Books Trimmer Industrial Machine Knives Diagnosis: A Condition-Based Maintenance Strategy Through Vibration Monitoring via Novelty Detection

Proceedings of the ASME 2022
International Mechanical Engineering Congress and Exposition
IMECE2022
October 30-November 3, 2022, Columbus, Ohio

IMECE2022-94547

BOOKS TRIMMER INDUSTRIAL MACHINE KNIVES DIAGNOSIS: A CONDITION-BASED MAINTENANCE STRATEGY THROUGH VIBRATION MONITORING VIA NOVELTY DETECTION

Luca Viale[1], Alessandro Paolo Daga[1], Luigi Garibaldi[1], Salvatore Caronia[2], and Ilaria Ronchi[2]

[1]Department of Mechanical and Aerospace Engineering, Politecnico di Torino,
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
[2]Tecnau SRL, Ivrea, Italy

ABSTRACT

In recent years, Artificial Intelligence (AI) is ever more exploited in all the scientific and industrial fields and is allowing significant developments in mechanical engineering too. An emblematic contribution was given in terms of safety and reliability since Machine Learning (ML) techniques permitted the monitoring and the prediction of the state of health of machinery, allowing the adoption of predictive maintenance strategies. In fact, data-driven models – based on acquisitions – attract considerable interest both thanks to its theoretical and application development. The evolution of diagnostic techniques is oriented towards Condition-Based Maintenance (CBM) strategies, thus allowing improvements in terms of safety enhancement, cost reduction and increased performances. This paper proposes the development and implementation of a diagnostic/prognostic tool applied to an automated books trimmer industrial machine, implementing condition monitoring by means of accelerometers which can be integrated into a Supervisory Control And Data Acquisition (SCADA) system. Given its use, the core components of this production line are three knives, subjected to significant impulsive forces. Therefore, the target of the work is to infer the wear of these three knives, as they are critical elements of the machinery and have a high impact on the quality of the final product. The project was carried out in collaboration with Tecnau - an industry-leading company - which made it possible to conduct experimentation and data acquisition on their machinery. An appropriate Design Of Experiments (DOE) and the use of inferential statistical techniques - such as the ANalysis Of VAriance (ANOVA) and the identification of significant effects - applied to the multivariate dataset allowed recognizing the most relevant features for Novelty Detection (ND). Both the Linear Discriminant Analysis (LDA) and the k-Nearest Neighbors (kNN) method permitted to correctly distinguish the patterns representing the health conditions of the machinery, classifying the data in the reduced multidimensional space according to the final product quality. The results obtained in terms of accuracy are very positive and promising. This means that the developed method is able to successfully identify the state of health of the blade in spite of varying functioning parameters (book thickness and size, paper type and characteristics) and operating conditions. The algorithm speed and its integration into the industrial line make a real-time condition-based maintenance strategy possible. This diagnostic method is suitable for applications oriented to the paradigm of Industry 4.0 and the digitalization of the industrial sector, which can be integrated with the Internet of Things (IoT) and cloud systems.

Keywords: machine diagnostics; books trimmer industrial machine; Novelty Detection; Condition Based Maintenance; Industry 4.0

NOMENCLATURE

AI	Artificial Intelligence
ANOVA	ANalysis Of VAriance
AUC	Area Under the Curve
CBM	Condition-Based Maintenance
CV	Cross-Validation
DOE	Design Of Experiments
FA	False Alarms
FF	Full Factorials
FN	Frobenius Norm
GA	Genetic Algorithm
IoT	Internet of Things
kNN	k-Nearest Neighbors
LDA	Linear Discriminant Analysis
MA	Missed Alarms
MANOVA	Multivariate ANalysis Of Variance
MCCV	Monte Carlo Cross-Validation
ML	Machine Learning
ND	Novelty Detection

OFAT	One-Factor-At-a-Time
RMS	Root Mean Square
ROC	Receiver Operating Characteristic
SCADA	Supervisory Control And Data Acquisition
VM	Vibration Monitoring

1. INTRODUCTION

In recent years, Artificial Intelligence (AI) has reaped great success and has found numerous applications in all scientific fields. AI has also fostered significant developments in mechanical engineering [1,2] and has generated a significant contribution in terms of safety and reliability. Indeed, Machine Learning (ML) techniques have allowed the monitoring and prediction of machinery state of health, allowing the adoption of predictive maintenance strategies. The evolution of these techniques is oriented toward Condition-Based Maintenance (CBM) strategies with the aim of improving safety, reducing costs and increasing performance [3,4]. Given that modern industrial systems have now reached a complexity level that makes traditional engineering approaches difficult to pursue (e.g., model-based), data-driven approaches arouse considerable interest as they make it possible to develop diagnostic and prognostic systems by exploiting acquisitions through non-destructive testing.

Among the various forms of condition monitoring, Vibration Monitoring (VM) is a particularly effective and used technique [5,6]. It allows the recognition of the machinery conditions by means of the measurement of its vibrations, which are the small oscillations around the equilibrium point that every mechanical device generates during its operating phase. VM requires the use of accelerometers, and this makes the diagnostic system relatively cheap, reliable, and flexible in relation to the different typologies of machinery.

The information obtained through the recordings of the accelerometric signals can be suitably processed and used to classify the health condition of machinery. A famous binary classification technique is Novelty Detection (ND), which allows recognizing a condition as novel when it deviates significantly from healthy reference data [7–9]. For instance, in [10] ND was integrated into a Genetic Algorithm (GA) optimizer for a complex machine diagnosis. Given that current ICT solutions allow the collection of large amounts of data, it is usually necessary to manage Big Data [11] in order to develop an optimal CBM strategy. In this regard, there are in the literature several strategies for features extraction and selection [12,13] and dimension reduction through statistical techniques [14,15].

This study proposes the development and implementation of a diagnostic/prognostic system on a complex machine for automatic book cutting. The conditions monitoring of the three blades - critical elements of the machinery and relevant to the quality of the product - is possible implementing a system of sensors for VM that could be integrated into a Supervisory Control And Data Acquisition (SCADA) system. The experimental campaign has been suitably investigated by means of a Design Of Experiments (DOE) [16,17] and the measured signals have been duly processed. The extracted features have been selected thanks to multivariate inferential statistical techniques. Finally, via the application of ML methods, it has been possible to distinguish the machinery conditions with high accuracy values. In particular, thanks to the implementation of classification techniques, the creation of a diagnostic method which allows improving the products' quality and the production line efficiency has been possible.

In detail, the article consists of the following sections. Section 2 describes the machinery under analysis, the experimental setup, the DOE, the statistical analysis, and the implementation of ML algorithms. Section 3 presents the obtained results and a related discussion. Finally, the conclusions of this study are reported in Section 4.

