Applied Acoustics 132 (2018) 167–181

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Online automatic diagnosis of wind turbine bearings progressive degradations under real experimental conditions based on unsupervised machine learning

Jaouher Ben Ali[a,*], Lotfi Saidi[a], Salma Harrath[a], Eric Bechhoefer[b], Mohamed Benbouzid[c,d]

[a] Université de Tunis, ENSIT, Laboratory of Signal Image and Energy Mastery (SIME), LR 13ES03, 1008, Tunisia
[b] Green Power Monitoring Systems, LLC, VT 05753, USA
[c] University of Brest, FRE CNRS 3744 IRDL, 29238 Brest, France
[d] Shanghai Maritime University, Shanghai, China

ARTICLE INFO

Keywords:
ART2
Feature extraction
Fault diagnosis
High speed shaft bearing
Wind turbines

ABSTRACT

As a critical component, failures of high-speed shaft bearing in wind turbines cause the unplanned stoppage of electrical energy production. Investigations related to naturally progressed defects of high-speed shaft bearings are relatively scarce and the online assessment in damage severities is rarely available in the literature. In this sense, this paper presents a new online vibration-based diagnosis method for wind turbine high-speed bearing monitoring. The adaptive resonance theory 2 (ART2) is proposed for an unsupervised classification of the extracted features. The Randall model is adapted considering the geometry of the tested bearing to train the ART2 in the offline step. In fact, the time domain, the frequency domain, and the time-frequency domain are investigated for a better bearing fault characterization. Indeed, the use of real measured data from a wind turbine drivetrain proves that the proposed data-driven approach is suitable for wind turbine bearings online condition monitoring even under real experimental conditions. This method reveals a better generalization capability compared to previous works even with noisy measurements.

1. Introduction

Onshore and offshore wind powers are one of the fastest growing sources of electricity generation in the world. By 2015, this growth has reached almost 23 percent in Canada, representing more than 3 billion dollars in investments and 10500 new jobs [1].

Despite its importance in clean energy production, wind turbine generators (WTGs) are subjected under extreme environmental conditions leading to several types of defects and unexpected production stops. Taking also into account the improper maintenance, WTGs often fail prior to their 20-year design life [2]. The damages in WTGs are dominated by failures in gears and bearings [2,3]. In fact, a statistical investigation has shown that gears faults present between 35.55% and 71.11% of all WTG failures [4] and bearings failures account for 64% of all gearbox failures [3,4].

The gearbox, which is a mechanical speed variator that adjusts the speed between the rotor and the generator, consists mainly of gears and bearings. Bearing-related failures (dominated by high-speed shaft bearings (HSSBs)) are about 64% of all gearbox failures, those of gears: (dominated by helical gears) present 25% and the remainder 11% (such as housing and cooling fan) are dominated by lubricants and filtration systems [4].

Recently, the development of load prediction, the reliability of wind gearbox has shown considerable improvement. Despite that, it remains considered as the most sensitive element of the wind turbine system, justifying the need to be well monitored. Due to bearing sensitivity and its high failure rate, one of the fundamental problems currently facing a wide range of industries is how to identify a bearing fault before it reaches a critical level and catastrophic failure [5]. This is important in that a high speed bearing can be replaced “up tower” as a scheduled maintenance event at little cost. A catastrophic failure will likely require a gearbox replacement, with the added cost and time required to marshal a crane.

In this sense, several good surveys on the existing fault diagnosis

Abbreviations: AGWN, Additive Gaussian White Noise; ANN, Artificial Neural Network; ART2, Adaptive Resonance Theory 2; D, degraded state; EEMD, Ensemble Empirical Mode Decomposition; EMD, Empirical Mode; EWMA, exponential weighted moving average; F, failure state; H, healthy state; HSSB, high-speed shaft bearing; IMF, intrinsic mode function; MDVD, Multi-Dimensional Variational Decomposition; SK, Spectral Kurtosis; STFT, Short-Time Fourier Transform; SVM, Support Vector Machines; VMD, Variational Mode Decomposition; WTG, wind turbine generator

* Corresponding author.
E-mail address: benallijaouher@yahoo.fr (J. Ben Ali).

https://doi.org/10.1016/j.apacoust.2017.11.021
Received 7 April 2017; Received in revised form 25 September 2017; Accepted 21 November 2017
Available online 25 November 2017
0003-682X/

techniques can be found in the literature. Among these surveys we quote the Refs [6–10] and we recommend the reading of the Refs [9,10]. Three different approaches were presented: The physical model developed by experts and validated on large sets, the rule based expert systems and the data-driven model. Authors in Ref [11] have reported that the most suitable approach for the diagnostic and prognostic of bearings is the data-driven one.

Several meaningful data acquisition techniques can be used for wind turbine bearing condition monitoring such as: Acoustic measurement, electrical effects monitoring, power quality, temperature monitoring, oil debris monitoring and vibration analysis. Vibration analysis is the most common technique used in the industry for any kind of rotating equipment and it is an effective tool for the bearing fault diagnosis [9].

Bearing vibration signals can be measured in all industrial systems and it contains the greatest information [5,11]. However, they are considered as non-stationary and non-linear [10]. Consequently, bearing fault diagnosis based on vibration signals needs to be studied by the most effective, robust and sensitive methods.

In this paper, we propose a new data-driven method for WTG diagnosis. For this, a real run-to-failure vibration history of HSSB was used. Our challenge is to define on the first hand, an accurate online technique to determine the health state of the tested HSSB. On a second hand, our challenge is to present an automatic decision capability without any human intervention or expert analysis. As such, three bearing states are defined: The healthy state, the degraded stated and the state of failure. Two steps are considered: The feature extraction step and the unsupervised feature classification step. An advantage of the proposed technique is that it does not need a learning step. In fact the learning and testing steps are done online at the same time as a result of the Adaptive Resonance Theory 2 (ART2) algorithm. An offline step is added to the ART2 neural network using Randall vibration model. This smart steps ensures that the first ART2 decision correspond to the healthy state. Experimental results validate that the proposed method is able to follow sensitively the regularity of bearing degradations even under variable conditions of speed and torque.

The remainder of this paper is organized as follows: a short literature review of bearing fault diagnosis based on vibration analysis is presented in Section 2. In Section 3, we detail the different steps of the proposed method. In this section, the feature extraction procedure is presented and the basic principles of the ART2 neural network are detailed. Experimental results are provided in the Section 4. Also, in this section a comparison of the performances of the proposed method with previous methodologies in literature is given. Finally, our conclusions and prospects are provided in Section 5.

2. Brief discussion of the reported diagnosis methods

To reduce the operating and maintenance costs of WTGs, reliable and effective bearing fault detection techniques are needed. In this sense, several efforts were presented in the literature using vibration monitoring. Most existing works tried to detect effectively the fault characteristic frequencies [10]. However, the passage from the time-domain to the frequency-domain requires the hypothesis of the stationary of the mechanical vibrations [12]. Unfortunately, HSSB vibration signals are considered non-stationary since WTG gearboxes often exhibit some speed and load variations in practice due to: tower shadow/wind shear, the variation of the speed and the quality of the wind (turbulence) [13,14].

To avoid this problem, Saidi et al. proposed to decompose bearing vibration signals into a number of stationary intrinsic mode functions (IMFs) based on the Empirical Mode Decomposition (EMD) method [12]. The first IMF is then used to compute the bi-spectrum and to detect the fault characteristic frequencies. The idea of this work was impressive despite that the authors haven't justified the use of the first IMF vs. other IMFs. Additionally, the experimental validation of the proposed method was based on Case Western Reserve University (CWRU) bearing data benchmark where the recordings were made at constant speed and load.

To deal with a wide range of speed and load fluctuations, authors in Ref [15] proposed a new method called Multi-Dimensional Variational Decomposition (MDVD). In fact, the Variational Mode Decomposition (VMD) is incorporated into convolutional Blind-Source Separation (BSS) to arrange with the load and speed variations. Experimental results based on some bearings with axial cracks in the outer race showed that the proposed method has great potential for WTG bearing fault diagnosis.

In addition to the problem of non-stationary, noises present a serious trouble in the study of bearing vibration signals [10]. In reality, these signals are always affected by the much stronger signals associated with the gears, shafts and rotor bars. In other words, bearing vibration signals are relatively weak (10e-3 gs) compared to other unintentionally recorded signals (10 s gs) [5]. To overcome this problem and to detect incipient faults at an early stage despite the high background noise, the cyclic bispectrum is used in [16]. By inheriting the advantages of the cyclic bispectrum, an alternative approach based on bispectrum and cyclostationarity analysis wad proposed. Despite all the efforts of authors, this method was not very sensitive to noises.

In [17], the fast kurtogram was combined with the genetic algorithm for resonance demodulation. Initial estimates and final optimization were more accurate than classical methods for the detection of fault characteristic frequencies. However, it is still very hard to distinguish incipient faults when vibration signals are smaller than the noise measurements. Also, this method, like the other works already quoted, requires an expert intervention analyze inthe frequency spectra. Thereby, the sensitivity and the specificity results of these works are highly dependent on the experience of the user.

