

Article

Monitoring the Damage Evolution in Rolling Contact Fatigue Tests Using Machine Learning and Vibrations

Luca Provezza ^{*}, Ileana Bodini , Candida Petrogalli, Matteo Lancini , Luigi Solazzi  and Michela Faccoli 

Department of Industrial and Mechanical Engineering, University of Brescia, Via Branze 38, 25123 Brescia, Italy; ileana.bodini@unibs.it (I.B.); candida.petrogalli@unibs.it (C.P.); matteo.lancini@unibs.it (M.L.); luigi.solazzi@unibs.it (L.S.); michela.faccoli@unibs.it (M.F.)

* Correspondence: luca.provezza@unibs.it; Tel.: +39-030-3715-539

Abstract: This study shows the application of a system to monitor the state of damage of railway wheel steel specimens during rolling contact fatigue tests. This system can make continuous measurements with an evaluation of damage without stopping the tests and without destructive measurements. Four tests were carried out to train the system by recording torque and vibration data. Both statistical and spectral features were extracted from the sensors signals. A Principal Component Analysis (PCA) was performed to reduce the volume of the initial dataset; then, the data were classified with the k-means algorithm. The results were then converted into probabilities curves. Metallurgical investigations (optical micrographs, wear curves) and hardness tests were carried out to assess the trends of machine learning analysis. The training tests were used to train the proposed algorithm. Three validation tests were performed by using the real-time results of the k-means algorithm as a stop condition. Metallurgical analysis was performed also in this case. The validation tests follow the results of the training test and metallurgical analysis confirms the damage found with the machine learning analysis: when the membership probability of the cluster corresponding to the damage state reaches a value higher than 0.5, the metallurgical analysis clearly shows the cracks on the surface of the specimen due to the rolling contact fatigue (RCF) damage mechanism. These preliminary results are positive, even if reproduced on a limited set of specimens. This approach could be integrated in rolling contact fatigue tests to provide additional information on damage progression.

Keywords: machine learning; rolling contact fatigue tests; damage assessment; metallurgical analysis



Citation: Provezza, L.; Bodini, I.; Petrogalli, C.; Lancini, M.; Solazzi, L.; Faccoli, M. Monitoring the Damage Evolution in Rolling Contact Fatigue Tests Using Machine Learning and Vibrations. *Metals* **2021**, *11*, 283. <https://doi.org/10.3390/met11020283>

Academic Editor: Luis Reis and Pedro Moreira

Received: 20 January 2021

Accepted: 2 February 2021

Published: 6 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Rolling contact fatigue (RCF) is typical damage in wheel and rail steels. The assessment of this typology of damage involves high costs and takes time, usually using destructive tests for the specimens adopted for the tests. Moreover, analytical models need much data to be calibrated and a great quantity of tests is required.

Accurate modeling on RCF has been studied until now: some studies have focused on the mechanical and metallurgical characterization of different steels for wheel and rail [1]; while other studies present the development of analytical models [2–5].

A twin-disk bench is a solution to characterize wheel and rail steel. The bench can operate in different controlled conditions: rotation speed, slip ratio, and load. In this way, the bench simulates the real operating condition of the wheel [6]. The traditional protocol [7] to characterize the damage of wheels and rails is to stop the tests after a predefined number of cycles during the tests, dismount the sample from the bench, and perform different tests: weighing measurements to evaluate the wear rate of the steel and acquire the surface state with a vision system [8,9]. At the end of the test, destructive measurements are performed to evaluate the effective damage of the steel.

The proposed study aims to quantitatively estimate the damage progress of a specimen during a test with continuous and non-contact measurements, using a non-supervised

machine learning technique [10–13], called k-means clustering [14], applied to vibrations and torque datasets. The idea was to validate a system developed in a previous work [15], to understand if the approach could be used to control the duration of the test. Preliminary RCF tests were carried out with the same wheel and rail steels to create a database used to train the algorithm. Then, a second set of tests were performed using the on-line system to evaluate the damage. The tests were stopped at different stages of the damage progression to confirm the results of the machine learning system with destructive measurements.

Vibrations and feedback signals are commonly used to classify the fault conditions of machines. Literature reports different works on this topic, chiefly applied to bearings of rotary machines. Convolutional Neural Network (CNN) is presented in many works as a valid algorithm to detect fault conditions of bearings parts [16–20]. In [21], the Remaining Useful Life (RUL) prediction of machining sensor data is performed with the combination of the Support Vector Machine (SVM) as a classification tool and Autoregressive and Integrated Moving Average (ARIMA) for the identification of the RUL. Meanwhile, in [22], a real-time classification of a cutting blade's status was developed. The model based on a logistic regression was verified with real measurement data from an industrial machine tool. Meanwhile, Bustillo et al. [23] have studied the experimental repetition of different machine learning algorithms, such as: regression trees, k-Nearest Neighbors (kNN), Artificial Neural Networks (ANN), bagging and Random Forest. The purpose of this work is to predict the tool-life in the face-turning operations of AISI 1045 steel discs, depending on different cooling systems and tool geometries. The results suggested that the use of raw experimental data, rather than their averaged values, can achieve machine-learning models of higher accuracy for tool-wear processes. In [24], acoustic emission features are correlated with corresponding natural wear of slow speed bearings through a series of laboratory experiments by using a neural network model and gaussian process regression. The authors in [25] propose a deep hybrid model that allows obtaining a higher accuracy than a single convolutional neural network. All these machine-learning algorithms are used to handle big amounts of raw data (until months of data recorded) and to detect fault conditions of the rotary machines. Having a high number of data is usually time-consuming and costly and it is important to find algorithms that need few data for training. In the area of unsupervised machine learning algorithms, k-means represent a powerful and simple algorithm to detect fault conditions of rotary machines [26–30]. K-means algorithm is normally used when the number of data is not excessive and usually matched with algorithms to reduce the number of features [13]. In [31], the k-means algorithm is used with a combination of vibration signals and thermal image features, to extract shape features using image segmentation.

In the current literature, many studies have been focused on quantifying the damage evolution and improvement of the fault condition, especially without stopping the tests [1,8,15,21,22].

The innovative contribution of this work is in the field of the small-scale RCF tests. Similar techniques have already been used to investigate variations in frequency responses of vibrations [32], but those approaches require operator-based manual tuning and extensive post-processing of the data. The proposed method is automatic and quantitatively estimates the damage in real-time during a test [33].

The paper is divided into the following sections: Section 2 presents the specimens, the instrumentation, the protocol and the data processing used to perform this experimental activity; Section 3 presents the experimental campaign carried out, divided into a training part and a validation part of the system, able to monitor the damage evolution of the wheel and rail steels. The discussion of the results is compared with the metallurgical analysis carried out on the specimen tested. In Section 4, some conclusions are discussed about this method for monitoring the damage evolution in rolling contact fatigue tests.

2. Materials and Methods

2.1. Specimens

The rolling contact tests were carried out with a two-disc test rig pairing a wheel steel disc as follower with a rail steel disc as driver. The diameter of the discs is 60 mm, and the contact surface width is 15 mm. The specimens were machined out of a new wheel rim and a new rail head with their axis perpendicular to the wheel tread and the rail longitudinal axis, respectively. The wheel discs are made of SANDLOS[®] S steel [7,34,35], which is type AAR CLASS C steel modified. SANDLOS[®] S steel has a higher content of Si than CLASS C and it is micro-alloyed with V and Nb. The last two chemical elements are added in accordance with maximum requirements ($V \leq 0.04\%$ and $Nb \leq 0.05\%$) to improve the cyclic yield strength and fracture toughness in the wheel tread. This steel is designed for interaction with head-hardened rails, also in the presence of sand or debris in the contact area. The rail discs are made of 350 HT EN 13674-1 steel, which is used for heat-treated rails. All steels are supplied by Lucchini RS. Their chemical composition and mechanical properties are shown in Table 1.

Table 1. Main chemical elements and mechanical properties of the wheel and rail steels.

Steels		SANDLOS [®] S	350 HT
Chemical Composition (%)	C	0.72	0.72
	Mn	0.80	1.04
	Si	0.87	0.30
	S	0.001	0.004
	P	0.010	0.015
Ultimate tensile stress (MPa)		1180	≥1175
Monotonic yield stress (MPa)		750	-
Elongation (%)		12	≥9
Hardness HB		347	345

2.2. Instrumentation

Figure 1 shows a schematic drawing of the twin-disk rig used in this work (University of Brescia, Brescia, Italy), which was developed internally as described in [2]. The shafts are independent, and a hydraulic actuator applies the imposed contact load. The rolling speed of the shafts is measured by encoders, and a load cell located at the piston head measures the contact load between specimens. The bench is equipped with a torque sensor with a full scale of 200 Nm positioned on the displaceable shaft.

Piezo-accelerometers (732a by Wilcoxon) with a full scale of 500 m/s², and a bandwidth of 26 kHz are fixed near each mandrel support, on the machine frame. Both the accelerometers and the torque sensors are acquired using a NI 9172 DAQ board (by NI), equipped with a NI9215 module (by NI) for the torque signal and a NI9233 module (by NI) for the accelerometers. The acquisition was performed using a custom software developed in Labview (by NI) running at a synchronous sampling frequency of 5 kHz for the whole duration of each test.