2. MATERIALS AND METHODS

This section briefly describes the machinery under analysis and the related choices regarding the sensors to implement on the test bench. Subsequently, the factors relevant for the analysis that are considered for the DOE development are highlighted. After describing the test scheduling, the choices made for the signals pre-processing and the features extraction are illustrated. Given the considerable amount of collected data, the statistical analyzes developed to select the most significant features to recognize the blade health condition are described. Finally, the ML techniques adopted for the classification are expounded.

2.1 Test bench

The machinery under analysis is an automated industrial books trimmer designed and launched on the market by Tecnau SRL [18]. As can be seen in FIGURE 1, it is mainly composed of three knives capable of cutting a wide range of books with different sizes and thicknesses at high speed. By following the instructions contained in and processed by the system software to automatically adjust the cutting measurements, the mobile guides correctly position each book. Before cutting the book sides except for the spine, each of them is compressed to remove the air contained between the pages. Thanks to the experience gained by the company both in terms of research and feedback received from customers, it is possible to assume at first that the blades are the critical elements of the machinery, and their wear is the factor that mainly determines the quality of the final product. However, given that the parameters varying during the machine operating phase are numerous (e.g., thickness, length and width of the book, type and grammage of the paper), this hypothesis has been verified using inferential statistical analyzes such as the ANalysis Of VAriance (ANOVA) and the identification of significant effects [19]. Furthermore, given that ND allows the fault to be directly detected only in the event of the exclusion of confounding influences (including operating and environmental conditions), the demonstration that the effect of the blade wear is the main significant factor is fundamental for being able to assert that the measured accelerometric signals correctly describe the health condition of the blades.

It was initially adopted a cDAQ-9185 chassis, two 4-channels NI-9234 modules with sampling frequency $f_s = 51.2$ kS/s/channel, three PCB 355B03 monoaxial accelerometers, and a PCB 356A16 triaxial accelerometer with range equal to $\pm 50g$ to perform the tests. In addition to these sensors, the signal of the two-channel incremental encoder integrated into the servomotors (with gear ratio i) that guide the blade crank mechanisms was acquired (from the kinematic analysis of the crank mechanisms it was possible to derive the equation $f(\theta)$ which describes the blade position as a function of the servomotors rotation angle θ). The described system is very expensive since it allows achieving a high degree of precision. Indeed, this acquisition system has been used only for an initial phase of investigation and understanding of the trimmer behavior. Subsequently, it will be possible to study an appropriate downgrade of the sensors in order to find out the best cost-performance trade-off.

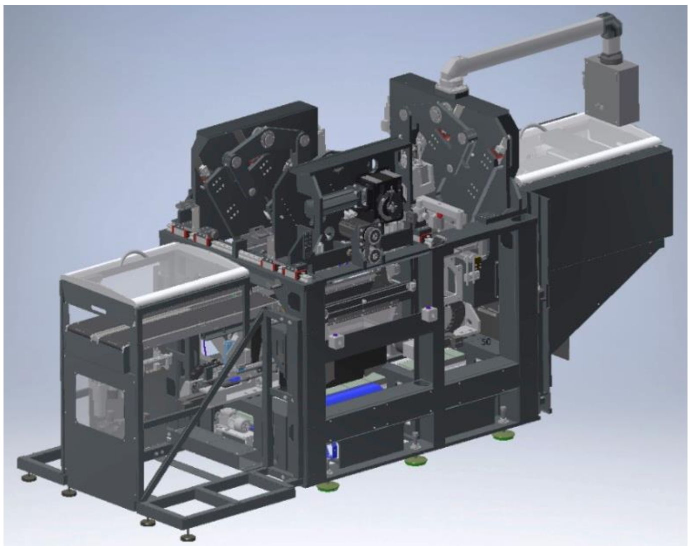

FIGURE 1: MACHINERY STRUCTURE.

Considering the machinery structure and referring to the concept of the *transmission path* minimization [19], accelerometers have been positioned as shown in FIGURE 2. In particular, the three monoaxial accelerometers have been positioned on the top of the structure (M1) to measure the effects generated by the machinery bending due to the cutting, on the mobile component of the blade (M2) and on its fixed corresponding part (M3) to measure the transmitted vibrations. On the other hand, the triaxial accelerometer (T) has been placed near the moving blade to obtain a three-axis measurement near the cut itself. All accelerometers have been fixed thanks to cyanoacrylate glue and duct tapes.

2.2 Design Of Experiments and identification of significant effects

The DOE was defined based on a Full Factorials (FF) model. Indeed, it allows both reducing the burden in terms of necessary experiments and also analyzing the effects of factors interactions unlike One-Factor-At-a-Time (OFAT) methods [20]. Two levels have been selected in this study both because they are sufficient for the analysis and in order to limit the number of experiments. Considering these conditions, 2^n different combinations are required for a Full Factorials analysis, where n is the number of selected factors. It should also be remembered that each test was performed three times in order to obtain a statistically relevant sample. Therefore, having identified five factors as relevant for the analysis, $3 \cdot 2^5 = 96$ tests have been planned.

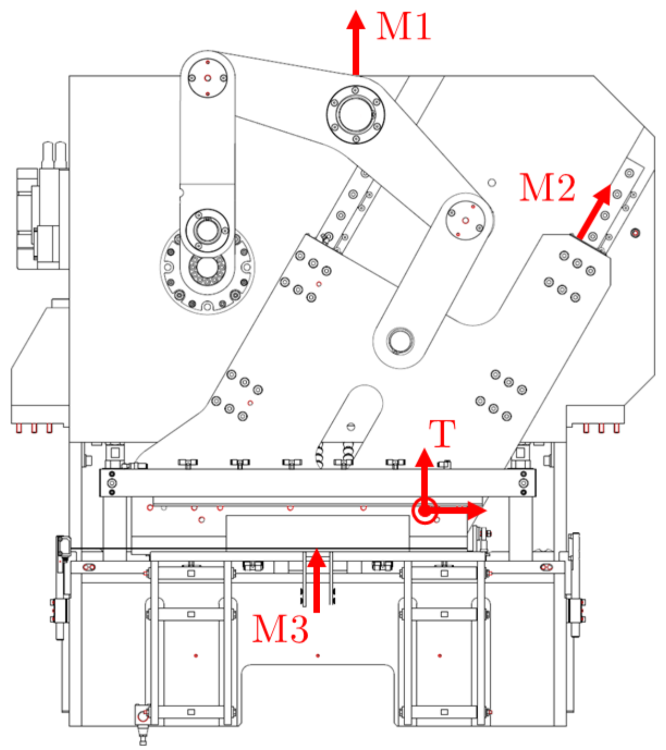

FIGURE 2: FRONT VIEW OF A KNIFE AND RELATED SENSORS POSITIONING.