Alternatively, several works suggest that bearing fault diagnosis is mainly a problem of pattern recognition [18]. The use of artificial intelligence techniques provides an expert system for automatic incipient fault detection without any human intervention [5]. In so doing, Artificial Neural Network (ANN) was used in [19] to classify some extracted features and to analyze ball bearing faults. The feature extraction step was based on the combination of EMD and Hilbert transform methods. The proposed strategy was interesting, that achieved high fault classification accuracy. Similarly, ANN has proved its effectiveness for bearing fault diagnosis in [5]. Time-domain and time-frequency domain features were extracted and initially evaluated using a statistical criterion. Then, the selected features were classified online to define the state of health of the studier bearing.

In [20], the curvilinear component analysis was used to select the most significant time-domain features on a first hand. The nonlinear behavior of the extracted features is visualized and well interpreted and explained physically. On a second hand, a hierarchical neural network was used to perform online the classification stage. This work has presented a powerful method even in different operating conditions. Similarly, the importance of the combination of time-domain features and ANN was validated also in [21].

In [22], a combination of higher order spectral features and Support Vector Machines (SVM) was presented. Authors initially computed the bi-spectrum of the recorded vibration signals and then extracted statistical frequency-domain features. Then, a "one-against-all" SVM strategy was adopted to classify the extracted features. The potential of this method has the processing capability to extract and classify the meaningful information coming from vibration signals with different frequency ranges and even with small energy.

Although that previously cited works have attempted to present a robust strategy for online bearing fault diagnosis, they suffer, with respect, from two common drawbacks:

- The feature extraction process was not justified. Three domains could be investigated: the time-domain, the frequency-domain and the time-frequency-domain. Unfortunately, previous works have

used one domain or a combination of two domains without any convincing explanation for this choice.

- The classification analysis was based on supervised classification algorithms where a training step was needed. In other words, for each bearing, the whole run-to-failure vibration signal was divided into parts: the first one for offline training and the second for online testing. Consequently, the authors assume that there is no critical defect resulting in a complete shutdown of the WTG in the first offline part, which is a strong hypothesis. We motion that bearings installed in WTGs are subjected to harsh environments during operation, including vibrations and sudden shocks under varying wind speed [23].

As strengths, in this work an effort is made to:

- Extract features based on time, frequency and time-frequency domains together. In this way, we ensure the obtaining of the maximum of meaningful information coming from vibration signals.
- Classify online the extracted features based on unsupervised algorithm. For this purpose, The Adaptive Resonance Theory 2 (ART2) neural network where no offline step for training is needed. More details will be given in the following sections.

3. Description of the proposed approach

This work aims to propose a new method for online bearing fault diagnosis in WTGs. As shown in Fig. 1, two online steps are proposed: the feature extraction step and the feature classification one. The next of this section deals with a detail description of these steps.

3.1. Feature extraction

In this step, the recorded part of the vibration signal is used to compute some features in the time, frequency and time-frequency domains. Historically, time-domain and frequency-domain analysis techniques have shown their potential in the analysis of mechanical vibrations [24]. Bearings generally operate under harsh environment conditions and are therefore subject to several types of faults. Numerous publications in academic journals, conference proceedings and technical reports, have noted the importance of the investigation of time-frequency-domain [25]. A detailed description of the feature extraction procedure in these tree domains is given as follow.

(a) Time-domain feature extraction

Time-domain analysis is a simple approach that consists mainly on the computation of traditional statistic features [24]. It is considered as a powerful tool, which characterizes the change of mechanical vibrations when faults occur [5]. With the intention of the characterization of the time information within bearings data, this paper proposes to use the traditional statistical features shown in Table 1 where x is the sampled time signal, i is the sample index and N is the number of samples.

Table.1
Proposed time-domain features.

Feature	Expression
RMS	$\left(\frac{1}{N}\sum_{i=1}^{N} x_i^2\right)^{\frac{1}{2}}$
Kurtosis	$\frac{1}{N}\sum_{i=1}^{N}\frac{(x_i-\bar{x})^4}{\rho^4}$
Skewness	$\frac{1}{N}\sum_{i=1}^{N}\frac{(x_i-\bar{x})^3}{\rho^3}$
Peak to peak (P-P)	$x_{\max}-x_{\min}$
Crest factor	$\frac{x_{\max}}{RMS}$
Shape factor	$\frac{RMS}{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert}$
Impulse factor	$\frac{x_{\max}}{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert}$
Margin factor	$\frac{x_{\max}}{\left(\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert\right)^2}$
Mean	$\bar{x}=\frac{1}{N}\sum_{i=1}^{N} x_i$
Standard deviation (Std)	$\sigma=\left(\frac{1}{N}\sum_{i=1}^{N}(x_i-mean)\right)^{\frac{1}{2}}$
Energy	$\sum_{i=1}^{N} x_i^2$
Energy entropy	$-\sum_{i=1}^{N} x_i\log(x_i)$

(b) Frequency-domain feature extraction

Frequency-domain is mainly based on the study of the spectral analysis. It is the most common method used in the industry for bearing fault diagnosis [24]. As rolling elements in bearings strike local faults on the outer or inner race, a shock is introduced that excites high frequency resonances of the whole structure between the bearing and the response transducer [11]. Hence, the shock impulsivity is most marked in the frequency-domain.

In this work, we propose firstly to use the Spectral Kurtosis (SK), which is defined as the Kurtosis of the signal spectral components. SK is considered as a powerful tool for detecting impulsive bearing signatures since they could be masked by other sources of vibration (gears, shafts, mechanical misalignments…) [26].

The SK of a signal $x(t)$ is defined as the normalized fourth-order spectral moment as given in the following Eq. (1) [26]:

$$SK(f)=\frac{\langle|X^4(t,f)|\rangle}{\langle X^4(t,f)\rangle^2}-2 \tag{1}$$

We note that $<\bullet>$ represents the time-frequency averaging operator, $X^4(t, f)$ and $X^2(t, f)$ are respectively the fourth-order and the second-order cumulants respectively of the band-pass filtered signal $x(t)$ around the frequency f.

For bearing fault diagnosis application, several filtering algorithms based on SK were proposed to reach the optimal frequency band for recovering the impulsive fault signatures masked in the raw vibration signals such using the kurtogram and the fast kurtogram [26,27].

In [23], authors have reported that the SK-derived features avoid the effect of noises by using the optimal Short-Time Fourier Transform

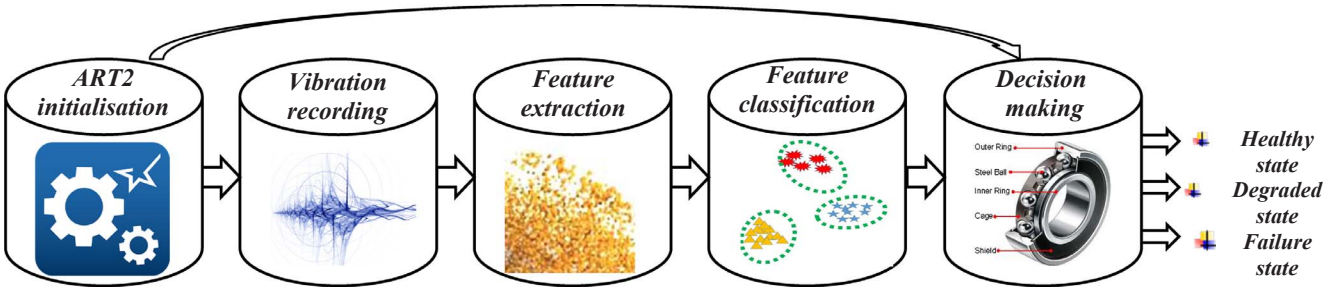

Fig. 1. Proposed framework for online bearing diagnosis.

(STFT) filter. These features are compared with classical time-domain features and judged as more robust. Thereby, we propose in this work to compute various window lengths of SK and selected the suitable frequency band where the kurtosis value was maximized. Then, based on the STFT plot, we propose to compute the same features presented in the Table 1 where x will be the sampled frequency signal. In this way, 12 frequency features are analyzed.

Secondly, we propose to compute the frequency domain based on the bispectrum. As previously noted, the bispectrum has high immuniction to noise. [22]. Theoretically, Gaussian noises become zero and non-Gaussian white noises are flat in the bispectrum [16,28]. The bispectrum can be estimated directly from the discrete Fourier transform N realization of a vibration signal as:

$$\widehat{B}(f_1 f_2) = \frac{1}{N}\sum_{i=1}^{N} X_i(f_1)X_i(f_2)X_i^*(f_1+f_2) \approx E\{X_i(f_1)X_i(f_2)X_i^*(f_1+f_2)\} \tag{2}$$

We note that f_1, f_2 are the frequency indices, X^* is the complex conjugate version of X, $X(f)$ is the discrete Fourier transform of the original time-domain signal x and E is the average over the N realizations. More detail about the computation of the bispectrum and its presentation in tree-dimensional plots are given in [22,16,28].