2.3. Protocol

The tests were performed under the same maximum contact pressure ($P = 1100$ MPa), rolling speed (500 rpm) and sliding/rolling ratio ($s = 1\%$). The contact pressure and the sliding speed conditions are typical of the wheel–rail contact on a curve. Before each test, the wheel and rail discs were cleaned in a bath of ethanol with ultrasonic vibrations and weighed by a precision balance with a resolution of 0.001 g. Both stepwise and continuous tests were performed: during the stepwise tests, the discs were periodically dismantled, ultrasonically cleaned, and weighed to evaluate their weight loss, while continuous tests were never interrupted, and the weight loss of the discs was measured only at the end of the tests. The curves of weight loss versus number of cycles were obtained for both the wheel and rail discs and the wear rates were calculated from the stepwise tests. At the end

of each test, the wheel disc was cut along the midplane orthogonally to the contact surface. The disc cross-section was ground, mechanically polished to a $1\ \mu\text{m}$ finish, etched with 2% Nital and examined with a Leica DMI 5000 M light optical microscope. The deformation under the contact surface and the crack morphology were investigated and the damage mechanisms were identified. The damage evolution of the wheel discs at an increasing number of cycles was documented. Vickers hardness tests were carried out on the cross-section of each wheel disc at varying distances from the contact surface to evaluate the steel work-hardening phenomenon and correlate it with the deformation beneath the contact surface. The tests were performed using a 1000 g load and a dwell time of 15 s, complying with ASTM E384 (2011).

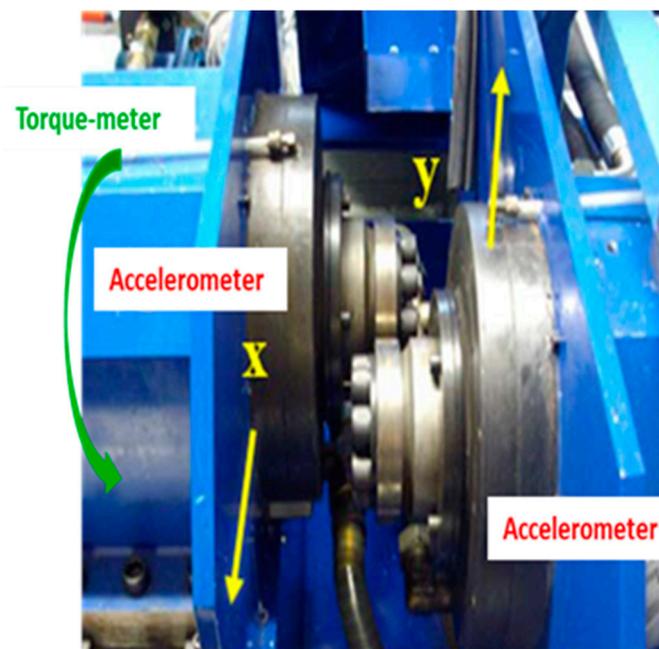
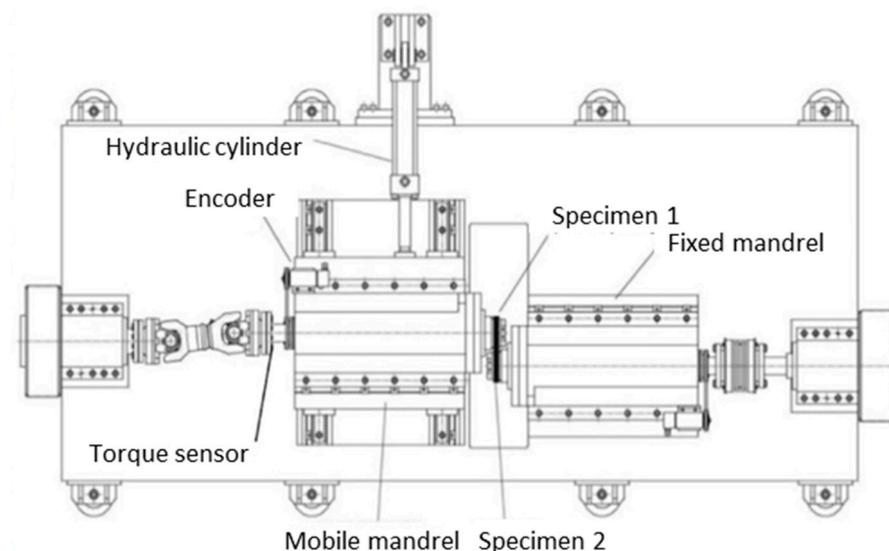


Figure 1. Schematic of the test rig (**top**) accelerometers and torque-meter positions (**bottom**).

2.4. Data Processing

The analog signals are recorded in packets of 0.2 s for faster post-processing, and further split in binary files representing 200 s to avoid the handling of big-size files. The

feature extraction and the relative weight of the features follow the same approach as in [15]. Software developed in Labview computed in real-time for all three signals (two vibration signals and one torque) and for each record, 44 features, aiming to get a detailed representation of the phenomenon every 0.2 s. The features extracted are, for each channel: mean, variance, root mean square, 75th percentiles, median, and the centroid of the power spectrum density, as shown in Table 2. Features depending on the combination of different signals were also computed for each channel pair: the maximum value of the cross-correlation function and its associated time delay, the centroid of the frequency response function, and the quartiles of the frequency response function cumulate as shown in Table 3.

Table 2. Features based on a single channel selected.

Feature	X Acceleration	Y Acceleration	Torque
Mean	Used	Used	Not used
Variance	Used	Used	Used
Root mean square	Used	Used	Used
Median	Not used	Used	Not used
75th Percentile	Not used	Used	Not used
PSD centroid	Used	Used	Used

Table 3. Features based on channel pairs selected.

	X Acceleration	Y Acceleration	Torque
X Acceleration		Frequency Response Function (FRF) centroid	Max Cross Correlation Time delay FRF centroid
		FRF 25th percentile	FRF 25th Percentile
		FRF median	FRF Median
		FRF 75th Percentile	FRF 75th Percentile
Y Acceleration	Max cross correlation		Max cross correlation
	Time delay		Time delay
	FRF centroid		FRF centroid
	FRF 25th percentile		FRF 25th percentile
	FRF median		FRF median
	FRF 75th percentile		FRF 75th percentile
Torque	FRF centroid	FRF centroid	
	FRF 25th percentile	FRF 25th percentile	
	FRF median	FRF median	
	FRF 75th percentile	FRF 75th percentile	Max cross correlation

The features were normalized following the Z-Score method: every feature was normalized by considering the mean and the standard deviation of the same feature considering all data available for training.

Then, a Principal Component Analysis (PCA) was computed to reduce the number of features using a limit of 90% of the variance explained, using the training data only. Then, the data described in the PC's space are computed by the k-means algorithm. The result of the k-means algorithm is a distance d_{ij} for each state, but this distance does not give quantitative information of belonging to a state. To solve this problem, the data are transformed in a probability through this formulation:

$$P_{ij} = \frac{1/(d_{ij})^2}{\sum_{k=1}^K 1/(d_{ik})^2} \quad (1)$$

where P_{ij} is the probability of the i -th record to belonging to a particular j state, d_{ij} is the distance of the i -th record of a j state, and d_{ik} is the distance of the other states calculated

by the k-means. In this way, the distance between the i -th data and a specific reference for that state is translated in a probability of belonging to that state, which is much easier to interpret than the simple distance.

2.5. Validation Approach

The validation of the proposed approach was carried out by using the real-time computation of the k-means algorithm. The real-time results of the k-means algorithm were used as a stop condition for the validation tests.

3. Experimental Results and Discussion

3.1. Test Performed

Table 4 shows the tests performed: four training tests with a predefined number of cycles, and three validation tests with the algorithm status used as a stop condition.

Table 4. Tests performed, and protocol followed.

Test ID	Protocol	Stop Condition	Cycles
T01	Stepwise	Predefined number of cycles	200,000
T02	Stepwise	Predefined number of cycles	200,000
T03	Continuous	Predefined number of cycles	200,000
T04	Continuous	Predefined number of cycles	200,000
V01	Continuous	Controlled by k-means results	10,000
V02	Continuous	Controlled by k-means results	25,858
V03	Continuous	Controlled by k-means results	100,000

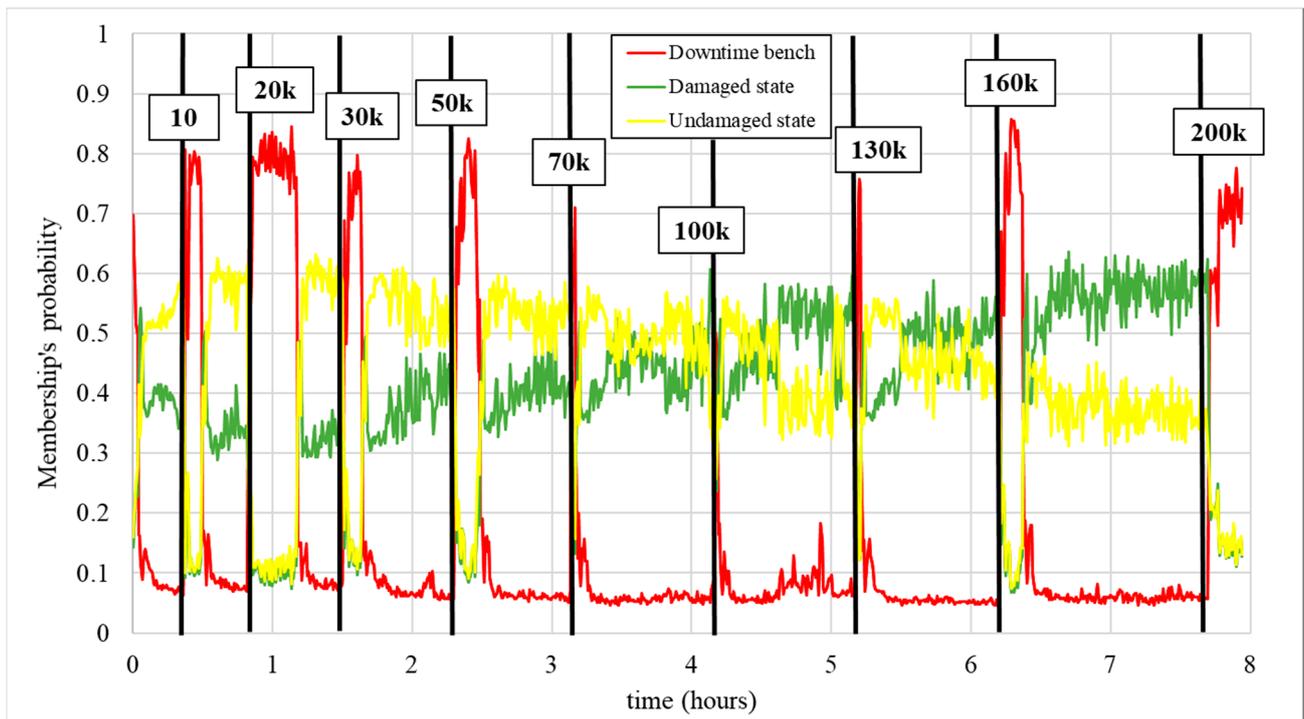
3.2. Training Tests

The training tests T01 and T02 were performed following with a stepwise protocol, while T03 and T04 followed a continuous protocol. Figures 2 and 3 show the results deriving from the elaboration of the datasets carried out with k-means: for each cycle, the probability of belonging to a cluster was calculated.

