Incremental Novelty detection and Fault Identification Scheme applied to a kinematic chain under non-stationary operation

***Abstract*—Classical methods for monitoring electromechanical systems lack two critical functions for effective industrial application: management of unexpected events and the incorporation of new patterns into the knowledge database. This study presents a novel, high-performance condition-monitoring method based on a four-stage incremental learning approach. First, non-stationary operation is characterised using normalised time-frequency maps. Second, operating novelties are detected using multivariate kernel density estimators. Third, the operating novelties are characterised and labelled to increase the knowledge available for subsequent diagnosis. Fourth, operating faults are diagnosed and classified using neural networks. The proposed method is validated experimentally with an industrial camshaft-based machine under a variety of operating conditions.**

***Keywords*—Condition monitoring; Data-driven modelling; Fault diagnosis; Non-stationary operation; Novelty detection**

1. INTRODUCTION

Condition Based Monitoring (CBM) has acquired strategic importance in the manufacturing sector. It is applied extensively in the field of electromechanical systems, formed by electrical machines coupled to rotatory and/or reciprocating mechanical components, which are of critical importance to multiple industrial processes [1,2]. However, as noted by H. Henao *et al.* [3], new areas of research and novel approaches are still needed for efficient CBM of electrical machines under operating conditions. Hybrid fault diagnosis, which combines multiple fault diagnosis methods, has emerged as a promising approach to leverage the strengths of classical methods such as model-based, signal-based, knowledge-based and active fault diagnosis [4,5].

The most common hybrid fault diagnosis approach integrates a signal-based method for data processing and a data-driven method for classification [6–8]. Although this approach provides satisfactory results, its applicability is limited, since specific processing and classification procedures are used to address predefined faults and operating conditions. Indeed, most condition monitoring studies in the literature consider a set of predefined faults for characterisation, which clearly differ from most industrial scenarios.

Consequently, one of the major challenges to the widespread industrial implementation of CBM is its capacity to manage unexpected events [9–12]. Fault conditions not previously considered, or even deviations from the expected nominal behaviour, are common operating scenarios that lead to diagnostic errors and false positives [13,14]. Indeed, only the nominal operating conditions are available for most industrial applications, so the use of classical diagnostic methods is unfeasible [3,15–17]. As such, the detection of operating novelties, or 'novelty detection', is considered an essential function for the next generation of CBM schemes.

An initial approach to address the lack of operating information, based on the combined use of novelty detection (unsupervised) and diagnosis (supervised) models, was proposed by Grbovic *et al.* [18] and is now being studied and implemented in a range of fields, including data stream analysis [19]. Nevertheless, this approach must be properly adapted to industrial electromechanical systems: accurate signal processing is required to highlight fault conditions, and relatively low ratios of available data must be considered.

There are few studies that describe the application of novelty detection models to electromechanical systems and fewer still that discuss the use of a complete novelty detection and fault identification scheme for condition based monitoring [20–25]. Notable among the published fault diagnosis methods is the proposal of Costa *et al.* [21], which applies a two-stage novelty detection and fault classification method to an industrial plant. Although the method offers satisfactory performance, it is significantly hampered by the amount of historical data required to properly calculate the data density functions and the use of an *ad hoc* signal processing procedure for the plant under test. Filev *et al*. [26] studied the feasibility of an autonomous machinery monitoring and diagnosis system, with special attention to the generalization capabilities of the structure in terms equipment under inspection. The results were promising but the algorithms are limited to the detection of two types of fault patterns. Finally, Wang *et al.* [15] applied a novelty detection scheme to machinery components in order to progressively redefine the initial characterisation boundaries of a set of available fault conditions. The support vector machine-based novelty detection approach shows a significant capacity to adapt boundaries when new information is available, but no consideration is given to incorporating new classes into the available knowledge.

Most CBM approaches dealing with novelty detection are subject to the same constraints: (i) the addition of new scenarios to the initial models is not considered, (ii) the processing stage focuses on the detection of specific faults, and (iii) the models require a large amount of data for characterisation. This study attempts to overcome these constraints through the analysis and validation of a hybrid fault diagnosis methodology that combines a signal processing stage and two data-driven models for novelty detection and fault diagnosis under an incremental learning scheme. The signal processing stage is compatible with non-stationary electric motor-based systems, thanks to the use of Normalised Time-Frequency Maps (NTFMs) of the stator currents. This approach allows numerical features to be estimated for stator current characterisation but also highlights deviations from the initial operating conditions. Next, a multi-modal novelty detection procedure supported by Multivariate Kernel Density Estimation (MVKDE) is proposed to detect novel scenarios and quantify their degree of novelty. The fault diagnosis stage uses classical Artificial Neural Networks (ANN) to identify the corresponding condition in the case of known operating scenarios. To enable incremental learning – that is, to increase the knowledge available to the condition-monitoring scheme – the data-driven novelty detection and diagnosis models are automatically retrained after novel data have been labelled.

The originality of this study derives from the multi-modal signal analysis, the combination of novelty detection and diagnosis models in a hybrid approach, and the potential applicability of the proposed scheme to industrial scenarios thanks to its simple configuration and high effectiveness. The performance of the proposed method is validated with an industrial camshaft-based machine, for which only the nominal operating conditions are initially available. Different fault scenarios are progressively introduced in order to analyse the detection and learning capabilities.

The remainder of this paper is organized as follows. Theoretical aspects of the proposed NTFMs and MVKDEs are described in Section II. The proposed methodology is described in Section III. The experimental tests used to validate the method is discussed in Section IV. Finally, concluding remarks are presented in the Section V.

2. SIGNAL PROCESSING AND MODELLING

2.1. Normalised time-frequency maps

Time-frequency analysis is the most suitable approach for examining non-stationary electric motor operation [3,27,28]. The Short-Time Fourier Transform (STFT) is a local fast-Fourier transform that combined with a sliding window, makes possible the analysis of the frequency content evolution over time. The STFT of a signal y is expressed as $Y(m,f)$, where m is the temporal index and f the spectral index. The magnitude squared of the STFT, $|Y(m,f)|^2$, is called the spectrogram and is expressed in dB as $20\log(|Y(m,f)|)$. Since modelling the system is a complex task it may be difficult to perform the STFT, but in the presence of novel patterns of operation the proposed objective is to detect changes between the STFT corresponding to healthy/nominal operation and the STFT corresponding to the novel operating conditions under analysis. As such, the normalisation of the STFT is proposed as a two-dimensional extension of the statistic-based method. The aim is to obtain a statistical reference from the corresponding healthy/nominal STFT by the computation of the average spectrogram, $M(m,f)$, and the standard deviation, $S(m,f)$, of each time-frequency point. The resulting normalised spectrogram – the NTFM, $Y_{CR}(m,f)$ – is computed according to Equation (1):

$$Y_{CR}(m,f) = \frac{|Y(m,f) - M(m,f)|}{S(m,f)} \tag{1}$$

The normalised spectrogram, $Y_{CR}(m,f)$, represents a standard normal distribution, $No(0,1)$, and the normalisation procedure can be assimilated to a Student's t-test. So, for each STFT computed from a new stator current acquisition, the resulting NTFM will exhibit values close to zero if similar to the reference STFT and values different to zero if novelties are detected. The detected values will be proportional to the difference from the reference. It is then possible to quantify differences considering the complete STFT or even specific regions of interest [29].

