Noname manuscript No.
(will be inserted by the editor)

The capacity of statistical features extracted from multiple signals to predict tool wear in the drilling process

Aitor Duo[1,2] · **Rosa Basagoiti**[1] · **Pedro J. Arrazola**[2] · **Javier Aperribay**[2] · **Mikel Cuesta**[2]

Received: date / Accepted: date

Abstract Industrial processes are being developed under a new scenario based on the digitalization of manufacturing processes. Through this, it is intended to improve the management of resources, decision-making, production costs and production times. Tool control monitoring systems (TCMS) play an important role in the achievement of these objectives. Therefore, it is necessary to develop light and scalable TCMS that can provide information about the tool status using the signals provided by the machine. Due to the lack of this type of systems in industrial environments, this work has two main objectives. First, the predictive capacity of statistical features in the time domain of internal and external signals for the prediction of tool wear in drilling processes was analysed. To this end, a methodology based on automatic learning algorithms was developed. Secondly, once the most sensitive signals to tool wear were identified, algorithms with signals of a certain tool geometry were trained and a model was obtained. Then, the model was tested using signals from two different tool geometries. The experiments were carried out on a vertical milling machine on a steel with composition 35CrMo4LowS under pre-established cutting conditions. The results show that the most sensitive signals to monitor the tool wear in the time domain are the feed force (external) and the z-axis motor torque (internal). The models created for the fulfilment of the second objective show a great capacity of prediction even when dealing with tools with different geometrical characteristics.

Keywords Tool wear · Drilling · Machine learning · Tool condition monitoring

✉ Aitor Duo
E-mail: aduo@mondragon.edu

Rosa Basagoiti
E-mail: rbasagoiti@mondragon.edu

Pedro J. Arrazola
E-mail: pjarrazola@mondragon.edu

Javier Aperribay
E-mail: japerribay@mondragon.edu

Mikel Cuesta
E-mail: mcuesta@mondragon.edu

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

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

1 Introduction

The digitalisation of industrial processes is leading to the creation of automatic control systems that can assist in decision making during machining. As the quality requirements of some industrial sectors are increasing (automotive, railway, naval, aeronautical, appliance or construction industries) the decision making in manufacturing processes is becoming more critical. Machining processes are a key factor in the manufacture of different parts for the above mentioned sectors, so it is necessary to develop intelligent systems to improve decision making.

The drilling process is usually carried out in the final steps of a machined part, in these stages a nearly finished part could be damaged [1], stopping the production line and consequently increasing the production cost. Therefore, the drilling process is one of the most critical machining operations for sectors that manufacture parts that rely on high quality holes. The hole quality is affected by tool wear, which is one of the phenomenon that affects hole surface integrity and hole dimensions [2]. In industrial environments the tool is changed based on a conservative estimation of tool life to

ensure product surface integrity, or to prevent premature tool failure [3]. For the above reasons an unmanned production is only possible with a tool control monitoring system [4].

Tool condition monitoring systems (TCMS) can be based on direct or indirect measurements. Direct measurements are considered more expensive than indirect measurements because they require more complex equipment and higher skilled workers. Thus, the trend is to use indirect measurements for tool wear detection. To determine tool life, the use of direct measurements are more extended [1, 5–8], to observe the evolution of the degradation of the tool both direct and indirect techniques are used [3, 9–14], and for tool wear detection purpose only indirect measurements are used [15].

In industrial environments the use of sensors is not widely extended, this is due to the vulnerability of these sensors to external noise, common in these environments. Some of sensors can be expensive or invasive. For these reasons external signals are more widely employed for research purposes than for industrial applications. There is requirement, therefore, to replace the external signals commonly used for fault detection with the internal signals provided by Computer numerical control (CNC) machine. This would facilitate the creation of non-invasive and easy-to-install fault detection systems.

In recent years, sensor technology has improved and the amount of data generated has increased. Therefore, machine learning techniques are used to enhance the machining process [16]. Machine learning algorithms allow the extraction of unknown information from previously collected data. These algorithms extract useful patterns (models) to be used with new data to make future predictions regarding the machining process.

This work has two main objectives. The first objective is to identify the most sensitive signals to predict tool wear. For this purpose, a comparison of the most widely used signals for tool wear detection was made from the point of view of machine learning. Two drill geometries were chosen and the same methodology was applied to the data recorded for each type of drill bit. The methodology used allows to identify the most sensitive signals for both drill bit configurations using the recorded raw data. The second objective is to predict tool wear of one tool geometry using the model created from the data from another tool geometry. Once the most sensitive signal and the most accurate algorithm were identified, the data acquired using one of the drill bits was used to predict the tool wear of the other drill bit configuration.

2 Related work

Tool state monitoring systems are very extended with research purposes. Regarding the industrial environments there are few applications that allows tool state monitoring in real time. In many cases it is necessary to install external sensors with the requirement of the sensor setup. The most extended signals for tool wear monitoring are vibrations, acoustic emissions, cutting forces and sound pressure. The spindle cutting power also is employed for tool wear detection.

The acquisition of signals to obtain relevant information about the cutting process results in large volumes of data. Therefore, the management and extraction of information must be done in an efficient way, selecting the features that best represent the physical quality of the cutting process to be controlled.

Tool geometry is a relevant factor in the aggressiveness of the cutting process [17]. That is why tool selection is an important aspect of the machining process. In any case, the modification of the cutting conditions is cheaper than changing the tool to increase the tool life. The system developed could support this kind of decision.

Returning to the signals used for such a system, an accelerometer is an easy to install sensor, so the vibration signals are widely extended for tool wear prediction. These types of signals represent the external variations of tool wear [18]. Regarding the statistical features of vibration signals for the detection of tool wear, several studies show the use of different signal features to meet the same objective. Kim and Choi [19] employed the amplitude of vibration signals as tool wear detection feature in turning operations, they stopped the process when the amplitude exceeded a preestablished limit. In contrast, Dimla Snr. [20] said that time domain features were more correlated with the modification of the cutting conditions than with tool wear. Frequency domain features of vibration signals showed better correlation with tool wear. Rmili et al. [21] shows how to identify the three phases of the tool wear (breaking-in, steady state, acceleration) using the mean power of vibration signals in turning operations, they report that the vibration in the x-axis is the most relevant signal. However, in the same way Krishnakumar et al. [22] used different time domain features to create a classification tree to classify three different levels of tool wear using vibration on the z-axis as the most relevant signal. Therefore, although the accelerometer is an easy to install sensor and not too invasive, the clamping, cantilevering and dynamics of the machine affects to the measurements made.

The acoustic emissions (AE) are transient elastic waves created by the plastic deformations and chip flow in machining operations [18]. AE signals are also widely employed for tool wear monitoring, the constant friction between tool and workpiece produces continuous signals while micro-cracks create transient signals [1]. Diniz et al. [12] identify the dispersion of RMS values of AE signals as a relevant feature for tool wear detection since tool wear increases the friction between the tool and the workpiece. As a result the amplitude of the AE signal increases, thus the RMS value shows

a higher dispersion. By contrast Bhuiyan et al. [13] say that this technique based on the RMS values of AE signals may have problems with loss of information about transient signals created by fractures on the tool surface. Locating the source of the continuous emissions is a difficult task. Therefore, the most common application is to detect transient signals [23]. In drilling processes with small diameter drills it is difficult to detect tool wear by using other types of sensors, so the use of AE signals can be of great help in this type of operation [24].