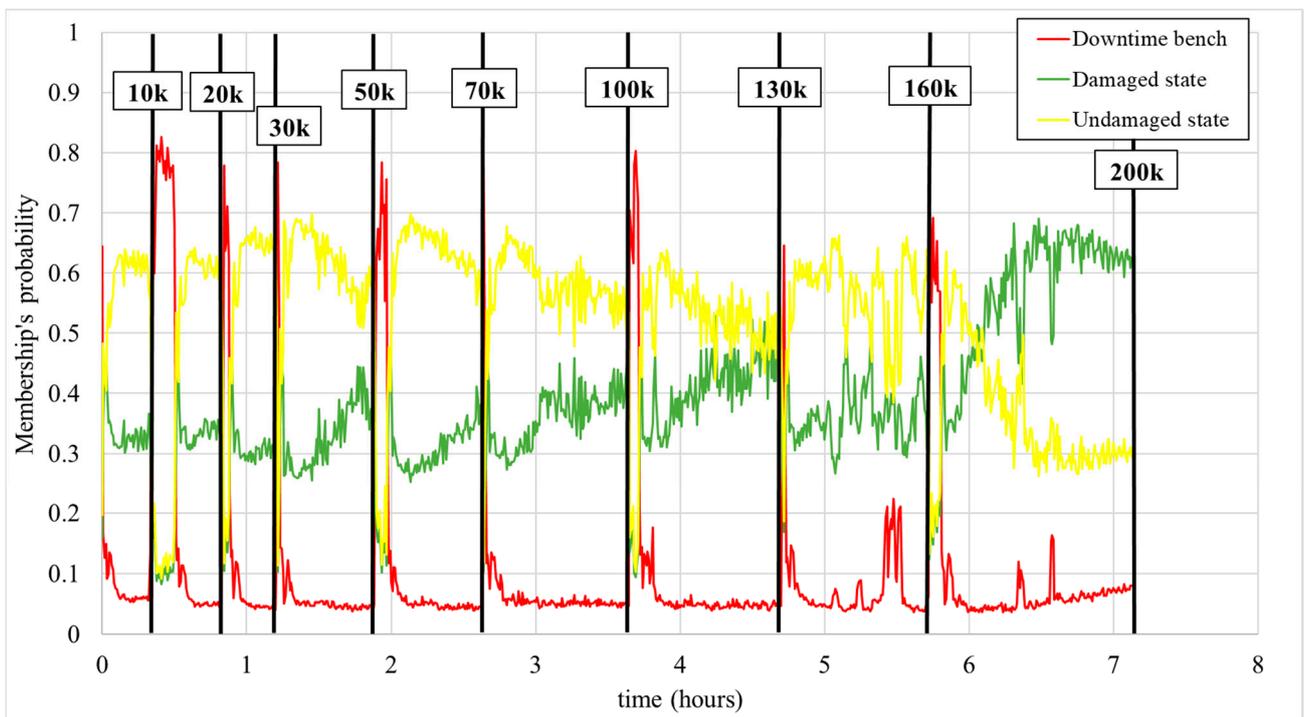
By analyzing the stepwise tests (Figure 2), three clusters were recognized from the k-means algorithm: cluster 1 (yellow) represents the undamaged state and it is the most frequently found during all the tests; cluster 2 (green) is associated with the damaged state and becomes prevalent only in the T02 test after 180,000 cycles. The least populated cluster, the number 3 (red), is the one that identifies the test rig downtime, occurring at 10,000, 20,000, 30,000, 50,000, 70,000, 100,000, 130,000, and 160,000 cycles.

The continuous tests (Figure 3) present a different trend. The undamaged state decreases after 70,000 cycles for T03, and after 20,000 cycles for the T04 tests. The damage state slowly increases its likelihood, reaching a condition of equal probability at 95,000 cycles for T03, and 25,000 for the T04 test. After that, the damaged state becomes the most likely until another change is detected by the algorithm, respectively, at 140,000 for T03, and the other change is detected at 100,000 for the T04 test. After these cycles, the undamaged state returns the most likely until the end of the test for the T04 test. For the T03 test, there is a third change that begins exactly at the end of the test with a return to the damaged state.

These trends found with machine learning analysis can be explained with the metal-lurgical analysis, in particular: optical microscope analysis, hardness tests, and weight loss with the calculated wear curves.

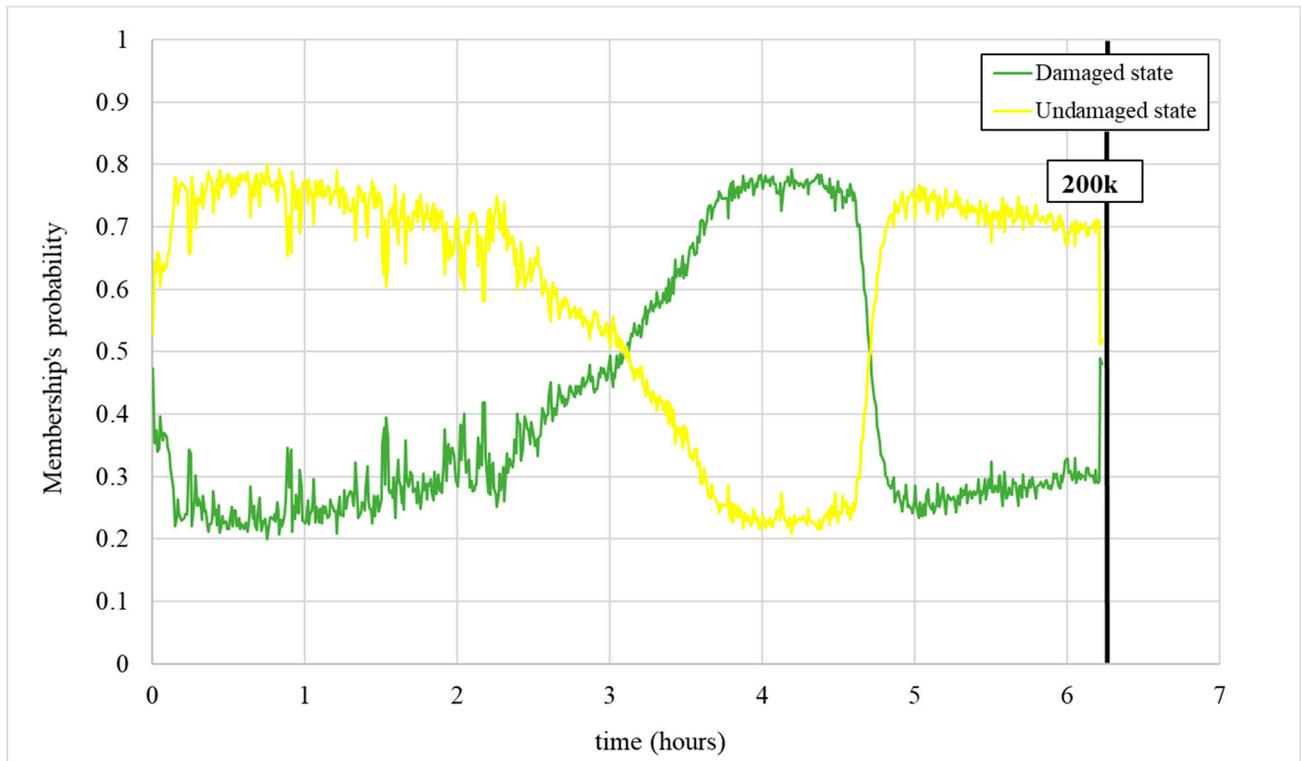


(a)