After establishing the bispectrum computation of the vibration signal, we propose to extract eleven features that have shown their effectiveness for the characterization signal regularities and irregularities [22]. Table 2 resumes the names and the mathematical expressions of these features.

As the bi-spectrum contains redundant information in several regions because its computation is based on the discrete Fourier transform, only the positive frequencies and of course only on the principal domain of $\mathfrak{F}$ is used :

$$\mathfrak{I} = \left\{(f_1 f_2): 0 \leqslant f_2 \leqslant f_1 \leqslant \frac{f_e}{2}, f_2 \leqslant -2f_1 + f_e\right\} \tag{3}$$

We note that L is the number of points within the non-redundant region given by Eq. (3), ϕ refers to the phase angle of the bi-spectrum, and I(•) is an indicator function which gives a value of 1 when the phase angle ϕ is within the range of bin Ψn given in the following Eq. (4):

$$\psi_n = \left\{\phi | -\pi + \frac{2\pi n}{N} \leqslant \phi \leqslant -\pi + \frac{2\pi(n+1)}{N}\right\} \tag{4}$$

Also, we note that where i and j are the two frequency bin indices in the non-redundant region expressed by Eq. (3). More details about the bi-spectrum features are presented in [22,29].

(c) Time-frequency-domain feature extraction

In the literature, popular method for time-frequency-domain feature extraction is the Empirical Mode Decomposition (EMD) [3,5,25]. EMD is one of the most powerful techniques and it has been extremely used in fault diagnosis of rotating machinery [25]. The EMD method decomposes the original nonstationary vibration signal into stationary signals called intrinsic mode function (IMFs) [12]. EMD is a self-adaptive signal processing technique that provides a number of IMFs indicating the natural oscillatory mode embedded in the signal [25]. However, the EMD lacking a theoretical foundation, suffers from: end effects, sifting stop criterion, and extremum interpolation, etc. [30–32]. To overcome these drawbacks, several derivate of the EMD were presented in the literature such as of Ensemble Empirical Mode Decomposition (EEMD), B-spline EMD (BS-EMD) and Variational Mode Decomposition (VMD) [25,33].

In this paper, we propose to use the VMD method because it is highlighted compared to EMD and EEMD in [33]. In fact, recorded vibration signals are decomposed into n stationary IMFs where n is variable between different records. This adaptive decomposition gives also a monotonic residue with a very low magnitude which is why it is not generally considered. Only the first five IMFs are considered for two reasons:

(1) Generally, the decomposition of bearing vibration signals provides more than seven IMFs [5]. In other words, we are sure that for any empirical decomposition five IMFs are found.
(2) The most important IMFs for the investigation of bearing vibration signals are the first five, especially the fourth and the fifth IMFs [34].

Dragomiretskiy and Zosso proposed improvements to the EMD overcoming the drawbacks of the classical EMD method: it adopts a non-recursive strategy to define the VMD. The main VMD problem is an optimization problem minimizing $\min_{u_n,\omega_n}\left\{\sum_n \|\partial t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_n(t)\right]e^{-j\omega_n t}\|_2^2\right\}$ to solve the mode-mixing problem of EMD. We note that $u_1 \ldots u_n$ are the n extracted IMFs, $\omega_1 \ldots \omega_n$ are the estimated center frequency of each IMF in the frequency spectrum, $*$ is the convolution operator δ and is the Dirac operator. Simply, the objective signal can be obtained by summing all IMFs together. The residue is considered negligible so that no little information is lost. The implementation of the VMD method can be done basically in six steps as follow:

Table.2
Proposed frequency-domain features based on bi-spectrum.

Feature	Expression
Sum of logarithmic amplitudes of the bispectrum	$F_1 = \sum_{f_1 f_2 \in \mathfrak{I}} \log(\lvert B(f_1 f_2)\rvert)$
Sum of logarithmic amplitudes of diagonal elements in the bi-spectrum	$F_2 = \sum_{f_k \in \mathfrak{I}} \log(\lvert B(f_k f_k)\rvert)$
kth-order spectral moment of amplitudes of diagonal elements in the bi-spectrum (k = 1: first order, k = 2: second order, k = 3: third order)	$H_k = \sum_{f_k \in \mathfrak{I}} k\log(\lvert B(f_k f_k)\rvert)$
Normalized bi-spectral entropy	$P_1 = -\sum_n \left\{\left(\frac{\lvert B(f_k f_k)\rvert}{\sum_{f_1 f_2} \lvert B(f_k f_k)\rvert}\right)\log\left(\frac{\lvert B(f_k f_k)\rvert}{\sum_{f_1 f_2} \lvert B(f_k f_k)\rvert}\right)\right\}$
Normalized bi-spectral squared entropy	$P_2 = -\sum_n \left\{\left(\frac{\lvert B(f_k f_k)\rvert^2}{\sum_{f_1 f_2} \lvert B(f_k f_k)\rvert^2}\right)\log\left(\frac{\lvert B(f_k f_k)\rvert^2}{\sum_{f_1 f_2} \lvert B(f_k f_k)\rvert^2}\right)\right\}$
Bi-spectrum phase entropy	$P_e = \sum_n \left(\frac{1}{L}\sum_{\mathfrak{I}} \mathrm{I}(\phi(B(f_1 f_2)))\log\left(\frac{1}{L}\sum_{\mathfrak{I}} \mathrm{I}(\phi(B(f_1 f_2)))\right)\right)$
First axe weighted center of bi-spectrum	$WCOB_1 = \frac{\sum_{\mathfrak{I}} iB(i,j)}{\sum_{\mathfrak{I}} B(i,j)}$
Second axe weighted center of bi-spectrum	$WCOB_2 = \frac{\sum_{\mathfrak{I}} jB(i,j)}{\sum_{\mathfrak{I}} B(i,j)}$
Mean magnitude of the bispectrum	$M_{ave} = \frac{1}{L}\sum_{\mathfrak{I}} \lvert b(f_1 f_2)\rvert$

i. Initialization (u and ω);
ii. Compute the frequency spectrum based on Hilbert transform and extract the unilateral frequency;
iii. Shift the unilateral frequency spectrum using the trigonometric identity strategy.
iv. Estimate the bandwidth of all IMFs;
v. Perform the classical Wiener filtering to achieve optimized IMFs in the Fourier frequency-domain.
vi. Pass to the time-domain and obtain the optimized IMFs using Lagrangian multiplier.

All details about the VMD method and the step-by-step algorithm implementation are well detailed in [35].

In this work, for the first five IMFs, the procedure of feature extraction was done as follow:

(1) Compute the energy as: $E_i = \sum_N (x_j)^2$ where $i = 1..0.5$ is the considered IMF and N is the number of samples in the considered IMF.
(2) Compute the total energy as: $E = \sum_{i=1}^{L} E_i$ where i is the considered IMF and L is the number of all IMFs.
(3) Compute the VMD energy entropy for each IMF as: $H_{en} = -\sum_L p_i \log(p_i)$ where $p_i = \frac{E_i}{E}$, $i = 1...5$ is the considered IMF and L is the number of all IMFs.

In this way, ten time-frequency-domain features are considered: the energy entropy of the first five IMFs and the energy rate of the first five IMFs also given by $\frac{E_i}{E}$.

The feature extraction step is considered as a very important to ensure the characterization of bearing fault signs. However, investigations related to the use of three domains are relatively scarce. Among the strengths of this work, the combination of time, frequency, and time-frequency domains that has never been considered in the literature. Consequently, a feature vector with 45 can be obtained online for each recorded vibration signal: 12 time-domain features, 23 frequency-domain features, and 10 time-frequency-domain features.

3.2. Unsupervised feature classification based on adaptive resonance Theory type 2 (ART2)

Several structures for ART2 have been proposed in the literature. According to Carpenter and Grossberg (the founders of the ART), the simplest structure is that given by Fig. 2 [36].

The six layers W, X, U, V, P and Q are formed by n neurons corresponding to the size of the input vector. Y is a competition layer where the winning neuron is the one with the greatest activity. This layer is formed by m neurons corresponding to the number of classes predefined by the operator [37].

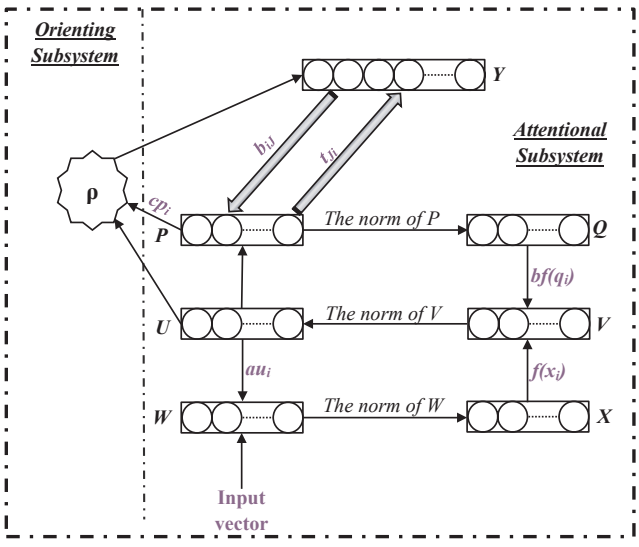

Fig. 2. Architecture of ART2 neural network.

