNCscope files. In particular, nominal current values from the spindle and machine axes (S, X, Y, and Z) were recorded, together with the mechanical power associated with the spindle and the X axis, as these are the only elements contributing motion to the cutting process. These current and power signals were provided by the internal TNCscope with a high temporal resolution, featuring a sampling interval of approximately 0.6 ms, enabling the capture of dynamic phenomena during machining. In addition, actual position data for all machine axes and the spindle were acquired, along with speed signals for the spindle and the X axis and the actual machining

feed rate. Position and speed signals were made available by the NC with a sampling interval of 3 ms, which is sufficient for detailed kinematic analysis of the process.

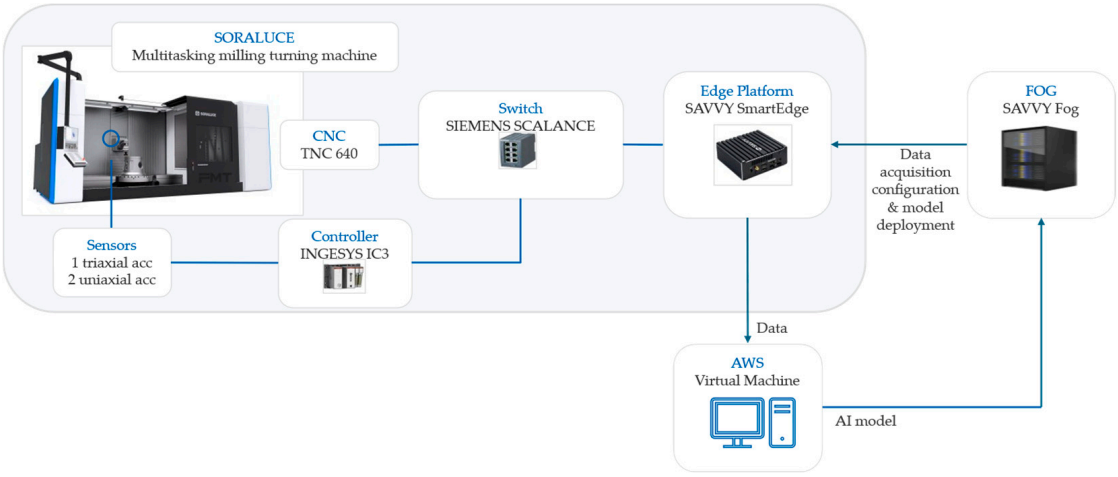


Figure 2. Schematic for the TCM system architecture.

On the other hand, to complement the internal CNC data with direct physical measurements, vibration signals were acquired using external accelerometers. A triaxial accelerometer Model 604B31 from PCB (Depew, NY, USA) was magnetically mounted behind the toolholder to acquire the accelerations generated during the cutting tests. Furthermore, the machine is equipped with a DAS (Dynamis Active Stabilizer) system, which incorporates two unidirectional accelerometers (PCB Model 602D01) mounted on the machine ram to measure accelerations along the X and Y directions. In total, five acceleration signals were acquired and managed by the INGESYS IC3 controller, which handles both data acquisition and preliminary signal processing before transmitting the data to the edge device. The location of the triaxial accelerometer and the DAS system on the machine ram can be seen in Figure 1.

The edge device serves as an integration layer that aggregates and synchronizes data streams coming from the Heidenhain TNC640 CNC and the INGESYS IC3 controller, enabling local storage for subsequent offline analysis or transmission to higher-level systems.

The virtual machine, deployed as a cloud-based server, is responsible for advanced offline data processing and the execution of the ML workflow. In this work, an AWS (Amazon Web Services) c4.4xlarge instance was employed due to its high computational capabilities, providing 16 vCPUs and 30.5 GB of RAM, which are well suited for compute-intensive data processing and model training tasks. Data acquired at the edge platform are transferred to this external server, where feature extraction and dataset preparation are performed to generate suitable inputs for the ML models. On this external PC, six ML models are trained with the objective of identifying the most suitable solution. The selected models are exported as a serialized .pkl file and subsequently deployed on the edge platform as an ensemble model for on-site inference during machining operations.

Finally, a fog platform is used to support the operation and orchestration of the edge device. The fog platform enables centralized configuration and supervision of the edge devices, including the definition and update of signal acquisition parameters. It also enables standardized data exchange between nodes through RESTful APIs, supporting an efficient and interoperable data flow. In addition, the fog platform supports the distribution and deployment of containerized ML models by managing Docker containers on the edge devices.

2.3. Tool Wear Characterization

After the execution of each groove, images from rake and flank faces of cutting tools were obtained through a Dino-Lite Edge digital microscope. Thanks to those images, flank wear measurements were performed in the three flank faces (frontal, left, right) of the cutting tools. Figure 3 shows the examples of the images and measurements performed on a cutting tool after a grooving test.

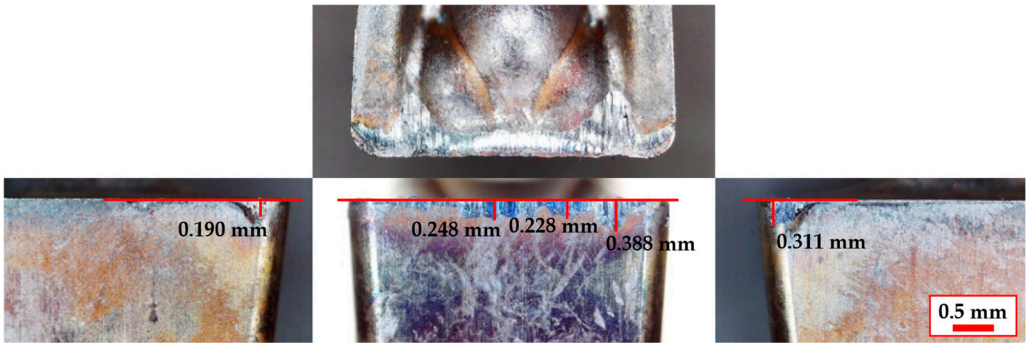


Figure 3. Example of images taken from the cutting tools and measurements performed on the flank faces.

For each cutting condition, sequential 4 mm (tool width) grooves were machined with a 15 mm depth. The tests were stopped after one of the flank wear measurements reached a value beyond 0.3 mm, the classical criterium for stopping a cutting test in turning operations according to the standard ISO 3685 [18]. Nevertheless, if the tool wear was close to this limit in one flank, while the other two remained in a low wear state, the test continued until the flank wear in one area was severe.

In the case of the second repetition of T4 (T4 R2), a wear value of 0.357 mm was obtained on the right flank after 4 grooves. However, since the wear on both the front and left flanks was very low, two more grooves were machined until the wear value at the right flank was severe.

2.4. Data Curation

As commented in the previous section, high-frequency data from different sources were obtained for each cutting test: raw monitoring data were gathered from the machine control in csv format, while acceleration data were acquired from INGESYS IC3 in binary format. These two acquisitions were not synchronized, so a preliminary data treatment step of the gathered data was required for synchronization. Thus, an algorithm was coded to detect the initial contact on the current from the machine spindle and the acceleration from the DAS system in the X direction. After this start was detected, the timestamp from both acquisitions was modified so this start point would be offset to the 0 s value. The finish of a grooving test was identified when the position of the X axis reached the maximum value of that test, i.e., the end of the groove. After the identification of these two events and the offset of the time stamps, the signals were cropped to start 3 s before the initial contact and end 3 s after the end of the groove. Figure 4 shows an example for the signals from the acceleration in the Z direction (Acc Z), Y axis current (Inom Y) and X axis positioning (Pos X) for a test after this preliminary data treatment.

