# Sensor signal selection for tool wear curve estimation and subsequent tool breakage prediction in a drilling operation

Aitor Duo[a]*, Rosa Basagoiti[a], Pedro J. Arrazola[b] and Mikel Cuesta[b]

[a] Data analysis and cybersecurity research group, Mondragon Unibertsitatea, Loramendi, 4; 20500 Arrasate - Mondragon (Gipuzkoa), Spain

[b] High-performance machining research group, Mondragon Unibertsitatea, Loramendi, 4; 20500 Arrasate - Mondragon (Gipuzkoa), Spain

Aitor Duo

E-mail: aduo@mondragon.edu

Data analysis and cybersecurity research group, Mondragon Unibertsitatea, Loramendi, 4; 20500 Arrasate Mondragon (Gipuzkoa), Spain

**16 digit Orcid:**
**Aitor Duo[1,2]:** 0000-0003-1001-7845; **Rosa Basagoiti[1]:** 0000-0001-8213-6152; **Pedro J. Arrazola[2]:** 0000-0001-5500-8956; **Mikel Cuesta[2]:** 0000-0001-6419-5394

***Acknowledgement***

This work has been developed by the intelligent systems for industrial systems group (IT-1357-19) and high-performance machining group (IT-1315- 19) of Mondragon Unibertsitatea supported by the Department of Education, Language policy and Culture of the Basque Government. This work has been carried out with the support of INTOOL II (KK-2020/00103) project from Elkartek Program.

# Sensor signal selection for tool wear curve estimation and subsequent tool breakage prediction in a drilling operation

Tool condition monitoring systems have an important role in machining processes to reduce defective component and ensure quality requirements. Stopping the process before the tool breaks or an excessive tool wear is reached can avoid costs resulting from that undesirable situation. This research work presents the results obtained in drilling process monitoring carried out on Inconel 718. Monitoring systems should be light and scalable. Following this idea, multiple sensors for external signal acquisition are used in this work (cutting forces, vibrations, and acoustic emissions) and several machine internal signals are collected. With all this data in hand, the main objective is to evaluate the capacity of each acquisition source for the reconstruction of the tool wear curve and subsequently detection of tool breakage. Given the difficulty of using all of these signals in a real system, the methodology used to analyse the data makes it possible to have a comparative analysis of the potential of each of these sources for tool wear monitoring during the drilling process. The results indicate cutting forces whether they come from internal signals or external signals can carry out this task accurately. At the same time of data acquisition, detailed tool wear measurements were added.

Keywords: Tool condition monitoring; Inconel 718; Drilling; Data mining; Machine learning

## 1 Introduction

Machining process monitoring has become a strong topic of study in recent years. With the steps being taken around industry 4.0. Data-driven models are being developed with a positive influence on sustainability, decision-making and increasing trend in production.

In the literature, different works can be found motivated by the tool wear prediction in drilling (Caggiano, Napolitano, et al. 2018), milling (Stavropoulos et al. 2016), or turning (Caggiano, Napolitano, and Teti 2017) processes. Roughness monitoring through vibration signals and singular spectrum analysis (García Plaza and Núñez López 2017) in

turning. Hole roundness quality inspection through vibration signals and wavelet packet transform based on feature extraction and artificial neural networks (Ranjan et al. 2020). Ad hoc process monitoring was developed by (Urbikain and López de Lacalle 2020) to continuously acquire and perform signal processing through NI devices with a chatter detection use case. However, despite the improvements observed in the last years, there is still a need to monitor and optimise machining processes to reduce costs and sustainability increasement.

According to Dudzinski et al. (2004), turning and drilling are the main operations in the manufacture of discs for gas turbine engines for the aerospace industry. These parts are produced from Nickel-based superalloys due to their high temperatures and corrosion resistance. Chen and Liao (2003) reported that during drilling of Nickel-based alloys, the tool wear is accelerated, being the abrasion wear the mechanism that predominates in the initial stages. They showed how the tool coating is abraded and flank wear increases rapidly. After this stage, the drill was chipped on the tool periphery leading to changes in cutting mechanisms, producing long serrated chip morphology. At this stage, during chip evacuation, there is a friction increase between the chip and the generated surface, producing low quality holes.

Drilling is the most productive process for hole making. It is considered one of the most critical processes since a poor finish can cause problems in the manufacturing chain of a component. This type of occurrence can be derived due to tool wear or tool breakage. The most common practice for tool wear or tool breakage assessment is the physical installation of sensors such as dynamometers, accelerometers, acoustic emissions, or potentiometers. Nevertheless, sometimes the installation of these devices can be intrusive regarding the cutting operation due to geometrical limitations, modification of the

machine stiffness or constant recalibration specifications. Nowadays, the machine CNC controllers allows the acquisition of internal variables/signals that could replace the sensor devices.

The main steps to be followed in a framework for process monitoring can be seen in Fig. 1: Acquisition of scientific and industrial variables. Processing and feature extraction to later apply Machine Learning (ML) for pattern recognition, select best suitable features according to the target and apply ML algorithms for automatic in-process prediction of industrial variables are the main steps to be followed.

*[Figure 1 near here]*

The main objective of the present work is to develop a system capable of monitoring tool wear/breakage following a methodology that seeks to identify the best sensors to achieve this goal. A methodology that compares the performance of different sensors on the tool wear/breakage estimation using ML algorithms in the drilling operation is presented. Following this methodology, it becomes possible to quantify the predictive capability of each of the deployed sensors.

Cutting forces and temperatures in machining are generated by the high thermomechanical loads needed to overcome the stresses to remove the material. During the process, the tool becomes worn due to adhesion, abrasion, diffusion, or oxidation mechanisms, leading to a modification in the tool microgeometry, e.g., a tool flank wear. Therefore, cutting forces and, especially, feed force is increased (Rahim and Sharif 2006). Vibrations arise from cyclic variations in dynamic cutting forces incorporating free, forced, periodic and random types of vibration (Bhuiyan and Choudhury 2014). Acoustic emissions are formed due to the irreversible change in the working material. In the deformation zones, the tension is released in the form of energy, friction, residual stresses,

chip break, chip strike, or phase transformations are other sources of acoustic emissions in machining processes (Karpuschewski 2001).

Dimla Snr. (2000) qualifies acoustic emissions as an additional method to a monitoring system. He concluded that cutting forces and vibration signals are the most suitable sensors for tool wear monitoring. However, cutting force signal measurements using a dynamometer may not be available in most machining applications, its use in the industrial environment is not widespread as it is an invasive and unpractical method. Stavropoulos et al. (2016) indicated that the spindle electrical current has a better relationship than the accelerometer to tool wear in milling processes as it is less vulnerable to ambient noise. This is an advantage over the use of sensors as the internal signals of the numerical control can now be accessed to measure spindle power or current. Opening a range of opportunities for a cost-effective method for tool condition monitoring.

From the signals collected by the sensors, several features are usually obtained in the time domain, either in the frequency domain or in the time-frequency domain. In the time domain, these features try to explain the distribution of the acquired data in a given space of time. These are generally mean, median, RMS, maximum, minimum, standard deviation, skewness, kurtosis, variance, standard error and amplitude (Caggiano, Angelone, et al. 2018; Caggiano, Napolitano, et al. 2018; Elangovan et al. 2011). In the frequency domain, the signals are transformed into frequency components generally using Fast Fourier Transform (FFT) and dominant spectral peaks (Sick 2002), signal power in specific frequency ranges (Sick 2002), energy in frequency bands and statistical features of the spectrum (Teti et al. 2010) are commonly obtained. The Time-frequency domain is obtained by applying Short Time Fourier Transform (STFT), Wavelet

transforms or Hilbert Huang Transform (Lauro et al. 2014).

The use of techniques that allow the selection and identification of the most relevant features for model creation is not a very common practice when it comes to tool wear monitoring. Only 15% of the analysed works use some procedure for the selection of representative features during tool condition monitoring (Teti et al. 2010; Sick 2002). According to Jemielniak (2019) each particular case needs a feature selection process without any human intervention considering as many features as possible. Applying feature selection can be beneficial for data comprehension, better generalisation and remove redundant features. (Caggiano, Angelone, et al. 2018; Caggiano, Napolitano, et al. 2018) used those features with a Spearman correlation coefficient greater than 0.7 to tool wear. Elangovan et al. (2011) compared decision tree-based feature selection and PCA based transformation and feature reduction showing that decision tree-based feature selection results in better prediction accuracy. Dheeraj Simon and Deivanathan (2019) compared the accuracy obtained with a different subsets for best suitable subset selection. However, Mehmood et al. (2012) observed that there is an interaction between the method and data properties.

Two categories can be found on supervised ML, (i) classification tries to predict discrete values or labels, while (ii) regression tries to predict continuous values. For tool condition monitoring, some works try to predict the measured wear curve (Caggiano, Napolitano, et al. 2018; Liu, Tseng, and Tran 2019) and face the problem as a regression. Others try to classify different ranges of the wear curve, and several criteria can be found in this strategy to establish tool wear ranges. Kilundu, Dehombreux, and Chiementin (2011) predict 3 levels of tool wear, new, light and advanced wear together with chip stuffing and non-cutting pass. According to Patel and Muthuswamy (2020), given the continuous

nature of tool wear it is preferred to solve as a regression problem.

The most widely known algorithm for regression is linear regression, however its limited ability and simplicity is often not applicable to real-life data. Therefore, the use of other types of algorithms capable of dealing with non-linear data should be employed. Neural networks are widely used for tool wear monitoring. Caggiano, Napolitano, et al. (2018) employed backpropagation artificial neural networks (ANN) for tool wear curve prediction in CFRP drilling operation, obtaining a minimum *RMSE*= 0.00023. They used time-domain features of thrust force and torque. In a similar work, Caggiano, Angelone, et al. (2018), in addition to the thrust force and the cutting torque they acquired the acoustic emissions, they applied principal component analysis (PCA) on the features with high correlation coefficient with tool wear curve always obtaining an *RMSE*<2.17E-03. Corne et al. (2017) compared different backpropagation algorithm performances for tool wear prediction in drilling Inconel 625 based on spindle power data, concluding that the Levenberg Marquart algorithm was the best option. Balazinski et al. (2002) made a comparison between three ML methods, ANN, a fuzzy decision support system (FDSS) and an artificial neural network-based fuzzy inference system (ANNBFIS) for tool wear curve reconstruction in turning process. They did not see any superior method to the rest, but they concluded that ANNBFIS was the most practical ML algorithm given its practicality. Using the same data and algorithms, Ren et al. (2010) added to the comparison the Takagi-Sugeno-Kang (TSK) fuzzy modelling based on subtractive clustering method, showing its superiority to the rest ML algorithms.

Jaini et al. (2020) used a gap sensor for tool condition classification in the drilling process. By obtaining the kurtosis and skewness features of the spindle movement, they can classify 11 drills with different damage. Dheeraj Simon and Deivanathan (2019) used

time-domain features of the vibration signal to detect the presence of wear in drilling processes, using the K-Star algorithm. Wu et al. (2017) compared algorithms based on ANN, SVM and RF, showing the superiority of RF over the other methods. Ao and Qiao (2010) used Logistic regression (LR) and autoregressive moving average (ARMA) models to predict the remaining useful life (RUL). Duo et al. (2019) compared different algorithm (J48, LMT, IBK and NB) performances to predict different tool wear levels to choose the most suitable signal based on time-domain features. Table 1 summarises the main works consulted on the subject.

Some works use a data partitioning method for the training and testing phase (i.e. 70% for training and 30% for testing) (Wang et al. 2013; Shankar, Mohanraj, and Rajasekar 2018). This practice should be avoided as tool wear is a phenomenon that increases with machining time. Indirectly, data leakage is induced in the testing phase, which can lead to better results than those obtained in a real system overestimating the produced model. The testing phase should be carried out in a completely new tool to obtain adequate results and validate the created model. The data leakage problem for tool wear curve estimation is illustrated in Fig 2.

*[Figure 2 near here]*

Sensor fusion can be unpractical in industrial applications with geometrical limitations, sensor positioning problems or sensor constant recalibration specifications. Thus, machine internal signals are an alternative to sensor limitations. Feature selection should be applied to generate simpler and lighter models to establish an accurate relationship with the measured tool wear as most of extracted features are not related to measured monitoring unit. The continuous indirect measurement of tool wear and assessment of the risk of irreversible tool damage in accordance with process inherent uncertainties are

highly valued for more sustainable and cost-effective machining.