First, the user of the ART2 algorithm sets the class number. Then, the user introduces its input vectors, one by one, and the unsupervised classification algorithm begins. Thereby, all the presented input vectors will be classified into m classes.

Compared to other machine learning algorithms, ART2 has two major advantages [38,39]:

(1) It is stable against noise and adaptive to new inputs;
(2) It learns new information without corrupting the memory of the old information already memorized.

In this work, the ART2 neural network is used as an unsupervised machine learning technique for the classification of the extracted feature vectors. By defining three classes for the heath state of bearings (Healthy state (H), Degraded state (D), Failure state (F)), the online feature extraction and classification procedure can provide the state of health accurately. More details are given in the Appendix A. The full details about the implementation of the ART2 algorithm are well presented in [36–39,50–54].

4. Experimental results

4.1. Experimental setup and data recording

The vibration data were collected thanks to the test conditions shown in Fig. 3. 50 days are considered for raw data recording. In fact, one raw acquisition per day at high sample rate equal to 100 kHz was done in 6 s. Thereby, 585936 samples are recorded each day.

The vibration signals are measured from a real world high speed shaft bearing installed in real WTG provided by the Green Power Monitoring Systems in the USA. The system that was used to acquire this data was designed and built by the team of the Green Power Monitoring Systems, LLC, VT 05753, USA. The sensor itself is a MEMS accelerometer (analog devices ADXL001, with a bandwidth of 32 kHz, and resonance at 22 kHz), with a sensitivity of 70 gs. This was sampled by an ADS1271, 24 bit, Sigma Delta ADC. The available sample rates were {97656, 48828, 24414, 12207, 6104, 3052, 1526} samples per second. This was part of a smart sensor, where the analysis was done at the sensor itself. The sensor had 32 MB of SRAM. The data word length is 32 bits, e.g. single precision. The available RAM was set up for a maximum of 22 s at the highest sample rate. The sensor was mounted radially to the high-speed shaft, on the gearbox bearing support. That is, the sensor was measuring vibration perpendicular to the shaft.

In the last day, the inspection of the tested bearing has validated an inner race fault. This bearing is typed 32222-J2-SKF tapered roller bearing. The tapered roller bearing is 200 mm (mm) in outer diameter, with a bore of 110 mm and a total length of 56 mm. It has 20 rolling elements at a 16° taper angle and weighs approximately 20 lb. The operating speed of the bearing is approximately 30 Hz.

4.2. Feature extraction

In this section, the described bearing previously ending with an inner race fault is explored. As the recording time of vibration signals was 6 s per day for 50 days, the raw run-to-failure history of the tested bearing shown in Fig. 4 is an association of the 50 measured slices, in ascending order of recording days. In other words, each day in Fig. 4 presents the recorded signal on this day during 6 s.

The computation of time-domain features based on Table 1 provides 50 values of each feature where Fig. 5 presents the evolution of some ones from the healthy bearing state to the failure state.

Besides to the twelve time-domain features, the frequency-domain investigated. Hence, the twelve features presented in Table 1 are

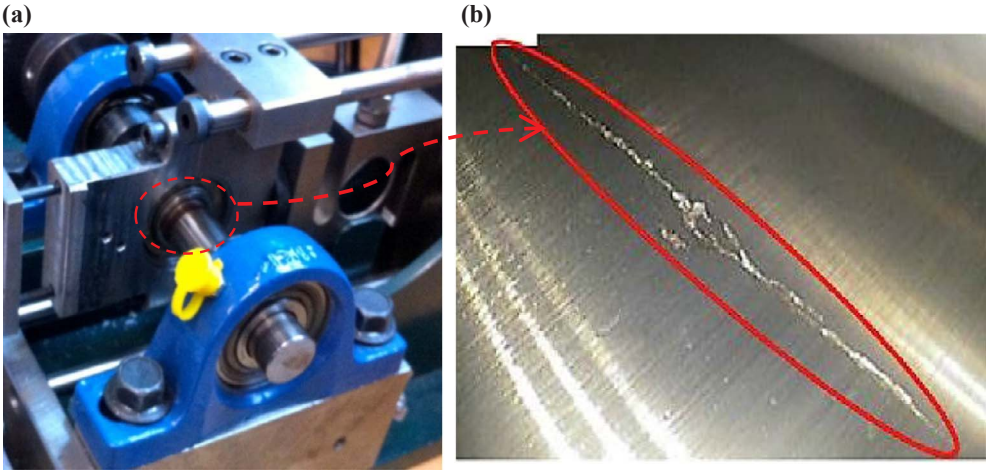

Fig. 3. (a) Bearing test set, (b) The inner race fault.

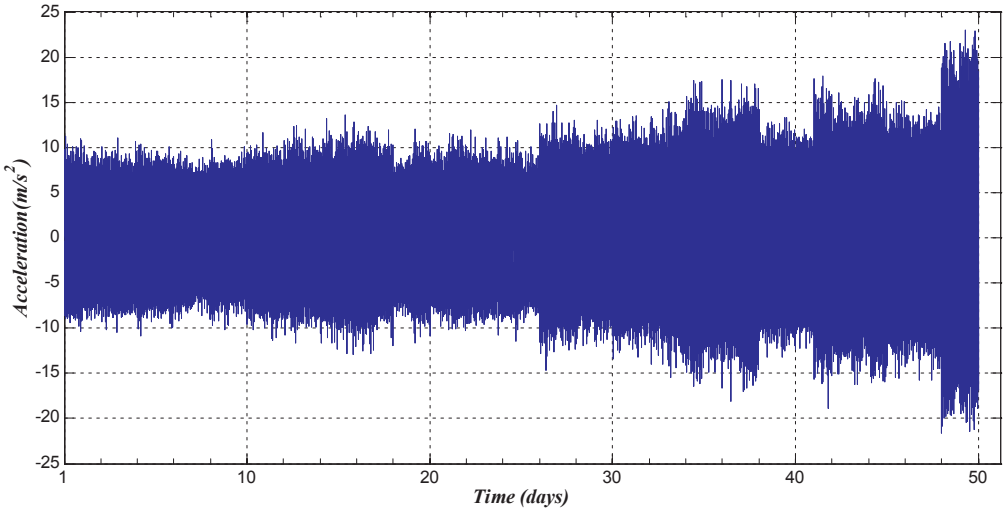

Fig. 4. Raw bearing run-to-failure vibration signals ending with an inner race fault.

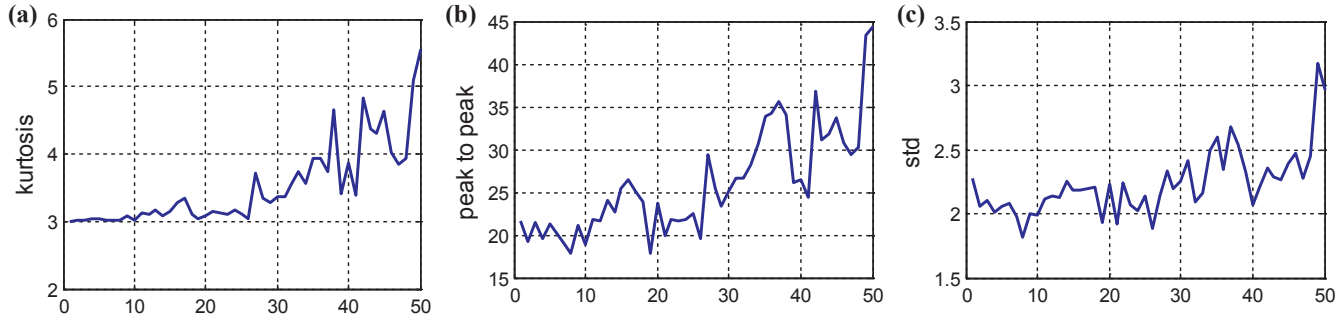

Fig. 5. The evolution of some time-domain features over the 50 days of vibration measurements; (a) Kurtosis, (b) peak-to-peak, (c) Standard deviation.

computed based on the SK where x will be the sampled frequency signal computed based on the 6 s time domain signal for each day. As a result, Fig. 6 shows that SK amplitude increases with the intensity of the fluctuations in the impulse amplitudes of bearing vibrations. In fact, Fig. 6 shows the SK results for three different kinds of recorded vibration signals: The normal (healthy) zone, the degradation zone, and the failure zone. It is clearly remarkable that the highest SK values are given in the failure zone. The decision whether a particular experimental run is likely to contain a faulty element or not, was based on the maximum value of the SK for the particular measurement. The idea behind this approach is the property of the SK, which states that the value of SK, defined by Eq. (1), increases with the intensity of the fluctuations in the impulse amplitudes [10]. Consequently, the SK amplitude can be used as an indication of the severity of the bearing damage. For more convincing, Fig. 7 shows the evolution of some extracted features based on the whole raw run-to-failure bearing history where the amplitude increases proportionally depending on the fault severity.