The factors considered in the DOE and the related levels are reported in TABLE 1. The factors can be both numerical and qualitative. For the latter, *low* and *high* levels were assigned in a purely arbitrary way. In addition, please note that the run order of the tests was randomized in such a way as to reduce as much as possible the effects related to time-dependent variables (e.g., the heating of machinery) that could affect the test results. The only factor excluded from this process is the blade wear because its disassembly, reassembly and adjustment would have excessively increased the cost and time required to carry out the experiments. Furthermore, it would generate a significant cross-correlation since this operation is time-consuming and the machine cools down. Finally, before carrying out the 96 scheduled tests it was necessary to carry out a few idling cycles to consider the machine operating in steady-state conditions and, consequently, to minimize the effects due to the transient phase of the machinery. Some example rows of the test matrix produced by means of DOE are reported in TABLE 2.

Among its advantages, this DOE allows deriving the effects in the subsequent phases, calculating the average difference of the results obtained between the two levels. In addition to the

TABLE 1: TEST-FACTORS AND THE RELATED LEVELS FOR 2-LEVEL FULL FACTORIALS ANALYSIS. THE NUMERICAL VALUES HAVE BEEN STANDARDIZED WITH RESPECT TO THE MAXIMUM DIMENSIONS REACHABLE BY THE MACHINE.

Factor	Name	Units	Low level	High level
A	Book thickness	mm	0.25	0.75
B	Book width	mm	0.25	0.75
C	Paper type	Material	Coated	Not coated
D	Paper grammage	g/m^2	Light	Heavy
E	Blade wear	State	Damaged	New

TABLE 2: TEST MATRIX ARRANGED ACCORDING TO THE "RUN ORDER". THE TWO LEVELS OF THE FACTORS ARE CODED (THE + SYMBOL REFERS TO THE HIGH LEVEL, WHILE - TO THE LOW LEVEL), AND THE OUTPUT VALUES HAVE BEEN ASSIGNED IN TERMS OF THE CUT QUALITY. THE ROWS HAVE NOT BEEN REPORTED ENTIRELY FOR THE SAKE OF BREVITY, AS REPRESENTED BY THE ELLIPSIS EXTENSION SYMBOL.

Run order	Factor A	Factor B	Factor C	Factor D	Factor E	Quality
1	+	–	–	+	–	3
2	–	+	+	–	–	3
3	+	+	+	–	–	1
…	…	…	…	…	…	…
94	+	+	–	+	+	5
95	–	–	–	–	+	4
96	–	–	–	+	+	4

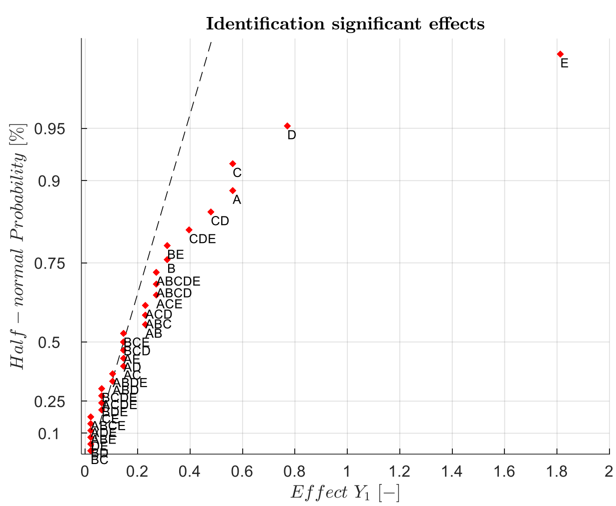

FIGURE 3: IDENTIFICATION OF SIGNIFICANT EFFECTS (LABELLED) THROUGH A COMPARISON WITH A HALF-NORMAL DISTRIBUTION (DASHED LINE).

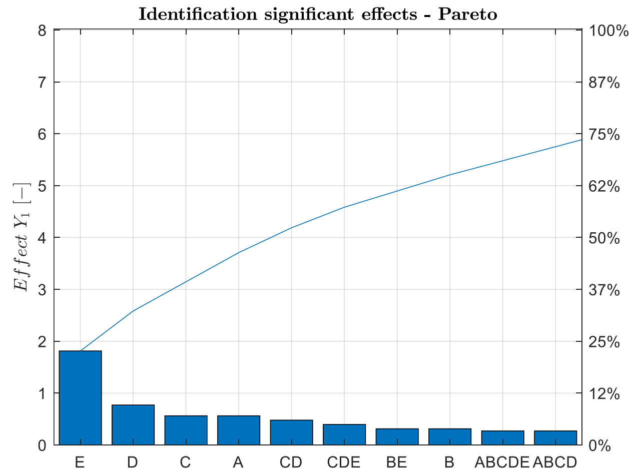

FIGURE 4: IDENTIFICATION OF SIGNIFICANT EFFECTS THROUGH A PARETO CHART. THE CONTINUOUS CURVE REPRESENTS THE CUMULATIVE VALUE OF THE EFFECTS.

effects of the factors considered independently (A, B, C, D, E), the Full Factorials analysis also allows examining their interactions in groups (e.g., AB, AC, AD, AE, BC, BD, ... , ABC, ABD, ...). The binomial coefficient represents the total number of interactions combinations. In particular, the effects due to the interaction of three or more factors are not considered particularly relevant for the study. For this effects analysis, the cut quality was selected as the output value since it is the most determining and relevant characteristic. This value was assigned thanks to the quality control of the final product based on a graduated scale from 1 to 5 (where 1 indicates the minimum quality and 5 indicates the maximum quality).

By comparing the results obtained and appropriately converted with respect to a half-normal distribution, as shown in FIGURE 3, it is possible to identify the factors that follow the half-normal (which are probably not significant in terms of effects) and those that differ visibly from it (which, therefore, are significant at a specific confidence interval). Effects that fall in a zone close to the normal (dashed line) evidently vary due to stochastic causes and, consequently, are presumably not significant. On the contrary, greater effects (e.g., E = blade wear) that fall far from the normal should be considered statistically significant. The effects can be studied similarly through a Pareto chart, as visible in FIGURE 4, where the ordered vertical bars represent the effects due to the related factor in absolute terms.

To avoid spurious outcomes and draw incorrect conclusions, it is necessary to verify that the identified effects are significant (E, D, C, A, CD, CDE, BE, B, ABCDE, ABCD, ACE, ACD, ABC, AB) through an ANOVA and an analysis of associated residual errors. In particular, it is necessary to calculate the sum of the squares of the major effects and that of the minor (residual) effects. From these, it is possible to calculate the related F-values. By comparing these with the reference distribution, it is possible to obtain consequently the probability that the factors - considered as significant – are such (i.e., the acceptable risk value to reject the null hypothesis). As shown in TABLE 3,

TABLE 3: 5-WAY ANOVA RESULTS CALCULATED WITH RESPECT TO THE MEASURED CUT QUALITY. ALL THE INTERACTIONS OF FACTORS (WHICH ALSO OBTAINED LOWER F-VALUES) HAVE BEEN AVOIDED FROM THE TABLE FOR THE SAKE OF CONCISENESS.