In the following sections, the experimental set-up and mounted sensors are first presented. Then, the feature extraction method used in this work is showed and how a suitable feature subset was selected. After that, the algorithms employed for tool wear curve regression and how the risk of irreversible tool damage was assessed is presented.

*[Table 1 near here]*

## 2 Methodology

### *2.1 Experimental set-up*

Drilling tests were carried out on a Lagun vertical milling machine tool with Fagor CNC8070 where the internal CNC variables may be accessed. 3 solid TiAlN coated carbide tools (MDS080SK) were used on Inconel 718 under the following cutting conditions; Vc = 15m/min, f = 0.1mm/rev, ø= 8mm and 6.5 mm depth throughout holes. The tool geometry can be seen in Fig. 3 a).

*[Figure 3 near here]*

During the tests, different sensors were installed to obtain signals related to the cutting process. In particular, an acoustic emission sensor (Kistler 8152C), a 3-component accelerometer (PCB J356A45) and a 4- component rotational dynamometer (Kistler 9123) were installed. Besides, several internal CNC signals shown in table 2 were collected. An analogue output (ao0 of the CNC) of the machine tool was used to obtain a signal at the acquisition time, allowing the simultaneous acquisition of internal and external signals. The surface of the workpiece is at Z=0 mm, when the position of the tool tip is Z=1 mm, a command is given to start the simultaneous acquisition on the NI USB 6361 and NI cDAQ 9178 acquisition cards. The experimental set-up used can be seen in Fig. 3 b). The internal signals of the machine were acquired at the CNC itself, while the external signals were acquired in an external PC.

*[Table 2 near here]*

## *2.2 Data preparation and wear prediction through several stages: from S1 to S6*

The presented methodology follows several steps with many interdependencies and will follow the next steps: a) gather data on different tools and many different sensors and internal signals for the drilling process, adding data relative to the wear. b) Extract features and select the best ones using data from different tools indistinctly. c) Choose the best algorithm between seven of them using data from different tools indistinctly. d) Use knowledge from steps b and c to predict the wear curve on each tool data separately. e) Use knowledge from steps b and c to predict tool breakage on each tool data separately.

### *2.2.1 Feature extraction*

The features were extracted from what was considered the most stable part of the cutting process. From the moment the tool tip is fully inserted into the material until the tool tip starts to emerge. This part can be seen in Fig. 4 represented as ”steady cutting”. This task was carried out using the tool tip position (POS_Z). Knowing where the tool tip is and performing a simultaneous acquisition of all the signals, it is possible to obtain the segment belonging to the stable machining zone that runs through the whole hole.

*[Figure 4 near here]*

The statistical features of the raw signals were first obtained. In addition to this, to obtain features of different frequency bands, a 3rd level wavelet packet decomposition was applied to each of the segments based on the work done by (Bombiński et al. 2016) and (Segreto, Karam, and Teti 2017) (Fig. 5). The transformation was applied only to all external signals. A total of 8 wavelet packages were obtained for each of the external signals covering one-eighth of the signal frequency spectrum.

*[Figure 5 near here]*

*Mean, RMS, standard deviation, maximum, minimum, signal amplitude, kurtosis, skewness,* and *variance* features were acquired from all the raw signals and all the obtained wavelet packages. [Signal name]_mean, [Signal name]_rms, [Signal name]_std, [Signal name]_max, [Signal name]_min, [Signal name]_peak, [Signal name]_kurt, [Signal name]_skew, [Signal name]_var were the names of each of the obtained feature to identify the feature of a specific signal. Overall, 720 features were considered.

### *2.2.2 Wear target estimation*

During the machining process and for the reconstruction of the wear curve, the tool condition was measured on each cutting edge every 10 holes made on the tool periphery (Fig. 6 a)), calculating then an average of the wear on the two cutting edges. Once the process was finished, using those estimated wear values, a curve was adjusted employing a 3rd-degree polynomial fitting shown in Fig. 6 b). The first hole in which a sign of break was observed on the periphery of the tool was also been identified and labelled. The criteria to finish the test was to achieve a wear value of $\mathrm{Vb} = 0.3\mathrm{mm}$ or 10 holes after the first breakage point was identified.

The third-degree polynomial curve estimated from the measurements made during the tests is used as the target for ML algorithms. Four phases have been identified during the evolution of the wear curve. (i) Break-in period is the phase where the tool wears out quickly. At the beginning of the process, for every tool, a $\mathrm{Vb} = 0.1\mathrm{mm}$ is reached after the first 10 holes (ii) Steady stage, the tool wear curve is smoothed and continues growing progressively. (iii) Failure region, where the tool suffers a breakage in the cutting edge (iv) Broken period, one or both cutting edges have suffered irreversible damage.

*[Figure 6 near here]*

During the tests, it has been observed in tool 1 that the first occurrence of a broken period appears in hole 90, while in tools 2 and 3, the broken period was observed in hole number 50. Even though the holes were drilled under the same cutting conditions, the tool periphery breakage was observed at different moments. Once the polynomial adjustment has been made, the datasets created with internal and external feature values and the estimated wear curve are joined as shown in Fig. 7 (Complete feature space) for each information source, including data for every tool.

### *2.2.3 Feature selection*

The non-relevant or redundant features must be removed from each dataset corresponding to one unique sensor for predicting the wear curve. The final decision regarding feature selection has been carried out through a voting scheme between different methods; a) two embedded methods, b) a wrapper method and c) a filter was used. The methods selected for this step have been Elastic Net (EN), Sequential Backward Search (SBS), Random Forest (RF) and Information Gain (IG). Every decision on selecting (1) or not (0) of a feature by an algorithm is added. In this way, each feature can obtain a maximum score of 4 (all the algorithms considered this specific feature as a candidate for tool wear curve prediction) and a minimum of 0. The selected features for each one of the sensors have been those that would have only features that obtain a score greater than 3. Ensuring that different criteria have been involved in the selection of these features. The process of feature selection is represented in Fig. 7 and carried out per each sensor separately.

*[Figure 7 near here]*

Specific names are used for the rest of this work to specify these reduced feature spaces,

dynamometer reduced feature space (*FC_RFS*), accelerometer reduced feature space (*ACC_RFS*) and internal signals reduced feature space (*INT_RFS*) were obtained.

### *2.2.4 Algorithm selection and wear curve prediction*

To analyse the impact of the variable selection process on the predictions made, the performance of some other algorithms predicting tool wear curve values was validated before and after the previous variable selection process using 10 fold 10 times cross-validation. This way, the behaviour of the selected algorithms for the prediction of the wear curve is evaluated under different input feature space.

The algorithms tested for the prediction of the wear curve were: (i) least absolute shrinkage and selection operator (LASSO), (ii) neural network (NNET), (iii) generalised linear model (GLM), (iv) K nearest neighbour (KNN), (v) M5 model rules (M5), (vi) multilayer perceptron (MLP) and (vii) Linear SVM for regression.

At this point, data has been used to extract the best feature or choose the best algorithms, but an open question is if at what point, data from one tool has the power to predict the status of other of the wear curve in another tool. To evaluate this point, the bootstrap sampling method was used. This method consists of selecting a part of the data for training (training samples) and another part for testing (out-of-bag samples). The training samples belong to two of the three tools used, and the out-of-bag samples belong to the remaining tool. After three iterations, the prediction of the wear curve for the three tools used in this work were obtained. To measure the behaviour of the model fitting against the measured tool wear curve, mean absolute error (MAE), root mean square error (RMSE) and R square ($R^2$) metrics were used, the mathematical notations of these metrics can be found on Table 3.

*[Table 3 near here]*

### *2.2.5 Tool breakage prediction*

In this section, the selected features previously selected during the feature selection process have been used to detect tool periphery breakage according to the levels established in subsection Wear target estimation. The algorithm used to create the models for detecting the tool periphery breakage was the Logistic Model Tree (LMT) because of its simplicity of interpreting the obtained model and identifying the features that allow the tool breakage detection. LMT combines logistic regression models with tree induction. A logistic model tree is a standard decision tree structure with logistic regression functions in the leaves (Landwehr, Hall, and Frank 2005). Unlike conventional decision trees, the leaves have an associated logistic regression function instead of a class label. At the leaves of the LMT, the functions $F(x)$ and $-F(x)$ determine the class membership probabilities by equations 1 and 2.

$$\Pr(\text{Without failure}) = \frac{e^{F(x)}}{e^{F(x)} + e^{-F(x)}}$$
$$\Pr(\text{Tool failure}) = \frac{e^{-F(x)}}{e^{F(x)} + e^{-F(x)}} \quad [1]$$

Where $F(x)$ is a linear model for the leave with the next form,

$$F(x) = a_0^j + \sum_{k=1}^{m} a_{v_k}^j \cdot v_k \quad [2]$$

Where $a_0^j$ is the intersection and $a_{v_k}^j$ is the coefficient of $v_k$ feature.

$-F(x)$ determines the probability of a tool failure (label =1), while $F(x)$ determines the probability of a tool without failure (label=0). When $F(x)$ has a high value, the probability of belonging to class 0 ”without failure” is higher. When $F(x)$ is closer to 0, the probability of belonging to one class or another is more doubtful, so this is the critical

area or failure region, where a tool can go from an ”without failure” (label:0) to a ”tool failure” (label:1) state. feasibility

Before the final model is created to predict the presence of breakage, several previous stages are used to impute values on unknown data that gradually control the uncertainty. A description of these stages can be seen in Fig. 8.

*[Figure 8 near here]*

**S1** The first strategy only considers those instances that have been correctly labelled for one unique tool. That is, without considering the transition zone where the tool is known to have suffered substantial and irreversible damage. This allows us to observe that the created model is effective to differentiate a bad state from a good state of the tool.

**S2** The second strategy is to use all the instances of all the tools together to do the same as in the first strategy.

**S3** The third strategy is to train the models with the well-labelled instances of a single tool and to test them in the instances belonging to the transition stage. In this way, the labels obtained in this step are imputed to the transition area.

**S4** The fourth strategy is to do the same as in the third strategy by considering all the instances of all the tools. This will also help in attributing the appropriate labels to the instances in the transition area.

**S5** With all the instances labelled, the models obtained are generalised, training in 70% of the data and testing in 30% on tools 2 and 3. To ensure the early detection of tool breakage the first instances have been introduced in the test partition.

**S6** With all the instances labelled, the models obtained are generalised, training in 70% of the data and testing in 30% on tools 1, 2 and 3. This last strategy will allow comparing the models created by introducing the data of one more tool.

There are more holes not presenting a break in the periphery of the tool than those that are performed with tool breakage, resulting in an imbalanced data problem. Having a few instances that belong to the tool failure class it is more difficult for the algorithm to learn what the decision boundary is. The training data has been balanced using the Synthetic Minority Oversampling Technique (SMOTE). It chooses two neighbouring minority instances and creates a new minority instance based on selected ones (Chawla et al. 2002). This allows the models to better generalise in terms of the minority class, in this case, the class that indicates tool failure. According to (Luque et al. 2019) for this type of imbalanced classification, the best evaluation metric results in the use of the Matthews correlation coefficient (MCC) since it takes into account the equilibrium ratios of the four categories in the confusion matrix (TP: true positives, TN: true negatives, FP: false positives, FN: false negatives) as seen in equation 3. The value of this metric is distributed from -1 to 1, being -1 the total disagreement between the predicted and the real value, 0 means random predictions and 1 the total agreement between the predicted and the real value.

$$\mathrm{MCC} = \frac{\mathrm{TP} \cdot \mathrm{TN} - \mathrm{FP} \cdot \mathrm{FN}}{\sqrt{(\mathrm{TP} + \mathrm{FP})(\mathrm{TP} + \mathrm{FN})(\mathrm{TN} + \mathrm{FP})(\mathrm{TN} + \mathrm{FN})}} \quad [3]$$

## 3 Results and discussion

### *3.1 Tool wear curve reconstruction*

The reconstruction of the wear curve has been carried out with the features of each of the

sensors separately. The results obtained with the acoustic emission signal have not been presented since they do not show any predictive capacity concerning tool wear curve. Different algorithms have been compared for each of the sensors used in this study. The models obtained were compared using the Complete Feature Space (CFS) and the Reduced Feature Space (RFS) for each of the sensors. In the case of the accelerometer, the reduced feature space is *ACC_RFS*, these features are those obtaining a score greater or equal to 3 during feature selection process.