Moreover to the extracted SK frequency-domain features, the bispectrum features presented in Table 2 are considered. Hence, eleven features are computed, from which Fig. 8 shows the evolution of some ones.

The investigation of time-frequency-domain is based on the VMD method in this paper. In fact, we propose to decompose each extracted measure (recorded during 6 s) into some IMFs. However, due to the large number of samples (585936) in each measure, the computational

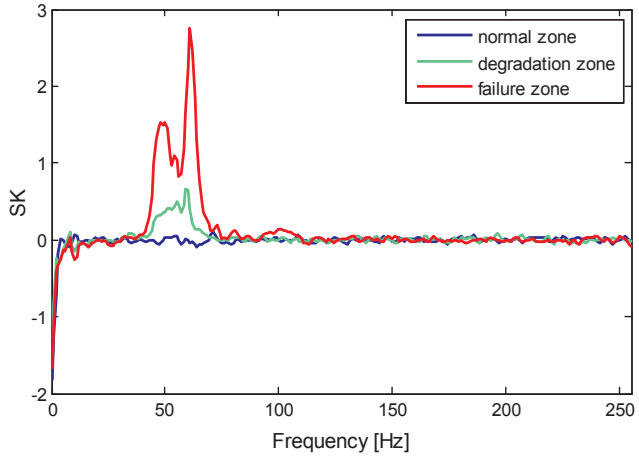

Fig. 6. Monotonically trend impulsiveness SK results over 6 s of acquisition time for three health status of HSSB.

time of the VMD was high (CPU time is about 40 min using core i3 computer). Several previous works have reported the problem of time consuming of the EMD method and its derived algorithms. In [5], the authors showed that the use of just 0.1 s recording time with a sampling frequency of 20 kHz is sufficient to preserve useful information on bearing vibration signal and to obtain meaningful IMFs. Consequently, we propose to decompose just the first part of each measure with duration of one second (97656 samples). Hence, the computational time is well reduced and the necessary IMF information is retained for the bearing health monitoring step. As an example, Fig. 9 shows the VMD results of the first part of the recorded vibration of the first day where 17 IMFs are obtained.

Each day, the VMD was done and the first five IMFs are considered for the time-frequency-domain feature extraction. The computation of the energy entropy and the energy rate of each IMF gives ten features. Fig. 10 shows the evolution of the two features extracted based on the first IMF. In this way, a feature vector is defined each day containing 45 features extracted from three domains. Then the ART2 neural network is adopted to classify these features and to define online the health state of the studied HSSB as explicated in the next section.

4.3. Unsupervised feature classification for online bearing diagnosis

In this work, one run-to-failure experience is considered ending with an inner race fault. For 50 days, a recorded vibration signal, during 6 s, is considered. By considering three domains for feature extraction, a feature vector with 45 components is computed each day. Then, the ART2 neural network is adopted as an unsupervised machine learning for the online classification of feature vectors. In fact, the number of neurons in W, X, U, V, P and Q layers is set at $n = 45$ corresponding to the size of the feature vectors. The number of neurons in the output layer Y is set as $m = 3$ corresponding to the number of desired outputs.

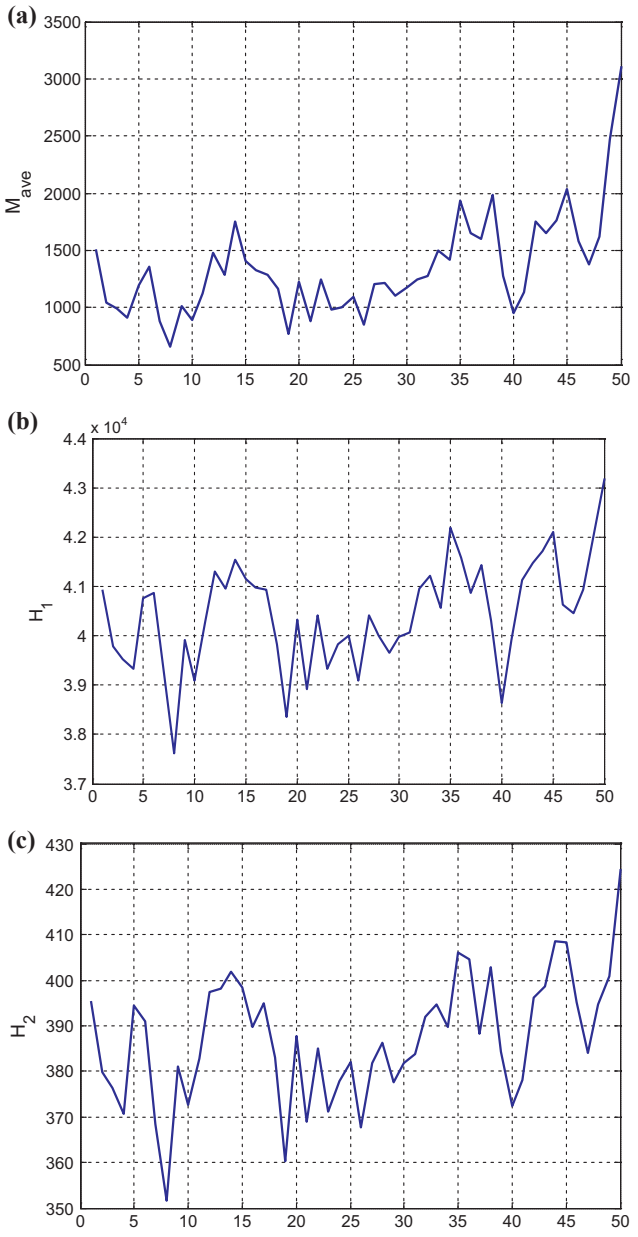

Fig. 8. The evolution of some bi-spectrum frequency-domain features over the 50 days of vibration measurements; (a) Mean magnitude of the bispectrum, (b) first order spectral moment, (c) second order spectral moment.

In this paper, we suggest three states for bearing health monitoring: the healthy state (H), the degraded state (D), and the failure state (F). The online classification results are shown in Fig. 11.

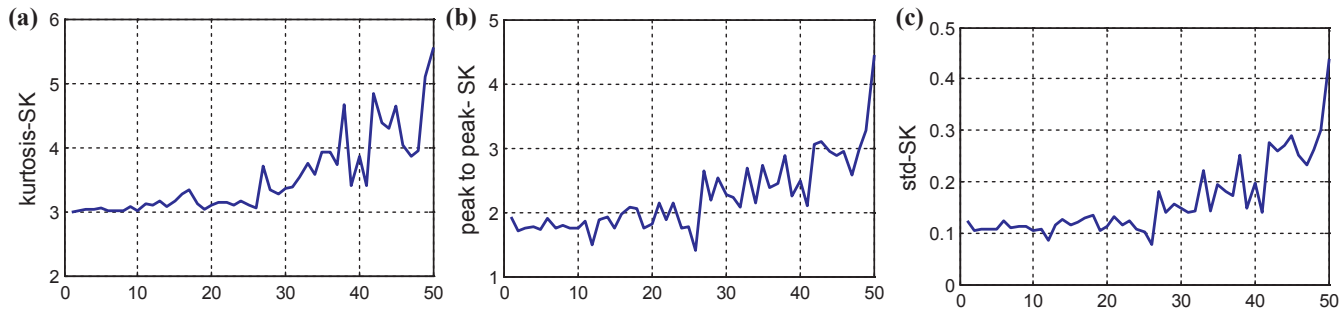

Fig. 7. The evolution of some SK frequency-domain features over the 50 days of vibration measurements; (a) Kurtosis-SK, (b) peak-to-peak-SK, (c) Standard deviation-SK.