2.2. Multivariate kernel density estimation

Multivariate kernel density estimation, also referred to as the Parzen-window or Parzen-Rosenblatt window method, is a flexible approach for estimating the densities of a multi-

dimensional data distribution [9]. Given a *d*-dimensional vector $\mathbf{X} = (X_1, \dots, X_d)^T$, where $X_1, \dots, X_d$ are one-dimensional variables, the vector $\mathbf{X_i}$ represents the *i*-th observation of the *d* variables: $\mathbf{X_i} = (X_{i1}, \dots, X_{id})$, where $i = 1, \dots, n$, and *n* is the total number of observations. The variable X_{ij} is the *i*-th observation of the variable X_j, where $j = 1, \dots, d$. The probability density function (PDF) of $\mathbf{X}$ is, then, given by the joint PDF of the random variables $(X_1, \dots, X_d)^T$ as shown in Equation 2:

$$f(\mathbf{X}) = f(X_1, \dots, X_d) \tag{2}$$

Kernel functions are applied to scale distances. For example, in a one-dimentional case where $u = (x - X_i)/h$, h is the smoothing parameter called bandwidth, and x is the observation under analysis. If we move to the multivariate form, the bandwidth for each distance $(x - X_i)$, can be set individually, resulting in a *d*-dimensional bandwidth $\mathbf{h} = (h_1, \dots, h_d)$. There are different approaches for forming a multi-dimensional kernel, $K(\mathbf{u}) = K(u_1, \dots, u_d)$, such as the multiplicative kernel, $K(\mathbf{u}) = K(u_1) \cdot \dots \cdot K(u_d)$. Using this approach, the density estimator is given as shown in Equation (3):

$$f_{\mathbf{h}}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n}\left\{\prod_{j=1}^{d} h_j^{-1} K\left(\frac{x_j - X_{ij}}{h_j}\right)\right\} \tag{3}$$

Since the PDF exhibits a high dependence on the selection of the bandwidth parameter vector [9], the least squares cross-validation represents a common approach to set the bandwidths. In this approach, each bandwidth h_j is selected with the objective to minimise the integrated mean square error resulting from the difference between the estimated and actual distributions, as shown in Equation (4):

$$IMSE(h_j) = \int \left\{f_{h_j}(x_j) - f(x_j)\right\} dx \tag{4}$$

3. METHODOLOGY

In real-world industrial applications, characteristic patterns of fault conditions are not commonly available, so the proposed condition-monitoring method has two main challenges:

- Significant features must be defined to characterise eventual operating conditions of the system without previous information on possible fault scenarios.

- The novelty detection and diagnosis models must be adapted to allow for the incremental incorporation of the new scenarios identified.

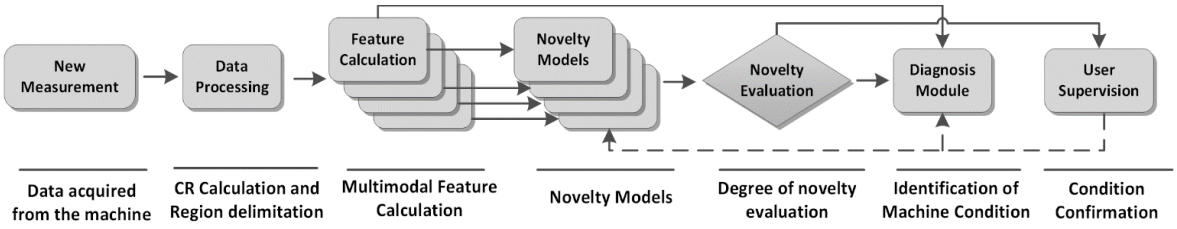

Fig. 1. Proposed incremental learning scheme for condition monitoring. Solid arrows represent the on-line procedure for novelty and diagnosis assessment of a new measurement. Dotted arrows represent the off-line incremental process for the incorporation of new patterns into the novelty detection and diagnosis models.

3.1. Measurements and processing

Malfunctions caused by misadjusted mechanical components in an electric motor-based machine are reflected in the shaft rotation effort pattern [30]. The kinematic chain and, by extension, the framework of the proposed study are representative of machinery subjected to torque variations throughout its working cycle. This is a well-established basis of operation in the industrial sector, where the rotational speed of the electric motor (i.e. the fundamental stator current frequency) is maintained to comply with a predefined manufacturing ratio but the rotational force (i.e. the required torque) follows a non-stationary pattern due, classically, to the driving of different cam profiles and their interrelation during the rotation cycle.

The non-stationary operation considered here are a challenging scenario for condition monitoring, first because the pattern of normality to be characterized must consider the characterisation of the whole operating cycle, and second (and most importantly) because potential new patterns would be revealed as sporadic and isolated variations of the torque in a specific part of the rotational cycle, without interfering with the rest of the stator components. Consequently, this study proposes a signal processing approach capable of monitoring the efforts as the shaft turns while keeping track of the shaft rotation. Indeed, misadjusted mechanical components in general, and mechanical cams in particular, driven by an electric motor may cause coherent mechanical perturbations reflected in the spectral distribution of the stator currents. Thus, a time-frequency method is proposed to monitor the full shaft turn and highlight eventual novelties in the spectral distribution. As described

above, the NTFM results from the STFT applied over the acquired signal but is normalised for the reference stator current signal; that is, the STFT of the signal acquired under healthy/nominal conditions. To take into account the persistence of eventual malfunctions, a time-window length corresponding to a full shaft turn is proposed for NTFM calculation. The resulting NTFM will show increments or decrements in accordance with its differences from the reference signal. In other words, an NTFM for healthy operating conditions will exhibit values close to 0 while an NTFM for novel operating conditions will exhibit values distant from 0 throughout the time-frequency representation.

To illustrate this procedure, Fig. 2a and Fig. 2b show an example of induction motor stator currents in time corresponding to healthy/nominal conditions and to fault conditions caused by the introduction of a mechanical perturbation, respectively. The corresponding STFTs are shown in Fig. 2c and Fig. 2d and the resulting NTFMs in Fig. 2e and Fig. 2f, respectively, where clear qualitative differences can be observed. It should be noted that the differences between the STFTs for the healthy/nominal and fault conditions are not obvious (Fig. 2c and Fig. 2d). Nevertheless, when the NTFMs are computed, the differences are highlighted in different parts of the spectrum.

3.2. Feature calculation and reduction

In order to numerically characterise the resulting NTFM, a segmentation approach that considers both time and frequency axes is proposed. It should be considered that the increase of the number of regions would improve the resolution but at the same time would increase the number of resulting features, which could lead to over-fitted models; by contrast, a small number of regions may not provide sufficient resolution to detect novelties in the operation. The number of regions should be defined according to the specific requirements of each application; however, a good trade-off between resolution and performance can be reached with a total of eight regions. For each region, two different sets of statistical time-frequency features are proposed, one to be used for novelty detection and the other for diagnosis. For novelty detection, since there is no *a priori* information about characteristic fault patterns, the root mean square (RMS) is estimated for each region. For diagnosis, a set of four statistical time-frequency features are proposed for each region, as shown in Table I: max. value, root mean square, crest factor and kurtosis. These features are performance

characteristics considered in multiple electric motor monitoring strategies presented in the literature.

The objectives of the novelty detection and the diagnosis are different. Novelty detection aims to identify deviations from the existing knowledge, classically known as the one-class problem approach, while diagnosis focuses on recognising previously characterised scenarios in a multi-class problem approach. Thus, the novelty detection feature set aims to provide a global characterisation of the existing patterns whereas the diagnosis feature set aims to provide further characteristic dimensions for the recognition of specific patterns.