Sound pressure is an additional option for tool wear detection. Measurement of sound pressure is versatile, easy to acquire and does not require direct contact with any element of the process. However, the variety of tools, the variability of cutting conditions and the conditions of production environments negatively affect the measurements made. Kothuru et al. [25] developed a TCMS based on machining operation audible sound signal using machine learning techniques. Pre-processing treatments are not sufficient to remove noise from measurements, so there are misclassifications. Therefore more sophisticated treatments are needed to remove noise from the signal. Seemuang et al. [26] tried to predict tool wear with the spindle audible sound signal, concluding that there is no relationship between the spindle audible sound signal spectrum bandwidth and tool wear. Nevertheless, the audible signal from the spindle makes it possible to identify changes in cutting conditions so that with this technique it is possible to acquire knowledge of the process and even combine it with some other type of instrumentation.

Cutting processes are operations with high plastic deformations and high stresses leading to a high cutting forces. There are different parameters that affect the cutting forces, the most relevant are cutting conditions, tool geometry and workpiece material properties. As the tool wears down the friction between the tool and workpiece is increased. As a result, the microgeometry of the tool is changed causing higher cutting forces [7, 17]. However, it is necessary to take into account the cutting conditions as well as the workpiece material properties to avoid false positives [11]. Stavropoulos et al. [14] proposes to measure the spindle current, it is less invasive and less sensitive to external noises.

Although all analyses are based on the extraction of features from the acquired signals, the use of automatic learning algorithms can help to understand and identify the most relevant features. Learning algorithms allows to extract unknown information from previously collected data. This is done by creating models capable of predicting past phenomena with new data. Different methods of detecting tool wear based on automatic learning have been seen in the literature. Corne et al. [27] study the possibility of obtaining a threshold before the catastrophic failure of the tool when drilling holes in Inconel using artificial neural networks based on data from the spindle power. Concluding that the tool is about to break at 120%-130% of the measured power when the first hole was drilled. However, it is only valid for the tool, material and cutting conditions used. Lee et al. [28] proposed a TCM system for milling operation based on RMS feature of cutting forces as input to Response Surface Methodology (RSM) algorithm obtaining good results. Even if good results are achieved, it is necessary to install a dynamometer for the measurement of cutting forces. The Suport vector machine (SVM) algorithm is an option for automatic learning as it represents a specific number of statistical features extracted from the signals in a larger dimension than the original feature space. Kothuru et al. [25] are based on audible signals from the cutting process for the extraction of statistical features that allow different levels of tool wear to be classified using the above mentioned algorithm. Pattern recognition, artificial neural networks, fuzzy logic, genetic algorithms and decision trees are widely used for predicting different phenomena. For the detection of tool wear many have been used with positive results.

As for commercial systems, such as the one offered by ARTIS [29]. Generally, they work like black boxes and need historical data in order to be trained to detect tool wear. However, if they are going to be employed in different cutting conditions, they need to be trained again. In the literature review there are not many applications that once the system has been trained in given cutting conditions (cutting speed, feed, tool geometry), the system will properly work not only in these cutting conditions but as well in different ones (in a limited process window) to detect tool wear.

During the literature review, most of the consulted methods only analyse a certain type of tool geometry and cutting condition, testing the developed model only to data gathered under the same conditions. The methodology proposed in this paper, evaluates the capability of a model trained on one tool geometry and cutting conditions to predict tool wear for other tool geometry under different cutting conditions, widening the scope of application for the use of a developed model. In addition, the methodology used during the development of these models, facilitates the comparison of the accuracy of the different signals at the time of predicting the tool wear for different geometries and cutting conditions using time domain statistical features.

In the following sections, first the methodology used is explained, detailing the setup used in this work. It is also explained the extraction of features from the signals acquired and their subsequent use with automatic learning algorithms. Afterwards, the results obtained are discussed and finally the conclusions obtained are shown.

3 Methods

In this section, the experimental setup is presented and the methodology used to perform sensitivity analysis of obtained signals from the point of view of machine learning is explained. A complete experimental setup was used to acquire the maximum amount of data of the drilling process. Thus, several sensors were installed to obtain cutting forces, vibrations, sound pressure and acoustic emissions in addition to acquiring signals from the CNC machining centre.

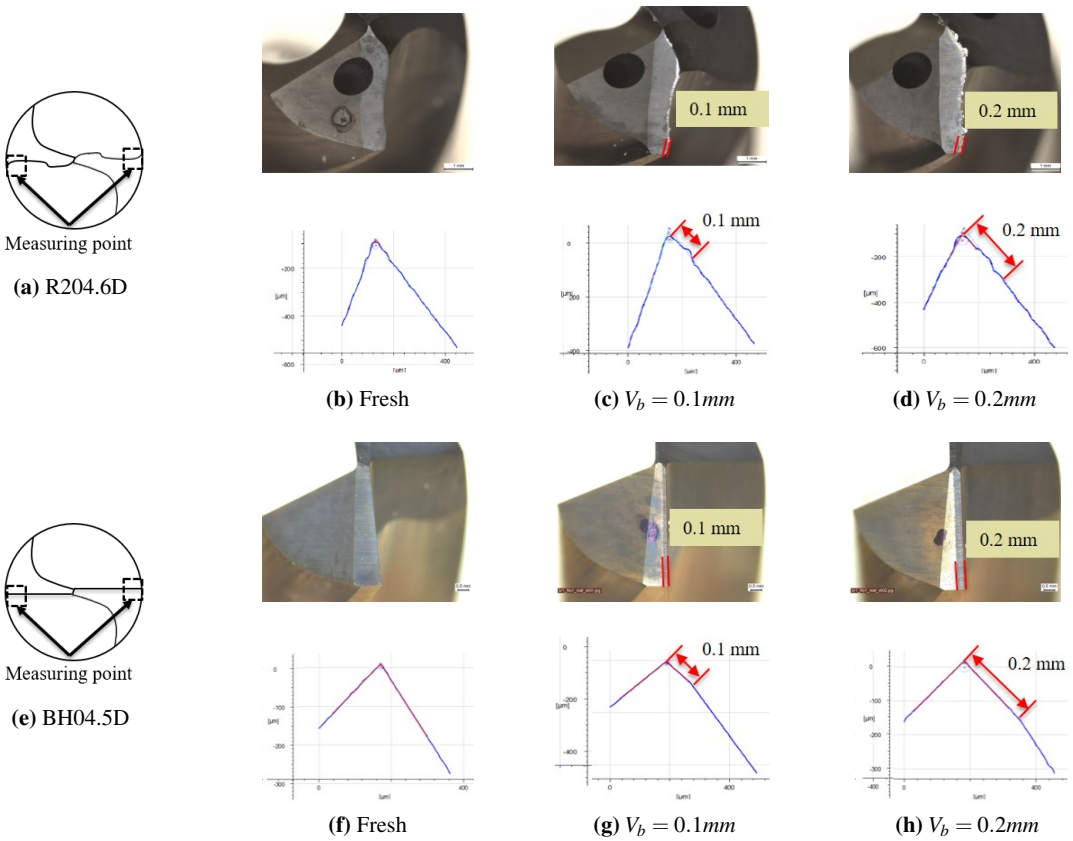
3.1 Experimental setup

The experiments were performed on a Lagun vertical CNC milling machine. The workpiece material was a steel of composition 35CrMo4LowS. As can be seen in Fig. 1 (a-e) two types of drills were used, Kendu R204.6D curved edge drill with a helix angle of 30°, and BH04.5D straight edge drill with a helix angle of 15°, both of ⌀8mm diameter. In the same image Fig. 1 (b-c-d-f-g-h) can be seen the different levels of wear considered in this work with their cutting edge profile.