Factor	Sum Sq.	D.O.F.	Mean Sq.	F	Prob>F
E	78.84	1	78.84	338.91	0
D	14.26	1	14.26	61.30	0
C	7.59	1	7.59	32.64	0
A	7.59	1	7.59	32.64	0
…	…	…	…	…	…
Error	18.84	81	0.23		
Total	150.16	95			

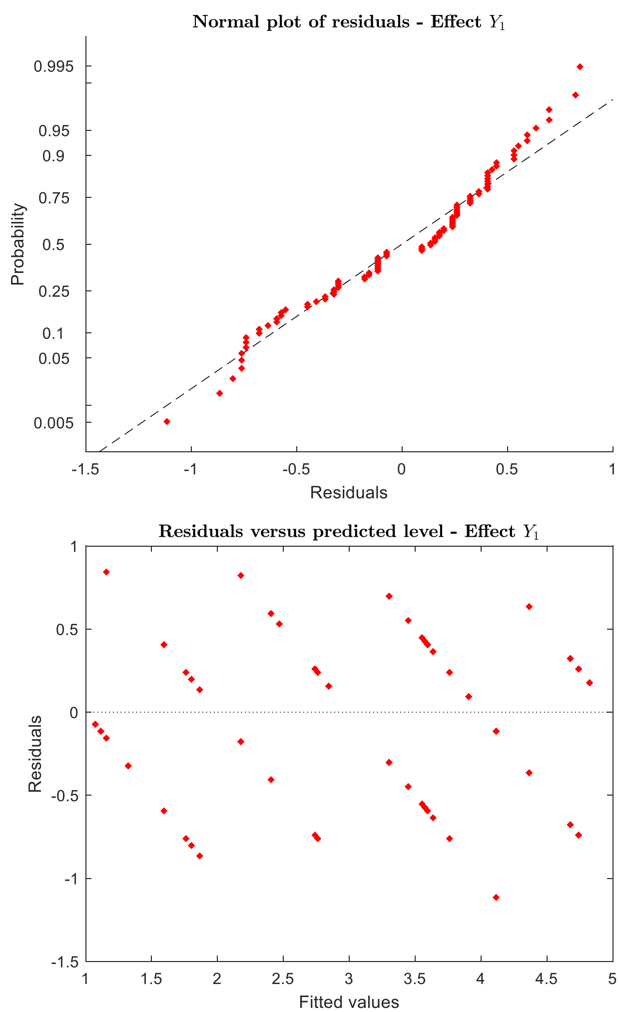

FIGURE 5: NORMAL PLOT OF RESIDUALS WITH RESPECT THE NORMAL DISTRIBUTION (DASHED LINE) AND COMPARISON BETWEEN RESIDUALS AND PREDICTED LEVELS FOR THE VALIDATION OF THE STATISTICAL ASSUMPTIONS (I.E., NORMALITY AND HOMOSKEDASTICI-TY).

carrying out a 5-way ANOVA, it is possible to demonstrate that all the effects identified as significant are such at a 5% confidence interval. However, in addition to a pure statistical analysis, it is necessary to consider that the interaction of three or more factors could be improbable to verify, as mentioned above.

The most important result of this analysis is not so much the identification of all the significant effects but consists of the detection of blade wear (factor E) as the first and principal significant effect. Indeed, it is possible to observe the weight of the blade wear compared to the other factors both in FIGURE 3 and thanks to Fisher's F-values in TABLE 3. Hence, this observation confirms that the blade wear is recognizable from the cut quality and, above all, validates the hypothesis that the features mainly describe this parameter. Finally, it should be remembered that the conditions of normality and homoskedasticity (i.e., the assumptions underlying the ANOVA) were checked thanks to the Q-Q plot, the Jarque-Bera and the Lilliepots test for the first and Bartlett's test for the second [21]. In particular, while homoskedasticity (i.e., property which characterizes a set of variables when they all have the same variance) has always been proved, not all tests on normality for each parameter have been verified (this may be mainly due to the presence of outliers). However, this ANOVA and the subsequent statistical analyses prove to be robust to the violations of these hypotheses, particularly if all the inspected groups show equal numerousness.

For the sake of completeness, FIGURE 5 shows the residual graphs (represented as a normal plot of residuals and residuals-predictions comparison) in order to validate the statistical assumptions. In particular, it can be observed that the residuals in the top graph follow rather faithfully the normal trend, while the residuals in the bottom graph are arranged uniformly and homogeneously with respect to the predicted levels. This means that the statistical hypotheses assumed that the residuals are normally distributed and have constant variance have been confirmed. For these reasons, the aforementioned analysis can be considered statistically valid.

2.3 Features extraction and selection

The features extraction only concerns the accelerometric signals. However, the channels recorded by the encoder are fundamental for the identification of the signal partition inherent to the actual cutting phase and, consequently, for the isolation of these data from the entire acquisition. For this reason, the two channels (A and B) of the incremental encoder (represented by a square wave measured in voltage) have been suitably processed to trace the equation of motion in terms of the servomotors' rotation angle θ and to subsequently obtain the blade position thanks to the kinematic relations of the crank mechanism.

Thanks to this processing, it is possible to obtain an average resolution on the blade position equal to about 0.04 mm. For instance, FIGURE 6 shows the trend of the normalized blade position with respect to its stroke during a cut. Thanks to this information, it is possible both to recognize the stroke end of the

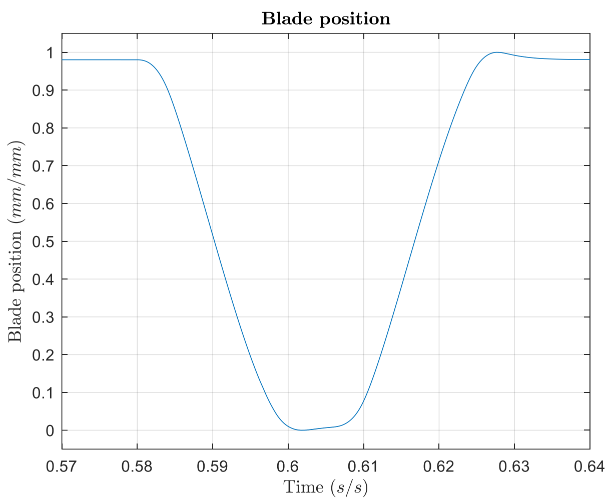

FIGURE 6: BLADE POSITION OBTAINED FROM THE ENCODER MEASUREMENTS DURING A CUTTING CYCLE AS A FUNCTION OF TIME. PLEASE NOTE THAT THE VALUES ARE NORMALIZED WITH RESPECT TO THE BLADE STROKE. PLEASE NOTE THAT BLADE POSITION AND TIME HAVE BEEN NORMALIZED WITH RESPECT TO THEIR MAXIMUM VALUES.