Max. value	$\hat{x} = \max(x)$	(5)
Root mean square	$RMS = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (x_i)^2}$	(6)
Crest factor	$CF = \frac{\hat{x}}{RMS}$	(7)
Kurtosis	$Kr = \frac{E[(x_i - \bar{x})^4]}{\sigma^4}$	(8)

Table I. Statistical time-frequency features for the characterisation of NTFM regions.

3.3. Novelty detection model

Following the NTFM time axis segmentation, a collaborative structure of novelty detection models is proposed, comprising one model for each of the shaft turn partitions. So, in the case of a full shaft turn with four partitions, each corresponding to 90 degrees, four novelty detection models would be proposed. Each model is trained with the set of features corresponding to the regions enclosed by each 90-degree partition. Therefore, in case of Fig. 2(e) or 2(f), the first novelty detection model would include the RMS values corresponding to regions 1 and 5, the second model would include the RMS values corresponding to regions 2 and 6, and so on. The novelty detection models provide quantification if a new sample diverges from the characterised distributions in which these models were trained. An MVKDE- based probabilistic approach is employed for this purpose, based on the calculation of the PDFs, $f_{\boldsymbol{h}}(\boldsymbol{X})$, where $\boldsymbol{X}$ is the training dataset characterised by the arrays of features.

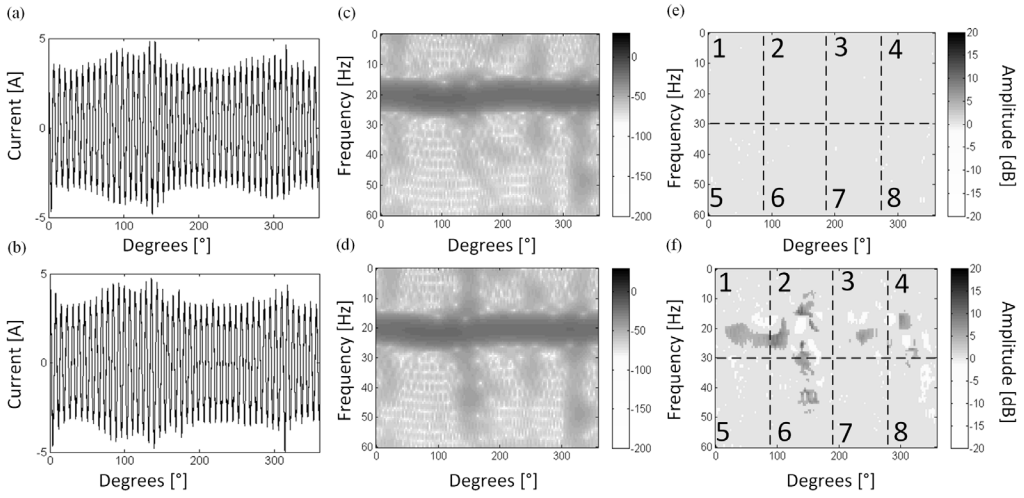

Fig. 2. NTFM calculation. (a) Example of a time-based stator current corresponding to healthy/nominal conditions. (b) Example of a time-based stator current under fault conditions. (c) Computation of the STFT of the stator current corresponding to healthy/nominal conditions. (d) Computation of the STFT of the stator current under fault conditions. (e) NTFM of the stator current corresponding to healthy/nominal conditions, and proposed region segmentation for feature calculation. (f) NTFM of the stator current under fault conditions, and proposed region segmentation for feature calculation.

The consideration of the PDFs to evaluate a new measurement, $\hat{\boldsymbol{x}}$, results in the degree of novelty, $f_{\boldsymbol{h}}(\hat{\boldsymbol{x}})$. Thus, low $f_{\boldsymbol{h}}(\hat{\boldsymbol{x}})$ denotes that the measurement under assessment diverges from the training data.

An evaluation procedure is used to interpret the novelty scores resulting from the set of novelty detection models. The proposed procedure assesses the presence of irregular patterns identified during the regions assessment, and determines a global degree of novelty for the acquisition under analysis. As such, a batch-type analysis is proposed during the on-line evaluation. If only one acquisition – for example, one electric motor shaft turn – is considered, the false alarm events may increase significantly because of outliers. Instead, and considering the field of application, tens of acquisitions are evaluated simultaneously by the novelty detection models. Thus, for each resulting NTFM, the degree of novelty is categorised under the label *Known*, *Uncertain* or *Novel* on the basis of the novelty score $f_{\boldsymbol{h}}(\hat{\boldsymbol{x}})$. Possible values range from zero to a maximum determined by the bandwidth of the models, the number of samples used for training and the selected kernel, thus $f_{\boldsymbol{h}}(\hat{\boldsymbol{x}}) \in \left[0 \ldots max\left(f_{\boldsymbol{h}}(\boldsymbol{X})\right)\right]$.

The proposed categorisation is determined by the novelty score as shown in Fig. 3. The label *Known* corresponds to measurements with an $f_h(\hat{x})$ equal to or greater than two-thirds of the maximum value of $f_h(X)$. The label *Uncertain* corresponds to measurements with a $f_h(\hat{x})$ between one-third and two-thirds of the maximum value of $f_h(X)$. Finally, the label *Novel* corresponds to measurements with an $f_{\mathbf{h}}(\hat{\mathbf{x}})$ equal to or less than one-third of the maximum value of $f_{\mathbf{h}}(\mathbf{X})$.

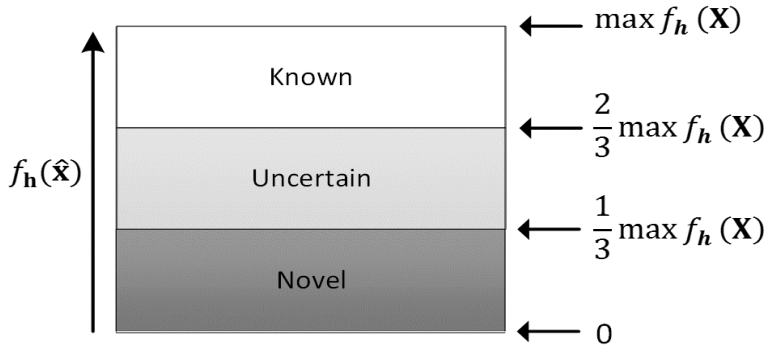

Fig. 3. Novelty degree categorisation according to the resulting novelty score.

The thresholds related to each of the labels can be adjusted to the application; however, considering that there is no information available for novel conditions, proportional thresholds represent a good trade-off between simplicity and performance. In fact, a higher value of the boundary between *Known* and *Uncertain* may generate false alarms, whereas a lower value may decrease the resolution of the novelty detection.

Using the proposed collaborative structure of novelty detection models, the degree of novelty of each measurement is categorised on the basis of the highest score returned by the different novelty detection models. So, for each full shaft turn, if at least one of the novelty detection models gives a categorisation of *Novel*, the measurement is assigned the label *Novel*, even if the other models give a result of *Known* or *Uncertain*. The label with the highest degree of novelty is assigned because each novelty detection model analyses a part of the shaft rotation and localised machinery faults must be detected and addressed.

Finally, once the set of measurements has been labelled, the condition of the machine is determined by majority vote rule. Thus, if 50% or more of the measurements are labelled *Known*, the diagnosis stage follows to identify the corresponding condition. If, however, more than 50% of the measurement are labelled *Uncertain* and/or *Novel* labels, the data are

stored and user supervision follows to assign the correct diagnostic label prior to the incremental learning procedure.