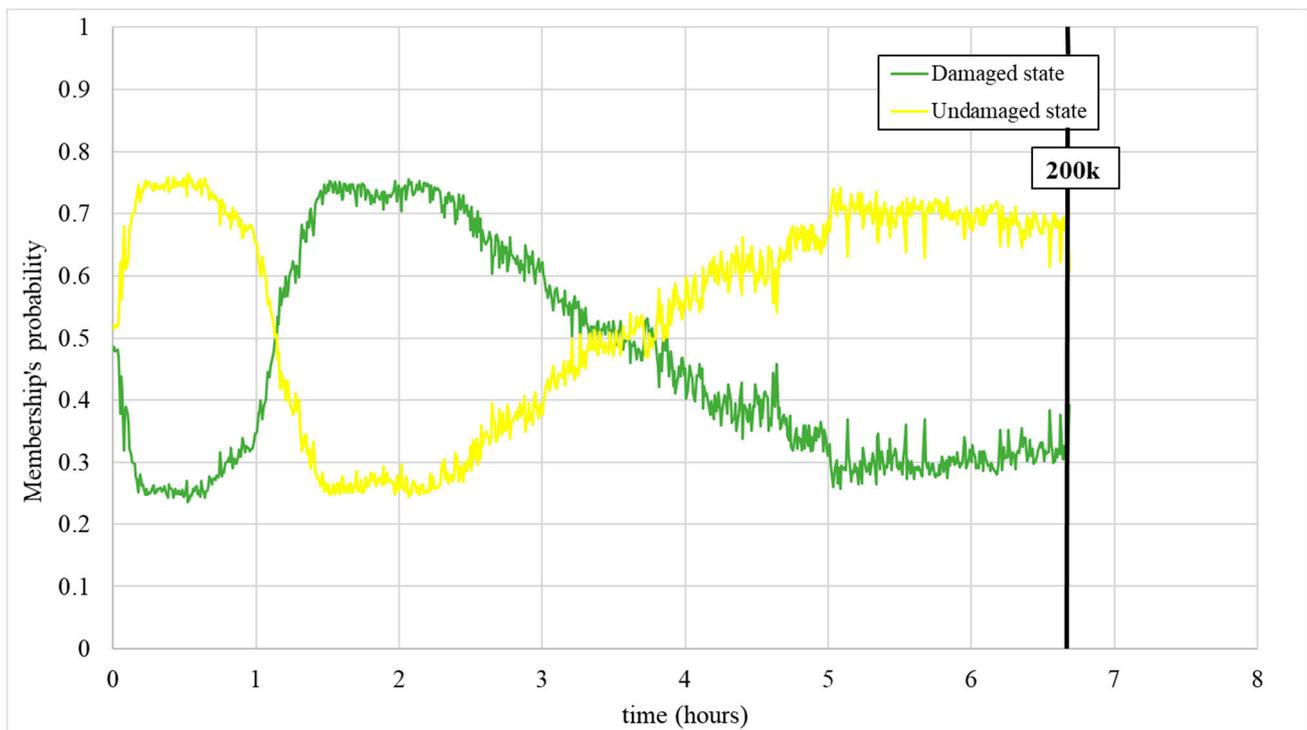


(b)

Figure 2. Probability of belonging to one of the three states (damaged, undamaged, downtime) deriving from the elaboration of the datasets carried out with k-means for the stepwise training tests T01 (a) and T02 (b).



(a)



(b)

Figure 3. Probability of belonging to one of the three states (damaged, undamaged, downtime) deriving from the elaboration of the datasets carried out with k-means for the continuous training tests T03 (a) and T04 (b).

Figure 4 shows the running surface of the wheel discs at the end of training tests T02 and T03. The running surfaces of the T02 and T03 specimens are completely damaged

with flaking appearing on the surface, which means that the damage is in progress. The machine learning analysis confirms the final state of damage and can suggest what could happen during the tests. For T02, the change in damage state occurs at 180,000 cycles, meanwhile, T03 has a rapid change in damage state at the end of the test.

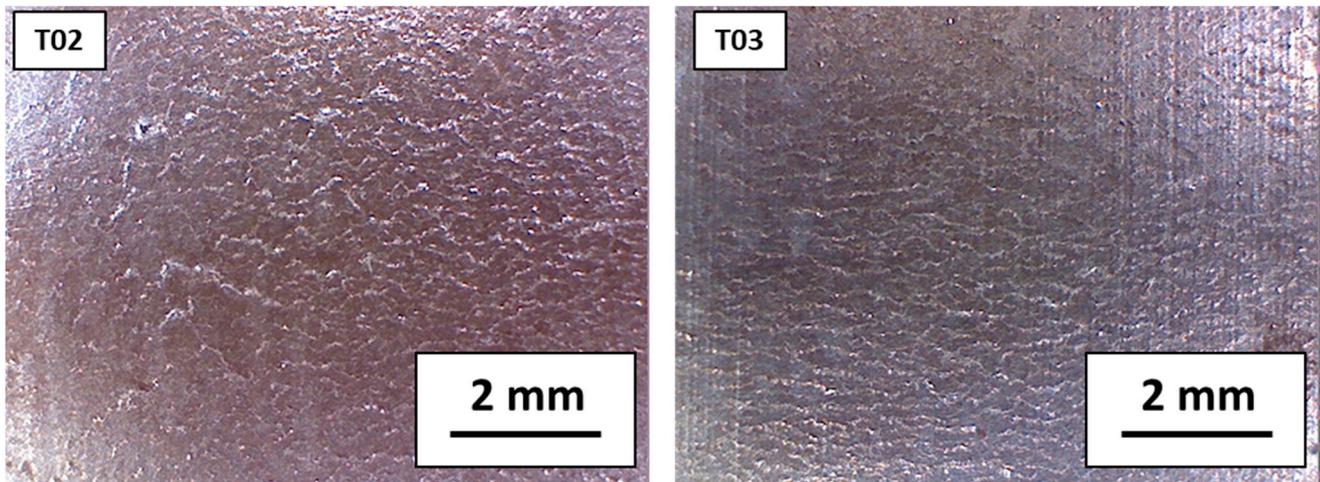


Figure 4. Running surfaces of the wheel discs at the end of the training tests (T02) and (T03).

The same conclusions can be drawn by considering the cross-section micrographs of Figures 5 and 6. In both cases, the cross-sections of the wheel discs T02 and T03 show surface cracks following the plastic deformed material.

An increase in hardness was observed under the contact surface in all tested discs. This result correlates well with the pattern of deformation observed metallographically in Figures 5 and 6, because the progressive accumulation of plastic strain under the contact surface during the test resulted in the steel hardening. The maximum hardening is close to the contact surface where the plastic deformation is more severe; then, the hardness gradually decreases at increasing distances from the surface because of the smaller deformation down to the value of the undeformed steel.

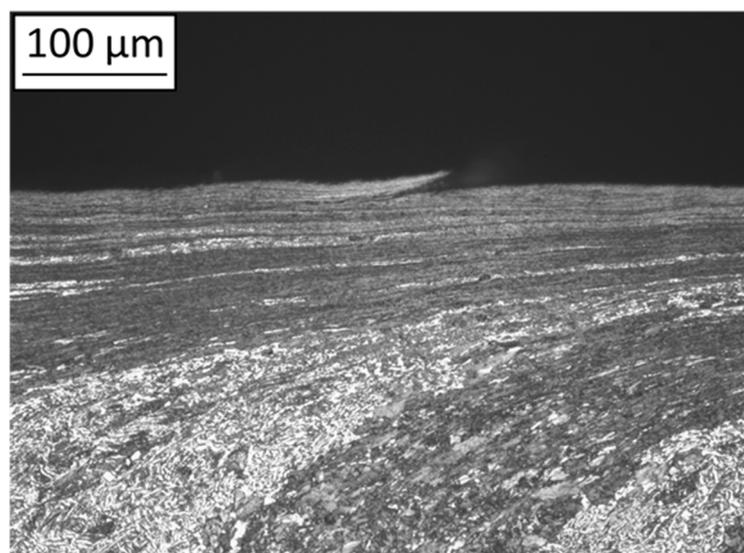


Figure 5. Cross-section of the wheel disc T02 at the end of the stepwise test (200,000 cycles).

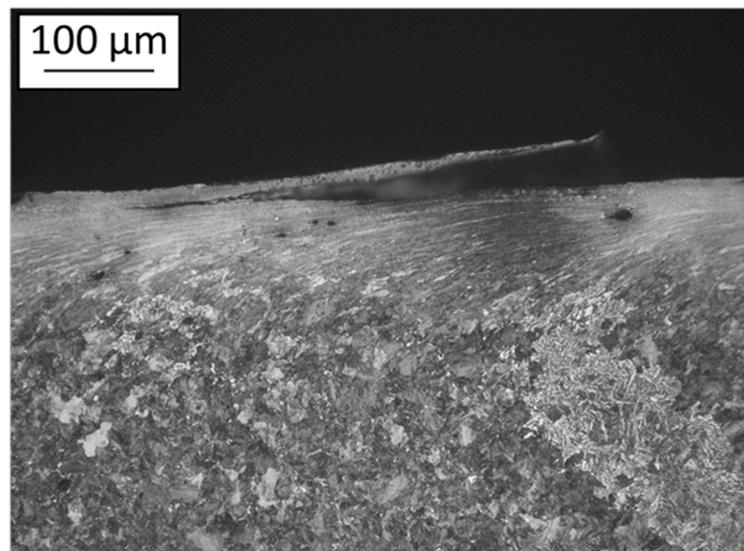


Figure 6. Cross-section of the wheel disc T03 at the end of the continuous test (200,000 cycles).

The accumulation of plastic deformation during the tests was also proved by the hardness profiles obtained on the disc cross-sections (Figure 7).

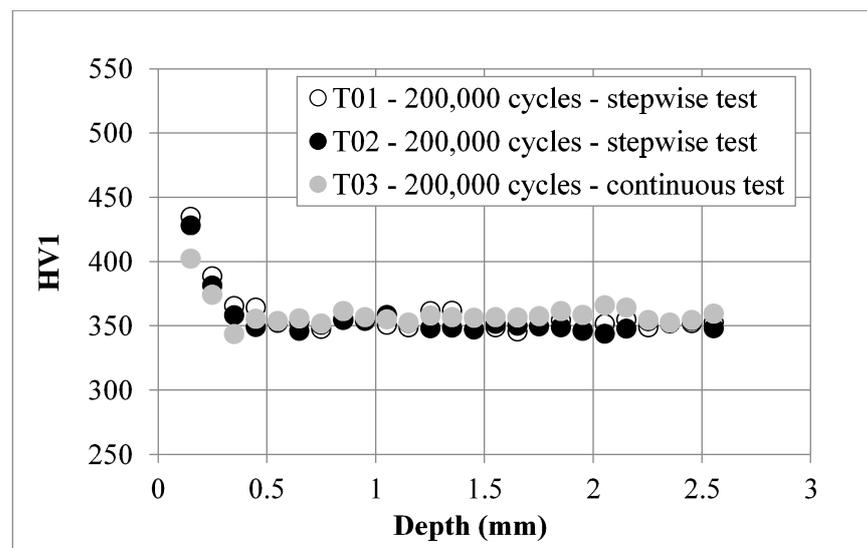


Figure 7. Vickers hardness as a function of the depth on the cross-section of the wheel discs for training tests.