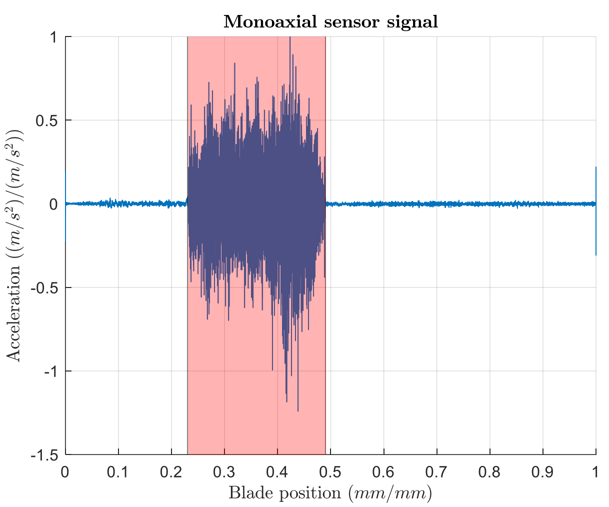

FIGURE 7: GRAPH REPRESENTING THE SIGNAL RECORDED BY A MONO-AXIAL ACCELEROMETER (RESPECTIVELY THE ONE POSITIONED ON THE KNIFE MATCHING PART) AS A FUNCTION OF THE BLADE POSITION. THE SIGNAL PARTITION INCLUDED BETWEEN THE CUT STARTING AND ENDING MOMENT HAS BEEN HIGHLIGHTED. PLEASE NOTE THAT BLADE POSITION AND ACCELERATION SIGNAL HAVE BEEN NORMALIZED WITH RESPECT TO THEIR MAXIMUM VALUES.

blade (coinciding with the moment in which the cut ended) and to discover the instant in which the cut started (knowing *a priori* the book thickness). FIGURE 7 shows a graph of an accelerometric channel in which the selected signal partition is highlighted and on which the features are extracted.

After having selected and isolated the signal partitions regarding the cut, it is possible to proceed with the features extraction. In a first analysis, the following features were selected as the most commonly used:

- Mean value:

$$MEAN = \frac{\sum_{k=1}^{N} y(t_k)}{N} \tag{1}$$

- Root Mean Square (RMS):

$$RMS = \sqrt{E[y(t_k)^2]} \tag{2}$$

- Skewness:

$$SKEW = E\left[\left(\frac{y(t_k)-\bar{y}(t_k)}{\sigma_y}\right)^3\right] \tag{3}$$

- Kurtosis:

$$KURT = E\left[\left(\frac{y(t_k)-\bar{y}(t_k)}{\sigma_y}\right)^4\right] \tag{4}$$

- Peak:

$$PEAK = \max(|y(t_k)|) \tag{5}$$

where $y(t_k)$ represents the acceleration signal as a quantity discretized by the acquisition process, and N is the number of samples.

To summarize, five features were considered for each signal collected by the six accelerometric channels (three monoaxial sensors and a triaxial one) for a total of 30 features for each of the 96 acquisitions.

Before proceeding with the extraction of the aforementioned features, the signals were further pre-processed to eliminate all the confounding channels by recognizing the outliers. In particular, given the high sensitivity to various considered factors (e.g., blade wear, grammage, and type of paper), it was observed that some tests led to the saturation of the uniaxial accelerometer positioned on the mobile blade. In this way, it was possible to remove these channels from the analysis, thus eliminating any related errors.

Once it has been verified that all the features are valid and not affected by errors, the features have been standardized as z-scores with respect to the healthy blade tests and a straightforward statistical analysis was subsequently defined to select the most relevant features in terms of identification of the blade wear. Student's T-tests [22] have been carried out to compare the distributions of the two wear conditions (i.e., new and damaged blade) and to verify their significant deviation. For this reason, having only two distributions, it is not necessary to carry out a one-way ANOVA (one-way since, in this case, the only effect to identify is the blade one). Furthermore, it will not even be necessary to carry out post-hoc ANOVA analyses (i.e., multivariate comparison techniques to investigate which groups show different mean values, such as Tukey's Honestly Significant Difference [23] or Fisher's Least Significant Difference [24]) since, if the null hypothesis is rejected, the deviation can only be revealed between the two distributions under analysis.

Once these 30 T-tests were performed, it was possible to recognize which of these features were statistically significant at a 5% confidence interval. Furthermore, by sorting the features by increasing p-values, it was possible to observe which are the most significant among those that allow rejecting the null hypothesis. TABLE 4 lists the features that allow rejecting the null hypothesis (i.e., the two distributions are not significantly distinguishable) sorted by increasing p-values. In addition, TABLE 4 shows that the first features (in terms of significance) are the mean values of the acquisitions obtained with monoaxial accelerometers. It was possible to recognize a preferential direction for the blade condition monitoring thanks to the triaxial accelerometer since its X-axis seems to perform properly.

TABLE 4: FEATURES THAT ALLOW DISTINGUISHING THE TWO DISTRIBUTIONS WITH NEW AND DAMAGED BLADES AT A 5% CONFIDENCE INTERVAL SORTED BY INCREASING P-VALUES (M1=MONOAXIAL ON THE STRUCTURE; M2=MONOAXIAL ON THE BLADE; M3=MONOAXIAL ON THE BLADE CORRESPONDING PART; TX=X-AXIS OF THE TRIAXIAL SENSOR).

Rank	Feature	Sensor	p-value
1	Mean	M1	$8.05 \cdot 10^{-15}$
2	Mean	M3	$7.73 \cdot 10^{-12}$
3	Mean	TX	$3.42 \cdot 10^{-6}$
4	Peak	M3	$1.18 \cdot 10^{-4}$
5	Peak	M1	$1.82 \cdot 10^{-4}$
6	Skewness	M2	$3.34 \cdot 10^{-3}$
7	Kurtosis	M2	$6.16 \cdot 10^{-3}$
8	Peak	M2	$9.44 \cdot 10^{-3}$
9	Kurtosis	M1	$9.74 \cdot 10^{-3}$
10	RMS	M1	$1.59 \cdot 10^{-2}$
11	RMS	M3	$4.23 \cdot 10^{-2}$

To conclude, FIGURE 8 shows an exemplary graph representing the *mean value* features of the accelerations measured by each sensor as a function of the run order. The vertical line represents the test division between the new (on the right) and damaged (on the left) blades. In a first analysis, the two conditions are rather effortlessly recognizable since data have a greater dispersion with the new blade.

For the sake of completeness, FIGURE 9 shows the histograms of the signal recorded by the monoaxial accelerometer (M1) positioned on the top of the structure (because it is very significant and has a non-zero mean during cutting) obtained by analyzing the entire acquisition and the solely partitioned samples concerning the actual cut. This comparison aims to demonstrate how the mean value is effectively zero throughout the entire acquisition for all accelerometers that are not subject to displacements, while this is not true when the solely cut partition is analyzed.