A tie of 50% *Known* and 50% *Uncertain* and/or *Novel* can be interpreted as an anomaly arising in the middle of the batch monitoring, therefore only half of the measurements are detected as *Known*. It is important to note that this is an incremental monitoring approach, therefore batches are analysed incrementally and if the anomaly arises in the middle of a batch it will be fully detected in the next batch.

3.4. Diagnosis

The aim of the diagnosis is to recognize, from previously characterized scenarios, the current condition of the machine under monitoring. The corresponding classification algorithm is required when the novelty detection has returned a majority of *Known* labels. Although different classifiers can be applied to perform the diagnosis, ANN is a data-driven self-adaptive information processing method inspired by biological systems and is the most common data-driven classification technique in the literature [3].

A simple NN-based classifier with a classical three-layer structure was used to obtain the diagnostic estimation of the conditions. The input layer is the set of statistical features obtained from each NTFM region. The hidden layer is proposed to have ten neurons following classical recommendations. Finally, the output layer is composed of the number of neurons corresponding to the available classes.

3.5. User supervision

User supervision is considered a mandatory step in labelled-based learning schemes prior to adaptation of the models to the novel patterns detected. Once the supervisor has inspected the results there are three possible scenarios: first, a new fault condition and corresponding diagnostic label are determined; second, the current scenario is correctly labelled but the patterns must be updated, and third, a false positive has been detected and the data are discarded.

4. EXPERIMENTAL RESULTS

The experimental platform used to analyse the effectiveness and performance of the proposed methodology replicates a typical non-stationary torque pattern widely present in

industrial applications [30]. The platform is composed of a 1.5 kW induction motor as drive linked to a 20:1 rated gearbox, which is coupled to a 120 cm camshaft fitted with two cycloidal cams. The induction machine is formed by six pairs of poles, has a rated speed of 1500 rpm and a rated torque of 20 Nm at 230 V_{AC}, and is controlled by means of a speed-loop-based vector control.

The measurement equipment is intended to acquire the stator current and shaft rotation position. A Tektronix A622 providing 100 mV/A output over AC/DC currents from 50 mA to 100 A-peak within a frequency range of DC to 100 kHz is used to measure a stator-phase current. The current probe is placed at the power converter stator phase output. A XCC1510P Schneider encoder, 360 points of resolution, is coupled to the camshaft to measure the shaft rotation. A DAQ NI 6143 device, 16-channels with 16 bits of resolution, is used to perform data acquisition. A 20 kHz sampling frequency is considered for stator motor current and encoder. The schematic representation of the experimental arrangement is depicted in Fig. 4. All processing procedures are implemented in Matlab on a standard i7-3.5 GHz computer.

In order to validate the proposed methodology, three different experimental scenarios were considered: healthy conditions, *Hc*, and two fault conditions created by inducing effort disturbances. The first fault condition, F_1, consists of a 25% decrease in the effort pattern of the first cam, C_1, through the adjustment of the thumbscrew for load grip using a dynamometric key. The second fault condition, F_2, consists of a 25% decrease in the effort patterns of both cams, C_1 and C_2, also through the adjustment of the thumbscrew for load grip. For each scenario – *Hc*, F_1 and F_2 – a batch of 30 measurements of full shaft turns is available for training and validation, and another batch of 100 measurements is available for testing purposes. Induced fault scenarios, emulating incipient deviation patterns due to wear and tear of mechanical fixation components under intensive and continuous operation have been considered.

To illustrate the proposed procedure, Fig. 5a and Fig. 5b, respectively, show the NTFMs in the presence of a stator current perturbation due to the first and second fault conditions, F_1 and F_2 . Although qualitative similarities are apparent between the two fault conditions, there are also significant changes in the spectral distributions in regions 4, 6 and 8.

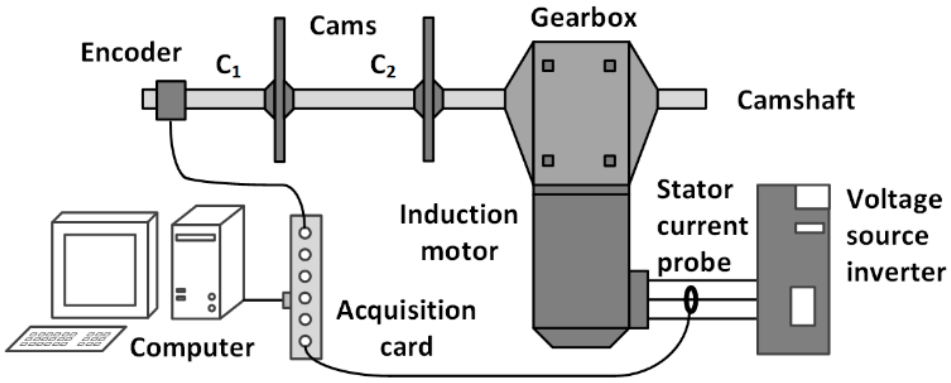

Fig. 4. Schematic representation of the experimental arrangement including an induction machine, a gearbox, a two-cam camshaft, a stator current probe, an encoder and an acquisition card. The C_1 and C_2 correspond to the disturbed cams in regard to the effort pattern to be provided by the electric motor.

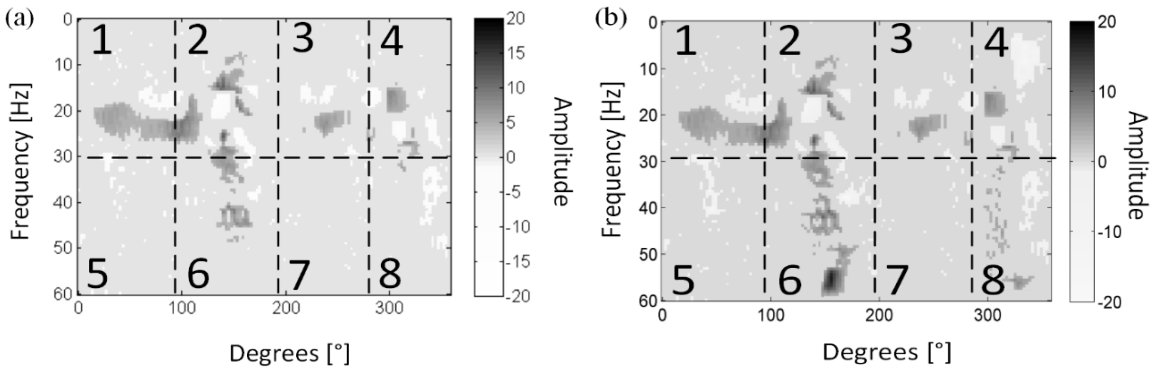

Fig. 5. Resulting NTFM and proposed regions. (a) First fault condition, F_1. (b) Second fault condition, F_2.

In order to highlight the effectiveness and performance of the proposed methodology, the experimental results are presented in four stages. In the first stage, the novelty detection models are initially trained and analysed with data corresponding to healthy conditions, *Hc*. Next, they are tested with data corresponding to a novel condition: fault condition F_1. In the second stage, the novelty detection models are updated to incorporate the patterns of fault condition F_1, and information from the novel fault condition, F_2, is considered to test he resulting models. In the third stage, once the models have been updated again to incorporate information for F_2, the diagnostic performance is analysed with data from the three scenarios, *Hc*, F_1 and F_2. Finally, to analyse the results of the proposed methodology relative to the state of the art, a classical fault diagnosis and novelty detection approach is implemented.