The monitored data were further curated to work as input to the regression models to be presented in the next section. Instead of feeding the whole data to the regression models, it was decided to generate statistical parameters for different time windows obtained from the monitored data. On the one hand, the variable sampling rate from the different data sources would create the need for a time interpolation of some data to generate the same

regular sampling on all the monitored signals. On the other hand, this is a common practice when applying ML techniques to TCM, as it reduces the input noise and the computational burden of the regression models [19].

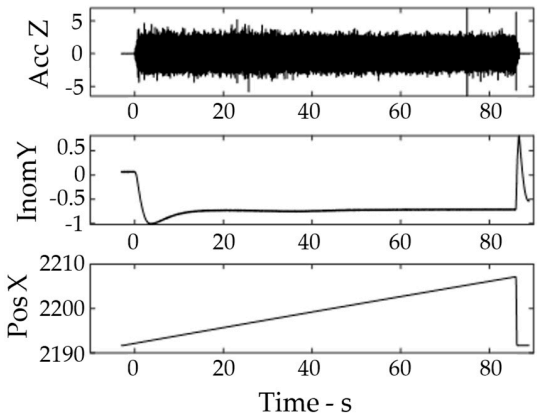


Figure 4. Example of curated monitoring signals for the machining of a groove.

The temporal data from the monitored signals were first divided into different time windows. In order to avoid non-stationary machining states, these time windows started 5 s after the grooving start and the last time window from a test would need to finish at least 1 s before the grooving end. The depiction of the identification of the time windows from a test can be seen in Figure 5 for the case of an acceleration signal in the Z direction. Notice that the time windows employed in the Figure are 10 s in length for ease of representation, while the time windows employed during the analysis were 1 s in length.

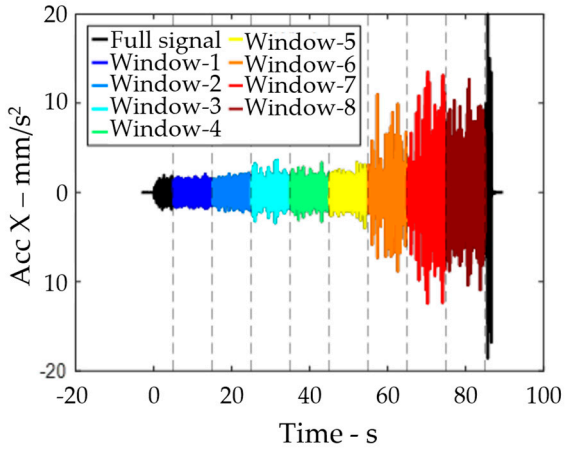


Figure 5. Example of the division of the temporal data from a test into sequential time windows. Window size in the Figure: 10 s.

Several statistics were computed from each of these windows for all the acquired signals (Table 2). After the calculation of the statistics for each time window, the calculated values were gathered sequentially to group all the windows extracted for all the passes performed for a given test. Figure 6 shows the evolution of the different statistics computed for two signals along a whole cutting test. The dashed lines indicate the end of each machined groove.

It can be observed that the statistics behave differently for each signal. In the case of an acceleration signal, the RMS, mean deviation and standard deviation statistics show an increasing trend along all the windows computed for a test. Nevertheless, some zones can be observed where these statistics have significantly increased values in the middle of the tests. These zones correspond to the appearance of chip entanglement issues during the tests.

Table 2. Statistics obtained from the monitoring data.

Statistic	Formula
Mean	$\mu = \frac{1}{n}\sum_{i}^{n} x_i$
Min	$x_{min} = \min(x)$
Max	$x_{max} = \max(x)$
Peak-to-Peak	$x_{p2p} = \max(x) - \min(x)$
RMS value	$x_{rms} = \sqrt{\frac{1}{n}\sum_{i}^{n} x_i^2}$
Mean deviation	$x_{dev} = \frac{1}{n}\sum_{i}^{n} \lvert x_i - \mu \rvert$
Standard deviation	$\sigma = \sqrt{\frac{1}{n}\sum_{i}^{n} (x_i - \mu)^2}$
Variance	$\sigma^2 = \frac{1}{n}\sum_{i}^{n} (x_i - \mu)^2$
Skewness	$sk = \frac{1}{n}\sum_{i}^{n} \left(\frac{x_i - \mu}{\sigma}\right)^3$
Kurtosis	$ku = \frac{1}{n}\sum_{i}^{n} \left(\frac{x_i - \mu}{\sigma}\right)^4$

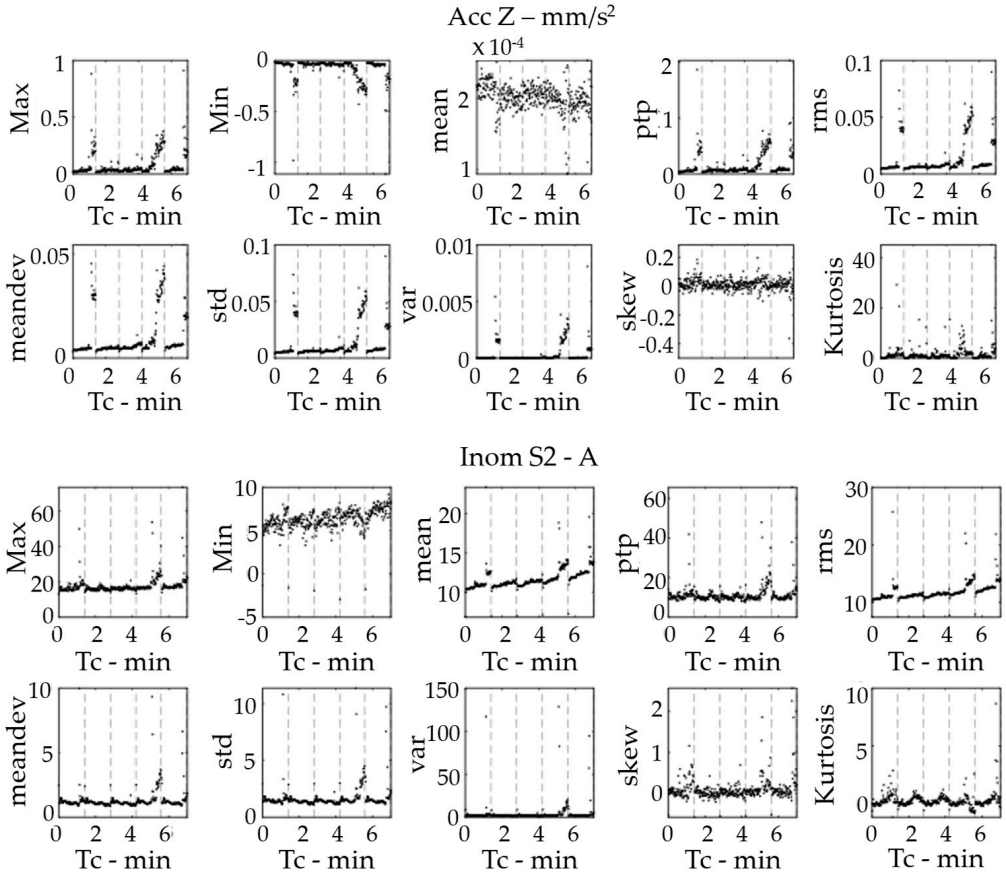


Figure 6. Evolution for calculated statistics along a whole test for acceleration in the Z direction and nominal current from the spindle.