*ACC_RFS= [WP5_ACCx_rms, WP5_ACCx_max, WP7_ACCx_rms, WP8_ACCx_rms, WP2_ACCy_rms, WP4_ACCy_kurt, WP3_ACCz_rms, ACCx_max, ACCx_skew, WP2_ACCx_rms, WP2_ACCx_max, WP2_ACCx_peak, WP4_ACCx_rms, WP6_ACCy_rms, WP7_ACCy_mean, WP3_ACCz_max, WP7_ACCz_kurt, WP8_ACCz_max, WP8_ACCz_peak]*

The complete feature space obtained from the accelerometer was of 243 features. Once the process of feature selection was applied, this number was reduced to 19 features.

In the case of the features obtained from the dynamometer, the feature space was reduced from 324 to 10 features (*FC_RFS*). Thus, having a considerable feature reduction.

*FC_RFS= [WP2_Fy_rms, Fx_mean, Fx_kurt, Mz_mean, WP2_Fy_min, WP2_Fy_skew, WP2_Fy_peak, WP2 Mz_rms,WP2_Mz_kurt, WP2_Mz_skew]*

The internal signals, from 81 feature space, it was reduced to 16 as seen in *INT_RFS* feature vector.

*INT_RFS= [TV50_mean, TV50_max, TV50F_std, TV51_max, TV2_mean, TV2_min, Vz_std, Vz_skew, TV50F_max, TV2_max, Vz_rms, Vz_max, JERKx_kurt, TV3_max, TV3_min, CV3_min]*

Table 4 summarize the results for each tested algorithm in the complete feature space and the reduced feature space of each sensor. The green square indicates the best result achieved with each sensor-algorithm pair in the complete feature space. The blue square points to the algorithm-sensor pair with the best result in the reduced feature space and the dashed line blue square are those with the same mean in accuracy as the best result achieved.

*[Table 4 near here]*

Considering the complete feature space of accelerometer, the KNN algorithm obtained the best case. Once the feature space was reduced the best result was obtained with the NNET algorithm. KNN, M5 and MLP obtain the same mean in accuracy with a p-value > 0.05 indicating that there is not a statistically significant difference between the results obtained by each of these algorithms. All of them outperforms the KNN algorithm in the complete feature space.

With the dynamometer signals better results were obtained in the reduced space using the *FC_RFS* features. The best result is obtained with the reduced space of features using the NNET algorithm, KNN and M5 algorithms show similar behaviour, as they do not show a statistically significant difference.

Internal signals do not require installing any sensor and it is, therefore, the most practical option for creating tool condition monitoring systems. In the case of internal signals, the KNN algorithm is the one that obtains the best results for the reconstruction of the wear curve in the reduced space of the features *INT_RFS* with a similar behaviour obtained by the M5 and NNET algorithms.

It is possible to reconstruct the wear curve both by using external signals of vibrations

and cutting forces and the machine internal signals. Slightly worse results were obtained with the accelerometer, although acceptable for environments that do not have the possibility of collecting internal signals or installing a dynamometer. Although several works show the ability of acoustic emissions to monitor tool wear, in this work, no trend has been observed that would allow good results to be obtained with this sensor. Once the comparison of the different algorithms has been made, it has been decided to use the NNET algorithm for the reconstruction of the wear curve. The process has been carried out by training with two of the tools and testing on the remaining tool. Table 5 shows the results obtained for each of the repetitions carried out. The wear curves obtained for each of the cases can be consulted in Appendix A.

*[Table 5 near here]*

Comparing the three cases, the best result is obtained by the dynamometer, since the lowest average error is obtained among the three tools. The internal signals tend to obtain a similar result with slightly higher errors. The accelerometer is in this case the one that obtains the worst result with the highest error. However, in general terms, good results are obtained in all cases and all options are valid for a tool wear curve reconstruction system.

### *3.2 Tool periphery breakage*

Once the wear curve has been reconstructed, the breakage of the tool periphery must also be detected not to damage or damage the minimum number of components possible.

During the first and second strategy (S1 and S2), it was not difficult to predict the state of the tool. In all cases it has been successfully tested that the models created from the features of any of the sensors are successful in differentiating a tool in good condition

from a tool in bad condition so the results have been satisfactory in 100% of the cases.

In the transition phase, where it is unknown at what point the tool has suffered irreversible damage, so the breakage or non-breakage label must be attributed on the pre-breakage and post-breakage labels. The S3 and S4 (3rd and 4th strategies) aim to assign labels to the instances in the transition area. Table 6 shows the label imputations made with each of the sensors. There can be seen the labels obtained with S3 and S4, the third strategy consists on training an algorithm with the features of each of the tools separately and the imputation was made based on each of the tools. The 4th strategy involves training the algorithm with both tools together and use the data of the holes from the transition area as testing.

*[Table 6 near here]*

The labels were imputed through an agreement between the different results. Thus, the results labels are those from the imputation row.

After allocating the labels corresponding to each of the tools, models were created using all the labels corresponding to tool 2 and tool 3 (those with a comparable tool life). 70% of the data was used to train the models and 30% to test them, the confusion matrices are shown in Fig. 9.

*[Figure 9 near here]*

Only considering tools 2 and 3, which have a comparable tool life and are broken in the same failure region, 100% accuracy is achieved with dynamometer and internal signals. The accelerometer sensor has the greatest difficulty for tool breakage detection.

Up to now, only the second and third tools have been tested for tool periphery breakage automatic detection. To see how the behaviour of the created models is altered, the

instances of tool 1 have been introduced in S6. Given the limited data available, training with two of the tools and testing with the remaining one is not possible as not all the variability of the data was collected. Given this situation, 70% of the data has been separated to train the model and 30% of the data to test it.

The results obtained with each sensor can be seen in Fig. 10. It shows the confusion matrix together with the wear curves indicating the correctly and incorrectly classified instances for each of the classes. False positives are those instances in the ”without failure” zone that are classified as ”tool failure” while false negative are those instances in the ”tool failure” zone classified as ”without failure”.

*[Figure 10 near here]*

The models obtained in S5 and S6 (strategies 5 and 6) can be consulted in Appendix B. Introducing a slightly different tool (with different life expectancy) more complicated models are obtained in which a greater number of features are involved. In the case of the dynamometer, the signal with the greatest impact with the highest weight coefficient is Mz mean for both strategies (S5 and S6). It can also be seen that in both models the same features are involved, however in S6 a greater number of features are present. In the case of the accelerometer the models obtained are not comparable, although the variables of strategy 5 appear in the model of strategy 6 the obtained models are of great complexity. Thus, it has little ability to generalise about the employed data. In the case of internal signals, they present the best results. The variable with the highest weight in the model in both cases is TV2_mean and it is remarkable that in the S6 only 7 features participate obtaining a good generalisation to testing data.

Finally, the results obtained by each of the sensors have been statistically compared. A sampling (70%-30%) was done 30 times and a t-test was applied between the results

obtained with each of the sensors. Table 7 shows the mean values of MCC of the models and the p values.

[Table 7 near here]

The results show that the internal signals can detect tool breakage with the best mean in MCC. Both the dynamometer and the accelerometer have significantly different averages with a p-value < 0.05.

The logistic regressions shown in the Appendix B belong to class 0 ”without failure” (F(x)). The logistic regressions for class 1 ”tool failure” are therefore the opposite of those shown (-F(x)). Simpler and more interpretable expressions were obtained with the dynamometer and the internal signals. The coefficients represent the effect of the variable per one unit of change in the predictor feature. Below is an example of the model achieved with the S6 for internal signals. Specifically, it is obtained the probability that the tool is not broken in holes 49 and 50 made with tool 2. It is in these holes where the tool has passed from ”without failure” to ”tool failure” state.

The model for internal signals and S6 shown in Appendix 2 is given by equation 4.

$$\begin{aligned} F(x) = 10.59 \; + \; & TV50_{mean} \cdot (-14) + TV51_{max} \cdot \; 0.61 \; + TV2_{mean} \\ & \cdot (-18.88) + Vz_{std} \cdot \; 1.42 \; + \; Vz_{skew} \cdot \; 2.41 \; + TV50_{max} \\ & \cdot (-5.78) \; + TV3_{max} \cdot (-1.82) \end{aligned} \qquad [4]$$

Substituting the values of the features by the normalised values (μ=0, σ=1) of hole 49 the expression is as follows,

$$\begin{aligned} F_{hole49} = 10.59 \; + \; & 0.13 \cdot (-14) + (-0.28) \cdot \; 0.61 \; + 0.4 \cdot (-18.88) + 0.39 \; * \; 1.42 \\ & +0.07 \cdot \; 2.41 \; + 0.10 \cdot (-5.78) \; + (-0.40) \cdot (-1.82) = 1.4 \end{aligned}$$

And the probability that in hole 49 the tool has not suffered a breakage is,

$$\Pr(\text{"without breakage"})_{\text{hole49}} = \frac{e^{1.4}}{e^{1.4} + e^{-1.4}} = 0.94$$

In hole 50, in the same way, making the substitution of the values the expression is,

$$F_{\text{hole50}} = 10.59 + 0.758 \cdot (-14) + 0.097 \cdot 0.61 + 1.194 \cdot (-18.88) + (-0.561) * 1.42 + (-1.431) \cdot 2.41 + 0.654 \cdot -5.78 + 0.317 \cdot (-1.82) = -31.13$$

And the probability that in hole 50 the tool has not suffered a breakage is,

$$\Pr(\text{"without breakage"})_{\text{hole50}} = \frac{e^{-31.13}}{e^{31.13} + e^{-31.13}} = 9.137e - 28$$

Seeing these results, when F(x) starts to get close to 0, the process should be stopped for premature tool change before the tool breaks.

Fig 11 shows the function F(x) obtained based on INT_RFS as a function of the holes drilled, the plot on the right margin shows the density plot of F(x) for each tool. The minimum value among the maximum of F(x) between the 3 tools before the breakage is -17 (the most conservative value), while the maximum value is -13. As can be seen, tools 2 and 3 have a similar tool life, while tool 1 has allowed a more significant number of holes to be drilled. This is because the inherent variability of the process, material properties, variability in tool manufacture or the degree of effectiveness of the coolant in reaching the cutting zone can cause these variabilities. The figure also shows the polynomial fit of each of the curves. In the equations obtained, it can be seen that as the first coefficient increases, the tool has a shorter life. Under the same conditions, there may be situations where the tool is prematurely changed in the hole 40 when double the number of holes could have been made. Therefore, the use of techniques to assess the condition of the tool based on the signals collected during the process is essential to avoid these situations.

*[Figure 11 near here]*

## 4 Conclusions

Three drilling repetitions were performed on Inconel 718 under fixed cutting conditions. The deployed set-up results in the use of different sensors for tool wear monitoring and subsequent tool periphery breakage detection. A methodology based on data mining and ML techniques allowed to select the most suitable sensor for tool wear monitoring, on the one hand, the reconstruction of the wear curve is addressed using features obtained from each of the sensors separately and comparing different algorithm performances. On the other hand, tool breakage is detected automatically using a methodology for the imputation of unknown tool breakage labels. The major contribution of the present work is the comparison of the predictive capability of four data sources (dynamometer, accelerometer, acoustic emissions and internal signals), each one individually, for the prediction of tool wear and the estimation of the risk that a tool may suffer irreversible damage. The feasibility of each of the sensors lies in (i) the implementation of the sensor in the work environment, (ii) the amount of data generated by each of the sensors, (iii) the cost of installing the sensor in the work environment and (iv) the prediction capacity provided. The following conclusions can be drawn:

- The external sensors could be replaced by internal signals without compromising the reliability of the predictions made. Thus, showing potential applications for big and small industries where a cost effective solution is required. In addition, the installation of sensors can lead to geometrical limitations or changes in the mechanical stiffness of the machine.

- In authors experience, the most practical signals for monitoring the tool condition are those provided by the machine itself (Internal signals). No sensor installation is required, the highest sampling frequency is 250Hz so it does not generate a

large amount of data and provides good predictive capabilities in terms of tool wear, both for the reconstruction of the wear curve and for the detection of tool breakage. In cases where it is not possible to collect internal signals, the most practical option is the accelerometer, as it results in a cheaper sensor than a dynamometer and less invasive. Nevertheless, the accelerometer doesn't give the capacity of detecting tool breakage, or at least in the feature space analysed in this work.