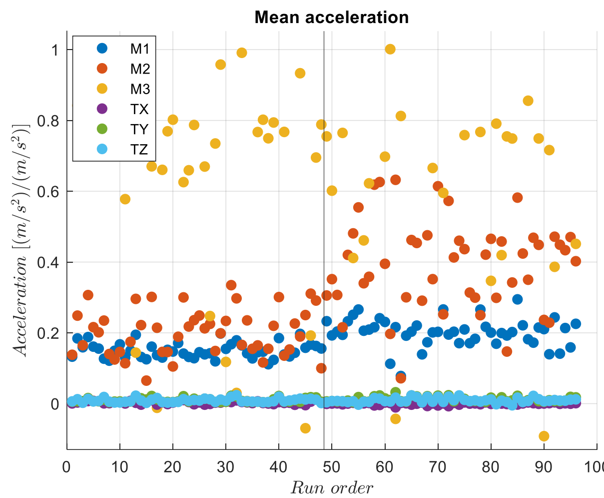

FIGURE 8: GRAPH REPRESENTING THE *MEAN VALUE* FEATURES OF THE ACCELERATIONS MEASURED BY EACH SENSOR AS A FUNCTION OF THE RUN ORDER. THE VERTICAL LINE REPRESENTS THE TEST DIVISION BETWEEN THE NEW (ON THE RIGHT) AND DAMAGED (ON THE LEFT) BLADES. PLEASE NOTE THAT THE VALUES HAVE BEEN NORMALIZED WITH RESPECT TO THE MAXIMUM VALUE.

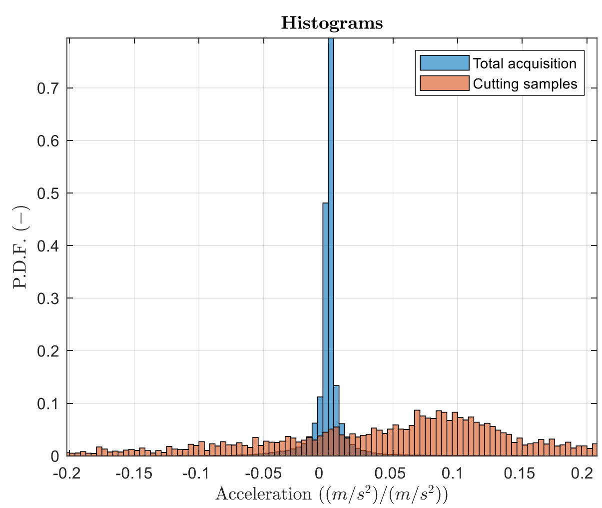

FIGURE 9: ENLARGEMENT OF HISTOGRAMS REGARDING THE ACCELERATION RECORDED BY THE MONOAXIAL SENSOR POSITIONED ON THE TOP OF THE STRUCTURE DURING TEST N.62 OF THE RUN ORDER REFERRED TO THE ENTIRE ACQUISITION (IN BLUE) AND THE PARTITIONED SAMPLES INHERENT TO THE ACTUAL CUT (IN RED).

2.4 Classification

After the features extraction and selection with a view to the most significant in terms of the blade wear identification, it is possible to proceed with the patterns recognition within the data to classify them in an appropriate multidimensional space. Before proceeding with this classification, it would be possible to carry out a more general MANOVA analysis (Multivariate ANalysis Of VAriance), which is a procedure to compare the sample means in a multivariate space [25]. This MANOVA analysis – in addition to Student's T-tests – would allow to consider all the input parameters simultaneously and, consequently, to compare and distinguish the distributions in a multivariate space. However, this MANOVA analysis would return an overall p-value, which would reveal the effective recognition of the variable under observation (i.e., the blade condition) without having further information about their classification. For this reason, it has been decided to carry out the classification directly, in such a way as to be able to study the difference in the sample means and, at the same time, to create and investigate the models as the features and the dataset vary.

The purpose of this analysis is to find and create a model that can recognize patterns with high accuracy (thanks to an initial training phase) and can subsequently predict them with a high degree of precision (thanks to a successive validation phase). Considering the number of available tests, it was decided to use the Cross-Validation (CV) technique to avoid overfitting problems. Since the K-fold CV technique consists of a random subdivision of data into K groups (one used as a validation dataset, while the remaining $K-1$ as training), the choice of these data groups can occur in different ways while maintaining the subgroups balanced with respect to the two classes. This means that each of these subdivisions can lead to different results because the models are generated in connection with the dataset, and they return varying, albeit slightly, accuracy values. Consequently, a Monte Carlo Cross-Validation (MCCV) [26] was applied to obtain more precise results thanks to the random effect reduction due to the division into $K = 5$ folds. It is possible thanks to the CV iteration for N times, in addition to the random subdivision of the samples into the training and validation subgroups. In particular, the same procedure described in [10] has been adopted with $N = 50$. However, given that a single type of damage was considered in the case under analysis, the results are calculated and interpreted mainly thanks to the accuracy indexes, Missed Alarms (MA), False Alarms (FA), Area Under the Curve (AUC) denominated Receiver Operating Characteristic (ROC), and the Frobenius Norm (FN). Thanks to this very effective MCCV method, the average accuracy and - in general - the performance indices considered tend to the theoretical value of the generated model. Please refer to [10] for a detailed description of the adopted MCCV method.

To conclude, it is recalled that the classification models that have been taken into consideration in this analysis are the Linear Discriminant Analysis (LDA) and the k-Nearest Neighbors (kNN) [27] with $k = 2$.

3. RESULTS AND DISCUSSION

It has been possible to adopt some supervised ML techniques to classify the blade health conditions, thanks to the DOE, the related statistical analysis for the identification of significant effects, and the features extraction and selection. Given that some of the 11 features selected based on their significance were obtained thanks to the uniaxial accelerometer M2 positioned on the mobile blade (which was subject to overload with consequent loss of information on some tests), it is necessary to circumvent this lack of information, albeit minimal. Therefore, this section contains the results obtained both (a) by keeping all the features and completely removing the overloaded tests, and (b) by only reducing the number of features employed for the classification and removing the channels inherent to the saturated accelerometer. In this way, it is possible to demonstrate whether it is more expedient to reduce the number of tests or features. The accuracy values obtained with the proposed method are reported in TABLE 5 and TABLE 6 as the features and tests considered vary.

An example of confusion matrices and ROC curves obtained during one out of $N = 50$ iterations carried out with MCCV are reported in FIGURE 10 and FIGURE 11 to highlight the behavior of the two classifiers.

TABLE 5: PERFORMANCE INDICES OBTAINED WITH LDA AND kNN CLASSIFIERS BY REMOVING THE TESTS IN WHICH SATURATION OCCURRED (a).

Features	Accuracy	MA	FA	AUC	FN
LDA	86,2%	3,1%	10,7%	0,99	0,36
kNN	91,2%	4,4%	4,4%	0,99	0,18

TABLE 6: PERFORMANCE INDICES OBTAINED WITH LDA AND kNN CLASSIFIERS BY REMOVING THE FEATURES RELATED TO THE SATURATED SENSORS (b).