Two types of drills (R204.6D and BH04.5D) were employed with three different tool condition states: fresh, tool flank wear of 0.1mm and tool flank wear of 0.2mm. While in R204.6D the tool flank wear of 0.1mm and 0.2mm was generated making preliminary tests aside of the experimental test presented in this paper, in BH04.5D tool flank wear was generated by sharpening.

To keep track of the changes in the microgeometry of the tools during the execution of the tests, all the tools were measured before and after performing the trials. These measurements were made in terms of cutting edge geometry. Thus, the flank wear was measured in a Leica DMS1000 macroscope and the cutting-edge radius and the cutting-edge angle were measured in an Alicona IFG4 3D profilometer. The measurements were taken at the periphery of the cutting edge, where the cutting speed reaches its maximum value as can be seen in Fig. 1 (a-e). No standard has been found for the measurement of tool wear in drilling processes so the measurements have been based on ISO 8688:1989 (tool-life testing in milling) and ISO 3685:1993 (Tool life testing with single-point turning tools) standards.

Table 1 shows the cutting conditions, tool geometry and tool identification. The main modified parameter was the



(a) R204.6D **(b)** Fresh **(c)** $V_b = 0.1mm$ **(d)** $V_b = 0.2mm$

(e) BH04.5D **(f)** Fresh **(g)** $V_b = 0.1mm$ **(h)** $V_b = 0.2mm$

Fig. 1 Drill bit geometries a-e) General geometry b-c-d-f-g-h) Considered tool wears: fresh, Vb=0.1 and Vb=0.2 with their cutting edge profile at measuring point

Table 1 Cutting conditions, cutting edge geometry and number of holes related to drill number

Drill identification		Cutting edge geometry						Cutting conditions				
		$V_b(mm)$		β (°)		$r(\mu m)$						
Drill n°	Drill ID	$\overline{V_b}$	$\sigma^2 \cdot 10^{-4}$	$\overline{\beta}$	σ^2	$\bar{r}$	σ^2	$Vc(m \cdot min^{-1})$	$f_n(mm \cdot rev^{-1})$	$L(mm)$	$\varnothing(mm)$	N° of holes
1	R204.6D_1_0	0	0	54.9	0.24	15	2	100	0.15	16	8	5
2	R204.6D_2_0	0	0	53.3	0.6	15	2	100	0.15	16	8	10
3	R204.6D_1_01	0.08	2.2	66.7	12.5	5.5	0.5	100	0.15	16	8	5
4	R204.6D_2_01	0.08	3.1	61.1	2.6	4	0	100	0.15	16	8	5
5	R204.6D_1_02	0.15	5.8	63.8	0.72	3.5	0.5	100	0.15	16	8	5
6	R204.6D_2_02	0.17	26.1	62.7	11.5	11	0	100	0.15	16	8	5
7	BH04.5D_1_0	0	0	79.6	0.32	6	2	40	0.07	5	8	5
8	BH04.5D_2_0	0	0	79.6	0.12	7.5	0.5	40	0.07	5	8	10
9	BH04.5D_1_01	0.14	0.5	91.5	0	7.5	4.5	40	0.07	5	8	5
10	BH04.5D_2_01	0.1	0	92.1	0.02	10.5	4.5	40	0.07	5	8	5
11	BH04.5D_3_01	0.12	2	91.5	0.08	9	2	40	0.07	5	8	5
12	BH04.5D_4_01	0.12	0	90.6	0.08	11	0	40	0.07	5	8	5
13	BH04.5D_1_02	0.28	0	90.2	0.02	11	18	40	0.07	5	8	5
14	BH04.5D_2_02	0.24	0.5	90.2	0.125	11.5	4.5	40	0.07	5	8	5
15	BH04.5D_3_02	0.2	0	90.5	0.125	8	0	40	0.07	5	8	5
16	BH04.5D_4_02	0.28	0.5	89.7	0.4	10.5	0.5	40	0.07	5	8	5
Total												90 holes

flank wear, which in this case is considered an input parameter to the cutting operation. Due to the lack of precision in the preparation of the tools, a unique value was assigned to represent the tool wear. Each of the tools was assigned an approximation of the wear value closest to those considered in this work (0, 0.1, 0.2 mm). The cutting speed, feed rate and hole depth were the same for each type of drill bit. A total of 90 holes was made.

To obtain a wide range of data from the drilling process, different sensors were installed in addition to the measurements of the signals provided by the CNC machining centre. Thus, a dynamometer, acoustic emission sensor, accelerometer and microphone were mounted on the setup. Fig. 2 shows the location of the installed sensors and the acquisition system used during the experiments. The internal signals were acquired from the CNC with a sampling frequency of $250Hz$, that is the frequency of the CNC running cycle. Furthermore, the dynamometer amplifier output signals were connected to the analogical input of the CNC. For this reason, the signals of the dynamometer were also recorded at $250Hz$. The acoustic emission signal was acquired at $1MHz$ and was mounted on the workpiece using a magnetic clamp. Finally, the vibration in the Z axis direction and sound pressure were recorded at $50kHz$. In [30] it can be observed the different relationships between the signals acquired in the present work. The duration of the holes varies depending on the length and the cutting conditions used to perform the tests. Those holes made with curved edge tools have lasted 1.6s while those made with straight edge tools have lasted 2.6s.

Table 2 shows the recorded signals, specifying where the signal was obtained, the type of the signal (Internal, external or measured wear), the description, the units in which the signal was measured and the sampling frequency. As can be seen, a wide variety of signals have been collected. In total there are 14 signals, of which 8 are internal signals, 5 are external signals and the last one is the tool wear measured directly before and after the tests.

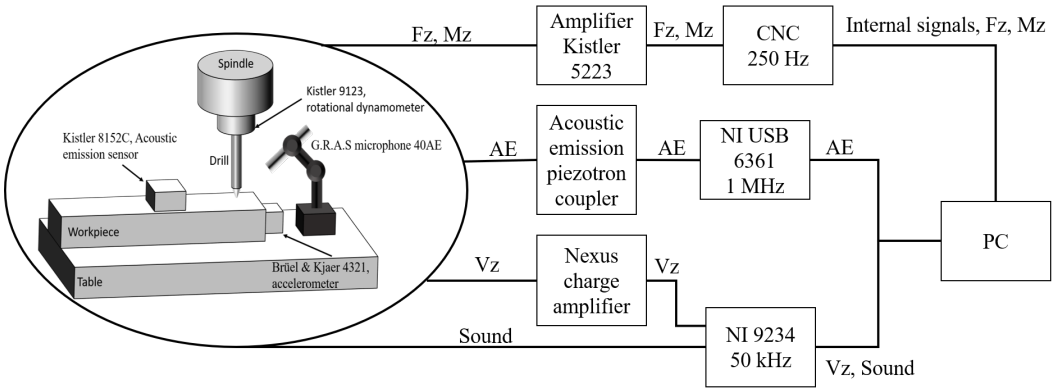


Fig. 2 Experimental setup and signal acquisition system used for the experiments

Table 2: Internal signals (I), external signals (E) and measured tool wear (W), and details about used sensors