The maximum hardness measured in the discs of the stepwise tests is higher than that in the disc of the continuous test and it suggested that the discs of the stepwise tests are in the initial stage of the damage, meanwhile, in the continuous test, the detachment of the surface material has already occurred with the consequence to have lower hardness near the surface in comparison with the stepwise tests.

Figure 8 shows the wear curves obtained from the continuous and stepwise tests. At first sight, the weight loss increases with the number of cycles because of progressive wear. In the stepwise tests, the wear curves show two different periods. There is an initial incubation period lasting 10,000 cycles where the wear rate is low and then steadily increases, followed by a steady-state period where the wear rate is higher and constant. The first period is probably due to the build-up of plastic strain in the disc surface before a threshold is reached and wear starts to take place [36]. Given the substantially linear

relationship between weight loss and number of cycles in the steady-state period of the curves, the wear rates of both wheel and rail discs were calculated for each test from the linear best fit of the experimental points.

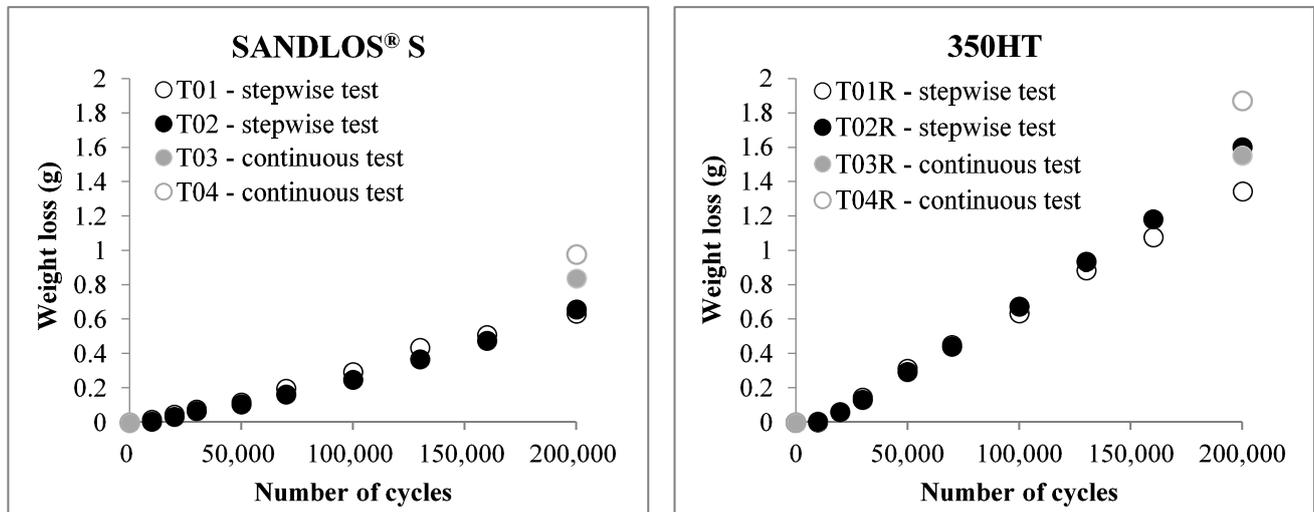


Figure 8. Weight loss of the wheel and rail discs in continuous and stepwise training tests.

The results are shown in Table 5: the wear rate of the SANDLOS® S discs is consistent with that obtained in a previous work carried out by the authors on the same steels in longer tests [7]. The weight loss of the wheel and rail discs measured during the two repeated stepwise tests can be superimposed, qualitatively indicating good reproducibility of the tests.

Table 5. Wear rate of the wheel and rail discs during the steady state period of the stepwise training tests.

Steel	Wear Rate (g/Cycle)
SANDLOS® S T01	3.34×10^{-6}
350HT T01R	7.10×10^{-6}
SANDLOS® S T02	3.08×10^{-6}
350HT T02R	7.90×10^{-6}

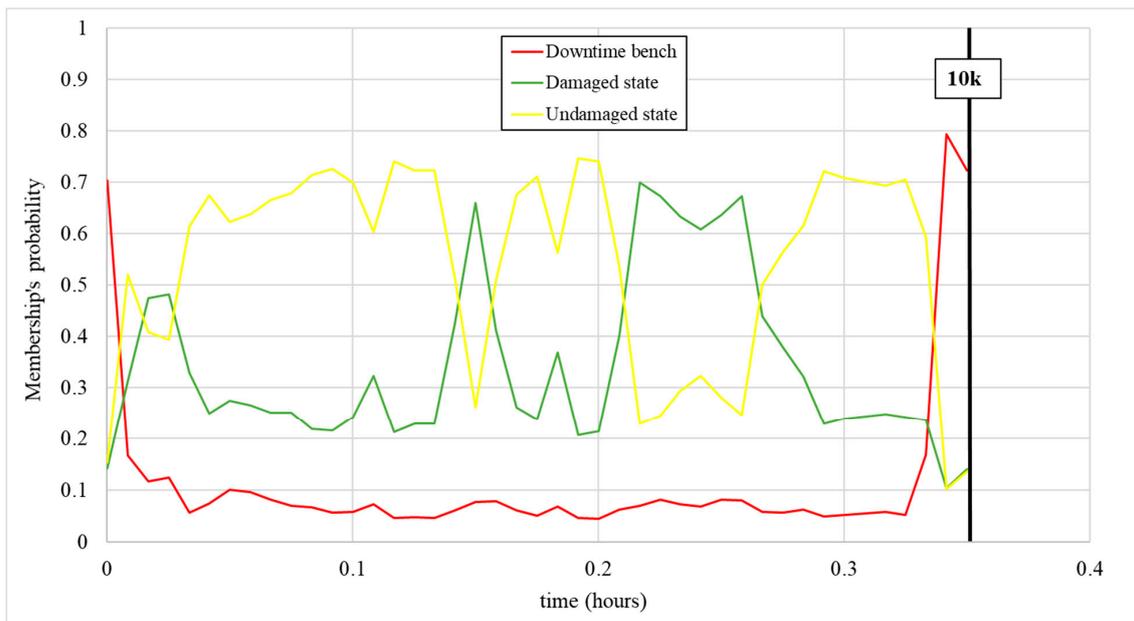
Another important consideration in the wear curves of the wheel discs is in the comparison of the weight loss between the continuous and stepwise tests at the end of tests. The weight loss of continuous tests is higher than the weight loss obtained in the stepwise tests. This confirms that the damage was more severe in the continuous tests.

These trends of training tests lead to an understanding that the RCF phenomenon is cyclic. At the initial stage of the test, the contact surface material accumulates plastic deformation that is proved by an increase in the hardness. In this state, the material is in the undamaged state. When the material reaches a limit in the accumulation of plastic deformation, the nucleation of surface cracks begins. The subsequent crack propagation leads to the detachment of material from the contact surface. In this case, the system is detecting a more likely damaged state. This condition remains until the material detachment leads to the complete removal of the plastic deformed material. After this, the process restarts from the accumulation of plastic deformation.

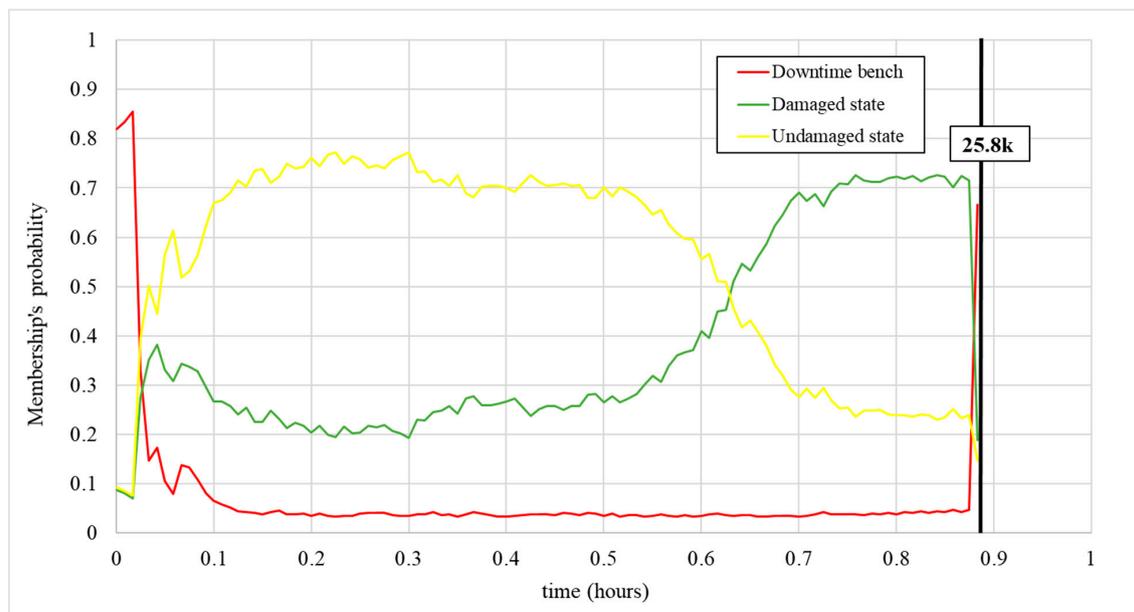
3.3. Validation Tests

Three validation tests were performed by using the real-time results of the k-means algorithm as a stop condition. The moments chosen to stop the tests were: in the undamaged state when the damage state begins to increase (Figure 9a), after the change when the

damage state becomes the most likely cluster (Figure 9b), and in the damage state at several cycles after the change (Figure 9c). V01 and V02 tests show a correct trend of the clusters: when the test is stopped or controlled by the k-means results, the cluster corresponding to the downtime bench becomes the most populated cluster. V03 test presents a moment when the cluster of the downtime bench becomes the most populated cluster at the initial stage of the test. This was not a downtime of the bench, but probably a moment where the algorithm recognizes a wrong cluster due to the vibration data with a low value. V01 presents a continuous change between the damage state (green) and undamaged state (yellow). This means that the two clusters are equiprobable and the damage is in an initial stage.