4.1. Novelty detection: healthy scenario

The NTFM corresponding to healthy conditions is computed first. It is then segmented in eight regions: four partitions on the shaft position axis, one each 90º, and two in the stator current bandwidth. In this case, the stator current bandwidth has been limited to the third harmonic of the electric motor, 60 Hz, so each partition encompasses 30 Hz. Next, the RMS statistic is calculated for each region in order to train the four resulting novelty detection models. Four PDFs are calculated for the novelty detection models, each of which corresponds to 90º division of the rotation. Thus, *pdf1* results from the estimated RMS from regions 1, dc to 30 Hz, and 5, 30 Hz to 60 Hz, both from 0º to 90º. Similarly, *pdf2* is obtained with the RMS from regions 2, dc to 30 Hz, and 6, 30 Hz to 60 Hz, both from 90º to 180º, and so on.

The training has been faced by means of MVKDE with multiplicative function and Gaussian function as kernel. The bandwidths of the MVKDE are established by means of least squares cross-validation. The *pdf1* resulting from such procedure is shown in Fig. 6, where the solid line describes the *Known* data boundary, and the dotted line describes the *Unknown* data boundary. So, measurements lying outside the *Unknown* boundary are considered *Novel* data.

As can be seen, all the information corresponding to the healthy operating conditions is concentrated near the zero value of both RMS axes. Similar behaviour is obtained for the rest of the PDFs. The contour plot in Fig. 6 represents the novelty score distribution, $f_{\boldsymbol{h}}(\hat{\boldsymbol{x}})$. The MVKDE bandwidths obtained were 0.653 and 0.651, which represent a good trade-off between avoiding false alarms and increasing detection resolution.

The next step is to assess the measurements for scenario F_1 over the resulting PDFs trained with measurements for healthy conditions. The projection of 30 measurements corresponding to F_1 over pdf_1 *and* pdf_2 is shown in Figure 7. Similar results are obtained for *pdf3* and *pdf4*. The measurements for F_1 data show a very low score for all four novelty maps: < 0.01 in all models, which implies that the fault has a generalised impact throughout the shaft rotation. As result, the 30 full shaft turns are labelled as *Novel.* Following the proposed procedure, user supervision is required to inspect novel behaviour and label the data; the models are then re-trained to include any new scenario detected.

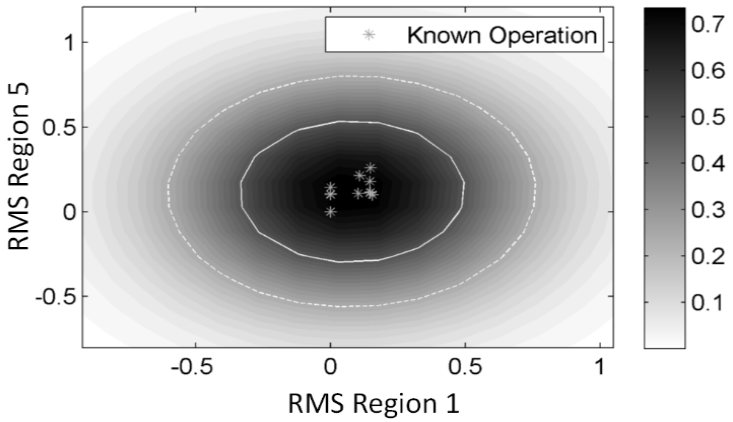

Fig. 6. Resulting pdf_1, considering regions 1, dc to 30 Hz, and 5, 30 Hz to 60 Hz, both from 0º to 90º, where * are measurements for validation. The boundary of *Known* data is represented by the solid line, while the boundary of *Uncertain* data is represented by the dotted line; the value of the resulting PDF is represented by the contour plot.

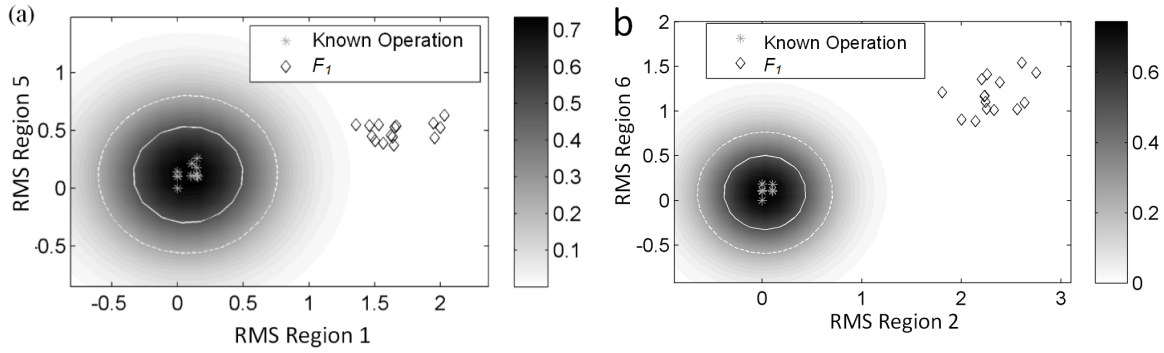

Fig. 7. Evaluation of a fault scenario on the probability densities obtained for the different regions. Boundary of *Known* data is represented by solid lines, while boundary of *Uncertain* data is represented by dotted lines. The contour plot represents the value of the resulting PDFs: (a) pdf_1, (b) pdf_2 .

4.2. Novelty detection: healthy and F_1 scenarios

The adaptability of the novelty maps to the novel fault scenario F_1 is depicted in Fig. 8, where the pdf_1 contour map can be seen. The same response results for pdf_2, pdf_3 and pdf_4. It should be noted that the consideration of more data distributions modifies the novelty score. For instance, pdf_1, formed initially by one single lobe, is restructured to enclose known distributions, thus, leading now to two lobes. Same contour maps of pdf_1 are represented in Fig 9, where the boundary planes corresponding to *Known*, *Uncertain* and *Novel* data are detailed.

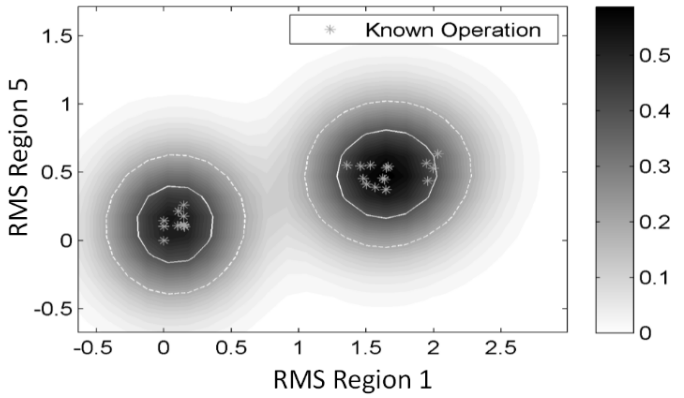

Fig. 8. Contour plot of pdf_1 after including F_1. Boundary of *Known* data is represented by solid lines, while boundary of *Uncertain* data is represented by dotted lines; the value of the resulting PDF is represented by the contour plot.