Regarding the spindle current signal, a similar trend to the one observed for some statistics from the acceleration is obtained for the mean and rms values. In these cases, deviations generated by the chip entanglement issues can also be seen, but of lower magnitude than those for the statistics obtained from the acceleration. In the case of the peak-to-peak, mean deviation and standard deviation, they show a flat baseline along a test, with the appearance of high values at the zones where chip entanglement occurs. This variable behavior observed for the different statistics can indicate that an expert system could be able to identify the appearance of chip entanglement and disregard its effect for the identification of the tool state.

In addition to the monitoring data gathered from both the machine control and the INGESYS IC3 controller, the tool wear data obtained from the tests needed to be further processed to provide an input to the regression models from the next section. These values were obtained after each groove from a given test. Thus, to generate values for all the time windows extracted from the monitoring data, linear interpolation of the tool flank wear values was performed along the execution of each test. The linear interpolation may smooth the evolution of the flank wear between consecutive measurements. However, flank wear usually progresses gradually during machining. For this reason, linear interpolation was used as a practical way to assign wear values to all the time windows while preserving the overall wear trend. As a result, while some local fluctuations in the target variable may have been reduced, the impact on the regression task is expected to be limited, and the generated labels remain aligned with the physical wear process.

Finally, in the case of the flank wear values for the first groove of a test, two options were evaluated: interpolating from zero to the wear after the first pass or keeping this last value along the whole first pass. In general, the initial wear of turning tools is not linear; rather, it follows very a steep initial curve and then stabilizes. With this in mind, it was decided to use the first measured value of tool wear for all windows in the first pass, rather than interpolating it from zero. An example of this interpolation of the flank wear values to all the time windows from a test can be seen in Figure 7.

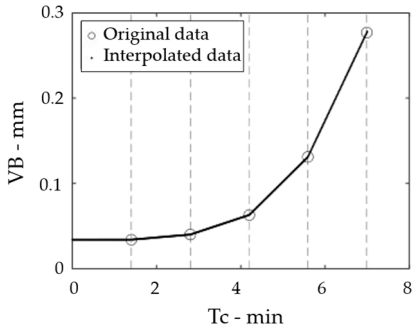


Figure 7. Flank wear data interpolation to all the time windows from a test.

Following the data curation process, a dataset was assembled comprising statistical features extracted from each monitoring signal for each time window, alongside the corresponding tool wear measurements. Specifically, the dataset included 10 statistical features derived from 11 distinct signals across 6356 time windows. For each time window, three flank wear values were provided as target variables.

2.5. Tool Wear Regression Analysis Through ML Techniques

To address the task of tool wear prediction based on the indicators extracted from acquired signals, this study employs a diverse set of six supervised regression algorithms: Linear Regression (LR), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbors Regressor (KNR), Gaussian Process Regressor (GPR) and Extreme Gradient Boosting (XGBoost)

Regressor. The models were implemented through Python 3.8 using the widely adopted Python packages scikit-learn 1.3.2 and xgboost 2.1.4.

LR assumes a linear relationship between input variables and target data, making it effective when such a relationship exists, though it struggles with non-linear patterns and outliers. In contrast, DTs are non-linear models that split data based on feature values to capture complex relationships, while an RF improves accuracy and stability by combining the predictions of multiple DTs. KNR is a flexible, non-parametric, memory-based model that predicts by averaging the target values of the nearest neighbors in the feature space, but it can be computationally expensive and sensitive to neighbor selection and distance metrics. GPR, a probabilistic, non-parametric model, incorporates prior knowledge and estimates uncertainty, making it well-suited for regression tasks where confidence is critical, but it may be computationally expensive for large or high-dimensional datasets. Lastly, XGBoost enhances model performance by combining multiple weak learners through a boosting process to minimize prediction errors [20,21].

The selection of these algorithms was motivated by the need to evaluate a broad spectrum of regression techniques, ranging from simple linear models to more complex, non-linear, and ensemble-based approaches. This variety allows for a robust and comparative analysis of model performance in predicting tool wear, a task characterized by potential non-linearities, noise, and varying data distributions. By applying and comparing these six models, the study aims to identify the most effective regression strategies for this specific industrial application and to provide insight into the relative strengths and limitations of each method within the context of TCM.

The analysis was carried out using three different input datasets, summarized in Table 3. Each dataset consists of features extracted from signals obtained from different data sources. In particular, DS1 is the largest dataset, containing the complete set of acquired data, while each of the subsequent datasets excludes one or more data sources. Each dataset was employed to independently model the three flank wear indicators—front, right, and left—treating each as a distinct target variable within the predictive framework.

Table 3. Datasets used for the analysis.

Dataset	Data Sources	Nº Features
DS1	Triaxial Accelerometer DAS Accelerometers TCN Scope (current, power)	110
DS2	DAS Accelerometers TCN Scope (current, power)	80
DS3	TCN Scope (current, power)	60

In each case, the dataset was randomly split for generating train-test sets, without separating the data by test, ensuring that the training and testing samples were mixed and not biased toward specific tests [22]. In particular, 80% of the data were considered for model training, and 20% for testing.

Grid search with K-fold cross-validation (GridSearchCV) is applied during model training to determine the optimal set of hyperparameters for each one. This technique involves partitioning the dataset into K subsets, ensuring that each instance is used for training and testing across different iterations, with no overlap between training and testing data within each fold. In this study, K = 5 was selected, as it is a widely accepted choice that offers a good balance between computational cost and validation reliability, particularly when data is limited [23,24]. Linear Regression was excluded from this process as it does not involve tunable hyperparameters.

3. Results

3.1. Grooving Tests Results

Table 4 shows an overview of the results concerning the cutting time and number of passes for each test, while Figure 8 shows the evolution of the three flank wear measurements performed for each test and repetition. The tests with most aggressive cutting conditions (T2) resulted in the lowest cutting times while, on the other hand, the tests performed for the most conservative conditions (T3) yielded the highest cutting times. The obtained results show a good agreement with the results reported in the bibliography, as the increase in both the cutting speed and the feed per turn generate lower cutting time results. Also, the effect of the cutting speed is significantly higher than the feed per turn. An increase of 17% on the cutting speed value (from 300 to 350 m/min) generates an average decrease of 44% in the cutting time, while an increase of 50% on the feed per turn (from 0.05 to 0.075) generates an average reduction of 38% on the cutting time. It is also noteworthy that most of the repetitions yield similar tool wear evolutions while some of them (front flank for T1 and T2 and right flank wear for T4) show significant deviations along the execution of the repetitions for a given test.

Table 4. Cutting time and number of performed grooves for each repetition in the performed tests.