N	Signal ID	Sensor type	Type	Description	Units	Sampling freq. (Hz)	Range	Sensitivity	Weigth
1	TV50	CNC	I	Spindle motor mechanical power	W	250	± 2147483647 W	×	×
2	TV51	CNC	I	Spindle motor electrical power	W	250	±100000 W	×	×
3	TV2	CNC	I	Z axis motor torque	N	250	±1000 % of the stall torque of the motor	×	×
4	V(X-Y-Z)	CNC	I	Tool speed in three axes	$mm \cdot s^{-1}$	250	×	×	×
5	ACCEL(X-Y-Z)	CNC	I	Tool acceleration in three axes	$mm \cdot s^{-2}$	250	×	×	×
6	JERK(X-Y-Z)	CNC	I	Tool jerkin three axes	$mm \cdot s^{-3}$	250	×	×	×
7	n	CNC	I	Feed rate	rpm	250	×	×	×
8	av	CNC	I	Feed rate	$mm \cdot m^{-1}$	250	×	×	×
9	Fz	Kistler 9123	E	Thrust force	N	250	± 20 $\cdot 10^3$ N	0.5 mV/Ibf	3Kg
10	Mz	Kistler 9123	E	Torque	$N \cdot m$	250	±200 $N \cdot m$	0.5 mV/N cm	3Kg
11	SP	G.R.A.S 40AE	E	Sound pressure	Pascal	$50 \cdot 10^4$	± 2dB (freq. range)	50 mV/Pa	6.50 g
12	V_z	Brüel & Kjaer 4321	E	Vibration in Z axis	$m \cdot s^{-2}$	$50 \cdot 10^4$	± 500 g	10 pC/g	55 g
13	AE	Kistler 8152C	E	Acoustic emissions	V	10^6	± 10 dB	48 dBref 1Vs/m	29 g
14	Vb	Leica DMS1000	W	(0, 0.1, 0.2) measured Vb	mm	×	×	×	×

The concept of power is same in the context of mechanical as well as electrical engineering applications. The mechanical power (TV50) is measured by the rate at which work is done. On the other hand, electrical power (TV51) is measured by the rate at which electrical energy is transformed. The concept differs in electrical and mechanical engineering because the expression of the work done or energy conversion is different in the two disciplines. In theory, electrical power should be transferred completely to mechanics. In practice, taking into account the efficiency of the machine, there will be power losses because of frictional factors.

3.2 Sensitivity analysis: machine learning approach

The analysis of the predictive capacity of the statistical features extracted from the signals was carried out using machine learning techniques. To that end, the following statistical features were calculated for each signal in time domain: mean, rms, standard deviation, maximum, minimum, kurtosis, skewness, variance and coefficient of variation. These statistical features explain the shape and distribution of the signal, so automatic learning is applied to them.

First of all, the data which constituted the dataset was selected. The rows of the dataset are called instances while the columns are attributes. Table 3 shows a generalized dataset of any of the acquired signals. In addition to the calculated statistical features some process parameters were added to the datasets: cutting speed (V_c), feed per revolution (f_n), drilled depth (L), the cutting-edge angle (β) and the cutting edge radius (r). All these attributes were used to predict the class, which is the tool wear (V_b) measured before the experiments. Vc and f_n variables were calculated acquiring a_v and n signals, which are penetration rate and spindle speed respectively, obtained from the drilling process. The mean of the signals was used to calculate V_c and f_n.

In order to evaluate the precision of each of the different acquired signals, separately, one different model has been trained for each one of the collected signals. Therefore, a dataset is related to only one signal. In total eleven datasets were created for each type of drill, one for each of the acquired signals. Among them, the ones with name TV50, TV51, TV2, Mz, Fz, SP and V_Z were created with 14 attributes and the class. In contrast V(X-Y-Z), ACCEL(X-Y-Z) and JERK(X-Y-Z) were built with 32 attributes and the class. This is due to the fact that signals composed of more than one component were put together in a dataset. The R204.6D drill type datasets had 35 instances while the BH04.5D type drill datasets had 55, having a total of 90 instances, one for each of the holes made.

Table 3 Generalized example of a signal dataset. Constituted by statistical features extracted from the signal, process input parameters, and the class

Statistical features									Process parameters					C
mean	rms	std. desv.	max.	min.	kurt	skew	var	coef. var	V_c	f_n	L	β	r	V_b
Hole1: 1st instance Hole2: 2nd instance ... HoleN: Nth instance														

In Fig. 3 can be seen different signals acquired depending on the wear of the tool used. The first row shows the Z axis motor torque (TV2) and the thrust force (Fz) for the different Vb levels considered in this work. The second row shows the mechanical power (TV50) of the head and the cutting torque (Mz). The graphs show signals for 3 different holes made with different flank wear. Specifically, the holes shown belong to the 1st hole of tools #7, #9 and #13. In order to create the dataset mentioned above, the parts belonging to the approach of the head to the work-piece material and the spindle retraction have been eliminated. The signal has been taken into account from the beginning of tool penetration up to the machined length.

Using these datasets, machine learning algorithms were used within a 10 folds cross validation process. Due to their differences in approximation at the time of generating the models, J48 [22, 31], LMT (Logistic Model tree) [24], IBk [22, 31] and NaiveBayes available in the Weka platform

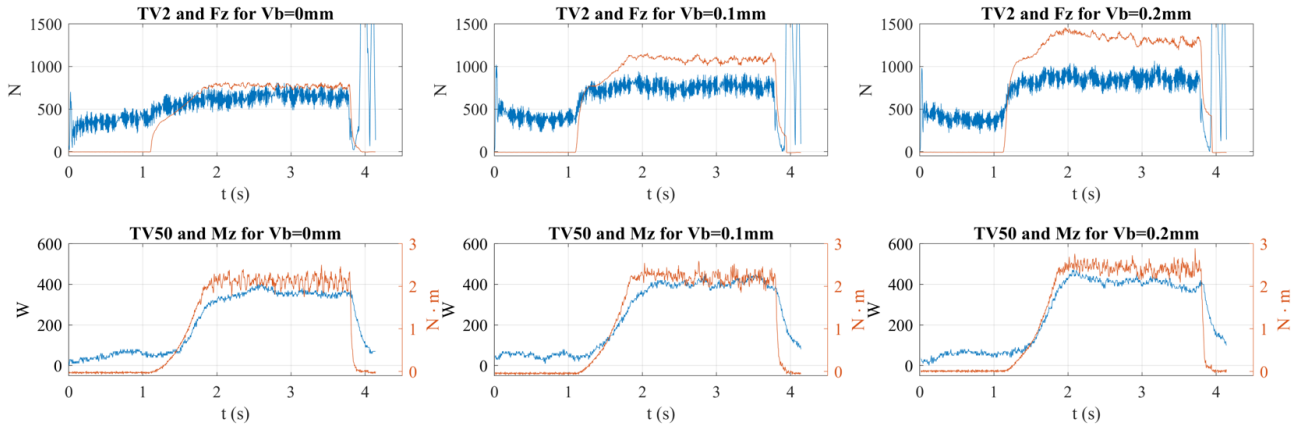


Fig. 3 Signals for different levels of Vb, TV2 and Fz, TV50 and Mz of BH04.5D type drill (1^{st} hole of Drill n° 7, 9 and 13, Table 1).

were tested. This process consists of performing 10 iterations where in each iteration the data is partitioned in 10 subsets. Then, the analysis (also called training) is performed in nine of these subsets, while the validation or testing uses the remaining subset. It is repeated 10 times rotating the validation subset until every instance has been validated once. Same validation process was applied to all datasets.