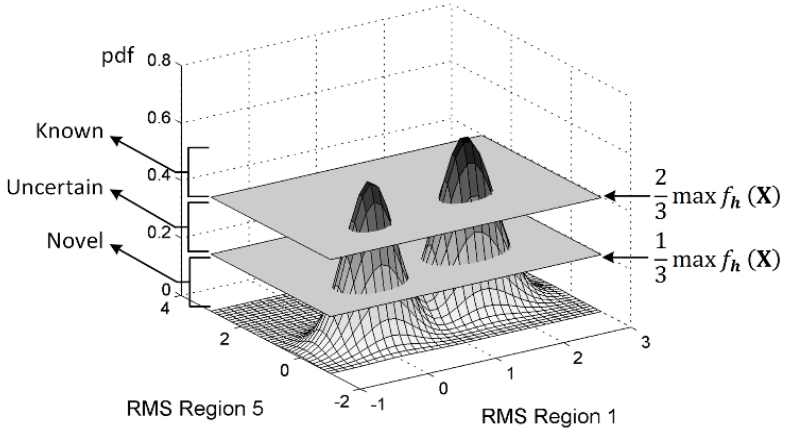

Fig. 9. Probability density function of regions 1 and 5 considering *Hc* and F_1 data as *Known* patterns. Three zones are considered to divide the feature space to categorise the novelty degree according to the novelty score.

Following experimental validation of the methodology, a batch of measurements for a second fault scenario, F_2, are introduced. As illustrated in Fig. 10, this scenario is most noticeably detected in the novelty detection models pdf_1 (Fig. 10.a) and pdf_2 (Fig. 10.b). The resulting novelty scores can be interpreted on the basis of the models that detected the novel behaviour. In this case the F_2 scenario only had an impact on the first half of the rotation cycle, from 0° to 180°. Twenty-nine of the 30 measurements are labelled as *Novel* in the second model; therefore, user supervision is required to inspect the new behaviour and label the data in order to further detect the cause of the fault. The models are then re-trained to incorporate the new scenario.

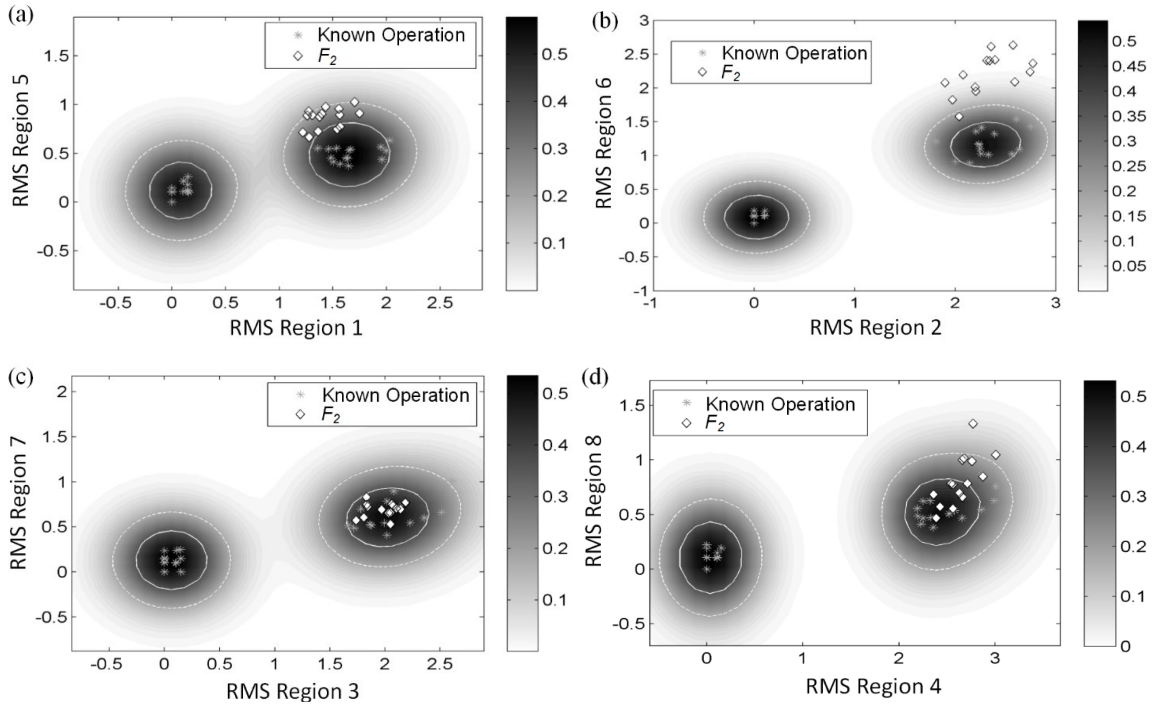

Fig. 10. Novelty detection models after the incorporation of the measurements corresponding to *Hc* and F_1, *, into the *Known* data and resulting projections of F_2 measurements, ◊, as a novel scenario. *Known* data boundary is represented by solid lines, while *Uncertain* data boundaries is represented by dotted lines. The value of the resulting PDFs is represented by the contour plot: (a) pdf_1, (b) pdf_2, (c) pdf_3, (d) pdf_4.

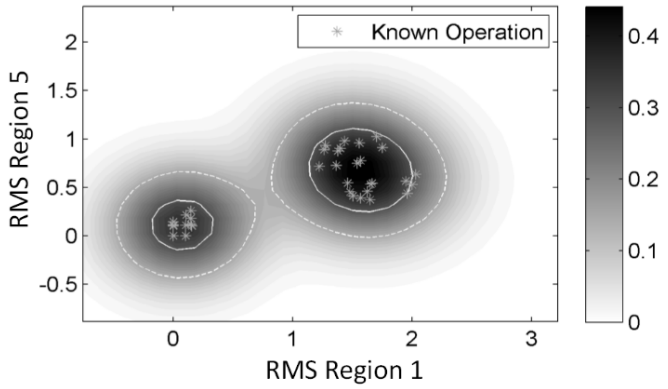

Fig. 11. Contour plot of pdf_1 after including F_2. Boundary of *Known* data is represented by solid lines, while boundary of *Uncertain* data is represented by dotted lines; the value of the resulting PDF is represented by the contour plot.

The adaptability of the novelty maps to the novel fault scenario F_2 is depicted in Fig. 11. The same response results for the rest of the PDFs. In this last case, the novelty detection models do not incorporate a new lobe in the feature space but update the nearest lobe, in this case, the one resulting from the previous fault F_1 data distribution. This implies that scenario

F_2 has similar effects on the RMS value to F_1, which is consistent with the induced malfunction, since F_2 encompasses the same cam, C_1, the malfunction considered in F_1 plus a malfunction in the second cam, C_2.

4.3. Diagnosis: healthy, F_1 and F_2 scenarios

In order to test the diagnosis model, 100 full shaft turns were analysed in scenarios *Hc*, F_1 and F_2. Of these 300 turns, 70% were for training and 30% for testing. This procedure was carried out five times with five randomly selected training-test sets. The ANN hidden layer is based on the sigmoid activation function, and the training procedure follows a classical backpropagation algorithm[25]. The results of the diagnosis stage are presented in Table II, which gives the global classification accuracy and the accuracy per class. A global classification accuracy of 98% for all scenarios is achieved. It must be noted that most of misclassifications are originated due to faults F_1 and F_2, which was expected considering that F_2 encompasses F_1.

Evaluation of diagnosis stage				
	Accuracy by class			**Global**
	Hc	***F₁***	***F₂***	
Training accuracy	1.00 (±0.00)	0.99 (±0.01)	0.98 (±0.02)	0.99 (±0.01)
Test accuracy	1.00 (±0.00)	0.98 (±0.01)	0.96 (±0.01)	0.98 (±0.01)

Table II. Diagnosis stage assessment where accuracy and its standard deviation are shown individually for each scenario and global.