Test	Repetition	Cutting Time	Nº of Grooves
T1	R1	11.7 min	7
	R2	11.5 min	6
T2	R1	8.1 min	7
	R2	7 min	5
T3	R1	24.5 min	12
	R2	18 min	8
T4	R1	15.7 min	11
	R2	9.8 min	6

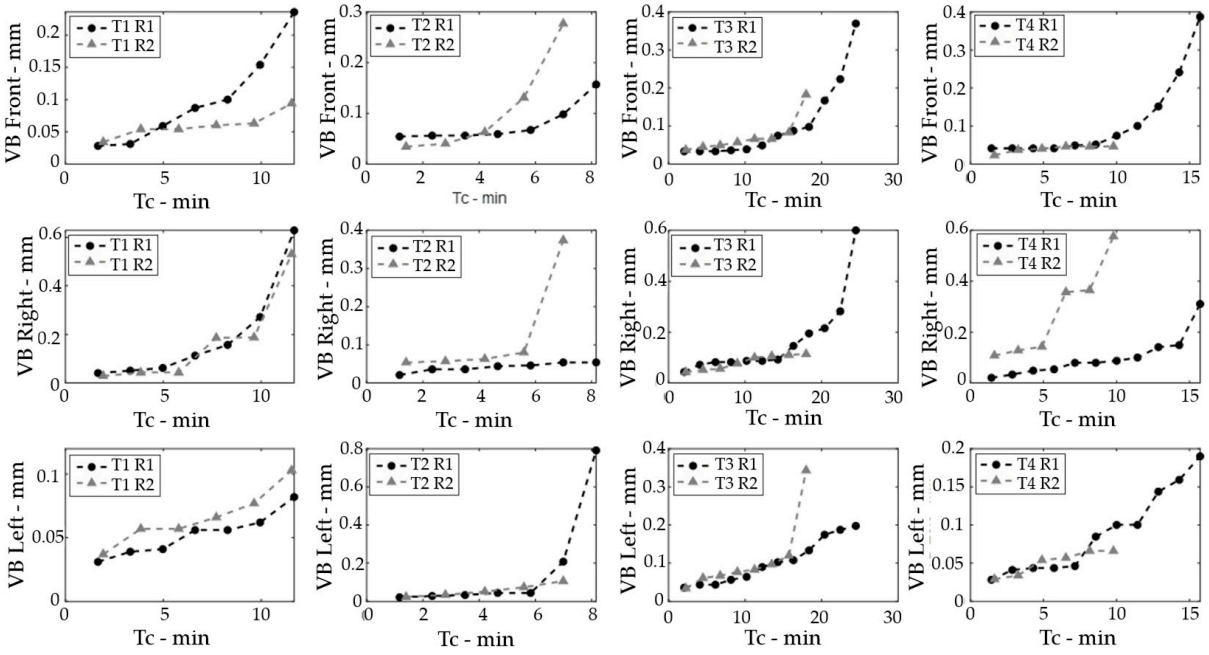


Figure 8. Tool flank wear measurements along the 2 repetitions performed for 4 cutting tests.

Concerning chip control during the tests, it must be noted that, due to the ferritic nature of the ST-52 steel, no chip breaking took place during any of the grooving tests

performed. This generated significantly large chip nests which would get entangled both on the toolholder and with the part clamping jaws. Examples of such chip entanglement issues can be seen in Figure 9. Due to the rotation of the turning table, the chip nests trapped at the jaws would intermittently impact against static elements of the machine. Also, the chips entangled at the toolholder would generate a continuous rubbing against the part. Such issues generated significant variations on the monitored signals. An example of this effect can be seen in Figure 10, where accelerations in the X direction (Acc X) and machine Y axis current (Inom Y) can be seen for the cases of a grooving test without chip control issues and another test with heavy chip entanglement. As can be observed, the amplitude of the acceleration in the X axis increases significantly when chip entanglement occurs, while in the case of machine Y axis current, variations in its mean value as well as noise are generated.

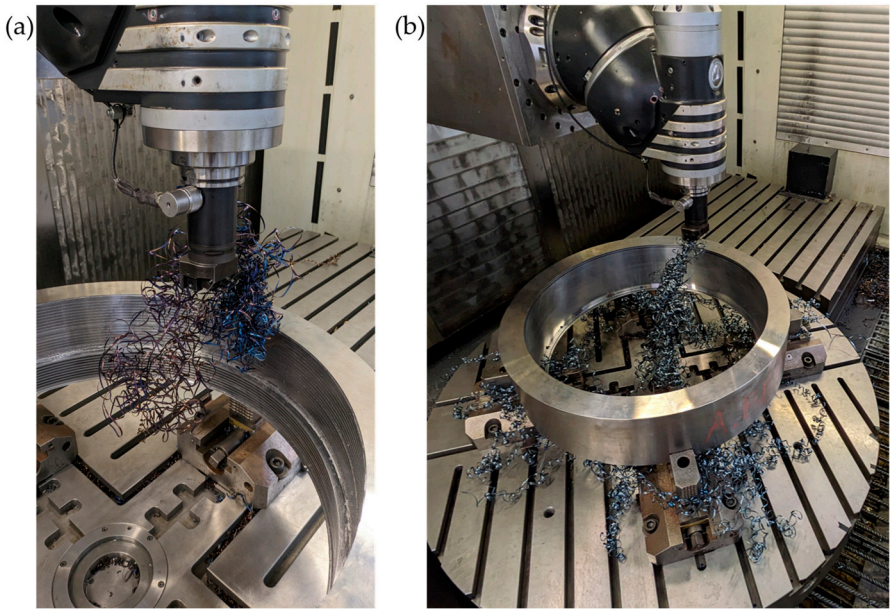


Figure 9. Examples of chip entanglement issues in (**a**) the toolholder, (**b**) both the toolholder and clamping jaws.

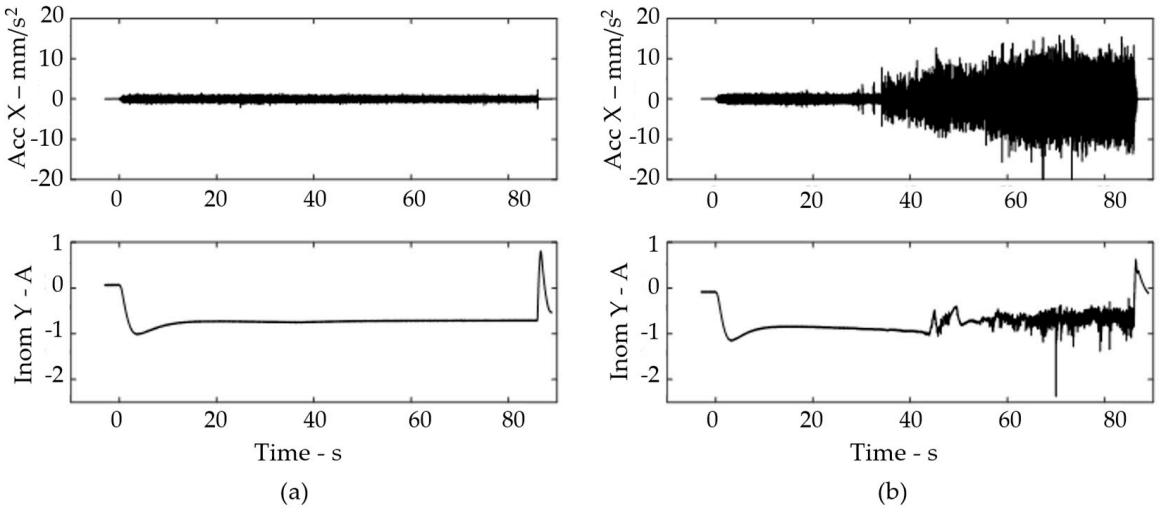


Figure 10. Example of accelerations and currents obtained for tests (**a**) without chip entanglement and (**b**) with severe chip entanglement.