Further statistical significance of the differences in accuracy of the results of these algorithms was also carried out at three level of difficulty. From the simplest classification to the more complex one, the three difficulty levels are listed below:

- **V1:** In the first version only the holes made with new drill bits ($V_b = 0$) and those with the maximum flank wear ($V_b = 0.2mm$) were considered. But, still the classification was in binary mode (false if the hole was made with a new drill bit, true if the hole was made with a worn drill bit)
- **V2:** Apart from holes made with new drill bits ($V_b = 0$) and those with the maximum flank wear ($V_b = 0.2mm$), instances corresponding to holes made with $V_b = 0.1mm$ drill bit were added. In this case, the classification was also carried out in binary mode (false if the hole was made with a new drill bit, true if the hole was made with a worn drill bit).
- **V3:** All flank wear levels were considered, $V_b = 0mm$, $V_b = 0.1mm$ and $V_b = 0.2mm$ in a multiclass classification process.

After the cross-validation process had been applied, a T-test was performed. The sensitivity analysis was performed using the WEKA software experimenter. For the paired T-test, the first signal is the dominant signal, while all other signals are compared against the dominant signal. Therefore, the most accurate signal was chosen as the dominant one. The objective of this test is to obtain the statistical significance of the signals analysed in this work with respect to the most accurate signal to predict tool wear.

3.3 Testing with different drill bits

To test the ability of one model created from one drill geometry to predict the tool wear of the other drill geometry, the models created with the most sensitive signals in drill R204.6D were used to test the wear on instances of both drills.

Fig. 4 shows a picture explaining the process. Firstly, statistical characteristics of selected signals and process parameters (V_c, f_n, L, β) along with V_b for R204.6D were used to train ①, and then the model was tested to predict V_b for both ②.

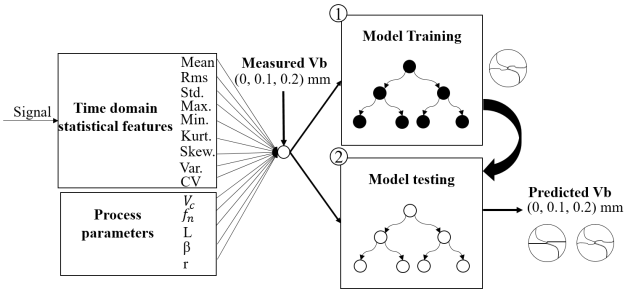


Fig. 4 Model training and testing using different geometries of drill bits

Further feature extraction processes made us distinguish between four different strategies explained below:

- **S1:** Calculates the statistical characteristics of the entire signal for each hole. The signal has been considered from

the beginning of the hole to the end of the hole, omitting any part that does not belong to the drilling process.

S2: Calculates the statistical characteristics for each 1mm depth of cut, segmenting the signal for one hole in as many segments as mm it has. As both tools are the same diameter, using this approach, statistical characteristics correspond to the same volume of material removed. For different cutting conditions, the 1mm window was calculated using equation equation 1. As a consequence, the number of instances available for the same hole could increase.

$$W_{1mm} = \frac{f_s \cdot 60}{a_v} \quad (1)$$

Where W_{1mm} is the $1mm$ window length ($Samples \cdot mm^{-1}$), a_v is the feed rate ($mm \cdot min^{-1}$) and f_s is the sampling frequency ($Samples \cdot seg^{-1}$).

S3: The predictions were made using the predictions made for each segment of the hole in a voting scheme. If the majority of the segments predicted one specific value for V_b, this will be the final decision for the predicted wear in this hole.

The process parameter changed during the experiments were: the type of drill and cutting conditions. None reference of the type of drill was included in the datasets. Regarding cutting conditions, the statistical features were extracted for 1 mm hole length, making them independent of the cutting speed for strategies S2 and S3.

4 Results and discussion

This section shows the results of the sensitivity analysis, comparing the ability of the different signals to predict the tool wear. It also shows the results of training a model with signals from one type of drill bit and predicting the wear of another type of drill bit.

4.1 Sensitivity of the signals to tool wear

As already mentioned, the sensitivity analysis was carried out based on a two-tailed paired T-test using 10-fold-cross-validation 10 times. In the three versions proposed(V1, V2, V3) and for both types of tools used in this work the most sensitive signals seem to be the thrust force (F_z) and the Z axis torque ($TV2$).

Table 4 shows the accuracy of each signal with respect to the algorithm used for the type BH04.5D drill bit. To perform this test $TV2$ signal was chosen as the most accurate signal to predict tool wear. Therefore, T paired test is applied against the $TV2$ signal. In the first version (V1), only new tools and tools with flank wear were classified. In all signals at least one or more algorithms fulfil the null hypothesis. In the second version (V2), although the classification is done in binary form, the degree of difficulty is increased, so the number of signals that fulfil the null hypothesis is considerably reduced. In the last version (V3), the only signal which fulfils the null hypothesis respect to $TV2$ that the difference of the means for the accuracy of the paired data is 0, is F_z. With regard to the other signals, none of them are statistically relevant. Therefore, they are not considered as signals with statistical features capable of predicting tool wear in the time domain better than $TV2$ or F_z do.

Table 4 Statistical significance of signals in a two tailed paired T-test for accuracy in predicting BH04.5D drill bit wear for three proposed versions. In white, the signals with the same mean in accuracy as $TV2$. In grey, the signals with different mean in accuracy

		Internal						External				
		TV2	TV50	TV51	V	ACCEL	JERK	Fz	Mz	AE	Sound	Vib
V1	J48	97	97	96	85	88	91	97	90	86	96	86
	LMT	99	91	91	84	84	87	100	96	82	96	84
	IBK	100	94	91	85	86	85	100	99	89	97	89
	NB	100	94	93	91	84	94	100	95	89	95	82
V2	J48	98	77	66	81	82	95	97	89	86	98	89
	LMT	98	69	72	85	80	89	100	87	82	95	86
	IBK	100	71	70	84	80	86	100	96	90	90	88
	NB	100	82	75	77	81	90	100	81	83	96	87
V3	J48	94	68	55	67	61	78	91	77	70	66	72
	LMT	94	64	69	75	72	80	94	80	73	70	57
	IBK	96	59	63	72	74	72	98	86	78	63	68
	NB	94	77	63	79	72	81	98	75	63	69	53

Regarding the prediction of the wear for the R204.6D type drill bit, Table 5 shows the corresponding results. The signal chosen as most accurate was Fz and the paired-T test wast applied against it. As in the previous case, the highest number of signals that accept the null hypothesis are found in the first version (V1) proposed. However, the same number of successes are not achieved as in the previous case. This may be due to differences in tool geometry. In the second version (V2), as expected, the number of successes is significantly reduced. In the last one results indicate that the $TV2$ signal is the one that makes most of the algorithms accept the null hypothesis.

Basically, the results match for both tool geometries. Given the two tool configurations, it can be seen from the results that the straight-edged tool (BH04.5D) has better results for some of the signals. This may be due to geometrical differences between the two tools or to the preparation prior to testing, as the BH04.5D tools were sharpened by the tool manufacturer leaving more uniform wear along the tool's flank.