Features	Accuracy	MA	FA	AUC	FN
LDA	89,8%	1,0%	9,2%	0,99	0,26
kNN	91,2%	5,3%	3,5%	1,00	0,18

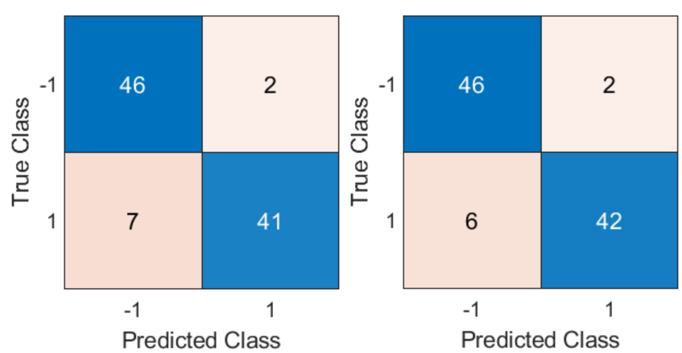

FIGURE 10: CONFUSION MATRIX INHERENT TO LDA (ON THE LEFT) AND kNN (ON THE RIGHT) MODELS DURING ONE OUT OF $N = 50$ ITERATIONS CONDUCTED VIA MCCV.

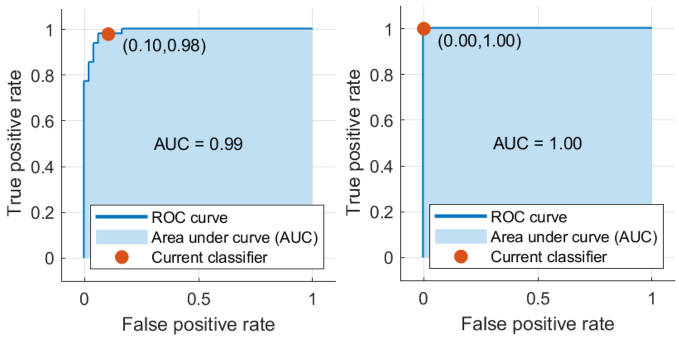

FIGURE 11: ROC CURVES INHERENT TO LDA (ON THE LEFT) AND kNN (ON THE RIGHT) MODELS DURING ONE OUT OF N = 50 ITERATIONS CONDUCTED VIA MCCV.

Finally, the parallel coordinates plot is shown to verify that the selected features can correctly distinguish the two conditions. This chart allows visualizing data with multiple dimensions on a single 2D diagram. In particular, the standardized values of each feature (i.e., z-scores) per each test are represented, taking the form of approximately horizontal lines. In this way, the distribution of the calculated data with respect to each feature is visible in a two-dimensional chart since the data projection regarding each feature is represented on their axes. This can help understand the relationships between features and identify the best features for separating classes. Indeed, the more the lines of each class are separable from each other, the more the blade conditions are distinguishable. All the 11 selected features have been reported in FIGURE 12 in such a way as to analyze which allows a better distinction of the blade conditions. It is possible to observe how the "bundles" of lines representing the new blade and the worn blade conditions are distinguishable thanks to the highest p-values features, while these intertwine more frequently, making it difficult to classify with the lowest p-values features. Therefore, it is possible to observe that the p-values obtained from the previous statistical analysis allow adequately selecting the most relevant features to recognize the state of health of the blade. Finally, it is possible to observe in FIGURE 13 how the scatter plot representing the tests in a 2D-reduced multidimensional space are distinguishable and separable.

To conclude, these supervised learning models allow recognizing the machinery state of health with good results both by varying the number and type of features and the dataset. Indeed, both the examined classifiers allowed reaching accuracy values close to 90%. LDA model could be preferred to kNN with the same accuracy values since the first can involve advantages in terms of simplicity and computing speed. In particular, this would be even more relevant if the possibility of continuously expanding the training dataset thanks to this monitoring system implementation on operating machinery is considered. This new data collection would allow generalizing the obtained results to different machines and distinct configurations of the initial setting at the same time. To reach this conclusion with the current dataset, it would be necessary to assume that the cloud of points obtained considering the initial conditions (i.e., with a new blade) is unaltered.

A second aspect to consider for the classification model choice concerns the cost-safety trade-off. Indeed, the error made by LDA models is principally caused by the failure to recognize borderline cases of the healthy blade as the higher FA percentages prove. On the contrary, kNN models mainly confuse some damaged blade cases, generating MA.

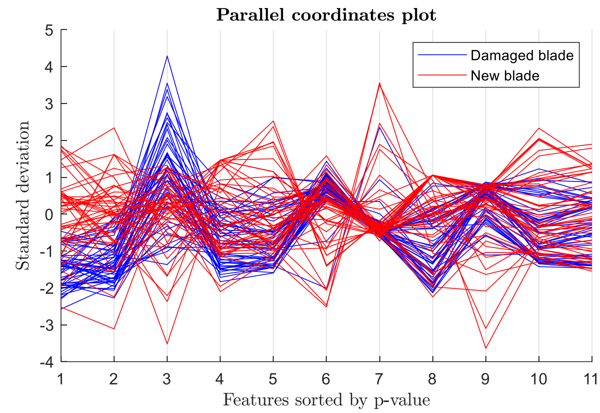

FIGURE 12: PARALLEL COORDINATES PLOT OF THE 11 SELECTED FEATURES SORTED BY p-VALUES AS IN TABLE 4.

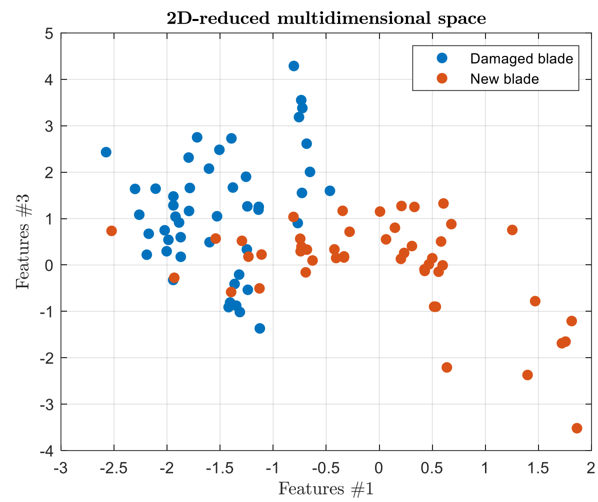

FIGURE 13: SCATTER PLOT REPRESENTING THE NEW BLADE AND DAMAGED BLADE TESTS IN A 2D-REDUCED MULTIDIMENSIONAL SPACE.

4. CONCLUSIONS

This work proposed the development and implementation of a diagnostic/prognostic tool applied to an industrial automated book trimmer in order to improve the safety, the reliability, the efficiency and the final quality of the production line under analysis. This was achieved by implementing a condition monitoring strategy by means of accelerometers that can be integrated into a SCADA system. Thanks to an in-depth preliminary study of the system and an ad hoc features extraction and selection, it was possible to apply some supervised ML techniques which permit the monitoring and prediction of the machinery state of health, allowing the adoption of predictive maintenance strategies.