3.2. ML Tool Wear Regression Results

In order to evaluate the performance of the algorithms, the following four metrics were chosen, since they are the most used in regression tasks [25]:

- Mean Absolute Error (MAE) computes the average of the absolute differences between the observed and predicted values.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|$$

- Root Mean Square Error (RMSE) is derived from Mean Square Error (MSE) by calculating its square root. MSE is the average squared distance between the observed and predicted values.

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

- Max Error (ME) indicates the maximum absolute error obtained from the predictions, i.e., the maximum observed vs. predicted difference.

$$\mathrm{ME} = \max\left(\left|y_i - \hat{y}_i\right|\right)$$

- R^2 Score, or coefficient of determination, is a measure of the proportion of variance in the dependent variable that can be explained from the independent variables.

$$R^2 = 1 - \frac{\sum_i \left(y_i - \hat{y}_i\right)^2}{\sum_i \left(y_i - \bar{y}_i\right)^2}$$

The best-performing algorithm is characterized by a high R^2 R^2 score and minimum error (MAE, RMSE and ME) values. An R^2 score of nearly 1 indicates that the model explains a vast majority of the variance in the target variable. Simultaneously, lower values for MAE and RMSE mean that model predictions are, on average, near to true ones, where RMSE gives higher weight to large errors. ME represents the largest single discrepancy between a prediction and an actual value.

Tables 5–7 show the MAE, RMSE, ME and R^2 scores obtained by each model when predicting the front, right and left flank wear for the three input datasets described in Table 3. It is noteworthy that, for each algorithm, the values for the metrics are approximately the same for the three datasets analyzed. This suggests that the features in DS1 and DS2, but not DS3, do not provide significant additional information for the prediction task, because their absence is not felt in the performance of the models. Therefore, it can be concluded that, based on features derived from current and power (NC control) signals, the variability of the target variable is explained as well as considering those features plus the ones derived from the signals extracted from DAS accelerometers and the triaxial accelerometer signals.

Looking at the results, it can be concluded that DT, RF, KNR and XGBoost algorithms yield the best performances for all three target variables and, therefore, are the top-performing algorithms in this case. These algorithms are capable of modeling complex, non-linear associations between the data: DT is capable of doing so since it can model non-linear relationships and hierarchical patterns through recursive partitioning without assuming any specific data distribution, while RF and XGBoost are ensemble methods that improve generalization and minimize overfitting by averaging the output of several decision trees. KNR gains from local averaging and as such is optimized for datasets having heterogeneous distribution.

LR and GPR algorithms have poorer performance. Linear Regression relies on a linear relationship between the inputs and the target variable, which is perhaps too simplistic to pick up on the correct patterns in tool wear. Gaussian Process Regression, though theoretically solid, is prone to scalability issues and is noisy and high-dimensional-data sensitive.

The four models, DT, RF, KNR and XGBoost, are the best-performing in all three flank wear regression tasks, each achieving an R^2 score higher than 0.9. It can be concluded that these algorithms offer strong and reliable capabilities across different wear types, consistently capturing the underlying patterns in the data with high accuracy.

Table 5. The performance of the six regression algorithms when predicting the front flank wear considering each dataset.

Algorithm	Dataset	MAE	RMSE	ME	R^2 Score
LR	DS1	0.0268	0.0375	0.2652	0.6300
	DS2	0.0268	0.0375	0.2652	0.6300
	DS3	0.0268	0.0375	0.2652	0.6300
DT	DS1	0.0019	0.0092	0.1732	0.9778
	DS2	0.0020	0.0100	0.2186	0.9738
	DS3	0.0019	0.0092	0.1732	0.9780
RF	DS1	0.0044	0.0097	0.1365	0.9752
	DS2	0.0043	0.0097	0.1340	0.9751
	DS3	0.0043	0.0097	0.1321	0.9751
KNR	DS1	0.0011	0.0061	0.1602	0.9904
	DS2	0.0011	0.0061	0.1602	0.9904
	DS3	0.0011	0.0061	0.1602	0.9904
GPR	DS1	0.0203	0.0434	0.3866	0.5051
	DS2	0.0203	0.0434	0.3866	0.5051
	DS3	0.0203	0.0434	0.3866	0.5051
XGBoost	DS1	0.0017	0.0055	0.1711	0.9920
	DS2	0.0017	0.0055	0.1711	0.9920
	DS3	0.0017	0.0055	0.1711	0.9920

Table 6. The performance of the six regression algorithms when predicting the right flank wear considering each dataset.

Algorithm	Dataset	MAE	RMSE	ME	R^2 Score
LR	DS1	0.0509	0.0715	0.3963	0.5869
	DS2	0.0509	0.0715	0.3963	0.5869
	DS3	0.0509	0.0715	0.3963	0.5869
DT	DS1	0.0084	0.0249	0.3477	0.9499
	DS2	0.0038	0.0201	0.5005	0.9674
	DS3	0.0068	0.0222	0.4075	0.9603
RF	DS1	0.0082	0.0187	0.2503	0.9719
	DS2	0.0081	0.0183	0.2444	0.9728
	DS3	0.0081	0.0184	0.2403	0.9726
KNR	DS1	0.0025	0.0125	0.2783	0.9875
	DS2	0.0025	0.0125	0.2783	0.9875
	DS3	0.0025	0.0125	0.2783	0.9875
GPR	DS1	0.0328	0.0744	0.6152	0.5541
	DS2	0.0328	0.0744	0.6152	0.5541
	DS3	0.0328	0.0744	0.6152	0.5541
XGBoost	DS1	0.0034	0.0126	0.2591	0.9872
	DS2	0.0034	0.0126	0.2591	0.9872
	DS3	0.0034	0.0126	0.2591	0.9872

Table 7. The performance of the six regression algorithms when predicting the left flank wear considering each dataset.

Algorithm	Dataset	MAE	RMSE	ME	R^2 Score
LR	DS1	0.0322	0.0548	0.5890	0.3515
	DS2	0.0322	0.0548	0.5890	0.3515
	DS3	0.0322	0.0548	0.5890	0.3515
DT	DS1	0.0028	0.0113	0.2556	0.9723
	DS2	0.0027	0.0106	0.2556	0.9760
	DS3	0.0021	0.0106	0.2572	0.9756
RF	DS1	0.0042	0.0095	0.1866	0.9807
	DS2	0.0042	0.0094	0.1838	0.9808
	DS3	0.0042	0.0094	0.1910	0.9809
KNR	DS1	0.0014	0.0081	0.2433	0.9859
	DS2	0.0014	0.0081	0.2433	0.9859
	DS3	0.0014	0.0081	0.2433	0.9859
GPR	DS1	0.0207	0.0429	0.6050	0.6036
	DS2	0.0207	0.0429	0.6050	0.6036
	DS3	0.0207	0.0429	0.6050	0.6036
XGBoost	DS1	0.0017	0.0051	0.1186	0.9943
	DS2	0.0017	0.0051	0.1186	0.9943
	DS3	0.0017	0.0051	0.1186	0.9943

3.3. Model Selection for TCM

In industrial ML applications, maintaining reliable performance over time is essential, as deployed models are exposed to evolving process conditions that may lead to model drift. A key factor influencing this degradation is overfitting, which reduces a model's ability to generalize beyond the training data. Therefore, controlling overfitting during model training is crucial to ensure robustness, limit susceptibility to drift, and sustain stable performance during on-site inference [26].