The results, as expected, showed that the F_z and $TV2$ signals were the most sensitive signal at the time of predict-

Table 5 Statistical significance of signals in a two tailed paired T-test for accuracy in predicting R204.6D drill bit wear for three proposed versions. In white, the signals with the same mean in accuracy as Fz. In gray, the signals with different mean in accuracy

		Internal						External				
		TV2	TV50	TV51	V	ACCEL	JERK	Fz	Mz	AE	Sound	Vib
V1	J48	96	79	91	60	65	74	99	83	71	54	69
	LMT	99	94	91	64	70	72	100	89	72	72	64
	IBK	96	88	84	73	65	70	100	85	49	73	61
	NB	96	91	91	69	61	57	96	86	66	65	65
V2	J48	93	72	73	68	71	71	100	70	58	52	75
	LMT	88	95	87	69	69	73	100	83	53	63	66
	IBK	79	78	76	75	70	71	100	73	58	53	69
	NB	75	77	77	70	69	71	97	72	53	54	68
V3	J48	85	56	63	55	59	59	91	63	54	46	65
	LMT	84	90	83	65	60	58	92	74	63	48	67
	IBK	74	69	59	73	56	55	95	63	57	52	67
	NB	87	75	72	74	63	55	91	70	66	38	69

ing the tool wear. Further analysis in the frequency domain could add feasible statistical characteristics, improving accuracy and robustness.

Sensors are fixed at all times, replicability of measurements is a difficult task. The location of the hole changes constantly, changing the source of the signal to be measured, so measurements are likely to vary as the hole changes location. In this case, the acoustic emissions sensor, the vibration sensor and the sound pressure sensor are affected by this phenomenon. A frequency-domain analysis would help to obtain the frequency bands that are least affected by the change of location of the hole and those that best correspond to the phenomena to be detected. This can also lead to more favourable results in all these signals.

4.2 Model testing for different tool geometries

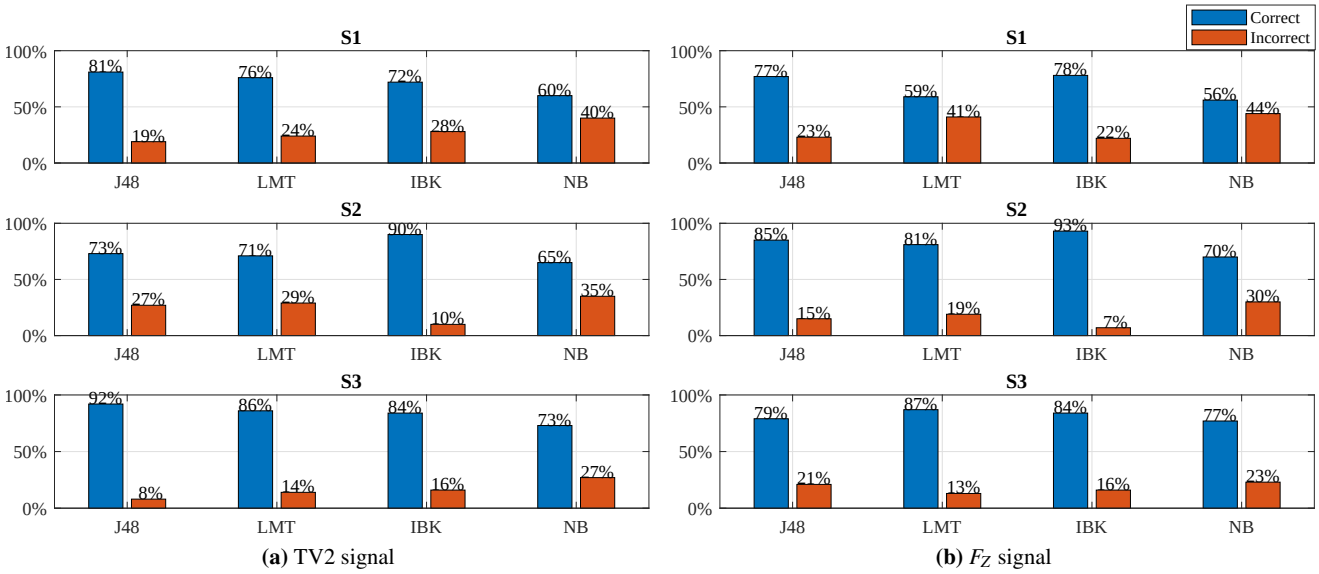
To check the impact of the tool geometry in the models used for tool wear detection, an algorithm was trained using R204.6D drill bits and then, tested with R204.6D and BH04.5D tool signals, using for that $TV2$ and F_z, the most sensitive signals for tool wear detection.

Fig. 5 shows the proportion of correct and incorrect predictions made for each of the proposed strategies and considered algorithms. The $TV2$ signal shows a good performance in terms of the first strategy (S1). However, despite the percentage of correct results obtained, other classification strategies have been carried out with the aim of improving the results obtained. The second strategy (S2) involves the segmentation of signals and therefore contains a larger number of instances for both the training and testing phases, the IBk algorithm has optimal performance when training with a larger number of instances. In the last strategy (S3), a voting system was added to the previous classification strategy. Comparing this strategy with the first one, it can be seen how it significantly improves the behaviour of all the algorithms. Thus, the segmentation of signals together with the voting system presents a better behaviour than using the complete signal.

Observing Fig. 5, the Fz signal shows a worse behaviour regarding the first strategy (S1). This could be due to the assembly and disassembly of the rotational dynamometer during the execution of the tests, the measurements made could vary in case of not having exactly the same assembly. The signal TV2 is acquired directly from the machine itself so the acquisition and this signal is always made in the same conditions. As for the second strategy (S2), the Fz signal appears to have a favourable effect on the number of correctly classified instances. The segmentation of the signal into independent instances seems to have a positive effect on the precision of the algorithms. It does not seem to have the same effect on the TV2 signal with which in this case the proportion of correctly classifieds decreases with respect to the first strategy (S1). In the last strategy (S3), predictions are made in terms of hole made, that is, the same as in the first strategy (S1). For both signals, the amount of correctly classified increases obtaining a considerable improvement in all the algorithms considered.

A confusion matrix shows the performance of a classification model, columns represent current values of a target (Measured Vb) and, in rows, the values predicted (Predicted Vb) by a model. the values in the cells represent the number of instances for which the Measured Vb has a certain Predicted Vb. Any value out of the diagonal of the confusion matrix correspond to the same number of wrongly classified instances. Precision is the fraction of all relevant instances divided among the obtained instances. The recall is the fraction of relevant instances obtained over the total number of relevant instances. The column on the far right of the plot shows the recall for each predicted class, while the row at the bottom of the plot shows the precision for each true class. The cell in the bottom right of the plot shows the overall accuracy.

In S2, for both signals, an optimal result is achieved using the IBk algorithm. Therefore, this algorithm benefits from the number of instances available for the training phase. Fig. 6 shows the confusion matrices for the predictions made using the IBk algorithm for the S2 and S3 strategies with the TV2 and Fz signals. Making use of S2 (Fig. 6 (a, c)) a good percentage of correctly classified instances is obtained, 90.4% for TV2 and 92.7% for Fz. For both signals in S2 the worst instances at the time of being classified are the worn tools to Vb=0.1 mm. This is because there are overlaps between the proposed labels for wear classification between the different statistical features extracted from the signals. This supposes a challenge at the time of assigning the labels