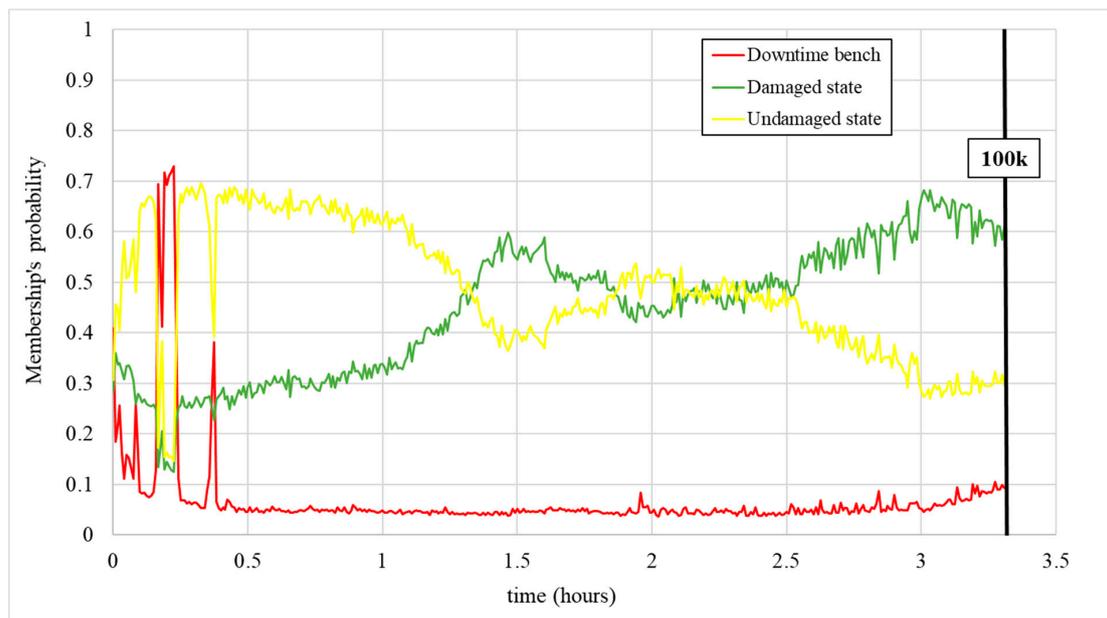


(a)



(b)

Figure 9. Cont.



(c)

Figure 9. Probability of belonging to one of the three states (damaged, undamaged, downtime) deriving from the elaboration of the datasets carried out with k-means for the validation tests V01 (a), V02 (b), and V03 (c).

The metallurgical analysis confirms the results found in the training tests. Figures 10 and 11 show the running surfaces and the cross-sections of the wheel discs at the end of the three continuous tests. The contact surfaces appear different at the end of the tests with different durations proving that the wheel disc damage changes during the test. In the early 10,000 cycles, the surface appearance did not change significantly; after 25,858 cycles, flaking appeared on the surface—this phenomenon progressively intensifies with the increasing of the number of cycles up to 100,000 cycles.

These observations can be correlated with ratcheting, surface crack nucleation and growth, and weight loss. The analysis with an optical microscope of the 10,000-cycle disc cross-section (Figure 11a) shows the presence below the contact surface of a plastic layer deformed in the friction direction; only one very short surface crack was found on the whole disc section. Therefore, the material was not significantly altered, and the disc weight loss was very limited, as shown in Figure 12. Ratcheting and consequent surface crack nucleation and growth were displayed on the cross-section of the 25,858- and 100,000-cycle discs (Figure 11b,c, respectively). These damage mechanisms are typical of wheel and rail steels in dry tests, as documented for example in [37]. The hardness profile confirms the analysis carried out with the optical microscope (Figure 13). The maximum hardness at 10,000 cycles is lower and proves the observation done in the cross-section of the wheel discs: the damage is in a preliminary stage and the surface did not change significantly.

After 25,858 cycles (V02), the hardness increases the symptom that the plastic layer below the surface is more deformed in the friction direction. After 100,000 cycles (V03), the hardness is still increasing proving that the plastic deformation of the material is more severe. In addition to this, the ratcheting continues with consequent surface crack nucleation and growth.

By analyzing the wear curves (Figure 12), the weight loss is similar to the weight loss found in the training tests. The weight loss of the validation tests V01 and V02 follows the same trend of the stepwise tests, because the material is in the initial stage of the damage and the difference between continuous and stepwise tests is not yet appreciable. The difference between continuous and stepwise tests is visible in V03 tests, where the weight loss is higher than the stepwise tests at the same number of cycles, and it is similar to

the trend found in the continuous training tests. This indicates a good reproducibility of the tests.

The machine learning analysis controlled by the k-means results confirms the trends found in the training tests. The metallurgical analysis proves the results found with the k-means algorithm. The observations with optical microscope, the hardness tests, and the wear curves prove that the material is being damaged with the typical damage mechanism of the wheel and rail steel in dry tests: wear, ratcheting, surface crack nucleation and growth.

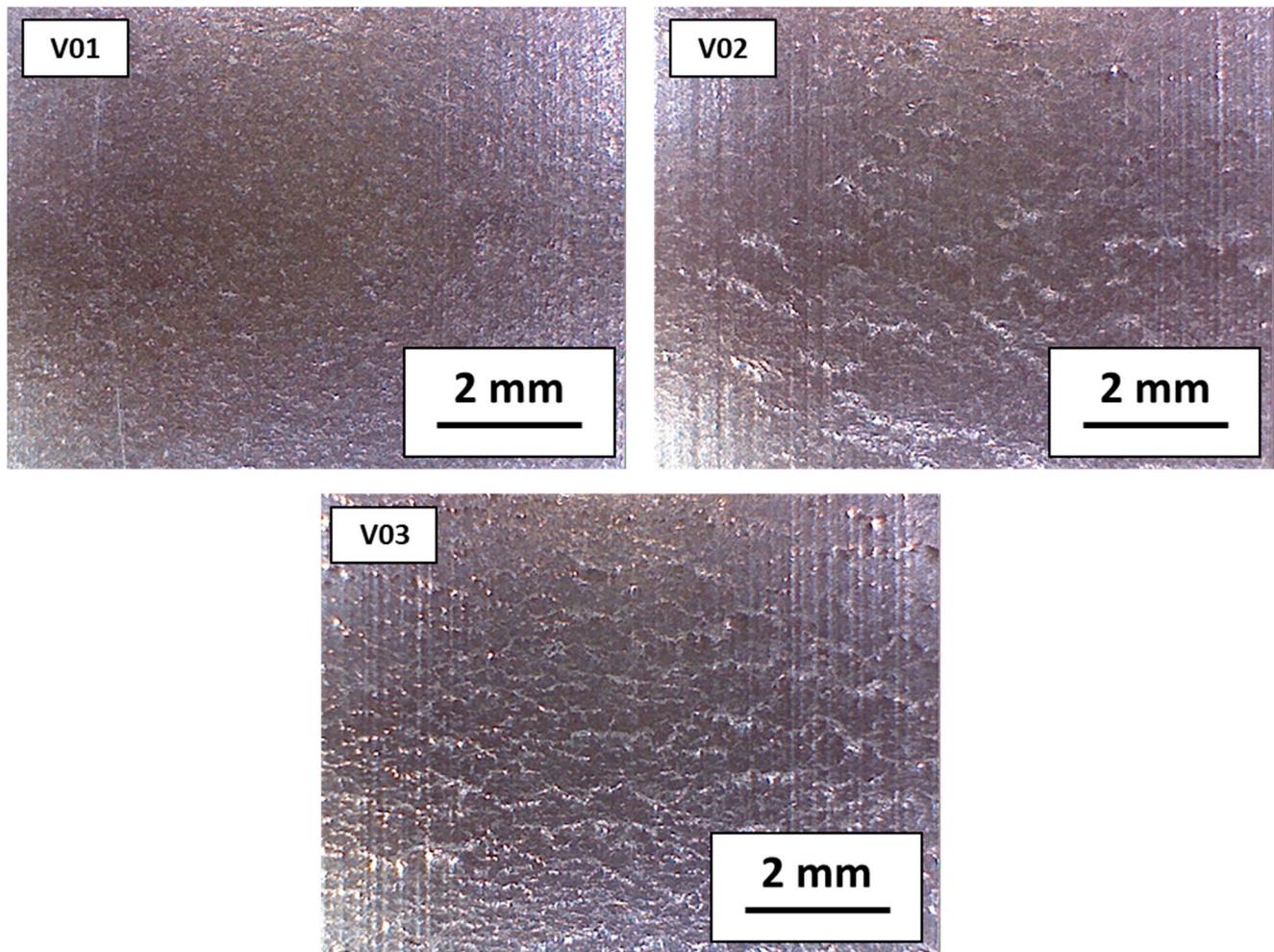


Figure 10. Running surfaces of the wheel discs at the end of the validation tests V01, V02 and V03.