To further assess the tendency of the evaluated algorithms to overfit, MAE, RMSE and R^2 values were compared between the training and test sets across the K-Fold cross-validation process. Overfitting was quantified by calculating the difference between the metrics obtained from the training and test sets. For MAE and RMSE, the percentage difference was used, computed as the absolute difference divided by the training metric value. In contrast, for the R^2 score, since it is a bounded metric, the absolute difference between training and test values was used. Models with smaller differences were considered to have a greater capability for generalization to unseen data. Tables 8–10 show these calculations for the different models but only taking into account dataset 3, as the previous analysis indicated that the additional sensors in DS1 and DS2 did not provide significant information for the models.

As can be seen, no clear trend can be observed for the case of the R^2 score. In the case of MAE and RMSE metrics, the lowest variation between the training and test sets was obtained for the RF model for the three flank wears considered, while the DT and XGBoost models yielded higher values. Although KNR achieved high performance in the evaluation metrics, the train–test comparison revealed a larger generalization gap than that observed for the remaining algorithms, indicating that the model fitted the training data more closely than it generalized to unseen samples. This behavior is consistent with overfitting, so care should be taken with the results obtained with this model as it is likely to be highly overfitted.

Taking into account both the performance and overfitting metrics shown in Tables 5–10, RF and XGBoost were selected as the best-performing models. RF, while yielding slightly lower performance metrics, appears to have the best generalization capacity for unseen data. XGBoost, on the other hand, obtains the best values for the regression performance metrics, although it may produce higher errors with data outside the training data set.

Table 8. The difference between the metrics obtained from training and test sets during the CV process for each algorithm to predict the front flank wear considering DS3 features.

Algorithm	MAE	RMSE	R^2 Score
DT	4.21×10^{0}	1.84×10^{0}	6.77×10^{-2}
RF	1.66×10^{0}	1.37×10^{0}	6.32×10^{-2}
KNR	2.03×10^{7}	1.05×10^{7}	5.02×10^{-2}
XGBoost	8.09×10^{0}	1.27×10^{1}	4.06×10^{-2}

Table 9. The difference between the metrics obtained from training and test sets during the CV process for each algorithm to predict the right flank wear considering DS3 features.

Algorithm	MAE	RMSE	R^2 Score
DT	1.62×10^{0}	1.45×10^{0}	9.94×10^{-2}
RF	1.89×10^{0}	1.59×10^{0}	7.33×10^{-2}
KNR	1.07×10^{7}	5.85×10^{6}	6.27×10^{-2}
XGBoost	1.07×10^{1}	1.72×10^{1}	6.29×10^{-2}

Table 10. The difference between the metrics obtained from training and test sets during the CV process for each algorithm to predict the left flank wear considering DS3 features.

Algorithm	MAE	RMSE	R^2 Score
DT	4.38×10^{0}	2.87×10^{0}	1.33×10^{-1}
RF	1.50×10^{0}	1.04×10^{0}	5.96×10^{-2}
KNR	1.85×10^{7}	9.21×10^{6}	1.01×10^{-1}
XGBoost	7.82×10^{0}	1.11×10^{1}	3.95×10^{-2}

The regression predictions and original data for the three flank wear evolutions obtained in the two repetitions of each test can be seen in Figure 11 for the RF model, while Figure 12 shows the same results for the XGBoost model. Looking at these figures, it can be observed that the regression results produced by the RF model are more scattered than the ones produced by the XGBoost model. Also, the local average values for the RF predictions differ slightly from the ground truth in several zones, while the predictions from the XGBoost model are significantly closer. Still, significant deviations can be seen in the regression results, as indicated by the Maximum Error (ME) results from Tables 5–7.

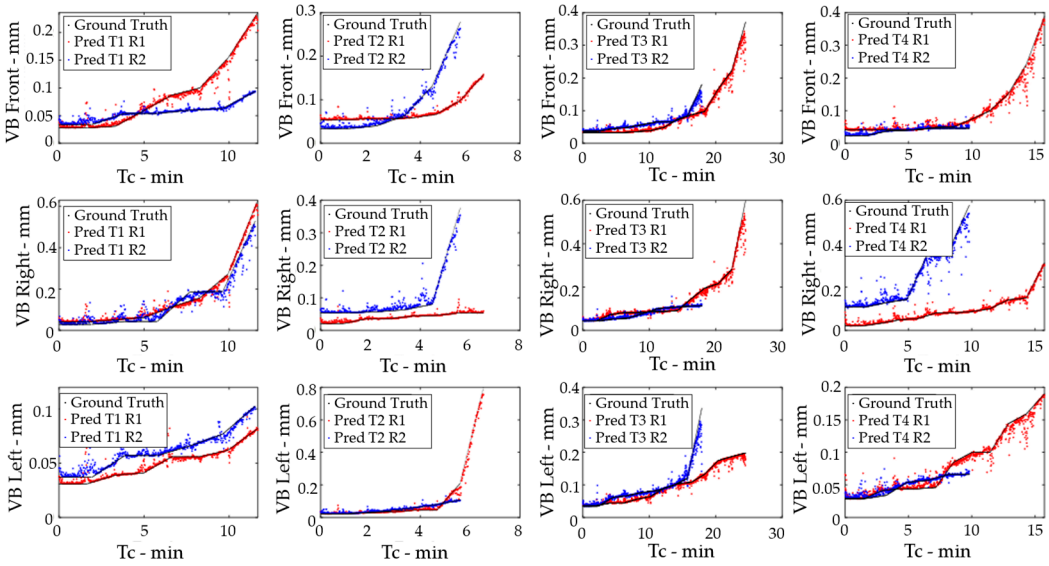


Figure 11. Regression results for the three flank wears considering Dataset 3 for the RF model.

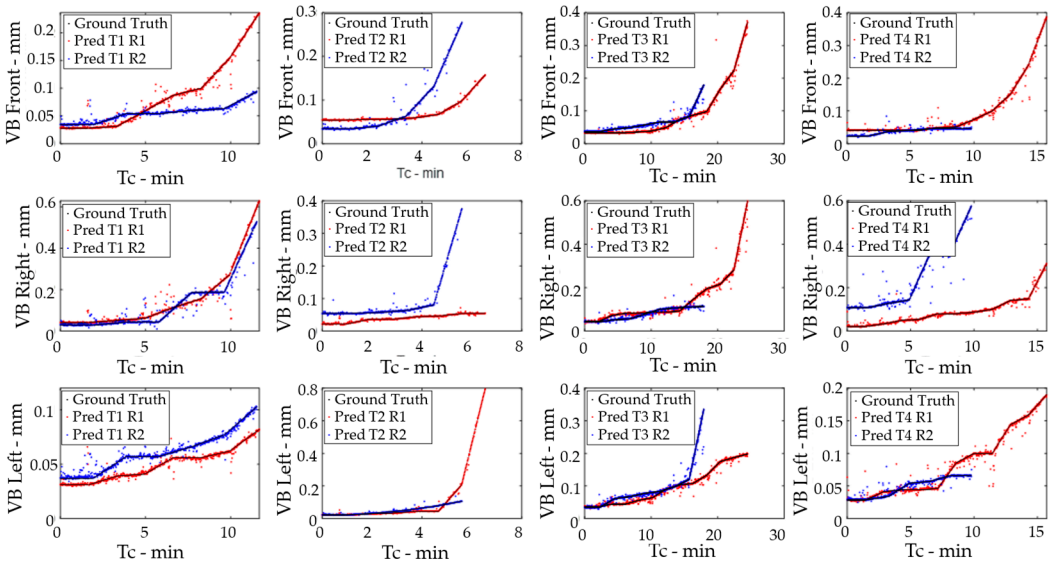


Figure 12. Regression results for the three flank wears considering Dataset 3 for the XGBoost model.