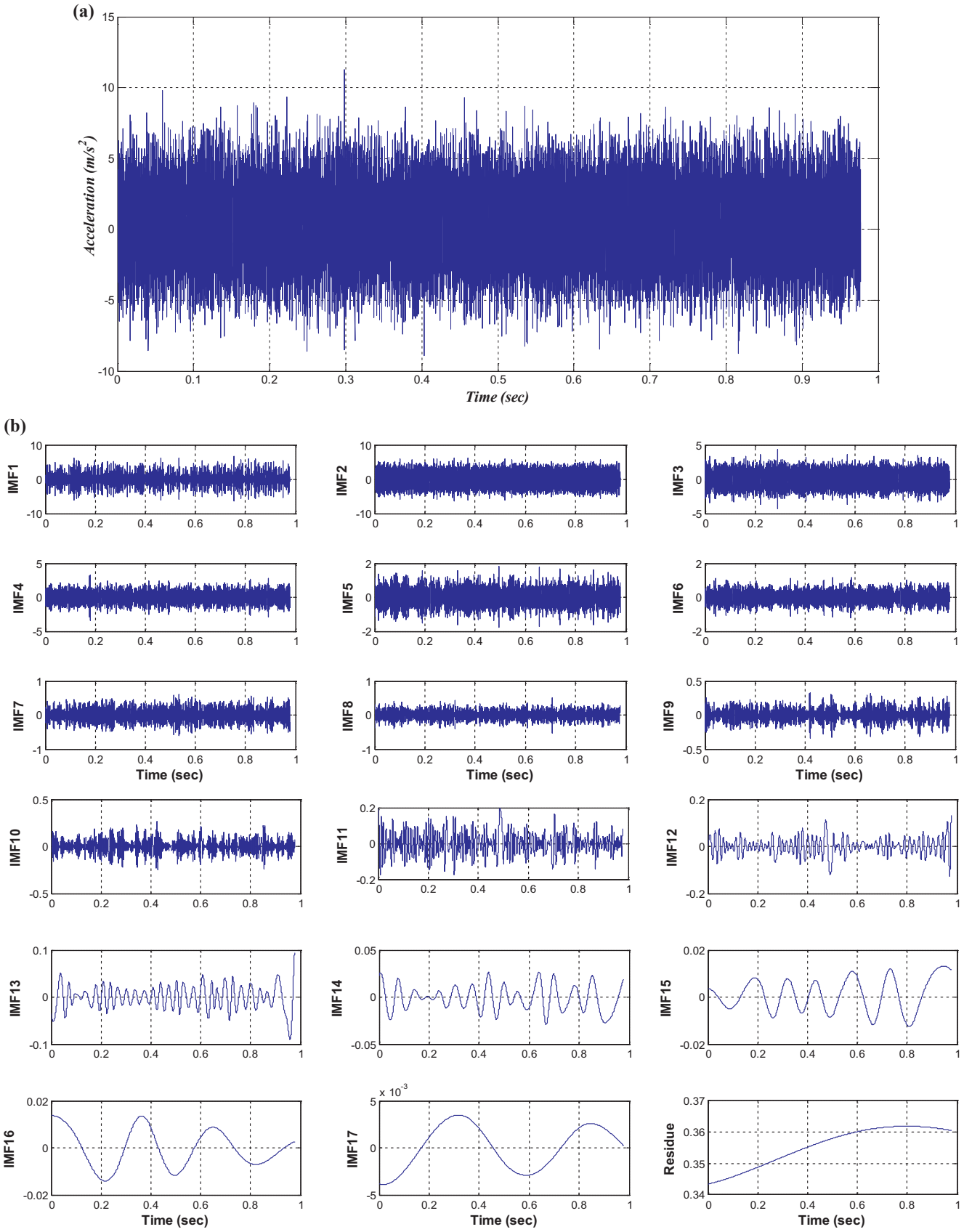

Fig. 9. VMD results of one second of the first day vibration signal; (a) Original raw vibration signal, (b) IMFs resulted from VMD.

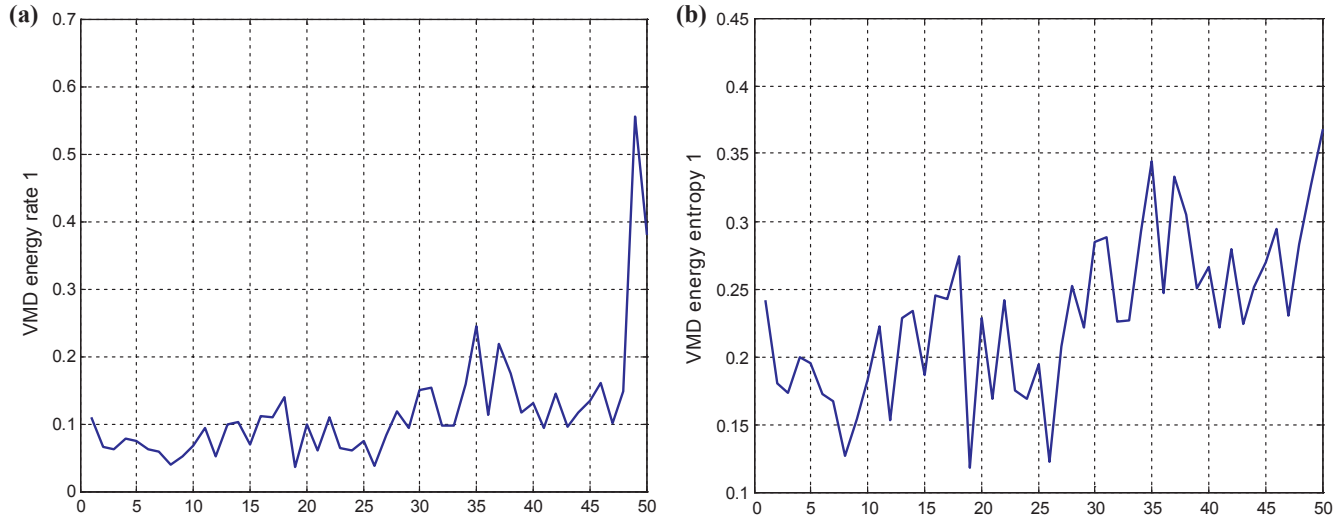

Fig. 10. The evolution of some VMD time-frequency-domain features over the 50 days of vibration measurements; (a) VMD energy rate of the first IMF, (b) VMD energy entropy of the first IMF.

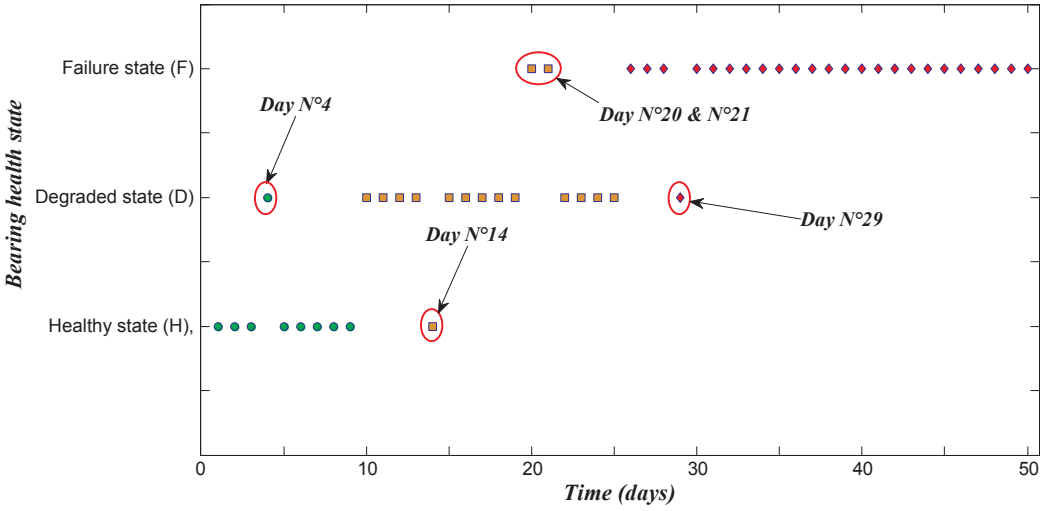

Fig. 11. Application of the proposed method for an online automatic decision making.

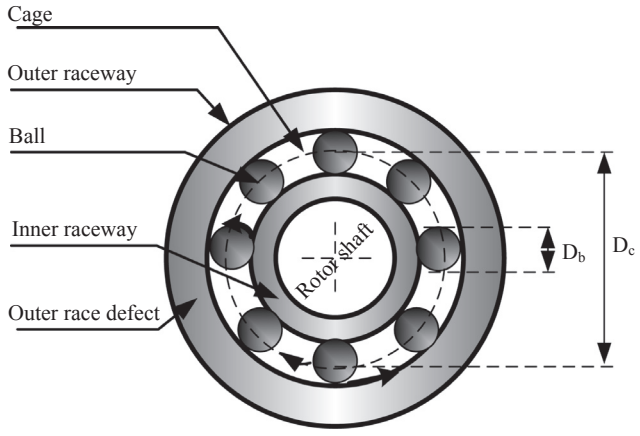

Fig. 12. Example of bearing geometry.

To be sure that the first neuron corresponds to the healthy state, the second one correspond to the degraded state and the third one correspond to the failure state, we propose an offline ART2 step. This step is based on the Randall bearing vibration model as explained in [10]. We have chosen this model because it is a relatively simple phenomenon to simulate. During operation, bearings generate a series of periodic shock-pulses where the repetition rate depends on bearing dimensions and rotational speed of the shaft. Shock-pulses are generated each time rolling elements strike the defected surface of the bearing outer and the inner races. Consequently, rolling elements excite resonances of the structure between the fault location and the vibration sensor [10]. Assuming a constant rotational speed and load of the bearing, the vibration signals generated by can be modeled by as [10,40]:

Table.3
Mathematical equations of bearing fault frequencies.