- The approach allowed both to select a subset of variables and to impute values to the unknown observations to later develop a possible solution for timely tool replacement. Strategy 6 (S6) involves 7 features in the models created with the dynamometer and the internal signals, allowing the creation of simpler models. Besides, the tool breakage detection method used in this work allows to stablish a threshold before tool breakage occurs even for tools with different tool lives.
- It has been observed that the same tool geometry under the same cutting conditions in the same material can vary in terms of tool life. It is therefore a challenge for the modelling of the wear curve. The LMT algorithm made it possible to observe from the beginning of the process that tool 1 would last longer than tool 2 and tool 3. For timely replacement of the cutting tool, both the estimation of the wear curve and the detection of a possible tool breakage are indispensable.

- In the present work the cutting conditions have been kept fixed. For a wider use of the proposed methodology and to obtain more robust models it is necessary to carry out a larger number of tests and obtain a larger amount of data.

**References**

Abu-Mahfouz, Issam. 2003. “Drilling Wear Detection and Classification Using Vibration Signals and Artificial Neural Network.” *International Journal of Machine Tools and Manufacture* 43 (7): 707–20. https://doi.org/10.1016/S0890-6955(03)00023-3.

Ao, Yinhui, and George Qiao. 2010. “Prognostics for Drilling Process with Wavelet Packet Decomposition.” *International Journal of Advanced Manufacturing Technology* 50 (1–4): 47–52. https://doi.org/10.1007/s00170-009-2509-6.

Balazinski, Marek, Ernest Czogala, Krzysztof Jemielniak, and Jacek Leski. 2002. “Tool Condition Monitoring Using Artificial Intelligence Methods.” *Engineering Applications of Artificial Intelligence* 15: 73–80.

Bhuiyan, M. S.H., and I. A. Choudhury. 2014. *Review of Sensor Applications in Tool Condition Monitoring in Machining*. *Comprehensive Materials Processing*. Vol. 13. Elsevier. https://doi.org/10.1016/B978-0-08-096532-1.01330-3.

Bombiński, Sebastian, Krzysztof Błazejak, Mirosław Nejman, and Krzysztof Jemielniak. 2016. “Sensor Signal Segmentation for Tool Condition Monitoring.” *Procedia CIRP* 46: 155–60. https://doi.org/10.1016/j.procir.2016.03.203.

Caggiano, Alessandra, Roberta Angelone, Francesco Napolitano, Luigi Nele, and Roberto Teti. 2018. “Dimensionality Reduction of Sensorial Features by Principal Component Analysis for ANN Machine Learning in Tool Condition Monitoring of CFRP Drilling.” *Procedia CIRP* 78: 307–12. https://doi.org/10.1016/j.procir.2018.09.072.

Caggiano, Alessandra, Francesco Napolitano, Luigi Nele, and Roberto Teti. 2018. “Multiple Sensor Monitoring for Tool Wear Forecast in Drilling of CFRP/CFRP Stacks with Traditional and Innovative Drill Bits.” *Procedia CIRP* 67: 404–9. https://doi.org/10.1016/j.procir.2017.12.233.

Caggiano, Alessandra, Francesco Napolitano, and Roberto Teti. 2017. “Dry Turning of Ti6Al4V: Tool Wear Curve Reconstruction Based on Cognitive Sensor Monitoring.” *Procedia CIRP* 62: 209–14. https://doi.org/10.1016/j.procir.2017.03.046.

Chawla, Nitesh v., Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. 2002. “SMOTE: Synthetic Minority over-Sampling Technique.” *Journal of Artificial Intelligence Research* 16 (Sept. 28): 321–57. https://doi.org/10.1613/jair.953.

Chen, Y. C., and Y. S. Liao. 2003. “Study on Wear Mechanisms in Drilling of Inconel 718 Superalloy.” *Journal of Materials Processing Technology* 140 (1-3 SPEC.): 269–73. https://doi.org/10.1016/S0924-0136(03)00792-1.

Corne, Raphael, Chandra Nath, Mohamed el Mansori, and Thomas Kurfess. 2017. “Study of Spindle Power Data with Neural Network for Predicting Real-Time Tool Wear/Breakage during Inconel Drilling.” *Journal of Manufacturing Systems* 43: 287–95. https://doi.org/10.1016/j.jmsy.2017.01.004.

Dheeraj Simon, Galipothu, and R. Deivanathan. 2019. “Early Detection of Drilling Tool Wear by Vibration Data Acquisition and Classification.” *Manufacturing Letters* 21: 60–65. https://doi.org/10.1016/j.mfglet.2019.08.006.

Dimla Snr., Dimla E. 2000. “Sensor Signals for Tool-Wear Monitoring in Metal Cutting Operations - a Review of Methods.” *International Journal of Machine Tools and Manufacture* 40 (8): 1073–98. https://doi.org/10.1016/S0890-6955(99)00122-4.

Dudzinski, D., A. Devillez, A. Moufki, D. Larrouquère, V. Zerrouki, and J. Vigneau. 2004. “A Review of Developments towards Dry and High Speed Machining of Inconel 718 Alloy.” *International Journal of Machine Tools and Manufacture* 44 (4): 439–56. https://doi.org/10.1016/S0890-6955(03)00159-7.

Duo, A., R. Basagoiti, P.J. Arrazola, J. Aperribay, and M. Cuesta. 2019. “The Capacity of Statistical Features Extracted from Multiple Signals to Predict Tool Wear in the Drilling Process.” *International Journal of Advanced Manufacturing Technology* 102 (5–8). https://doi.org/10.1007/s00170-019-03300-5.

Elangovan, M., S. Babu Devasenapati, N. R. Sakthivel, and K. I. Ramachandran. 2011. “Evaluation of Expert System for Condition Monitoring of a Single Point Cutting Tool Using Principle Component Analysis and Decision Tree Algorithm.” *Expert Systems with Applications* 38 (4): 4450–59. https://doi.org/10.1016/j.eswa.2010.09.116.

García Plaza, E., and P. J. Núñez López. 2017. “Surface Roughness Monitoring by Singular Spectrum Analysis of Vibration Signals.” *Mechanical Systems and Signal Processing*. https://doi.org/10.1016/j.ymssp.2016.06.039.

Jaini, Siti Nurfadilah Binti, Deug Woo Lee, Seung Jun Lee, Mi Ru Kim, and Gil Ho Son. 2020. “Indirect Tool Monitoring in Drilling Based on Gap Sensor Signal and Multilayer Perceptron Feed Forward Neural Network.” *Journal of Intelligent Manufacturing*, no. 0123456789. https://doi.org/10.1007/s10845-020-01635-5.

Jemielniak, Krzysztof. 2019. “Contemporary Challenges in Tool Condition Monitoring.” *Journal of Machine Engineering* 19 (1): 48–61. https://doi.org/10.5604/01.3001.0013.0448.

Karpuschewski, Bernhard. 2001. *Sensoren Zur Prozessüberwachung Beim Spanen*. VDI-Verlag.

Kilundu, Bovic, Pierre Dehombreux, and Xavier Chiementin. 2011. “Tool Wear Monitoring by Machine Learning Techniques and Singular Spectrum Analysis.” *Mechanical Systems and Signal Processing* 25 (1): 400–415. https://doi.org/10.1016/j.ymssp.2010.07.014.

Krishnakumar, P., K. Rameshkumar, and K. I. Ramachandran. 2015. "Tool Wear Condition Prediction Using Vibration Signals in High Speed Machining (HSM) of Titanium (Ti-6Al-4V) Alloy." *Procedia Computer Science* 50: 270–75. https://doi.org/10.1016/j.procs.2015.04.049.

Landwehr, Niels, Mark Hall, and Eibe Frank. 2005. "Logistic Model Trees." *Machine Learning* 59 (1–2): 161–205. https://doi.org/10.1007/s10994-005-0466-3.

Lauro, C. H., L. C. Brandão, D. Baldo, R. A. Reis, and J. P. Davim. 2014. "Monitoring and Processing Signal Applied in Machining Processes - A Review." *Measurement: Journal of the International Measurement Confederation* 58: 73–86. https://doi.org/10.1016/j.measurement.2014.08.035.

Liu, Meng-kun, Yi-heng Tseng, and Minh-quang Tran. 2019. "Tool Wear Monitoring and Prediction Based on Sound Signal." *The International Journal of Advanced Manufacturing Technology*.

Luque, Amalia, Alejandro Carrasco, Alejandro Martín, and Ana de las Heras. 2019. "The Impact of Class Imbalance in Classification Performance Metrics Based on the Binary Confusion Matrix." *Pattern Recognition* 91: 216–31. https://doi.org/10.1016/j.patcog.2019.02.023.

Mehmood, Tahir, Kristian Hovde Liland, Lars Snipen, and Solve Sæbø. 2012. "Review of Input Variable Selection Methods for Artificial Neural Networks." *Chemometrics and Intelligent Laboratory Systems* 118: 62–69. https://doi.org/10.1016/j.chemolab.2012.07.010.

Mohsen, Marani, Zeinali Mohammadjavad, Jules Kouam, Victor Songmene, and Chris K. Mechefske. 2020. "Prediction of Cutting Tool Wear during a Turning Process Using Artificial Intelligence Techniques." *The International Journal of Advanced*

*Manufacturing Technology* 32 (2): 174–82. https://doi.org/10.1080/0951192X.2018.1550681.

Patel, Zeel Bharatkumar, and Sreekumar Muthuswamy. 2020. "A Machine Learning Scheme for Tool Wear Monitoring and Replacement in IoT-Enabled Smart Manufacturing." In *Lecture Notes in Mechanical Engineering*, 439–47. https://doi.org/10.1007/978-981-15-2696-1_43.

Rahim, Erween Abd., and Safian Sharif. 2006. "Investigation on Tool Life and Surface Integrity When Drilling Ti-6Al-4V and Ti-5Al-4V-Mo/Fe." *JSME International Journal Series C* 49 (2): 340–45. https://doi.org/10.1299/jsmec.49.340.

Ranjan, Jitesh, Karali Patra, Tibor Szalay, Mozammel Mia, Munish Kumar Gupta, Qinghua Song, Grzegorz Krolczyk, Roman Chudy, Vladislav Alievich Pashnyov, and Danil Yurievich Pimenov. 2020. "Artificial Intelligence-Based Hole Quality Prediction in Micro-Drilling Using Multiple Sensors." *Sensors (Switzerland)* 20 (3): 1–14. https://doi.org/10.3390/s20030885.

Ren, Qun, Marek Balazinski, Luc Baron, and Krzysztof Jemielniak. 2010. "TSK Fuzzy Modeling for Tool Wear Condition in Turning Processes : An Experimental Study." *Engineering Applications of Artificial Intelligence* 24 (2): 260–65. https://doi.org/10.1016/j.engappai.2010.10.016.

Segreto, Tiziana, Sara Karam, and Roberto Teti. 2017. "Signal Processing and Pattern Recognition for Surface Roughness Assessment in Multiple Sensor Monitoring of Robot-Assisted Polishing." *International Journal of Advanced Manufacturing Technology* 90 (1–4): 1023–33. https://doi.org/10.1007/s00170-016-9463-x.

Shankar, S, T Mohanraj, and R Rajasekar. 2018. "Prediction of Cutting Tool Wear during Milling Process Using Artificial Intelligence Techniques." *International Journal of*

*Computer Integrated Manufacturing* 00 (00): 1–9. https://doi.org/10.1080/0951192X.2018.1550681.

Sick, Bernhard. 2002. "On-Line and Indirect Tool Wear Monitoring in Turning with Artificial Neural Networks: A Review of More than a Decade of Research." *Mechanical Systems and Signal Processing* 16 (4): 487–546. https://doi.org/10.1006/mssp.2001.1460.

Stavropoulos, P., A. Papacharalampopoulos, E. Vasiliadis, and G. Chryssolouris. 2016. "Tool Wear Predictability Estimation in Milling Based on Multi-Sensorial Data." *International Journal of Advanced Manufacturing Technology* 82 (1–4): 509–21. https://doi.org/10.1007/s00170-015-7317-6.

Teti, R., K. Jemielniak, G. O'Donnell, and D. Dornfeld. 2010. "Advanced Monitoring of Machining Operations." *CIRP Annals - Manufacturing Technology* 59 (2): 717–39. https://doi.org/10.1016/j.cirp.2010.05.010.

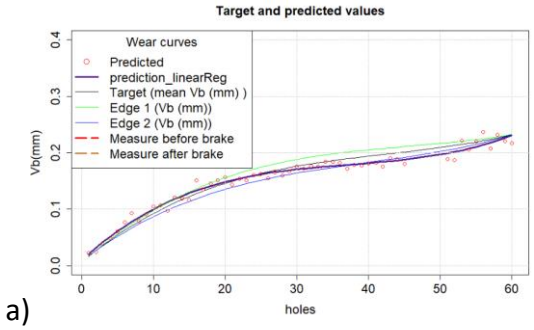
Urbikain, G., and L. N. López de Lacalle. 2020. "MoniThor: A Complete Monitoring Tool for Machining Data Acquisition Based on FPGA Programming." *SoftwareX* 11: 100387. https://doi.org/10.1016/j.softx.2019.100387.