3.4. Containerized Deployment of the TCM Model

Based on the model selection results discussed in Section 3.3, where RF and XGBoost demonstrated the best trade-off between predictive performance and generalization capability, the final TCM solution is defined as an ensemble model that combines both approaches. The predictions of the two models are aggregated through an averaging strategy, yielding a single output that leverages their complementary strengths. This ensemble formulation enhances robustness and reduces sensitivity to individual model deviations under varying machining conditions.

Consequently, the trained models are serialized into a separate .pkl files and packaged within a Docker container together with the Python-based inference implementation. This container encapsulates the model artifacts, the Python code responsible for loading the serialized models and performing predictions, and all required software dependencies.

The generated Docker container is deployed on the edge platform via the fog layer, which acts as an intermediary for distributing containerized applications across the edge infrastructure. Through this mechanism, the TCM model can be efficiently transferred to the target edge device, where it is instantiated and executed locally.

Once deployed, the containerized TCM model operates continuously at the edge. Each time new data streams are received from the CNC, the incoming signals are processed and forwarded to the model, which performs on-site inference and generates a set of tool condition predictions. The inference results are then communicated back to the machine PLC, where they are used to trigger notifications or warnings at the CNC, enabling timely awareness of tool condition states during machining operations. This deployment strategy ensures low-latency response, reduces reliance on external connectivity, and supports robust and scalable TCM functionality directly at the edge.

The following results present the response of the proposed ensemble model during operation, analyzing the combined predictions of the Random Forest and XGBoost models under representative machining conditions. To evaluate the performance of the ensemble model for on-machine tool wear evaluation, an additional machining test was carried out. The selected cutting conditions were the ones for T3, as it yielded the longest cutting times in the previous tests. After the execution of each grooving pass from this test, NC data are ingested by the edge platform. The inference service, running in a Docker container, consumes the data via a REST API, preprocesses them, and subsequently executes the

ensemble model for tool wear evaluation. The results obtained through this test are shown in Figure 13, while the performance metrics obtained are shown in Table 11.

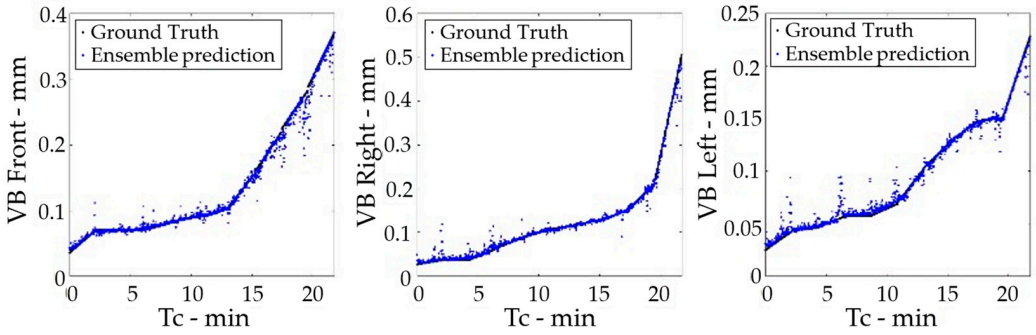


Figure 13. Regression results from the containerized model ensemble.

Table 11. Performance metrics for the containerized model ensemble.

Dataset	MAE	RMSE	ME	R^2 Score
VB Front	0.0048	0.0101	0.0912	0.9869
VB Right	0.0042	0.0094	0.1100	0.9905
VB Left	0.0030	0.0059	0.0522	0.9866

It can be observed that the containerized model ensemble can provide tool wear regression results very close to the ones obtained for the original data used for the model training. This shows the capability of the proposed containerized TCM model to provide accurate results when fed with unseen data.

4. Discussion

The present study evaluated the development and deployment of a data-driven TCM system for grooving operations under severe chip control issues. The results support the working hypothesis that supervised learning can provide accurate wear estimates in turning environments where signals are non-periodic and affected by intermittent disturbances. Consistent with previous TCM studies, non-linear regressors achieved markedly higher accuracy than a linear baseline, indicating that the relation between statistical monitoring features and wear progression is strongly non-linear under the tested conditions.

Chip entanglement produced transient excursions in both vibration and CNC internal signals, corroborating observations in the literature that chip formation dynamics can dominate monitoring variability in turning. In this context, tree-based ensembles (RF and XGBoost) delivered the most reliable performance across the three flank wear targets, combining high predictive accuracy with improved tolerance to heterogeneous noise and feature interactions. Importantly, the cross-validation analysis suggests a trade-off between peak accuracy (favoring XGBoost) and generalization (favouring RF), which is critical for maintaining stable inference when process conditions vary over time.

A further contribution of this work is the sensor relevance outcome: the DS3 dataset, based solely on CNC-derived current and power, achieved performance very similar to datasets including external accelerometers. This points to an attractive industrial pathway, as it reduces hardware integration effort and facilitates the scaling of a TCM solution across machines using signals already available in the NC. In addition, the containerized deployment on an edge device demonstrates a practical route from offline model development to a portable inference service, enabling low-latency monitoring and controlled model updates through orchestration mechanisms.

These findings should be interpreted considering the study scope. The applicability of the proposed strategy to different materials and machining conditions remains uncertain as experiments were conducted for a specific tool–workpiece system (ST-52) and a single insert/holder configuration. Also, wear labels were obtained from discrete measurements interpolated at window level, which may affect local prediction errors. Future work will focus on evaluating the proposed approach across different materials, cutting regimes and machining set-ups. Furthermore, the possibility of reducing occasional large deviations through temporal filtering and decision logic and explicitly detecting chip-entanglement episodes to mitigate their influence on inference will also be tackled. In this context, improvements in the model monitoring strategy are required, moving from the current batch-based data processing method toward continuous, high-frequency data acquisition and streaming-based monitoring, in order to enable a more responsive and reliable detection of tool condition changes.

Finally, strategies for model monitoring and periodic re-calibration will be required to address drift during long-term industrial operation. Although the deployment results are encouraging, the validation was performed using a limited number of additional tests. Therefore, the reported performance should be considered representative only of the operating conditions evaluated in this study. In long-term industrial applications, variations in cutting conditions, machine behavior, tool characteristics, or sensor performance may affect model accuracy over time. Consequently, periodic performance monitoring and, if necessary, model retraining with newly acquired data may be required to maintain reliable predictions. Further validation under a broader range of operating conditions will be addressed in future work.

5. Conclusions

The following conclusions can be drawn from the presented work:

- The power and current data from machine control can be used to train regression models capable of achieving a high degree of accuracy in the predictions of tool wear. On the other hand, at least for the tests presented here, the use of accelerometers does not provide additional improvement over the use of power and current variables. This is a significant finding as it could simplify the implementation of TCM systems, reducing their cost and complexity.
- From the models tested in the present work, the RF, KNR and XGBoost models are capable of providing R^2 score values above 0.97 and RMSE values below 0.02. Nevertheless, the KNR model seems to train with significant overfitting, casting doubts on its performance when unseen data would be fed to it.
- The industrial edge device has been validated as a suitable platform for the deployment and execution of ML models, demonstrating its capability to support real-time inference in production environments.
- The containerized ensemble model based on Random Forest and XGBoost models has shown robust predictive performance, delivering stable results with no observed signs of model drift under the evaluated conditions.
- Regarding the RF and XGBoost models, they yield the most robust regression performance, providing limited errors and the capability for a consistent output when facing unseen data. Nevertheless, the models still yield significant errors when looking at single samples, a point which should be taken into consideration when building an expert system for TCM during grooving operations.