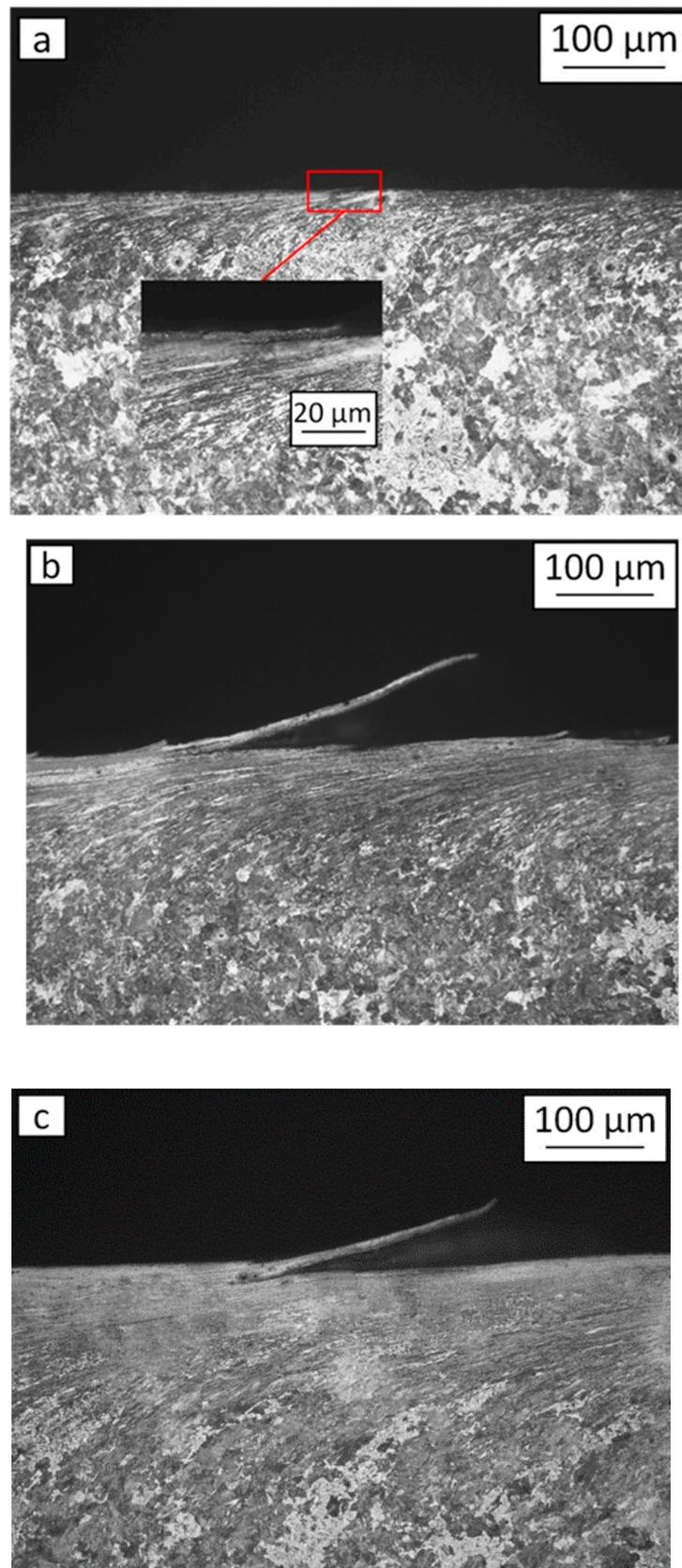


Figure 11. Cross-sections of the wheel discs at the end of the continuous validation tests: (a) V01 (10,000 cycles), (b) V02 (25,858 cycles), and (c) V03 (100,000 cycles).

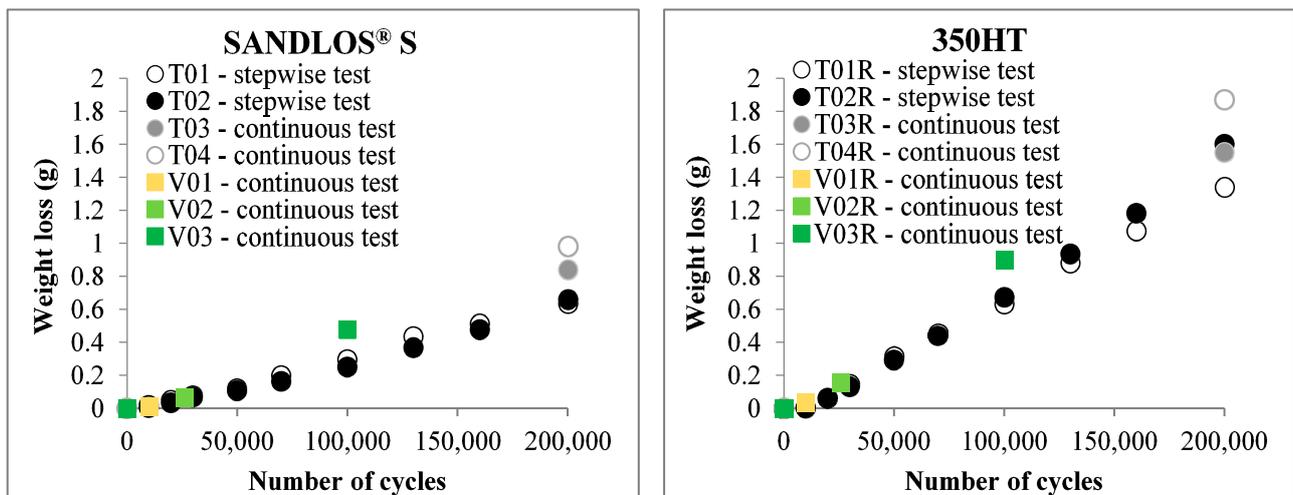


Figure 12. Weight loss of the wheel and rail discs in validation tests and comparison with training tests.

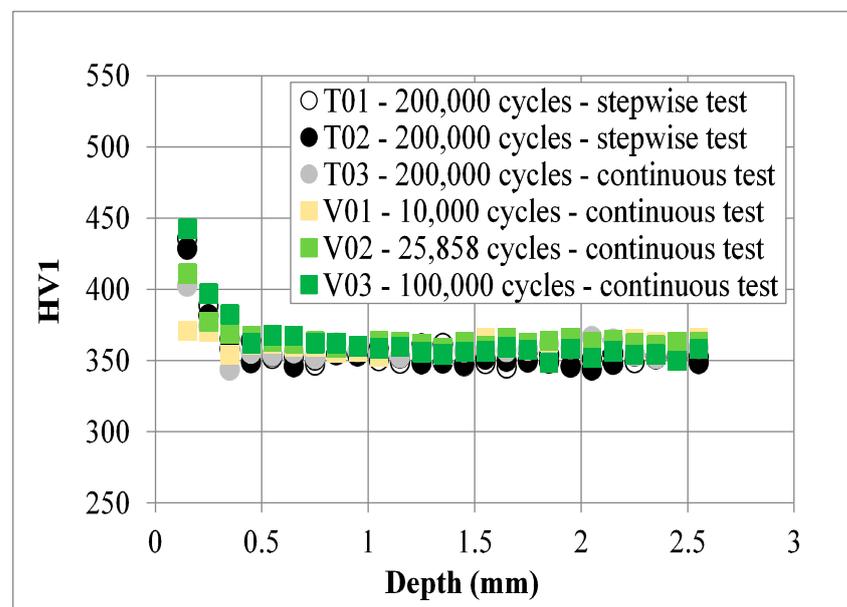


Figure 13. Vickers hardness as a function of the depth on the cross-section of the wheel discs for validation tests and comparison with training tests.

4. Conclusions

The present work shows a new technique to evaluate the evolution of damage in RCF tests with the use of a k-means algorithm used for clustering the vibrations and torque data. The classical and easy method found in the literature to evaluate the damage of a steel wheel is to perform a small-scale test with a twin disk bench and stop the test at an intermediate number of cycles applying destructive analysis to evaluate the damage.

The application of the k-means algorithm to small-scale RCF tests is a novelty for these typologies of tests. Indeed, this technique allows us to avoid the intermediate stop of a test to evaluate the damage of a steel wheel. The damage is quantitatively estimated mathematically with a distance between the current data and a specific reference, called cluster-centers. The cluster-centers are directly correlated with the traditional damage in RCF (wear, ratcheting, etc.). The membership probability calculated from these distances represents an effective way to quantitatively estimate the severity of damage during the tests.

The validation tests controlled by the k-means results confirm the trends obtained in the training tests. The metallurgical analysis shows the ability of the algorithm to detect the different stages of the traditional mechanism of damage in RCF and quantitatively estimate its severity, even if the metallurgical analysis is not a comparative measure to allow validation from a metrological point of view. When the cluster representing the damage state reaches a value of membership probability higher than 0.5, the metallurgical analysis show clearly the cracks on the surface of the specimen due to the RCF damage mechanism.

These preliminary tests are encouraging and show the ability of the algorithm to detect every stage of the damage in wheel and rail steels with several limitations. A limitation of these results is the number of specimens tested. Indeed, a further number of tests are needed to confirm these preliminary results from a metrological point of view to have a better train of the k-means algorithm. The limitation to perform further tests is in the duration of the tests. The duration of a test depends on the steel analyzed and the operating conditions: the time of a test is from one day to a week in dry conditions. Another limitation of these tests is the difficulty in creating other specimens. Indeed, the specimens are directly extracted from a real wheel with a difficult method of extraction. This makes it difficult to have a high number of specimens for material and reduce the possibility to perform the correct number of results to have validation from a metrological point of view.

With a correct number of results to have validation from a metrological point of view, constants and parametric models representing the damage evolution could be found to give further information to designers for a better evaluation of damage in railway wheels.

Author Contributions: Conceptualization, M.L., I.B., L.S. and M.F.; methodology, M.L., I.B. and C.P.; software, M.L. and L.P.; investigation, L.P. and C.P.; data curation, I.B., M.F. and L.P.; writing—review and editing, M.L., M.F., L.S., C.P., and L.P.; supervision, M.L.; project administration, M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the “PRD/SMOCCCT” project of the Department of Industrial and Mechanical Engineering of the University of Brescia.