Fault frequency	Equation
Inner race fault frequency	$f_I = \frac{f_r}{2} N_b \left(1 + \frac{D_b \cos\beta}{D_c}\right)$
Outer race fault frequency	$f_O = \frac{f_r}{2} N_b \left(1 - \frac{D_b \cos\beta}{D_c}\right)$
Cage fault frequency	$f_C = \frac{f_r}{2} \left(1 - \frac{D_b \cos\beta}{D_c}\right)$
Ball fault frequency	$f_B = \frac{f_r}{2} \frac{D_c}{D_b} \left(1 - \left(\frac{D_b \cos\beta}{D_c}\right)^2\right)$

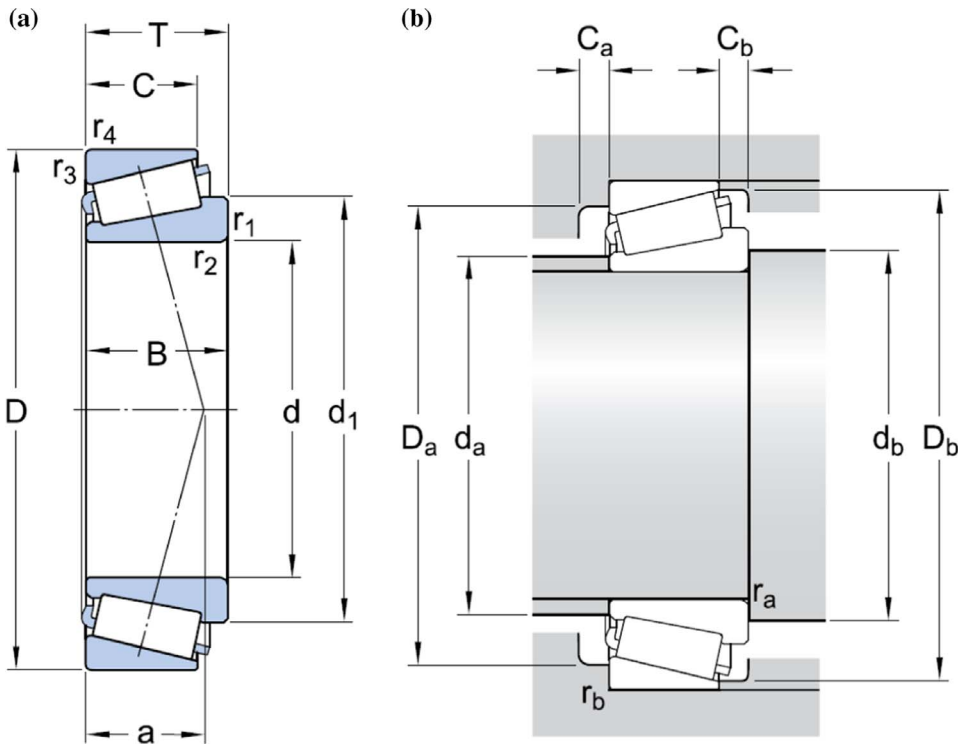

Fig. 13. 32222-J2-SKF tapered roller single row bearings [42]; (a) geometry and dimension, (b) mounting dimensions.

Table.4
Characteristics of the used HSSB end bearing typed 32222-J2-SKF [42].

Index	Size	Index	Size
d	110 (mm)	d_a Maximum	127 (mm)
D	200 (mm)	d_b minimum	124 (mm)
T	56 (mm)	D_a minimum	170 (mm)
d1	151.53 (mm)	D_a Maximum	188 (mm)
B	53 (mm)	D_b minimum	190 (mm)
C	46 (mm)	C_a minimum	6 (mm)
$r_{1,2}$ minimum	3 (mm)	C_b minimum	10 (mm)
$r_{3,4}$ minimum	2.5 (mm)	r_a Maximum	3 (mm)
a	46 (mm)	r_b Maximum	2.5 (mm)
Number of rolling elements	20	taper angle	16 (°)

$$x(t) = \sum_i \omega\left(t - i\frac{1}{F_{char}} - \tau_i\right) + n(t) \tag{5}$$

where $\omega(t)$ is the waveform generated by a single impact (related to resonance frequencies of the system), τ_i is an independent and identically distributed random variable and $n(t)$ is an additive background random noise. Now, it has to be stated that τ_i introduces influence of the rolling elements slips into the model. F_{char} is the frequency characteristic.

Generally, any bearing fault model is based on bearing fault signatures [10,12]. For a better explanation, Fig. 12 shows that bearings consist mainly of outer and inner raceways, the balls and the cage. Bearing single-point defects may enhance vibration and noise level and can be localized and classified into:

- outerracewaydefect;
- inner racewaydefect;
- ball defect.

The single-point defect may be seen by fault frequencies appearing in the machine vibration spectrum record. The frequencies for these components can be accurately. Therefore, there are different fault frequency characteristics associated with each component among the four parts of the bearing [10,12,41]. These frequencies are :

- f_I: inner race fault frequency;
- f_O: outer race fault frequency;
- f_C: cage fault frequency;
- f_B: ball fault frequency.

The mathematical equations of bearing fault frequencies are given in the Table 3 where:

- f_r : rotor shaft frequency;
- N_b : number of rolling elements;
- D_b : ball diameter;
- D_c : pitch diameter;
- β : ball contact angle.

Corresponding to the size and geometry of the used bearing typed 32222-J2-SKF, summarized Fig. 13 and Table 4, the HSSB fault frequencies considering a speed of 2400 r/min (the reference speed according to the manufacturer) are:

- f_I = 755.37 Hz;
- f_O = 44.62 Hz;
- f_C = 2.23 Hz;
- f_B = 4.55 Hz.

In this work, we have considered the frequency characteristic of the cage because it is the small one compared to another frequency. In other words, the fault signatures are repeated few times compared to other fault types. We motion that the proposed method don't consider the fault type: just the severity of the fault has been under study. The proposed method is based on online feature extraction and classification where derived indices from vibration signals are investigated. Thereby, a simulated Randall fault model following Eq. (5) is used for ART2 neural network initialization to ensure that the first neuron correspond to the healthy state, the second one correspond to the degraded state and the third one correspond to the failure state. The used

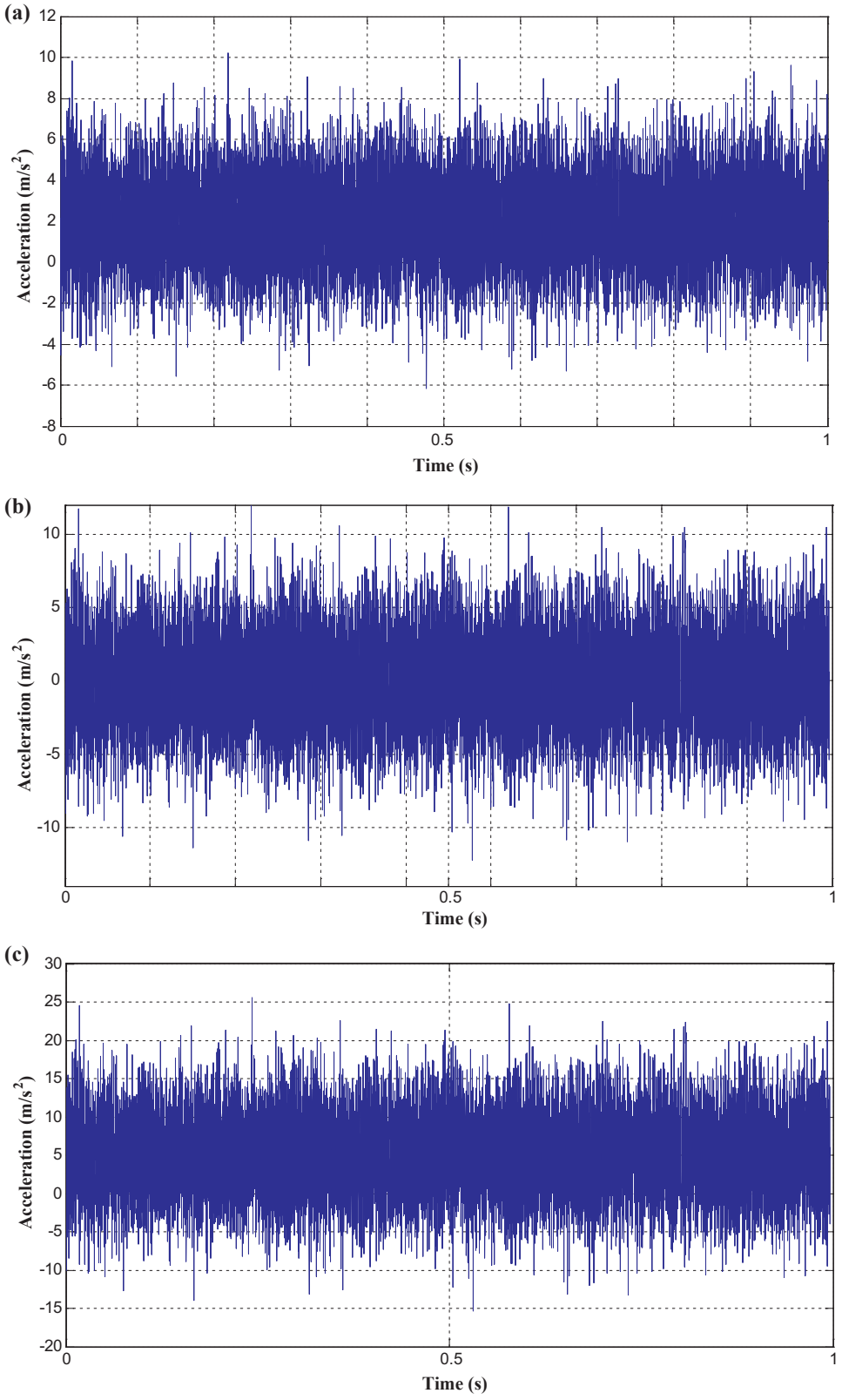

Fig. 14. Simulated Randall model; (a) healthy bearing, (b) degraded bearing, (c) failer bearing.