The results obtained are very positive and promising. This means that the proposed method is able to successfully identify the blade state of health despite the operating parameters and working conditions vary. The algorithm speed and its integration into the industrial line permit a maintenance strategy based on real-time condition monitoring. This diagnostic method is suitable for applications oriented to the Industry 4.0 paradigm and to the digitalization of the industrial sector, which can be integrated with Internet of Things (IoT) and cloud systems.

The effectiveness of these results was obtained using a dataset consisting of 96 tests (equivalent to a Full Factorial design with five factors, repeated three times), statistically sufficient for a complete analysis and development of a diagnostic model. It exists the awareness that these promising results may slightly vary by extending the analysis and, consequently, the available dataset.

On this basis, it is possible to continue the study – for instance – by expanding the analysis to a wider machines population to generalize the model and to investigate different initial settings and further parameters. Furthermore, it would be interesting to study the evolution of the parameters considered during the entire life of the machine in order to distinguish its conditions with greater precision. Finally, it will be necessary to industrialize the entire diagnostic/prognostic system with an appropriate analysis inherent to the sensors downgrade and the automation of acquisitions.

ACKNOWLEDGEMENTS

The authors wish to thank Tecnau SRL for this collaboration and for making it possible to conduct experimentation and data acquisition on their machinery.

REFERENCES

[1] Patel, A. R., Ramaiya, K. K., Bhatia, C. V., Shah, H. N., and Bhavsar, S. N., 2021, "Artificial Intelligence: Prospect in Mechanical Engineering Field—A Review," *Data Science and Intelligent Applications*, K. Kotecha, V. Piuri, H.N. Shah, and R. Patel, eds., Springer, Singapore, pp. 267–282.

[2] Worden, K., and Dulieu-Barton, J. M., 2004, "An Overview of Intelligent Fault Detection in Systems and Structures," Structural Health Monitoring, **3**(1), pp. 85–98.

[3] Jardine, A. K. S., Lin, D., and Banjevic, D., 2006, "A Review on Machinery Diagnostics and Prognostics Implementing Condition-Based Maintenance," Mechanical Systems and Signal Processing, **20**(7), pp. 1483–1510.

[4] Accorsi, R., Manzini, R., Pascarella, P., Patella, M., and Sassi, S., 2017, "Data Mining and Machine Learning for Condition-Based Maintenance," Procedia manufacturing, **11**, pp. 1153–1161.

[5] Castellani, F., Garibaldi, L., Daga, A. P., Astolfi, D., and Natili, F., 2020, "Diagnosis of Faulty Wind Turbine Bearings Using Tower Vibration Measurements," Energies, **13**(6), p. 1474.

[6] Daga, A. P., and Garibaldi, L., 2019, "Machine Vibration Monitoring for Diagnostics through Hypothesis Testing," Information, **10**(6), p. 204.

[7] Pimentel, M. A., Clifton, D. A., Clifton, L., and Tarassenko, L., 2014, "A Review of Novelty Detection," Signal Processing, **99**, pp. 215–249.

[8] Japkowicz, N., Myers, C., and Gluck, M., 1995, "A Novelty Detection Approach to Classification," *In Proceedings of the Fourteenth Joint Conference on Artificial Intelligence*, pp. 518–523.

[9] Daga, A. P., Fasana, A., Garibaldi, L., and Marchesiello, S., 2020, "On the Use of PCA for Diagnostics via Novelty Detection: Interpretation, Practical Application Notes and Recommendation for Use," *PHM Society European Conference*, pp. 13–13.

[10] Viale, L., Daga, A. P., Fasana, A., and Garibaldi, L., 2022, "From Novelty Detection to a Genetic Algorithm Optimized Classification for the Diagnosis of a SCADA-Equipped Complex Machine," Machines, **10**(4), p. 270.

[11] Daga, A. P., Fasana, A., Garibaldi, L., and Marchesiello, S., 2020, "Big Data Management: A Vibration Monitoring Point of View," *2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT*, IEEE, pp. 548–553.

[12] Kim, H. E., and Jeon, T. H., 2016, "Evaluation of Feature Extraction Techniques for Intelligent Fault Diagnostics of High-Pressure LNG Pump," *Proceedings of the 10th World Congress on Engineering Asset Management (Wceam 2015)*, K.T. Koskinen, H. Kortelainen, J. Aaltonen, T. Uusitalo, K. Komonen, J. Mathew, and J. Laitinen, eds., Springer-Verlag Berlin, Berlin, pp. 553–562.

[13] Chandrashekar, G., and Sahin, F., 2014, "A Survey on Feature Selection Methods," Computers & Electrical Engineering, **40**(1), pp. 16–28.

[14] Carreira-Perpinán, M. A., 1997, "A Review of Dimension Reduction Techniques," Department of Computer Science. University of Sheffield. Tech. Rep. CS-96-09, **9**, pp. 1–69.

[15] Li, L., 2010, "Dimension Reduction for High-Dimensional Data," *Statistical Methods in Molecular Biology*, H. Bang, X.K. Zhou, H.L. van Epps, and M. Mazumdar, eds., Humana Press, Totowa, NJ, pp. 417–434.

[16] Anderson, M. J., and Whitcomb, P. J., 2017, *DOE Simplified: Practical Tools for Effective Experimentation*, CRC press.

[17] Antony, J., 2014, *Design of Experiments for Engineers and Scientists*, Elsevier.

[18] "Tecnau SRL," Tecnau [Online]. Available: https://www.tecnau.com/. [Accessed: 14-Apr-2022].

[19] McFadden, P. D., and Smith, J. D., 1986, "Effect of Transmission Path on Measured Gear Vibration."

[20] Czitrom, V., 1999, "One-Factor-at-a-Time versus Designed Experiments," The American Statistician, **53**(2), pp. 126–131.

[21] Kozak, M., and Piepho, H.-P., 2018, "What's Normal Anyway? Residual Plots Are More Telling than Significance Tests When Checking ANOVA Assumptions," Journal of agronomy and crop science, **204**(1), pp. 86–98.

[22] Witte, R. S., and Witte, J. S., 2017, *Statistics*, John Wiley & Sons.
[23] Abdi, H., and Williams, L. J., 2010, "Tukey's Honestly Significant Difference (HSD) Test," Encyclopedia of research design, **3**(1), pp. 1–5.
[24] Williams, L. J., and Abdi, H., 2010, "Fisher's Least Significant Difference (LSD) Test," Encyclopedia of research design, **218**, pp. 840–853.
[25] Warne, R. T., 2014, "A Primer on Multivariate Analysis of Variance (MANOVA) for Behavioral Scientists.," Practical Assessment, Research & Evaluation, **19**.
[26] Xu, Q.-S., and Liang, Y.-Z., 2001, "Monte Carlo Cross Validation," Chemometrics and Intelligent Laboratory Systems, **56**(1), pp. 1–11.
[27] Sammut, C., and Webb, G. I., 2011, *Encyclopedia of Machine Learning*, Springer Science & Business Media.