Future research will address techniques to reduce the high deviations for single samples observed when using these models. On the one hand, a weighted ensemble for the results from these models could reduce such deviations. On the other hand, considering

an expert system capable of providing on-line tool wear values based on streaming of monitoring data, the use of simple time filtering for the regression results could enhance the performance of such an expert system by reducing the deviations of single samples in a viable time-frame.

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

Funding: This work was supported by the European Commission and the Spanish Government through the AIMS 5.0 project (HE-KDT-JU-2022-01-IA-101112089), funded under the Chips Joint Undertaking.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

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

References

1. Tlusty, J.; Andrews, G.C. A Critical Review of Sensors for Unmanned Machining. *CIRP Ann.* **1983**, *32*, 563–572. [CrossRef]
2. Jemielniak, K. Contemporary challenges in tool condition monitoring. *J. Mach. Eng.* **2019**, *19*, 48–61. [CrossRef]
3. Li, X.; Liu, X.; Yue, C.; Liang, S.Y.; Wang, L. Systematic Review on Tool Breakage Monitoring Techniques in Machining Operations. *Int. J. Mach. Tools Manuf.* **2022**, *176*, 103882. [CrossRef]
4. Munaro, R.; Attanasio, A.; Del Prete, A. Tool Wear Monitoring with Artificial Intelligence Methods: A Review. *J. Manuf. Mater. Process.* **2023**, *7*, 129. [CrossRef]
5. Sadhukhan, C.; Mitra, S.K.; Biswas, R.; Naskar, M.K. Tool Condition Monitoring: Unscented Kalman Filter for Tool Flank Wear Estimation in Turning of Inconel 718. *Mach. Sci. Technol.* **2021**, *25*, 331–348. [CrossRef]
6. Çakir, F.H.; Er, Ü.; Tekkalmaz, M. Monitoring the Wear of Turning Tools with the Electromechanical Impedance Technique. *J. Intell. Mater. Syst. Struct.* **2023**, *34*, 1341–1352. [CrossRef]
7. Bhuiyan, M.S.H.; Choudhury, I.A.; Dahari, M. Monitoring the Tool Wear, Surface Roughness and Chip Formation Occurrences Using Multiple Sensors in Turning. *J. Manuf. Syst.* **2014**, *33*, 476–487. [CrossRef]
8. Denkena, B.; Bergmann, B.; Stiehl, T.H. Wear Curve Based Online Feature Assessment for Tool Condition Monitoring. *Procedia CIRP* **2020**, *88*, 312–317. [CrossRef]
9. Colantonio, L.; Equeter, L.; Dehombreux, P.; Ducobu, F. Confidence Interval Estimation for Cutting Tool Wear Prediction in Turning Using Bootstrap-Based Artificial Neural Networks. *Sensors* **2024**, *24*, 3432. [CrossRef] [PubMed]
10. Kim, G.; Yang, S.M.; Kim, D.M.; Choi, J.G.; Lim, S.; Park, H.W. Developing a Deep Learning-Based Uncertainty-Aware Tool Wear Prediction Method Using Smartphone Sensors for the Turning Process of Ti-6Al-4V. *J. Manuf. Syst.* **2024**, *76*, 133–157. [CrossRef]
11. Peng, R.; Xiao, Z.; Peng, Y.; Zhang, X.; Zhao, L.; Gao, J. Research on Multi-Source Information Fusion Tool Wear Monitoring Based on MKW-GPR Model. *Measurement* **2025**, *242*, 116055. [CrossRef]
12. Pashmforoush, F.; Ebrahimi Araghizad, A.; Budak, E. Physics-Informed Tool Wear Prediction in Turning Process: A Thermo-Mechanical Wear-Included Force Model Integrated with Machine Learning. *J. Manuf. Syst.* **2024**, *77*, 266–283. [CrossRef]
13. Kotha Amarnath, S.; Inturi, V.; Rajasekharan, S.G.; Priyadarshini, A. Combining Sensor Fusion and a Machine Learning Framework for Accurate Tool Wear Prediction During Machining. *Machines* **2025**, *13*, 132. [CrossRef]
14. Ferrando Chacón, J.L.; Fernández de Barrena, T.; García, A.; Sáez de Buruaga, M.; Badiola, X.; Vicente, J. A Novel Machine Learning-Based Methodology for Tool Wear Prediction Using Acoustic Emission Signals. *Sensors* **2021**, *21*, 5984. [CrossRef] [PubMed]
15. Webb, G.I.; Hyde, R.; Cao, H.; Nguyen, H.L.; Petitjean, F. Characterizing Concept Drift. *Data Min. Knowl. Discov.* **2016**, *30*, 964–994. [CrossRef]
16. Savvy Data Systems. Available online: https://www.savvydatasystems.com/en/ (accessed on 20 May 2026).
17. Uhlmann, E.; Polte, M.; Heper, M. *Open-Source Framework for Automated Milling Experiments: G-Code Adaptation and NC Data Acquisition with Siemens Industrial Edge*; Technische Universität Berlin: Berlin, Germany, 2025. [CrossRef]
18. *ISO 3685:1993*; Tool-Life Testing with Single-Point Turning Tools. International Organization for Standardization: Geneva, Switzerland, 1993. Available online: https://www.iso.org/es/contents/data/standard/00/91/9151.html (accessed on 19 May 2026).

19. Jones, T.; Cao, Y. Tool Wear Prediction Based on Multisensor Data Fusion and Machine Learning. *Int. J. Adv. Manuf. Technol.* **2025**, *137*, 5213–5225. [CrossRef]
20. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; ACM: San Francisco, CA, USA, 2016; pp. 785–794.
21. Cosenza, D.N.; Korhonen, L.; Maltamo, M.; Packalen, P.; Strunk, J.L.; Næsset, E.; Gobakken, T.; Soares, P.; Tomé, M. Comparison of Linear Regression, k-Nearest Neighbour and Random Forest Methods in Airborne Laser-Scanning-Based Prediction of Growing Stock. *For. Int. J. For. Res.* **2021**, *94*, 311–323. [CrossRef]
22. Domínguez-Monferrer, C.; Fernández-Pérez, J.; De Santos, R.; Miguélez, M.H.; Cantero, J.L. Machine Learning Approach in Non-Intrusive Monitoring of Tool Wear Evolution in Massive CFRP Automatic Drilling Processes in the Aircraft Industry. *J. Manuf. Syst.* **2022**, *65*, 622–639. [CrossRef]
23. Chollet, F. *Deep Learning with Python*; Manning Publications Co.: Shelter Island, NY, USA, 2018.
24. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013.
25. Ferrisi, S.; Guido, R.; Lofaro, D.; Zangara, G.; Conforti, D.; Ambrogio, G. A CNN Architecture for Tool Condition Monitoring via Contactless Microphone: Regression and Classification Approaches. *Int. J. Adv. Manuf. Technol.* **2025**, *136*, 601–618. [CrossRef]
26. Sheppert, A.P. Techniques for Mitigating Overfitting in Machine Learning: A Comprehensive Review, Taxonomy, and Practical Guide. *Front. Artif. Intell.* **2026**, *9*, 1794271. [CrossRef] [PubMed]

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