4.4. Comparison of CBM and a classical approach

To highlight the advantages of the proposed multi-modal novelty detection method, a comparison was performed with a classical approach consisting of a single novelty detection model trained with eight-dimensional vectors resulting from the feature calculation over the eight regions considered. In this case, the classical One-Class Support Vector Machine (OC-SVM) with a Gaussian kernel was implemented. The novelty output categorisation was simplified to *Novel* and *Known*.

The comparison was performed for the second stage of the novelty detection previously analysed, in which the *Known* scenarios are *Hc* and F_1 and a *Novel* scenario F_2 is presented. One hundred measurements for each of the three scenarios were used, so a total of 300 measurements were considered. The three scenarios were grouped in three sets: training set, *Known* set and *Novel* set. For the training set, which contained the *Hc* and F_1 data, 140

measurements were used; for the *Known* set, which contained the *Hc* and F_1 data, 60 measurements were used, and for the *Novel* set, which contained the F_2 data, 100 measurements were used. This same procedure was carried out five times considering random selection of the training-test set distributions. The following metrics were considered for performance comparison:

- *Total accuracy*: The ratio between the measurements correctly classified from the novel and known sets, and the total number of measurements in order to estimate the global performance of the novelty detection model.
- *Novel set accuracy*: The ratio between measurements correctly classified from the novel set, and the total number of measurements in such set in order to estimate the true positive ratio (TPR).
- *Known set accuracy*: The ratio between measurements correctly classified from the known set, and the total number of measurements in such set in order to estimate the true negative ratio (TNR).

The results obtained from the novelty detection stage with the proposed method and the classical method are presented in Table III.

Comparison of novelty detection results		
	Classical method	**Proposed method**
Total accuracy	0.84(±0.05)	0.94(±0.05)
***Novel* set accuracy**	0.87(±0.07)	0.93(±0.04)
***Known* set accuracy**	0.79(±0.04)	0.96(±0.04)

Table III. Comparison of the proposed multi-modal novelty detection method and a classical single novelty detection model.

As can be seen, the proposed method delivers better results overall than the classical method: total accuracy increases by approximately 12%, *Novel* set accuracy by approximately 6%, and *Known* set accuracy by approximately 17%. After a five-fold cross-validation to adjust the parameters of the OC-SVM, the classical approach tends to overfit the bimodal distribution of the data, leading to compact boundaries; therefore, the *Known* set accuracy decreases considerably while a high *Novel* set accuracy is obtained. The accuracies for both sets are improved under the proposed method. The novelty evaluation module and the multi-modal feature calculation approach showed greater robustness when configuration

parameters were tuned for each model according the data distribution of each part of the full shaft turn rather than applying a single model to characterise the complete turn.

4.5. Comparison under different conditions

Several tests were performed to further assess the proposed method under different conditions. First, the rotation speed of the induction motor attached to the camshaft was increased. To allow for the greater range of frequencies resulting from the increased rotation speed, the frequency band analysis was expanded from [0 to 60 Hz] to [0 to 90 Hz]. The results are shown in Table IV. Similar results were obtained under both conditions, although the change of operating conditions required a different region selection (in this case a higher frequency band) because the cam fault patterns were reflected differently.

Comparison of working conditions		
	Novelty detection	**Diagnosis**
1500 rpm	0.93(±0.05)	0.98(±0.01)
3000 rpm	0.92(±0.02)	0.96(±0.03)

Table IV. Comparison of the proposed multi-modal method results at different rotation speeds.

Another test was introduced to assess the classification model used for diagnosis. In this case another classifier was tested which had similar results to the proposed neural network. The results shown in Table V, in addition to the cross-validation, indicate that misclassification errors are caused by the similarity of the faults, as explained in the manuscript, rather than by an overfitting of the models or incorrect selection of configuration parameters.

Comparison of diagnosis stage		
	Neural network	**Multi-class SVM**
Training accuracy	0.99 (±0.01)	0.98 (±0.02)
Test accuracy	0.98 (±0.01)	0.96 (±0.02)

Table V. Comparison of multi-fault classifiers over the diagnosis stage.

To tune down the fault detectability of the method, a new test is included to assess the capability to detect the effort disturbances induced by the faults. It should be noted that the faults considered in this study (F_1 and F_2) correspond to a 25% decrease in the effort patterns of the cams (C_1 and C_2) by adjusting the thumbscrew related to the load grip using a dynamometric key. A test with lower effort patterns was included (15% and 10%), from which it can be concluded that the method (with the same configuration parameters and region selection) cannot efficiently detect changes in the effort pattern of less than 15%. This limitation has two primary causes: the low degree of reflection of the effort patterns in the NTFMs, and the fact that the feature calculated from the regions is not significantly different to the known conditions, as a result of which the novelty detection models are not capable of detecting them and the fault classifier cannot distinguish between them. The results are presented in Table VI.

Comparison of effort patterns in F1 and F2		
	Novelty detection	**Diagnosis**
25% decrease	0.93(±0.05)	0.98(±0.01)
15% decrease	0.83(±0.05)	0.88(±0.04)
10% decrease	0.52(±0.12)	0.66(±0.16)

Table VI. Comparison of different effort patterns on the cams to tune down the fault detectability of the method.

5. CONCLUSIONS

This paper proposes a novel framework for incremental learning of a condition-monitoring and diagnosis method applied to a camshaft-based machine by analysing the stator current signatures of the electric motor. The new method has four key features: (i) It considers the NTFMs as a non-stationary signal analysis; the NTFM-based approach allows the estimation of high-performance features in order to detect and quantify deviations from healthy operation of the camshaft machine. (ii) It uses a successful novelty detection structure based on probabilistic models to detect two novel scenarios, F_1 and F_2. (iii) The learning procedure enables novel scenarios to be incorporated into the models to upgrade the knowledge available. (iv) A simple neural network algorithm is used for classification, which provides highly reliable results, since only known signatures are considered for diagnosis. The

experimental validation has been supported by three different conditions which represent an important set of scenarios: healthy (*Hc*), a single fault (F_1), and two faults combined (F_2). In all cases, the diagnosis results are satisfactory. In regard with the novelty analysis, half of the novelty detection models identified the novel fault scenarios as *Known*. This fact revealed as a clear example of the novelty detection difficulties under a practical application framework. When facing *Unknown* fault scenarios, it is important to consider multiple feature approaches, as proposed with the novelty detection model scheme, since the lower the number of features considered, the higher the risk of misidentification. Also, it must be considered that human supervision is critical once a novel scenario has been detected to check and identify the root-cause. It has been prioritized during the design of the proposed methodology the visualization of the underlying physical phenomena of the machine condition, thus, the proposed analysis is constrained to two-dimensional representations. It is in this sense that a supervisor is vital in such industrial machinery monitoring schemes since novel scenarios must be properly labelled. The proposed methodology shows a diagnostic accuracy of 98% and an increase of 12% in total novelty detection accuracy compared to a classical approach, which are high performance ratios. The obtained performance suggest that the proposed methodology could be useful for other rotating mechanical component faults and operating conditions.

ACKNOWLEDGMENTS

This work was supported in part by the by the Spanish Ministry of Economy and Competitiveness under the project TRA2016-80472-R.

REFERENCES

[1] Climente-Alarcon V, Antonino-Daviu JA, Strangas EG, Riera-Guasp M. Rotor-bar breakage mechanism and prognosis in an induction motor. IEEE Trans Ind Electron 2015;62:1814–25. doi:10.1109/TIE.2014.2336604.

[2] Liu Y, Bazzi AM. A review and comparison of fault detection and diagnosis methods for squirrel-cage induction motors: State of the art. ISA Trans 2017;70:400–9. doi:10.1016/J.ISATRA.2017.06.001.