Wang, Guofeng, Chang Liu, Yinhu Cui, and Xiaoliang Feng. 2013. "Tool Wear Monitoring Based on Cointegration Modelling of Multisensory Information." *International Journal of Computer Integrated Manufacturing*, no. August 2014: 37–41. https://doi.org/10.1080/0951192X.2013.814162.

Wu, Dazhong, Connor Jennings, Janis Terpenny, Robert X. Gao, and Soundar Kumara. 2017. "A Comparative Study on Machine Learning Algorithms for Smart Manufacturing: Tool Wear Prediction Using Random Forests." *Journal of Manufacturing Science and Engineering* 139 (7): 071018. https://doi.org/10.1115/1.4036350.

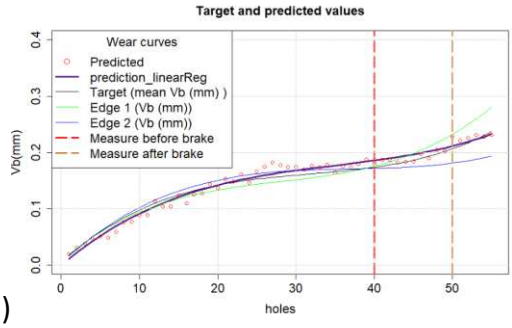
## Appendix A: Tool wear reconstruction results for each case

This appendix shows the results obtained from the bootstrapping process with each of the sensor for predicting the wear curve.



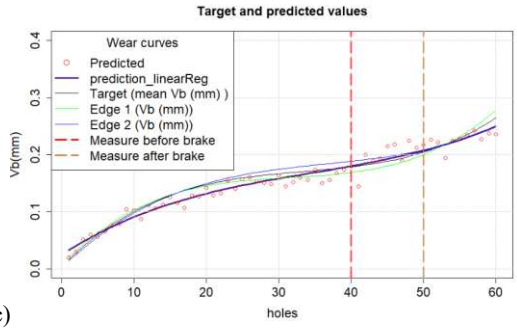

a)

MAE= 0.0069; RMSE= 0.0078; R2= 0.996




b)

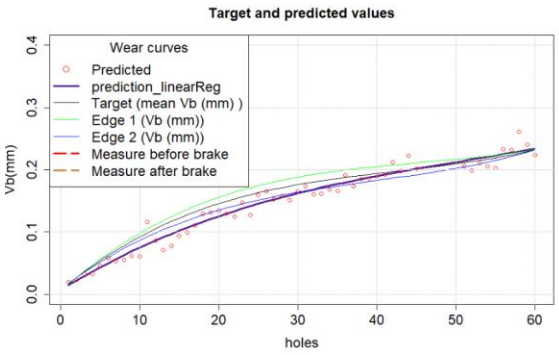
MAE= 0.0066; RMSE= 0.0074; R2= 0.995



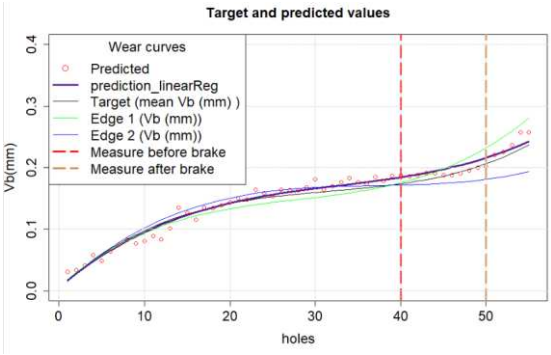

c)

MAE= 0.0076; RMSE= 0.0092; R2= 0.989

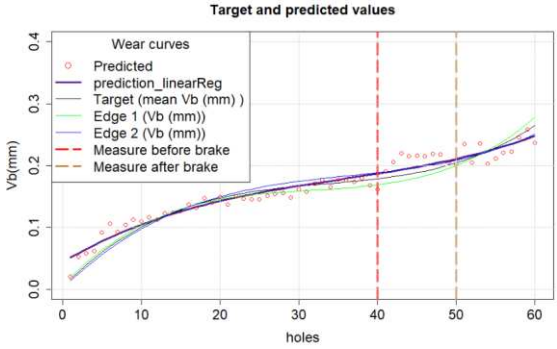
***Fig A. 1*** *Tool wear reconstruction for all the tools using the reduced space features FCRFS and NNET algorithm a) training samples from 2 and 3 tools and 1st tool wear curve reconstruction b) training samples from 1 and 3 tools and 2nd tool wear curve reconstruction c) training samples from 1 and 2 tools and 3rd tool wear curve reconstruction*




MAE= 0.011; RMSE= 0.013; R2= 0.992

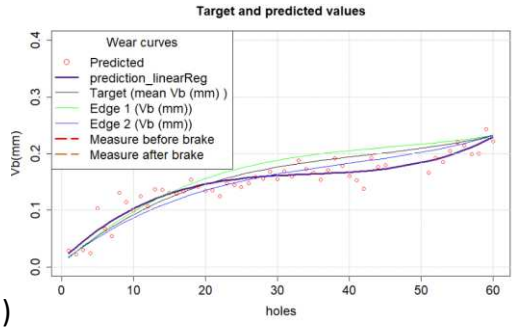



MAE= 0.0056; RMSE= 0.0067; R2= 0.998

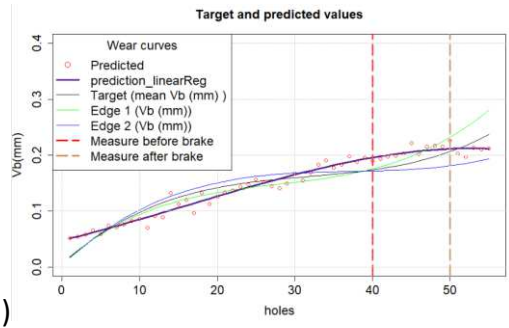



MAE= 0.0071; RMSE= 0.0098; R2= 0.989

***Fig A. 2*** *Tool wear reconstruction for all the tools using the reduced space features INTRFS and NNET algorithm a) training samples from 2 and 3 tools and 1st tool wear curve reconstruction b) training samples from 1 and 3 tools and 2nd tool wear curve reconstruction c) training samples from 1 and 2 tools and 3rd tool wear curve reconstruction*



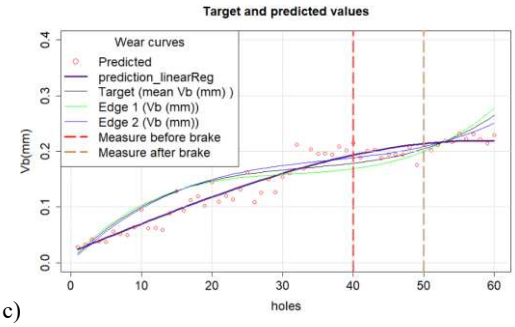

a)




b)

MAE= 0.0136; RMSE= 0.0159; R2= 0.985

MAE= 0.0137; RMSE= 0.0155; R2= 0.955




c)

MAE= 0.0178; RMSE= 0.0210; R2= 0.955

***Fig A. 3*** *Tool wear reconstruction for all the tools using the reduced space features ACCRFS and NNET algorithm a) training samples from 2 and 3 tools and 1st tool wear curve reconstruction b) training samples from 1 and 3 tools and 2nd tool wear curve reconstruction c) training samples from 1 and 2 tools and 3rd tool wear curve reconstruction*

**Appendix B: Models obtained in strategy 5 and 6**

This appendix shows the models obtained from strategies 5 and 6 respectively.

***Cutting forces***

| | |
|---|---|
| S5 | F(x)=13.11 + [Fx_mean] * 3.22 +[Mz_mean] * -15.75 +[WP2_Fy_peak] * 0.9 |
| S6 | F(x)=5.42 + [WP2_Fy_rms] * 3.58 +**[Fx_mean]** * 1.1 +**[Mz_mean]** * -17.53 +[WP2_Fy_min] * -0.23 +**[WP2_Fy_peak]** * 0.56 +[WP2_Mz_rms] * 1 +[WP2_Mz_skew] * -1.77 |

***Accelerometer***

| | |
|---|---|
| S5 | WP2_ACCx_rms <= -0.337507: LM_1:1/2 (31)<br><br>WP2_ACCx_rms > -0.337507<br><br>\| WP4_ACCx_rms <= -0.677682: LM_2:1/3 (12)<br><br>\| WP4_ACCx_rms > -0.677682: LM_3:1/3 (38)<br><br><br>LM_1:<br><br>F(x)= 1.24 + [WP3_ACCz_max] * -0.19<br><br><br>LM_2:<br><br>F(x)= -0.55 + [WP2_ACCy_rms] * -1.43 +[WP4_ACCx_rms] * 0.47<br><br>+[WP3_ACCz_max] * -0.3<br><br><br>LM_3:<br><br>F(x)= 0.71 + [ACCx_max] * -0.45 +[WP4_ACCx_rms] * 0.47<br><br>+[WP3_ACCz_max] * -0.3 |
| S6 | WP2_ACCy_rms <= 0.719511 |

| | |
|---|---|
| | \| WP2_ACCy_rms <= -0.814449: LM_1:41/123 (31)<br>\| WP2_ACCy_rms > -0.814449<br>\| \| WP4_ACCx_rms <= -0.68583: LM_2:38/161 (35)<br>\| \| WP4_ACCx_rms > -0.68583: LM_3:41/164 (80)<br>WP2_ACCy_rms > 0.719511: LM_4:36/77 (29)<br>Number of Leaves : 4<br>Size of the Tree : 7<br>LM_1:<br>F(x)= 24.96 + [WP7_ACCx_rms] * -0.33 +[WP8_ACCx_rms] * -0.54 +[WP2_ACCy_rms] * -0.55 +[WP4_ACCy_kurt] * -1.02 +[WP3_ACCz_rms] * 7.56 +[ACCx_max] * 0.16 +[ACCx_skew] * -262.45 +[WP2_ACCx_rms] * -2.4 +[WP2_ACCx_max] * 0.27 +WP2_ACCx_peak] * -6.32 +[WP4_ACCx_rms] * 0.84 +[WP6_ACCy_rms] * -0.39 +[WP7_ACCy_mean] * 1.54 +[WP3_ACCz_max] * -1.08 +[WP7_ACCz_kurt] * -57.62 +WP8_ACCz_max] * 1.87<br><br>LM_2:<br>F(x)= -5.18 + [WP5_ACCx_rms] * 0.57WP5_ACCx_max] * -2.68 +[WP7_ACCx_rms] * -0.33 +[WP8_ACCx_rms] * -1.26 +WP2_ACCy_rms] * 0.3 +[WP4_ACCy_kurt] * -1.2 +WP3_ACCz_rms] * 1.78 +[ACCx_max] * 0.16 +[ACCx_skew] * -101.13 +WP2_ACCx_rms] * -0.35 +[WP2_ACCx_max] * 15.74 +[WP2_ACCx_peak] * -0.03 +WP4_ACCx_rms] * 0.95 +[WP6_ACCy_rms] * -0.45 +[WP7_ACCy_mean] * -0.19 +[WP3_ACCz_max] * -1.17 + WP7_ACCz_kurt] * 33.87 +[WP8_ACCz_max] * -0.2 +[WP8_ACCz_peak] * 0.13 |

| | |
|---|---|
| | LM_3:<br>F(x)= 12.04 + [WP5_ACCx_rms] * 0.57 +[WP5_ACCx_max] * -3.12 +[WP7_ACCx_rms] * -0.33 +[WP8_ACCx_rms] * -1.26 +[WP2_ACCy_rms] * 1 + [WP4_ACCy_kurt] * -2.9 +[WP3_ACCz_rms] * 2.58 +[ACCx_max] * 0.16 +[ACCx_skew] * -160.26 +[WP2_ACCx_rms] * -0.64 +[WP2_ACCx_max] * 0.12 +[WP2_ACCx_peak] * -0.75 +[WP4_ACCx_rms] * 1.89 +[WP6_ACCy_rms] * -0.97 +[WP7_ACCy_mean] * -0.48 +[WP3_ACCz_max] * -1.17 +[WP7_ACCz_kurt] * 3.33 +[WP8_ACCz_max] * -0.2 +[WP8_ACCz_peak] * 0.13<br><br>LM_4:<br>F(x)= 21.16 + [WP7_ACCx_rms] * -0.47 +[WP2_ACCy_rms] * 0.34 +[WP3_ACCz_rms] * 1.1 +[ACCx_max] * -0.1 +[ACCx_skew] * -30.22 +[WP2_ACCx_rms] * -0.38 +[WP2_ACCx_max] * 0.27 +[WP2_ACCx_peak] * 0.1 +[WP4_ACCx_rms] * 0.43 +[WP6_ACCy_rms] * -0.39 +[WP7_ACCy_mean] * -0.08 +[WP3_ACCz_max] * -0.61 +[WP7_ACCz_kurt] * 1.53 +[WP8_ACCz_max] * -0.2 +[WP8_ACCz_peak] * 0.01 |