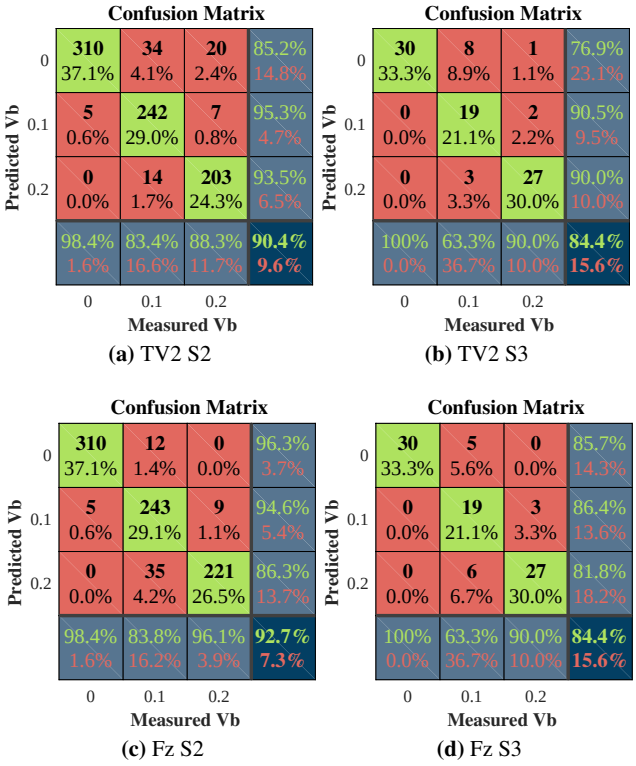
(a) TV2 signal **(b)** F_Z signal

Fig. 5 Performance of all the strategies and algorithms tested

based on the measurements made in the Leica DMS1000 macroscope and Alicona profilometer.

Observing the precision and recall in Fig. 6 (a, b). In general, the precision increases from strategy S2 to strategy S3. However, the recall is affected and decreases. As a result of increasing the precision of the model, the recall is affected and decreases. For Vb = 0 mm in S3, 100% of the cases are correctly classified. Nevertheless, there are several false positives that are classified as new tools. On the other hand, in Fig. 6 (c, d), in the case of the Fz signal, the precision and recall decrease from S2 to S3. However, in the case of S2 it can be seen that the results are closer to the class to which they should belong.

Fig. 7 compares the performance of the J48 algorithm using the TV2 for strategies 1 and 3. The improvement is remarkable. The final classification is the same in both strategies. On the one hand the statistical characteristics extracted from the complete signal (S1) are taken into account. On the other hand, the classification is made on the basis of the segments obtained (S3). So the difference is in the use that is made of the data obtained from the process. The 5 holes made with 0.1mm wear tools classified as new tools in S3 belong to a single tool, the tool labelled with ID: BH04.5D 2 01 . It is the tool with the lowest wear measured directly on the tool, making it difficult to distinguish them from the new ones. In general, it can be observed that precision and recall increase from S1 to S3. The number of false positives is reduced, observing a considerable improvement when performing a classification based on signal segments. This way of handling the data allows to increase the probability of success of a certain model, since in addition to the current output it takes into account previous system outputs.



(a) TV2 S2 **(b)** TV2 S3

(c) Fz S2 **(d)** Fz S3

Fig. 6 Confusion matrices applying strategies 2 and 3 for the IBK algorithm and the $TV2$ and Fz signals

5 Conclusions

In this work a comparative study of the sensibility of the most commonly used signals for tool wear detection in drilling

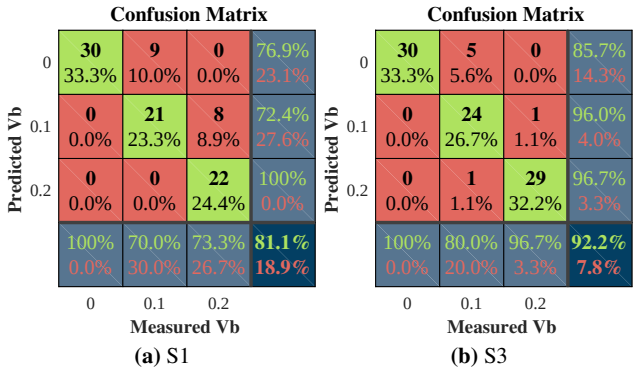


Fig. 7 Confusion matrix: J48 algorithm for $TV2$ signal

processes has been carried out. Tests were performed with two different types of drill bits under a complete setup in which multiple signals were acquired for an evaluation of their sensitivity to tool wear in drilling processes. The methodology used has made it possible to identify the most sensitive signals in the time domain, with the Z-axis motor torque and thrust force being the most sensitive to tool wear. In addition, based on the training phase of the algorithms with the $TV2$ signal of R204.6D type drill bits, tool wear of BH04.5D type drills has been predicted.

- The methodology used has made it possible to identify the variables of interest for this study. This methodology has helped to identify the extent to which the different signals collected are capable of detecting tool wear.
- The most sensitive signals are $TV2$ and F_z. Once the sensitivity analysis has been carried out, it has been possible to see that the most affected signals by tool wear are the thrust force and the torque of the Z axis motor. With the above mentioned signals, quite precise results have been achieved by predicting tool wear with both types of drill bits.
- The $TV2$ signal shows great predictability in terms of other tools. The model created with the $TV2$ signal from R204.6D drill bits shows great accuracy in predicting tool wear for BH04.5D drill bits. The F_z signal does not perform as well as TV2 in this respect.
- The strategy of segmenting the signals with a 1 mm hole depth window means an increase in the number of instances available for the algorithm training phase. In both cases ($TV2$ and F_z) a better result is achieved with the added voting system than in the first pursued strategy.
- The signals have been analysed in the time domain, they need to be analysed further to obtain accurate indicators of tool wear. It is necessary to apply different treatments to the signals to obtain better indicators for the detection of tool wear.
- With this work it has been possible to identify the $TV2$ signal as the most predictive internal signal for a given material, two different types of tool geometry and two different cutting conditions.
- In case of wanting to develop a real wear detection system, it would be based on signals selected from the results of this study, but taking into account the convenience of the use of each sensor in each specific case. In addition, models that consider more than one signal at a time to improve accuracy and reduce false positives or true negatives should participate in the final system.

Acknowledgements This work has been developed by the data analysis and cybersecurity research group and high-performance machining research group supported by the Department of Education, Language policy and Culture of the Basque Government. The authors thank the SMAPRO project (KK-2017/00021) for funding the research presented in this paper.

References

1. Jindal A, Others (2012) Analysis of Tool Wear Rate in Drilling Operation Using Scanning Electron Microscope (SEM). Journal of Minerals and Materials Characterization and Engineering 11(01):43
2. Radhakrishnan T, Wu SM (1981) On-Line Hole Quality Evaluation for Drilling Composite Material Using Dynamic Data. Journal of Engineering for Industry 103(1):119, DOI 10.1115/1.3184452
3. Arrazola PJ, Garay A, Fernandez E, Ostolaza K (2014) Correlation between tool flank wear, force signals and surface integrity when turning bars of Inconel 718 in finishing conditions. International Journal of Machining and Machinability of Materials 15(1/2):84, DOI 10.1504/IJMMM.2014.059193, URL http://www.inderscience.com/link.php?id=59193
4. Jantunen E (2002) A summary of methods applied to tool condition monitoring in drilling. International Journal of Machine Tools and Manufacture 42(9):997–1010, DOI 10.1016/S0890-6955(02)00040-8
5. Sharman ARC, Amarasinghe A, Ridgway K (2008) Tool life and surface integrity aspects when drilling and hole making in Inconel 718. Journal of Materials Processing Technology 200(1-3):424–432, DOI 10.1016/j.jmatprotec.2007.08.080
6. Rahim EA, Sharif S (2007) Tool failure modes and wear mechanism of coated carbide tools when drilling Ti-6Al-4V. Int J Precision Technology Vol. 1 No.(1):pp.30–39, DOI 10.1504/IJPTECH.2007.015342
7. Rahim EA, Sharif S (2006) Investigation on Tool Life and Surface Integrity when Drilling Ti-6A1-4V and Ti-5Al-4V-Mo/Fe. JSME International Journal Series C (2)
8. Yingfei G, de Escalona PM, Galloway A (2017) Influence of Cutting Parameters and Tool Wear on the Surface Integrity of Cobalt-Based Stellite 6 Alloy When