[3] Henao H, Capolino G-A, Fernandez-Cabanas M, Filippetti F, Bruzzese C, Strangas E, et al. Trends in Fault Diagnosis for Electrical Machines: A Review of Diagnostic Techniques. IEEE Ind Electron Mag 2014;8:31–42. doi:10.1109/MIE.2013.2287651.

[4] Gao Z, Cecati C, Ding SX. A survey of fault diagnosis and fault-tolerant techniques-part II: Fault diagnosis with knowledge-based and hybrid/active approaches. IEEE Trans Ind Electron 2015;62:3768–74. doi:10.1109/TIE.2015.2419013.

[5] Xue X, Zhou J. A hybrid fault diagnosis approach based on mixed-domain state features for rotating machinery. ISA Trans 2017;66:284–95. doi:10.1016/J.ISATRA.2016.10.014.

[6] Seshadrinath J, Singh B, Panigrahi BK. Vibration Analysis Based Interturn Fault Diagnosis in Induction Machines. IEEE Trans Ind Informatics 2014;10:340–50. doi:10.1109/TII.2013.2271979.

[7] He D, Li R, Zhu J. Plastic Bearing Fault Diagnosis Based on a Two-Step Data Mining Approach. IEEE Trans Ind Electron 2013;60:3429–40. doi:10.1109/TIE.2012.2192894.

[8] Ebrahimi BM, Roshtkhari MJ, Faiz J, Khatami S V. Advanced Eccentricity Fault Recognition in Permanent Magnet Synchronous Motors Using Stator Current Signature Analysis. IEEE Trans Ind Electron 2014;61:2041–52. doi:10.1109/TIE.2013.2263777.

[9] Pimentel MAF, Clifton DA, Clifton L, Tarassenko L. A review of novelty detection. Signal Processing 2014;99:215–49. doi:10.1016/j.sigpro.2013.12.026.

[10] Langone R, Alzate C, De Ketelaere B, Vlasselaer J, Meert W, Suykens JAK. LS-SVM based spectral clustering and regression for predicting maintenance of industrial machines. Eng Appl Artif Intell 2015;37:268–78. doi:10.1016/j.engappai.2014.09.008.

[11] Dong L, LIU S, ZHANG H. A method of anomaly detection and fault diagnosis with online adaptive learning under small training samples. Pattern Recognit 2017;64:374–85. doi:10.1016/j.patcog.2016.11.026.

[12] Jamali S, Farsa AR, Ghaffarzadeh N. Identification of optimal features for fast and accurate classification of power quality disturbances. Measurement 2018;116:565–74. doi:https://doi.org/10.1016/j.measurement.2017.10.034.

[13] Ding X, Li Y, Belatreche A, Maguire LP. An experimental evaluation of novelty detection methods. Neurocomputing 2014;135:313–27. doi:10.1016/j.neucom.2013.12.002.

[14] Sankavaram C, Kodali A, Pattipati KR, Singh S. Incremental classifiers for data-driven fault diagnosis applied to automotive systems. IEEE Access 2015;3:407–19. doi:10.1109/ACCESS.2015.2422833.

[15] Wang S, Yub J, Lapirac E, Lee J. A modified support vector data description based novelty detection approach for machinery components. Appl Soft Comput J 2013;13:1193–205. doi:10.1016/j.asoc.2012.11.005.

[16] Lemos A, Caminhas W, Gomide F. Adaptive fault detection and diagnosis using an evolving fuzzy classifier. Inf Sci (Ny) 2013;220:64–85. doi:10.1016/j.ins.2011.08.030.

[17] Li C, Sanchez RV, Zurita G, Cerrada M, Cabrera D, Vásquez RE. Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals. Mech Syst Signal Process 2016;76–77:283–93. doi:10.1016/j.ymssp.2016.02.007.

[18] Grbovic M, Li W, Subrahmanya NA, Usadi AK, Vucetic S. Cold Start Approach for Data-Driven Fault Detection. IEEE Trans Ind Informatics 2013;9:2264–73. doi:10.1109/TII.2012.2231870.

[19] Precup R-E, Angelov P, Costa BSJ, Sayed-Mouchaweh M. An overview on fault diagnosis and nature-inspired optimal control of industrial process applications. Comput Ind 2015;74:75–94. doi:10.1016/j.compind.2015.03.001.

[20] Bezerra CG, Costa BSJ, Guedes LA, Angelov PP. An evolving approach to unsupervised and Real-Time fault detection in industrial processes. Expert Syst Appl 2016;63:134–44. doi:10.1016/j.eswa.2016.06.035.

[21] Costa BSJ, Angelov PP, Guedes LA. Fully unsupervised fault detection and identification based on recursive density estimation and self-evolving cloud-based classifier. Neurocomputing 2015;150, Part:289–303. doi:10.1016/j.neucom.2014.05.086.

[22] Lazzaretti AE, Tax DMJ, Vieira Neto H, Ferreira VH. Novelty detection and multi-class classification in power distribution voltage waveforms. Expert Syst Appl 2016;45:322–30. doi:10.1016/j.eswa.2015.09.048.

[23] Hu Y, Baraldi P, Di Maio F, Zio E. A Systematic Semi-Supervised Self-adaptable Fault Diagnostics approach in an evolving environment. Mech Syst Signal Process 2016;88:0–1. doi:10.1016/j.ymssp.2016.11.004.

[24] Salahshoor K, Kordestani M, Khoshro MS. Fault detection and diagnosis of an industrial steam turbine using fusion of SVM (support vector machine) and ANFIS (adaptive neuro-fuzzy inference system) classifiers. Energy 2010;35:5472–82. doi:10.1016/j.energy.2010.06.001.

[25] Adouni A, Chariag D, Diallo D, Ben Hamed M, Sbita L. FDI based on Artificial Neural Network for Low-Voltage-Ride-Through in DFIG-based Wind Turbine. ISA Trans 2016;64:353–64. doi:10.1016/J.ISATRA.2016.05.009.

[26] Filev DP, Chinnam RB, Tseng F, Baruah P. An industrial strength novelty detection framework for autonomous equipment monitoring and diagnostics. IEEE Trans Ind Informatics 2010;6:767–79. doi:10.1109/TII.2010.2060732.

[27] Langone R, Mauricio Agudelo O, De Moor B, Suykens JAK. Incremental kernel spectral clustering for online learning of non-stationary data. Neurocomputing 2014;139:246–60. doi:10.1016/j.neucom.2014.02.036.

[28] Antonino-Daviu J, Aviyente S, Strangas EG, Riera-Guasp M. Scale invariant feature extraction algorithm for the automatic diagnosis of rotor asymmetries in induction motors. IEEE Trans Ind Informatics 2013;9:100–8. doi:10.1109/TII.2012.2198659.

[29] Picot A, Zurita D, Cariño J, Fournier E, Régnier J, Ortega JA. Industrial machinery diagnosis by means of normalized time-frequency maps. Proc. - SDEMPED 2015 IEEE 10th Int. Symp. Diagnostics Electr. Mach. Power Electron. Drives, 2015, p. 158–64. doi:10.1109/DEMPED.2015.7303684.

[30] Lassaad W, Mohamed T, Yassine D, Fakher C, Taher F, Mohamed H. Nonlinear dynamic behaviour of a cam mechanism with oscillating roller follower in presence of profile error. Front Mech Eng 2013;8:127–36. doi:10.1007/s11465-013-0254-x.