***Internal signals***

| | |
|---|---|
| S5 | F(x)= 0.56 + [TV2_mean] * -0.72 |
| S6 | F(x)= 10.59 + [TV50_mean] * -14 +[TV51_max] * 0.61 +[**TV2_mean]** * -18.88 +[Vz_std] * 1.42 +[Vz_skew] * 2.41 +[TV50F_max] * -5.78 +[TV3_max] * -1.82 |

***Table 1*** *Drilling tool wear monitoring works based on machine learning algorithms*

| Ref. | Sensor | | | | | | | Process | | Input | | | Method | Output | | | | Feature selection |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Dynamometer | AE | Accelerometer | Current | Sound | Gap sensor | Strain | Drilling | Others | Time-domain | Frequency domain | Time-frequency domain | | Wear curve | Wear levels | Wear types | Remaining useful time | |
| (Caggiano, Napolitano, et al. 2018) | x | | | | | | | x | | x | | | • ANN | x | | | | Spearman correlation |
| (Caggiano, Angelone, et al. 2018) | x | x | | | | | | x | | x | | | • ANN | x | | | | Spearman correlation |
| (Elangovan et al. 2011) | | | x | | | | | | x | x | | | • C4.5 | | x | | | PCA-decision tree |
| (Wang et al. 2013) | x | | | | | | | | x | x | | | • Regression • model | x | | | | - |
| (Abu-Mahfouz 2003) | | | x | | | | | x | | x | x | | • ANN | | | x | | - |
| (Jaini et al. 2020) | | | | | | x | | x | | x | | | • MLP | | x | | | - |
| (Shankar, Mohanraj, and Rajasekar 2018) | x | | | | x | | | | x | | | | • ANN | x | | | | - |
| (Dheeraj Simon and Deivanathan 2019) | | | x | | | | | x | | x | | | • K star | | x | | | Training with different subsets |
| (Krishnakumar, Rameshkumar, and Ramachandran 2015) | | | x | | | | | | x | x | | | • J48 • ANN | | x | | | J48 |
| (Kilundu, Dehombreux, and Chiementin 2011) | | | x | | | | | | x | | x | | • Bayesian network • KNN • Decision tree • ANN | | x | | | Discriminant analysis |
| (Wu et al. 2017) | x | x | x | | | | | | x | x | | | • ANN • SVM • RF | x | | | | - |
| (Ao and Qiao 2010) | | | | x | | | | x | | | | x | • LR • ARMA | | | | x | Discriminant analysis |
| (Mohsen et al. 2020) | x | | | | | | | | x | x | | | • ANFIS | x | | | | - |
| (Duo et al. 2019) | x | x | x | x | | | | x | | x | | | • J48 • LMT • IBK • NB | | x | | | - |

***Table 2*** *Acquired signals, sampling frequency and sensor details*

| Source | Signal ID | Description | Units | Range | Sensitivity | Fs (Hz) |
|---|---|---|---|---|---|---|
| Internal | TV50 | Spindle motor power feedback | W | ±2147483647 | | 250 |
| | TV51 | Active power supplied by the drive | W | ±100000 W | | |
| | TV2 | Z-axis motor torque | N | ±1000 % of the stall torque of the motor | | |
| | TV3 | Power percentage used with respect to the maximum power available in the servo system | | 0..3276.3% | | |
| | POS_(X-Y-Z) | Tool tip position | mm | | | |
| | V_(X-Y-Z) | Tool tip speed | $mm \cdot s^{-1}$ | | | |
| | ACCEL_(X_Y_Z) | Tool tip acceleration | $mm \cdot s^{-2}$ | | | |
| | JERK_(X_Y_Z) | Tool tip jerk | $mm \cdot s^{-3}$ | | | |
| | SREAL | Spindle speed | rpm | | | |
| | FREAL | Feed rate | $mm \cdot m^{-1}$ | | | |
| Kistler 9123 | F(x-y-z) | Cutting force in three axes | N | ±20e3 N | 0.5 mV/Ibf | 10e3 |
| | Mz | Torque | N·m | ±200 N·m | 0.5 mV/N cm | 10e3 |
| PCB J356A45 | ACC_(x-y-z) | Vibration in three axes | $m \cdot s^{-2}$ | ±50 g pk | 100 mV/g | 25.6e3 |
| Kistler 8152C | AE | Acoustic emissions | v | 10 dB | 48 dBref 1Vs/m | 1e6 |
| LeicaDMS1000 | Vb | Tool wear | mm | | | |

***Table 3:*** *Model evaluation metric used for regression*

| | |
|---|---|
| MAE | $\frac{1}{n}\sum_{j=1}^{n}\left\|y_j - \widehat{y_j}\right\|$ |
| RMSE | $\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \widehat{y_j}\right)^2}$ |
| $R^2$ | $1 - \frac{\sum\left(y_j - \widehat{y_j}\right)^2}{\sum\left(y_j - \overline{y_j}\right)^2}$ |

***Table 4*** *Results of algorithms using each sensor signals applying 10 folds 10 times cross-validation in complete feature space (CFS) and reduced feature space (RFS). The green square indicates the best result achieved CFS. The blue square indicates the best result achieved with the RFS and the dashed line blue square indicates those algorithms with the same mean in results to the best algorithm (p-value>0.05).*

| | | **Accelerometer** | | | **Dynamometer** | | | **Internal signals** | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **MAE** | **RMSE** | **R2** | **MAE** | **RMSE** | **R2** | **MAE** | **RMSE** | **R2** |
| **LASSO** | CFS | 0.112 | 0.319 | 0.393 | 0.039 | 0.049 | 0.581 | 0.034 | 0.058 | 0.605 |
| | RFS | 0.028 | 0.040 | 0.678 | 0.013 | 0.0155 | 0.941 | 0.019 | 0.024 | 0.844 |
| **NNET** | CFS | 0.029 | 0.040 | 0.645 | 0.018 | 0.022 | 0.870 | 0.022 | 0.027 | 0.808 |
| | RFS | **0.024** | **0.030** | **0.747** | **0.011** | **0.014** | **0.945** | 0.188 | 0.023 | 0.859 |
| **GLM** | CFS | 6.870 | 15.821 | 0.0922 | 0.804 | 0.993 | 0.103 | 0.057 | 0.119 | 0.463 |
| | RFS | 0.028 | 0.039 | 0.612 | 0.012 | 0.015 | 0.94 | 0.019 | 0.024 | 0.846 |
| **KNN** | CFS | **0.027** | **0.034** | **0.676** | 0.016 | 0.021 | 0.894 | 0.020 | 0.028 | 0.811 |
| | RFS | 0.024 | 0.031 | 0.734 | 0.011 | 0.015 | 0.941 | **0.016** | **0.021** | **0.882** |
| **M5** | CFS | 0.034 | 0.048 | 0.480 | **0.012** | **0.017** | **0.915** | **0.016** | **0.021** | **0.879** |
| | RFS | 0.023 | 0.033 | 0.740 | 0.011 | 0.015 | 0.940 | 0.017 | 0.022 | 0.877 |
| **MLP** | CFS | 0.032 | 0.040 | 0.606 | 0.031 | 0.038 | 0.683 | 0.026 | 0.033 | 0.732 |
| | RFS | 0.025 | 0.031 | 0.760 | 0.013 | 0.016 | 0.936 | 0.020 | 0.025 | 0.838 |
| **SVM** | CFS | 0.076 | 0.097 | 0.519 | 0.094 | 0.097 | 0.867 | 0.082 | 0.086 | 0.793 |
| | RFS | 0.047 | 0.061 | 0.588 | 0.074 | 0.076 | 0.931 | 0.055 | 0.061 | 0.836 |

***Table 5*** *Results obtained from bootstrapping process for each tool and the mean value.*

| | **FC_RFS** | | | **ACC_RFS** | | | **INT_RFS** | | |
|---|---|---|---|---|---|---|---|---|---|
| | MAE | RMSE | R2 | MAE | RMSE | R2 | MAE | RMSE | R2 |
| **Tool 1** | 0.0069 | 0.0078 | 0.996 | 0.0136 | 0.0159 | 0.985 | 0.011 | 0.013 | 0.992 |
| **Tool 2** | 0.0066 | 0.0074 | 0.995 | 0.0137 | 0.0155 | 0.955 | 0.0056 | 0.0067 | 0.998 |
| **Tool 3** | 0.0076 | 0.0092 | 0.989 | 0.0178 | 0.0210 | 0.955 | 0.0071 | 0.0098 | 0.989 |
| **Mean** | 0.0070 | 0.0081 | 0.993 | 0.0150 | 0.0174 | 0.965 | 0.0079 | 0.0098 | 0.993 |

***Table 6*** *Label imputation for transition phase during tool breakage a) 3rd strategy (S3) b) 4th strategy (S4). H2 correspond to the labels imputed to the transition phase of the 2nd tool while H3 corresponds to the 3rd tool.*

| a) | **H2** | **H3** |
|---|---|---|
| **FC_RFS** | 0000000001 | 0001110111 |
| **ACC_RFS** | 0000000000 | 0001111111 |
| **INT_RFS** | 0000000000 | 0111111111 |
| **[FC_RFS, ACC_RFS, INT_RFS]** | 0000000000 | 0111111111 |
| b) | **H2** | **H3** |
| **FC_RFS** | 0000000101 | 0101111101 |
| **ACC_RFS** | 0000000000 | 1001000000 |
| **INTR_FS** | 0000000001 | 0111111111 |
| **[FC_RFS, ACC_RFS, INT_RFS]** | 0000000001 | 0111101111 |
| **Imputation** | **0000000001** | **0111111111** |

***Table 7*** *Mean in MCC for each of the sensors and p values of t-test results.*

| | MCC | Internal signals | Dynamometer |
|---|---|---|---|
| Internal signals | 0.9807 | | |
| Dynamometer | 0.8818 | 7.796e-08 | |
| Accelerometer | 0.3899 | 2.2e-16 | 2.2e-16 |