Machined Under a Dry Cutting Environment. Journal of Materials Engineering and Performance 26(1):312–326, DOI 10.1007/s11665-016-2438-0
9. Zhao Q, Qin X, Ji C, Li Y, Sun D, Jin Y (2015) Tool life and hole surface integrity studies for hole-making of Ti6Al4V alloy. International Journal of Advanced Manufacturing Technology 79(5-8):1017–1026, DOI 10.1007/s00170-015-6890-z
10. Eckstein M, Vrabe M, Maková I (2016) Tool wear and surface roughness evolution in hole making process of Inconel 718. Materials Science Forum 862:11–17, DOI 10.4028/www.scientific.net/MSF.862.11, URL https://www.scopus.com/inward/record.uri?eid=2-s2.0-84979986226{&}doi=10.4028{%}2Fwww.scientific.net{%}2FMSF.862.11{&}partnerID=40{&}md5=3ebdd464d8da3873e14e2755ff79f364
11. Shah M, Unanue L, Bidare P, Galfarsoro U, Iriarte Luis Mªand Karunakaran KP, Arrazola PJ (2010) Tool control monitoring applied to drilling. Proceedings of the 6th MUGV conference, Cluny, France pp 1–10
12. Diniz AE, Liu JJ, Dornfeld da (1994) Correlating tool life, tool wear and surface roughness by monitoring acoustic emission in finish turning : Wear, Vol. 152, No. 2, pp. 395407 (Jan. 1992). NDT & E International 27:213–214, DOI http://dx.doi.org/10.1016/0963-8695(94)90499-5, URL http://www.sciencedirect.com/science/article/pii/0963869594904995
13. Bhuiyan MSH, Choudhury IA, Dahari M, Nukman Y, Dawal SZ (2016) Application of acoustic emission sensor to investigate the frequency of tool wear and plastic deformation in tool condition monitoring. Measurement: Journal of the International Measurement Confederation 92:208–217, DOI 10.1016/j.measurement.2016.06.006
14. Stavropoulos P, Papacharalampopoulos A, Vasiliadis E, Chryssolouris G (2016) Tool wear predictability estimation in milling based on multi-sensorial data. International Journal of Advanced Manufacturing Technology 82(1-4):509–521, DOI 10.1007/s00170-015-7317-6
15. Rafezi H, Behzad M, Akbari J (2012) Time Domain and Frequency Spectrum Analysis of SoundSignal for Drill Wear Detection. International Journal of Computer and Electrical Engineering 4(5):722
16. Ademujimi TT, Brundage MP, Prabhu VV (2017) A Review of Current Machine Learning Techniques Used in Manufacturing Diagnosis. In: Lödding H, Riedel R, Thoben KD, von Cieminski G, Kiritsis D (eds) Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing, Springer International Publishing, Cham, pp 407–415
17. Wei Y, An Q, Ming W, Chen M (2016) Effect of drilling parameters and tool geometry on drilling performance in drilling carbon fiber-reinforced plastic/titanium alloy stacks. Advances in Mechanical Engineering 8(9):1–16, DOI 10.1177/1687814016670281
18. Wu D, Jennings C, Terpenny J, Kumara S, Gao R (2017) Cloud-Based Parallel Machine Learning for Prognostics and Health Management: A Tool Wear Prediction Case Study. Journal of Manufacturing Science and Engineering (c), DOI 10.1115/1.4038002, URL http://manufacturingscience.asmedigitalcollection.asme.org/article.aspx?doi=10.1115/1.4038002
19. Kim JD, Choi IH (1996) Development of a tool failure detection system using multi-sensors. International Journal of Machine Tools and Manufacture 36(8):861–870, DOI 10.1016/0890-6955(96)00115-0
20. Dimla Snr DE (2002) The correlation of vibration signal features to cutting tool wear in a metal turning operation. International Journal of Advanced Manufacturing Technology 19(10):705–713, DOI 10.1007/s001700200080
21. Rmili W, Ouahabi A, Serra R, Leroy R (2016) An automatic system based on vibratory analysis for cutting tool wear monitoring. Measurement: Journal of the International Measurement Confederation 77:117–123, DOI 10.1016/j.measurement.2015.09.010, URL http://dx.doi.org/10.1016/j.measurement.2015.09.010
22. Krishnakumar P, Rameshkumar K, Ramachandran KI (2015) Tool wear condition prediction using vibration signals in high speed machining (HSM) of Titanium (Ti-6Al-4V) alloy. Procedia Computer Science 50:270–275, DOI 10.1016/j.procs.2015.04.049, URL http://dx.doi.org/10.1016/j.procs.2015.04.049
23. Grosse CU, Linzer LM (2008) Signal-Based AE Analysis. In: Grosse C, Ohtsu M (eds) Acoustic Emission Testing: Basics for Research - Applications in Civil Engineering, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 53–99, DOI 10.1007/978-3-540-69972-9_5, URL https://doi.org/10.1007/978-3-540-69972-9{_}5
24. Kilundu B, Dehombreux P, Chiementin X (2011) Tool wear monitoring by machine learning techniques and singular spectrum analysis. Mechanical Systems and Signal Processing 25(1):400–415, DOI DOI10.1016/j.ymssp.2010.07.014
25. Kothuru A, Nooka SP, Liu R (2017) Application of audible sound signals for tool wear monitoring using machine learning techniques in end milling. The International Journal of Advanced Manufacturing Technology pp 1–12, DOI 10.1007/s00170-017-1460-1, URL http://link.springer.com/10.1007/

s00170-017-1460-1
26. Seemuang N, McLeay T, Slatter T (2016) Using spindle noise to monitor tool wear in a turning process. International Journal of Advanced Manufacturing Technology 86(9-12):2781–2790, DOI 10.1007/s00170-015-8303-8, URL http://dx.doi.org/10.1007/s00170-015-8303-8
27. Corne R, Nath C, El Mansori M, Kurfess T (2017) Study of spindle power data with neural network for predicting real-time tool wear/breakage during inconel drilling. Journal of Manufacturing Systems 43:287–295, DOI 10.1016/j.jmsy.2017.01.004, URL http://dx.doi.org/10.1016/j.jmsy.2017.01.004
28. Lee J, Choi HJ, Nam J, Jo SB, Kim M, Lee SW (2017) Development and analysis of an online tool condition monitoring and diagnosis system for a milling process and its real-time implementation. Journal of Mechanical Science and Technology 31(12):5695–5703, DOI 10.1007/s12206-017-1110-4
29. MARPOSS (2018) Artis MARPOSS. URL http://www.artis.de/en/
30. Duo A, Basagoiti R, Arrazola PJ, Aperribay J (2018) A comparative study between internal and external signals for tool wear detection in drilling processes. 14Th International Conference on High Speed Machining (1):1–4
31. Ferreiro S, Sierra B, Irigoien I, Gorritxategi E (2011) Data mining for quality control: Burr detection in the drilling process. Computers and Industrial Engineering 60(4):801–810, DOI 10.1016/j.cie.2011.01.018, URL http://dx.doi.org/10.1016/j.cie.2011.01.018