Data Availability Statement: Not applicable.

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

References

- Mazzù, A.; Petrogalli, C.; Lancini, M.; Ghidini, A.; Faccoli, M. Effect of Wear on Surface Crack Propagation in Rail-Wheel Wet Contact. *J. Mater. Eng. Perform.* **2018**, *27*, 630–639. [\[CrossRef\]](#)
- Donzella, G.; Faccoli, M.; Ghidini, A.; Mazzù, A.; Roberti, R. The competitive role of wear and RCF in a rail steel. *Eng. Fract. Mech.* **2005**, *72*, 287–308. [\[CrossRef\]](#)
- Selech, J.; Andrzejczak, K. An aggregate criterion for selecting a distribution for times to failure of components of rail vehicles. *Maint. Reliab.* **2020**, *22*, 102–111. [\[CrossRef\]](#)
- Vaičiūnas, G.; Bureika, G.; Steišūnas, S. Rail vehicle axle-box bearing damage detection considering the intensity of heating alteration. *Maint Reliab.* **2020**, *22*, 724–729. [\[CrossRef\]](#)
- Konowrocki, R.; Kalinowski, D.; Szolc, T.; Marczewski, A. Identification of safety hazards and operating conditions of the low-floor tram with independently rotating wheels with various drive control algorithms. *Maint. Reliab.* **2021**, *23*, 21–33.
- Ekberg, A.; Kabo, E. Fatigue of railway wheels and rails under rolling contact and thermal loading—an overview. *Wear* **2005**, *258*, 1288–1300. [\[CrossRef\]](#)
- Faccoli, M.; Petrogalli, C.; Lancini, M.; Ghidini, A.; Mazzù, A. Effect of desert sand on wear and rolling contact fatigue behaviour of various railway wheel steels. *Wear* **2018**, *396–397*, 146–161. [\[CrossRef\]](#)
- Bodini, I.; Sansoni, G.; Lancini, M.; Pasinetti, S.; Docchio, F. A novel optical apparatus for the study of rolling contact wear/fatigue based on a high-speed camera and multiple-source laser illumination. *Rev. Sci. Instrum.* **2016**, *87*, 083701. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lancini, M.; Bodini, I.; Petrogalli, C.; Provezza, L.; Faccoli, M.; Sansoni, G.; Solazzi, L.; Mazzù, A. Damage phenomena characterization in RCF tests using image analysis and vibration-based machine learning. In Proceedings of the ISMA, Leuven, Belgium, 17–19 September 2018; pp. 3935–3946.
- Portugal, I.; Alencar, P.; Cowan, D. The use of machine learning algorithms in recommender systems: A systematic review. *Expert Syst. Appl.* **2018**, *97*, 205–227. [\[CrossRef\]](#)
- Jain, A.K.; Murty, M.N.; Flynn, P.J. Data clustering: A review. *ACM Comput. Surv.* **1999**, *31*, 264–323. [\[CrossRef\]](#)
- Jain, A.K.; Duin, P.W.; Mao, J. Statistical pattern recognition: A review. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 4–37. [\[CrossRef\]](#)

13. Kanungo, S.; Bhatnagar, V.V. Beyond generic models for information system quality: The use of interpretive structural modeling (ISM). *Syst. Res. Behav. Sci.* **2002**, *19*, 531–549. [[CrossRef](#)]
14. Boutsidis, C.; Magdon-Ismail, M. Deterministic Feature Selection for K-Means Clustering. *IEEE Trans. Inf. Theory* **2013**, *59*, 6099–6110. [[CrossRef](#)]
15. Chiesa, M.B.; Bodini, I.; Petrogalli, C.; Provezza, L.; Faccoli, M.; Mazzù, A.; Solazzi, L.; Sansoni, G.; Lancini, M. K-means clustering approach for damage evolution monitoring in RCF tests. *J. Phys.* **2018**, *1065*, 102018.
16. Elforjani, M. Diagnosis and prognosis of slow speed bearing behavior under grease starvation condition. *Struct. Health Monit.* **2018**, *17*, 532–548. [[CrossRef](#)]
17. Janssens, O.; Slavkovikj, V.; Vervisch, B.; Stockman, K.; Loccufer, M.; Verstockt, S.; Van de Walle, R.; Van Hoecke, S. Convolutional Neural Network Based Fault Detection for Rotating Machinery. *J. Sound Vib.* **2016**, *377*, 331–345. [[CrossRef](#)]
18. Hayashi, S.; Asakura, T.; Zhang, S. Study of machine fault diagnosis system using neural networks. In Proceedings of the 2002 International Joint Conference on Neural Networks, Honolulu, HI, USA, 12–17 May 2002; pp. 956–961.
19. Kankar, P.K.; Sharma, S.C.; Harsha, S.P. Fault diagnosis of ball bearings using continuous wavelet transform. *Appl. Soft Comput. J.* **2011**, *11*, 2300–2312. [[CrossRef](#)]
20. Prieto, M.D.; Cirrincione, G.; Espinosa, A.G.; Ortega, J.A.; Henao, H. Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks. *IEEE Trans. Ind. Electron.* **2013**, *60*, 3398–3407. [[CrossRef](#)]
21. Kozłowski, E.; Mazurkiewicz, D.; Zabiński, T.; Prucnal, S.; Sęp, J. Machining sensor data management for operation-level predictive model. *Expert Syst. Appl.* **2020**, *159*, 113600. [[CrossRef](#)]
22. Kozłowski, E.; Mazurkiewicz, D.; Zabiński, T.; Prucnal, S.; Sęp, J. Assessment model of cutting tool condition for real-time supervision system. *Maint. Reliab.* **2019**, *21*, 679–685. [[CrossRef](#)]
23. Bustillo, A.; Reis, R.; Machado, A.; Pimenov, D. Improving the accuracy of machine-learning models with data from machine test repetitions. *J. Intell. Manuf.* **2020**, *5–8*, 1–19. [[CrossRef](#)]
24. Elforjani, M.; Shanbr, S. Prognosis of Bearing Acoustic Emission Signals Using Supervised Machine Learning. *IEEE Trans. Ind. Electron.* **2018**, *65*, 5864–5871. [[CrossRef](#)]
25. Shaheryar, A.; Yin, X.-C.; Yousuf, W. Robust Feature Extraction on Vibration Data under Deep-Learning Framework: An Application for Fault Identification in Rotary Machines. *Int. J. Comput. Appl.* **2017**, *167*, 37–45. [[CrossRef](#)]
26. Danner, C.; Gov, K.; Xu, S. Condition Monitoring Using Accelerometer Readings. Available online: http://cs229.stanford.edu/proj2015/102_report.pdf (accessed on 20 January 2021).
27. Amruthnath, N.; Gupta, T. Fault class prediction in unsupervised learning using model-based clustering approach. In Proceedings of the International Conference on Information and Computer Technologies ICICT 2018, DeKalb, IL, USA, 23–25 March 2018; pp. 5–12.
28. Jia, F.; Lei, S.; Xing, S.; Lin, J. A method of automatic feature extraction from massive vibration signals of machines. In Proceedings of the 2016 IEEE International Instrumentation and Measurement Technology Conference, Taipei, Taiwan, 23–26 May 2016; pp. 1–6.
29. Mekaroonkamon, T.; Wongsu, S. A comparative investigation of the robustness of unsupervised clustering techniques for rotating machine fault diagnosis with poorly-separated data. In Proceedings of the 8th International Conference on Advanced Computational Intelligence ICACI 2016, Chiang Mai, Thailand, 14–16 February 2016; pp. 165–172.
30. Bouguerne, A. K-Means Clustering Based on Vibration Signal for Classification Induction Machine Faults. In Proceedings of the International Conference on Automatic control, Telecommunications and Signals (ICATS15), Annaba, Algeria, 16–18 November 2015; pp. 1–5.
31. Hidayat, A.Y.; Widodo, A.; Haryadi, G.D. Fault Diagnostic System Bearing Centrifugal Pump Using K-Means Method for Thermography Image and Signal Analysis Vibrations. *MATEC Web Conf.* **2018**, *159*, 6.
32. Lancini, M.; Bodini, I.; Vetturi, D.; Pasinetti, S.; Mazzù, A.; Solazzi, L.; Petrogalli, C.; Faccoli, M. *Using Vibration Measurements to Detect High Wear Rates in Rolling Contact Fatigue Tests*; Acta IMEKO: Prague, Czech Republic, 2015; pp. 66–74.
33. Bodini, I.; Sansoni, G.; Lancini, M.; Pasinetti, S.; Docchio, F. Feasibility study of a vision system for on-line monitoring of rolling contact fatigue tests. *J. Phys. Conf. Ser.* **2017**, *778*, 012007. [[CrossRef](#)]
34. Ghidini, A.; Faccoli, M.; Mazzù, A.; Petrogalli, C. An innovative steel grade family for forged-rolled solid wheels designed for sandy environments. *Ing. Ferro.* **2018**, *73*, 729–742.
35. Faccoli, M.; Petrogalli, C.; Ghidini, A. On Mechanical Properties of New Railway Wheel Steels for Desert Environments and Sand Caused Wheel Damage Mechanisms. *J. Mater. Eng. Perform.* **2019**, *28*, 2946–2953. [[CrossRef](#)]
36. Lewis, R.; Dwyer-Joyce, R.S. Wear mechanisms and transitions in railway wheel steels. *Proc. Inst. Mech. Eng. Part J J. Eng. Tribol.* **2004**, *218*, 467–478. [[CrossRef](#)]
37. Mazzù, A.; Solazzi, L.; Lancini, M.; Petrogalli, C.; Ghidini, A.; Faccoli, M. An experimental procedure for surface damage assessment in railway wheel and rail steels. *Wear* **2015**, *342*, 22–32. [[CrossRef](#)]