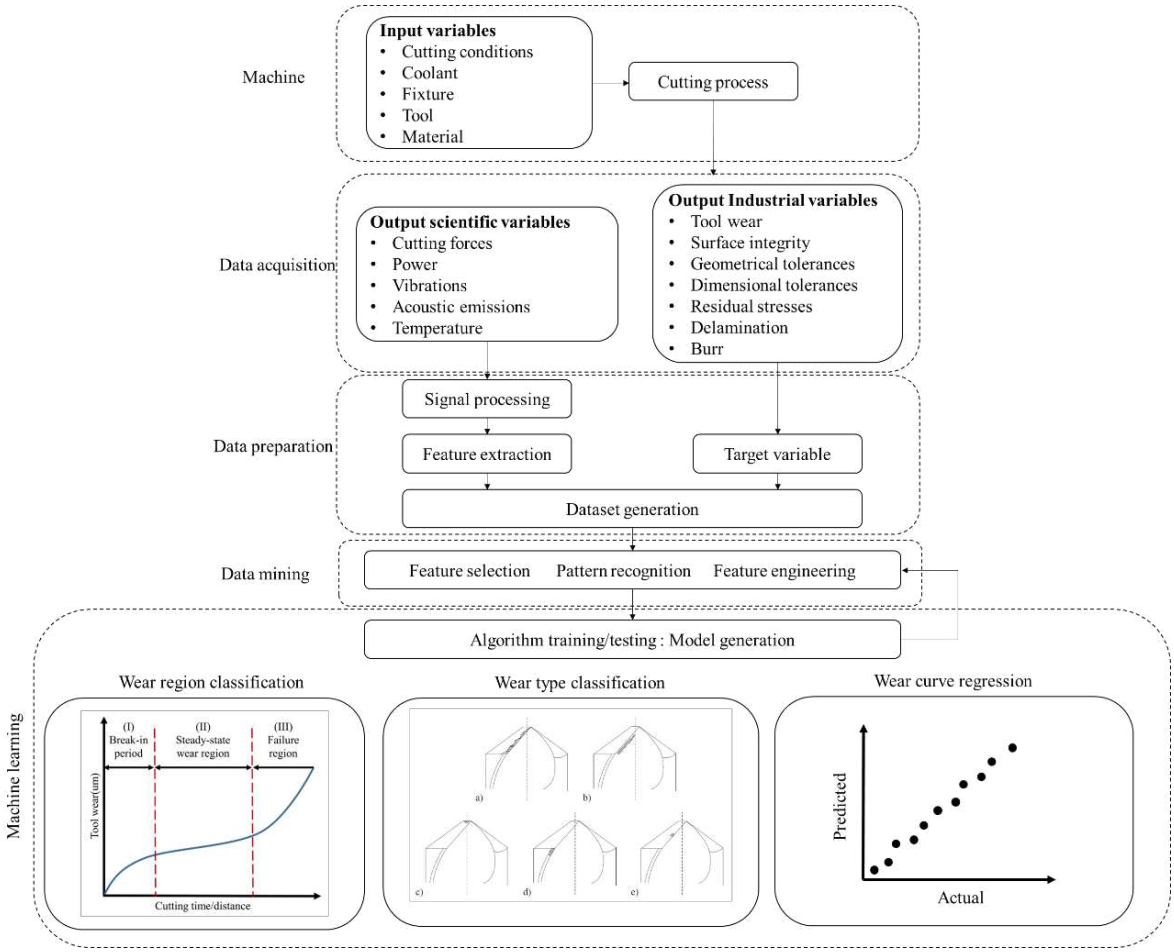



***Fig 1*** *Process monitoring framework and different approaches for tool condition monitoring*

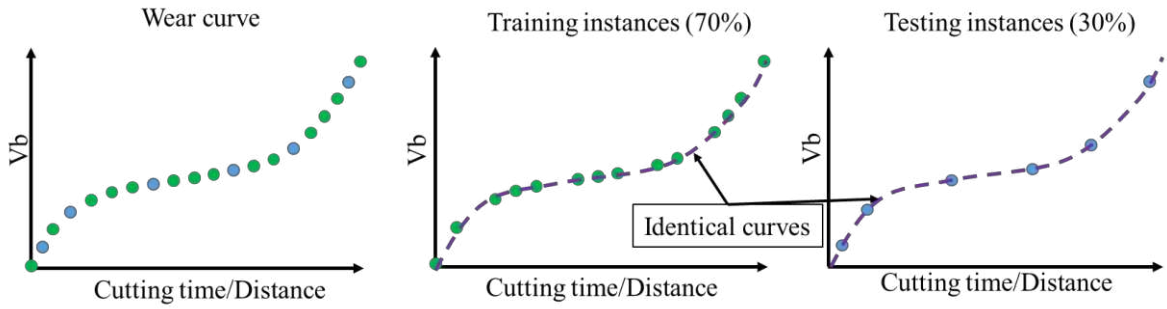



***Fig 2*** *Data leakage problem in tool wear prediction*

a) 

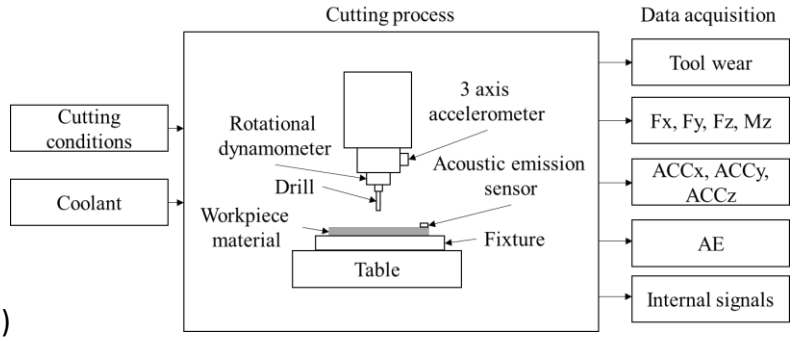
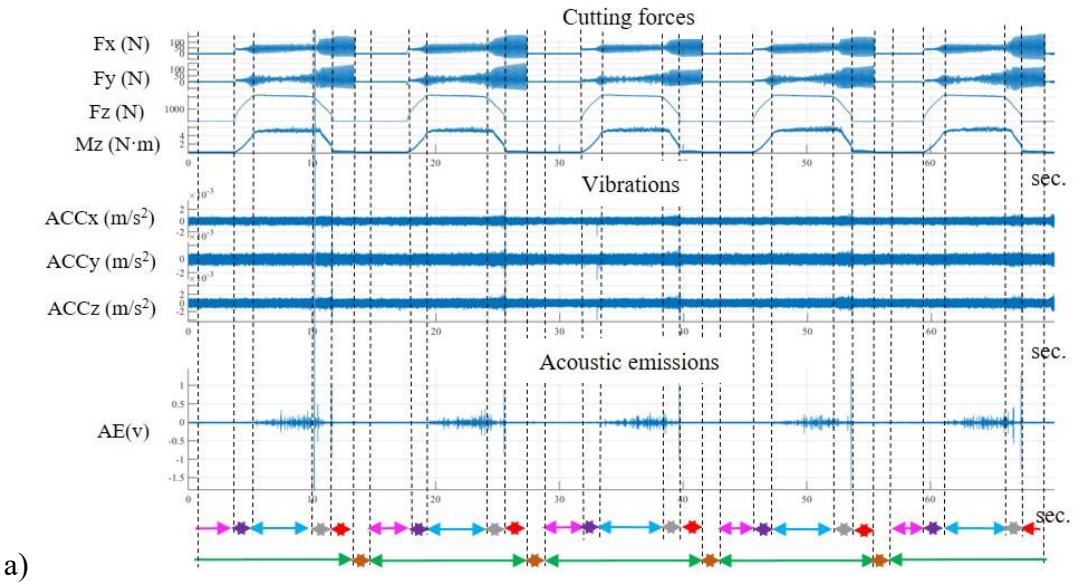
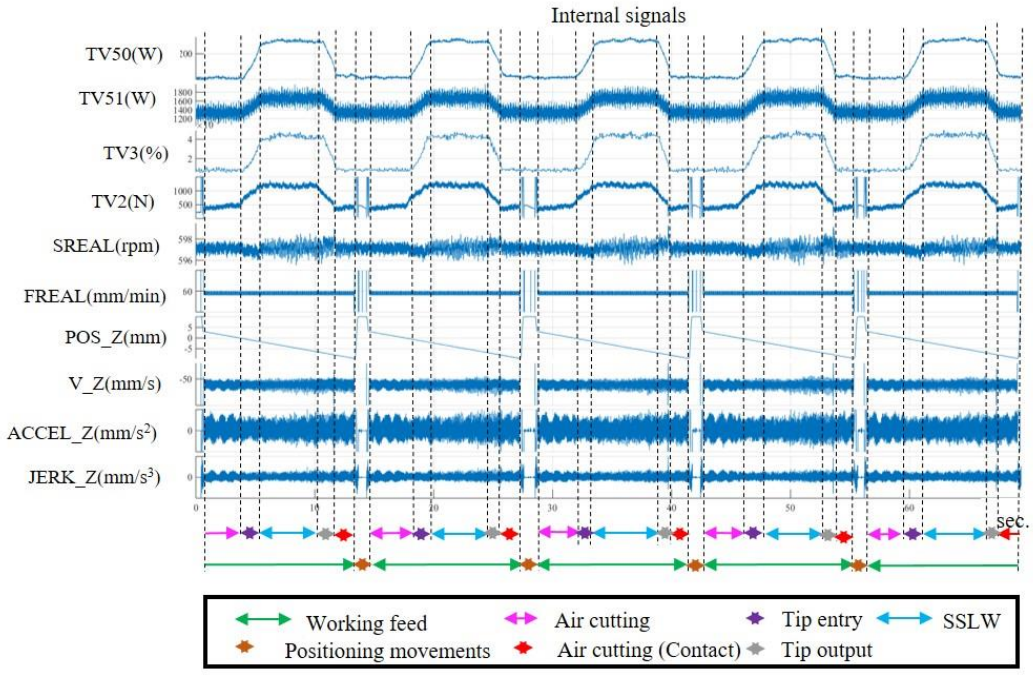
b) 


***Fig 3*** *Cutting tool used in experiments and experimental setup a) Flank face b) experimental set-up*




a)




b)

***Fig 4*** *Acquired external and internal signals and applied segmentation a) External signals b) Internal signals*

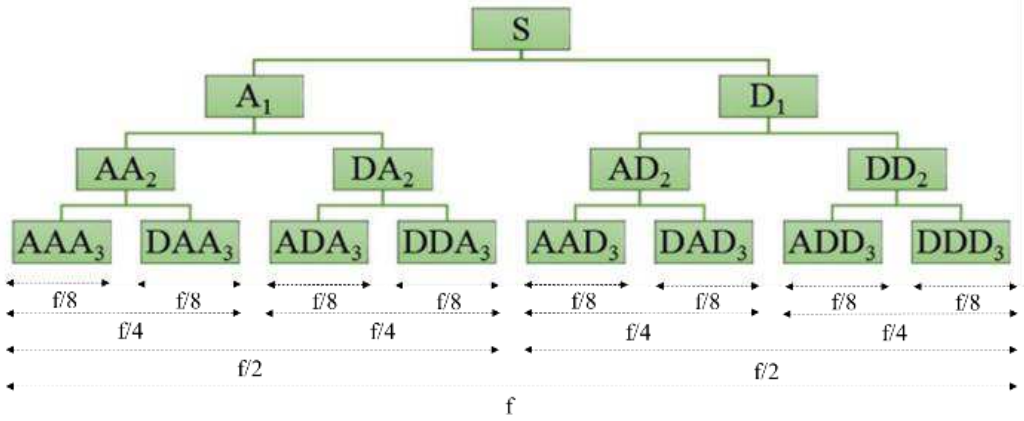



***Fig 5*** *Decomposition wavelet tree up to the third level. S original signal* (Segreto, Karam, and Teti 2017)

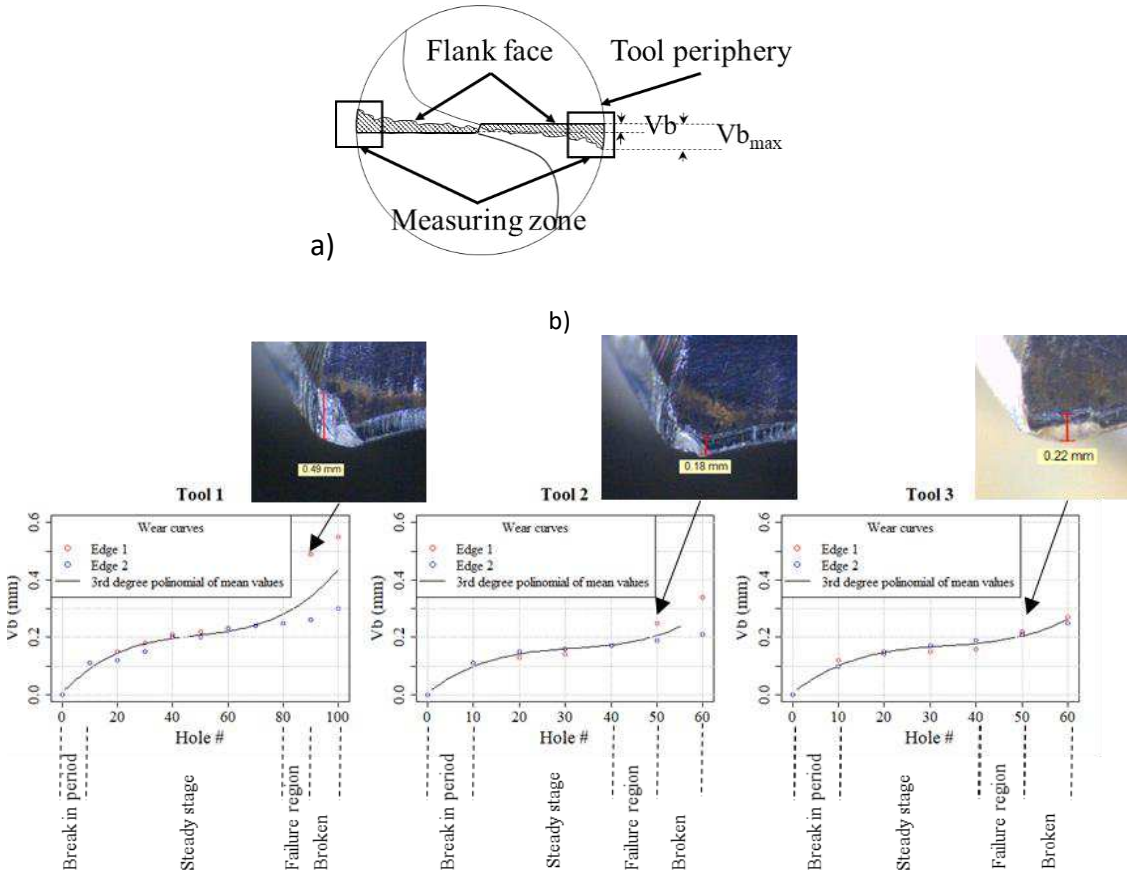



***Fig 6*** *Tool wear curve a) Tool wear measuring point on $Vc_{max}$ section b) Tool wear measurements and third-degree polynomial for each repetition and tool breakage measured hole for each repetition*

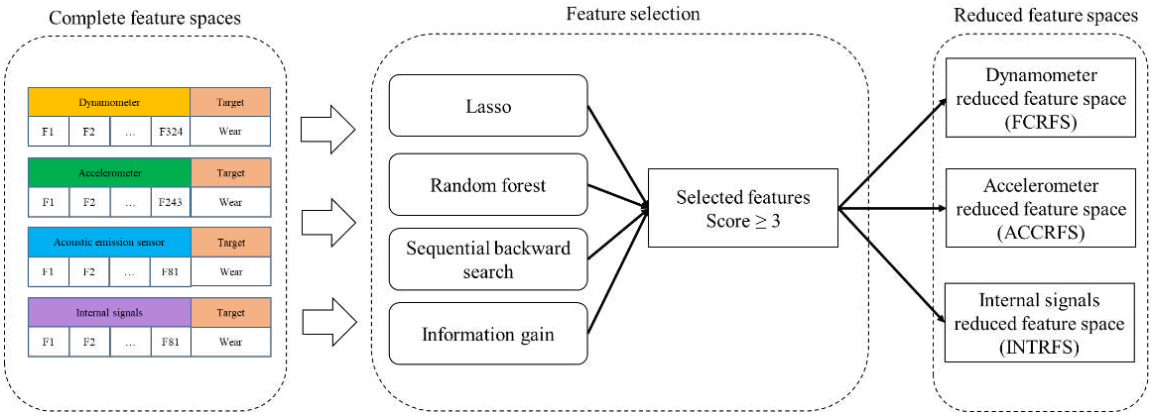



***Fig** 7 Feature selection process*

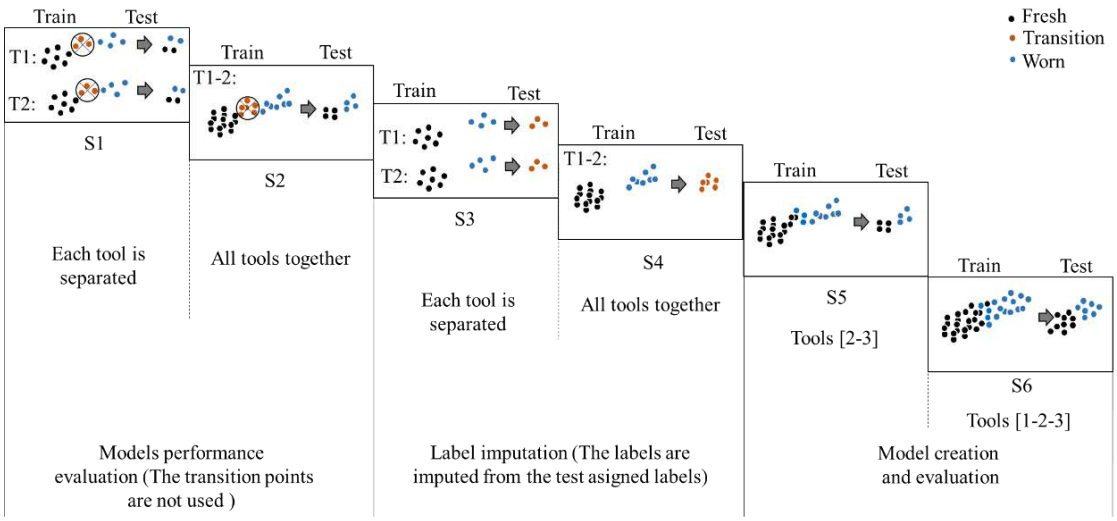
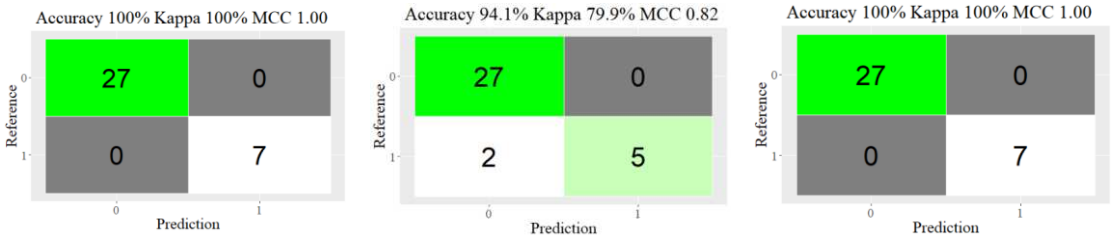
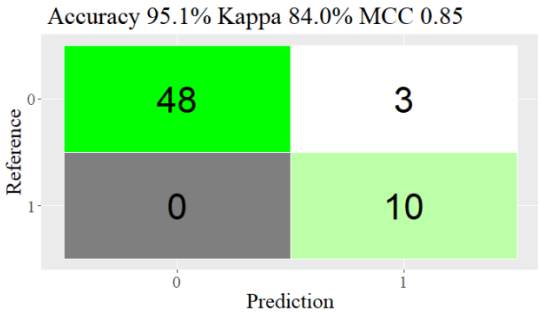



***Fig 8*** *Followed methodology for tool periphery breakage detection*

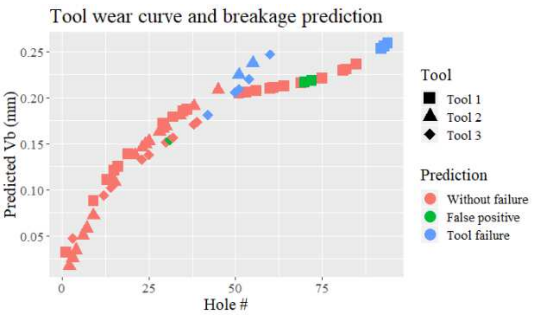



a) b) c)

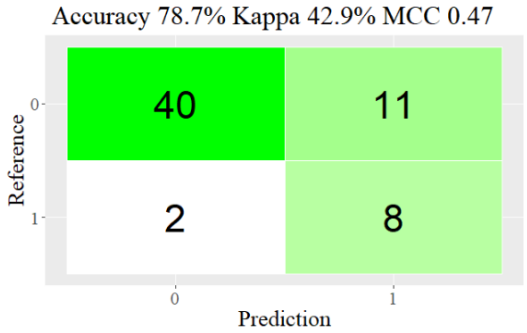
***Fig 9*** *Confusion matrices for each sensor for S5 (strategy 5). a) FC_RFS b) ACC_RFS c) INT_RFS*



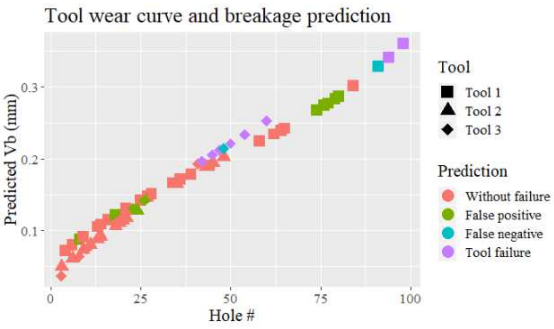

a) Confusion matrix with FC_RFS



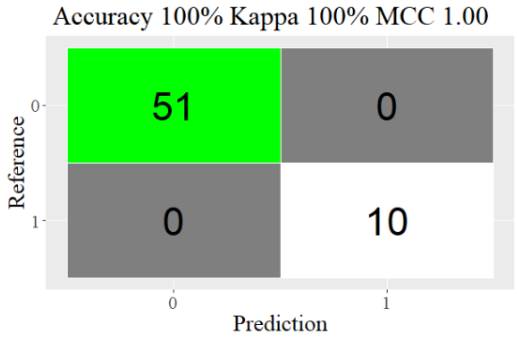

d) Wear curve/breakage prediction with FC_RFS



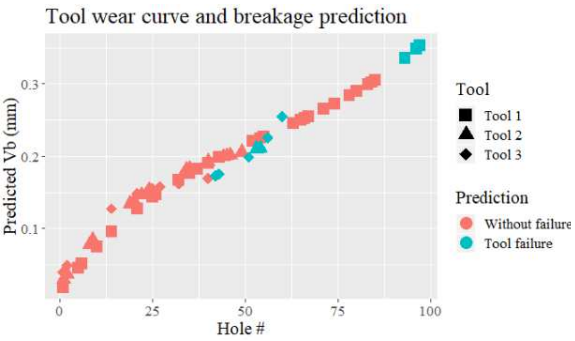

b) Confusion matrix with ACC_RFS




e) Wear curve/breakage prediction with ACC RFS




c) Confusion matrix with INT_RFS




f) Wear curve/breakage prediction with INT_RFS

***Fig 10*** *Results obtained in S6 (strategy 6) for each of the sensors in testing subset*

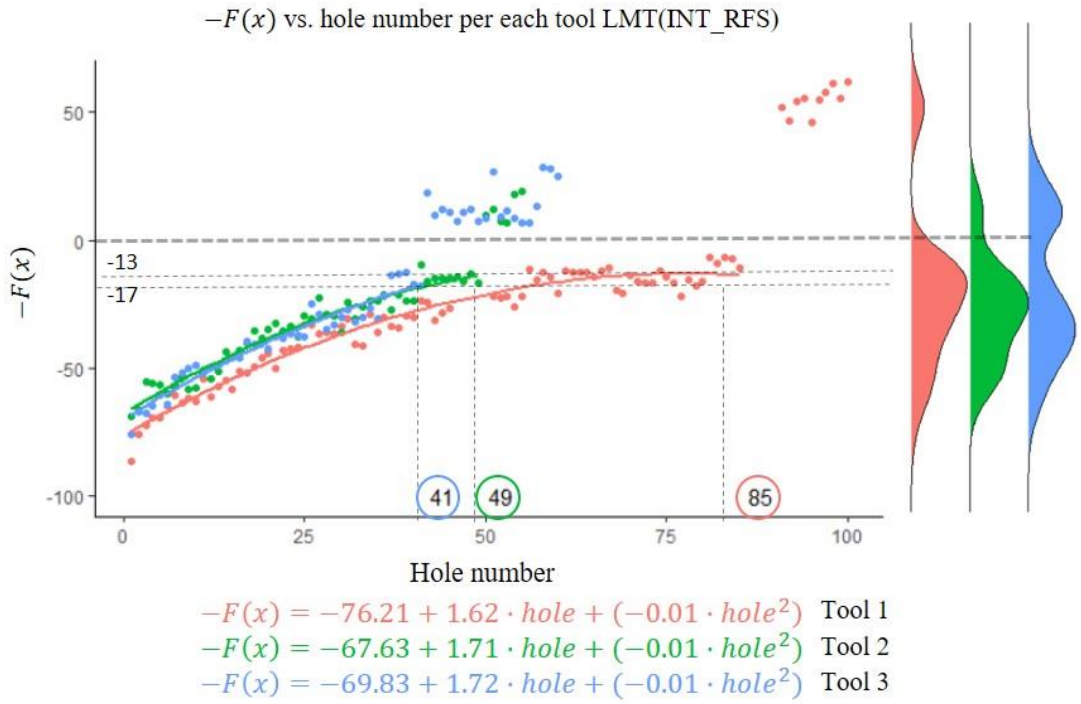



***Fig 118:*** *Logistic model function based on INT_RFS. On the horizontal axis are marked the holes where tool breakage has been observed for each tool. On the vertical axis are marked the F(x) values at which tool breakage has been observed*