Randall modal is shown in Fig. 14.

The ART2 learning process is based on an adaptive learning where new information is added the old memory without corrupting it. This advantage, compared to other machine learning techniques, is achieved thanks to the update process for the winning neuron J. The structure of the ART2 is always varied and updated to reach a good stability and generalization to the neural network without changing the number of neurons in the output layer [38,39]. In fact, the update process is based on Eqs. (6) and (7) where α is the learning rate; J is the index of the winning neuron and c and d are two constants ($cd/(1\text{-}d) \leq 1$ and the standard values are $d = 0.9$ and $c = 0.1$).

$$t_{Ji} = \alpha d u_i + \{1 + \alpha d(d-1)\} t_{Ji} \quad (6)$$

Table.5
Comparison with some previous works.

Literature	Features	Classifier	Learning type	Accuracy [%]	False alarm [%]
[45]	Frequency domain.	v-SVM	Supervised.	100	*NA*
[5]	Time-frequency domain	ANN	Supervised.	93	7
[44]	Time-frequency domain	PNN and SFAM neural networks	Supervised	97.82	2.18
[48]	Time-frequency domain	Improved Particle swarm optimization (PSO) and least squares SVM	Supervised	97	*NA*
[49]	Time and frequency domains	(PSO) and SVM	Supervised	97.5	*NA*
This work	Time, frequency and time-frequency domains	ART2 neural network	Unsupervised	90	10

**NA*: It exists however its value is not provided by authors (not available).

$$b_{iJ} = \alpha d u_i + \{1 + \alpha d(d-1)\} b_{iJ} \quad (7)$$

The experimental results of the proposed method for online diagnosis of HSSB are illustrated in Fig. 11. In fact, the proposed method is able to indicate qualitatively a clear distinction between the healthy and the failure bearing states. In addition, incipient fault has occurred in the day number 10 and the severity of this failure has increased remarkably in the day number 25. Besides, the separation between the healthy state, the degraded state and the failure state was done automatically and without any human intervention. However, some false alarms are unfortunately provided in day's number 4, 14, 20, 21 and 29. This may be due to a variation in load, such as the passage of weather front or some other environmental factor. Logically, when the failure state is established and the severity of the fault has increased, it is impossible mechanically to reach the healthy and the degraded states except after repair. It is true that the rate of false alarm presents only 10%, however, this is a drawback of the proposed method that should be addressed in next works.

4. Discussion and comparison with some previous works

Studies related to naturally induced and naturally progressed defects in the rolling element bearings are relatively scarce. In fact, online detection of the rolling element bearing degradation for naturally progressing defect into damage stages is not well reported in the literature [43,44].

In [43], some indices derived from time-domain, frequency-domain, and time–frequency-domain were computed. Then, the Gram–Schmidt orthogonalization process is used for a meaningful calculation of the Mahalanobis distance and thereby the bearing state can be estimated. Also, a discussion about the variation of the derived indices versus time was included in a way to define empirically the best ones.

In [44], the authors have presented a relatively complex method by combining two types of neural networks with the linear discriminant analysis method. Unfortunately, this method showed an important rate of false alarms. Consequently, a smart step that calculates the highest occurrence of the fault severity was introduced. This step minimizes the rate of false alarms and unfortunately it causes a decision delay depending on the size of the sliding window chosen for the fitting.

In [45], the one class support vector machine was used to detect characteristic changes of the monitored bearing vibration signals in order to detect REB defects. To validate the proposed non-destructive diagnostic method, real bearing run-to-failure-histories were exploited.

In [46], the combination of locality preserving projections and exponential weighted moving average (EWMA) was presented. The non-monotonic EWMA performance quantification index shows undesirable results, in terms of false alarm creation. Motivated by the Gaussian mixture model and the negative log likelihood probability advantages, experimental results were improved, but they still have false alarms [47].

Several other works have been reported in the literature for bearing fault diagnosis. For a more suitable comparison, Table 5 summarizes some previous works on bearing faults identification. As advantage, the proposed method was based on unsupervised learning mode where no training mode is needed. However, previous works were based on supervised algorithm where run-to-failure vibration histories were divided in two parts: one for training and one for testing. Unfortunately, nothing ensures that a sudden degradation or a bearing shock cannot be done in the training phase. Consequently, the proposed method is considered very consistent and more reliable compared to previous works. Besides, in this work we have considered three domains for feature extraction, while previous works have used one domain or combined two domains. Hence, several benefits are ensured by the proposed methodology especially for a real-time implementation.

With respect to the reported works previously in this section they still suffer from several drawbacks. Namely, these works were based on laboratory recorded vibration signals, lacking noise associated with variable industrial conditions. We note that the relatively weak bearing signals are always masked by quite stronger signals (gears, shafts, etc.) [10]. Besides, these works does not consider the WTG diagnosis where HSSB vibration signals are non-stationary due the variable wind load and wind speed. Hence, the proposed method is this paper is highlighted in terms of reality, efficiency and robustness of the proposed methodology.

To convince readers of the viability of the proposed method, we propose to add Additive Gaussian White Noise (AGWN) to all recorded vibration signals during the 50 days. The ART2 neural network was able to perform the same experimental results as shown in Fig. 11 even with the added noise. Despite the AGWN effect, the online bearing monitoring was still effective and the ART2 decision making still also accurate.

5. Conclusions and prospects

This work develops a new strategy that classifies the degradation of wind turbine high-speed bearings into three states (healthy, degraded, and failure). The online bearing monitoring task was accurately performed under real variable experimental conditions and the proposed method has presented several advantages such as the effectiveness against noise and extreme industrial conditions. In fact, recorded vibration signal was processed and classified to immediately estimate the degradation level instead of observing trends for a long period. Experimental validation based on real measured data from a wind turbine drivetrain proves that the proposed data-driven approach achieves very encouraging results and is suitable for wind turbine bearings online condition monitoring.

For future work, two prospective investigations will be considered:

- Although the proposed method contains the feature extraction step, another one will be considered to eliminate non-significant features that drive bearing defect decision making;
- Failure prognostic should be considered to anticipate the failure date by predicting the future health state of and its remaining useful life (RUL).

Appendix A. Processing summary of the ART2

It has been a long time when Carpenter and Grossberg have raised the compromise of stability-plasticity. This compromise can be summarized in three questions [36–39,50–54]:

(1) How can a learning system remain adaptive (plastic) in response to significant input, yet remain stable in response to irrelevant input?
(2) How does the system known to switch between its plastic and its stable modes?
(3) How can the system retain previously learned information (stable) while continuing to learn new things?

To answer these questions, Carpenter and Grossberg have defined the Adaptive Resonance Theory (ART) the origin of several neural networks including supervised and non-supervised learning strategies such as ART1, ART2, ARTmap, Fuzzy ARTmap, SFAM, PFAM, IPFAM, etc.

In this Appendix A, we will focus only on the ART2 neural network. In the following, we present two tables where Table A defines the used parameters and Table B summarizes the needed steps for the real implementation of the ART2 process.

Table A
ART2 parameters.

n	The number of components of the input vectors
m	The number of classifiers in the competition layer Y
a,b	Two constants proposed by Carpenter and Grossberg. Generally, they set a = b = 10
c	Constant proposed by Carpenter and Grossberg used in the reset mechanism. Generally, they c = 1
d	The activation of the winning classifier should respect the following condition: cd/(1-d) ≤ 1. The standard values are d = 0.9 and c = 0.1. This condition prevents the appearance of a reset during the resonance. The ratio should be chosen around 1 to achieve a wide range effective to vigilance
e	A very small added constant defined to prevent a division by zeros The choice of this parameter is intuitive. However, it can cause a digital instability for network
θ	The noise cancellation threshold where $\theta = 1/\sqrt{n}$
α	The learning coefficient
ρ	Vigilance parameter where $\rho \geqslant 0.7$
$t_{ji}(0)$	The initial state of the low–high synaptic weight matrix where $t_{ij}(0) = 0$
$b_{ij}(0)$	The initial state of the high-low synaptic weight matrix where $b_{ij}(0) \leqslant 1/[(1-d)\sqrt{n}]$

Table B
ART2 algorithm.

Step 1	Parameters initialization (a, b, c, d, e, α, θ, ρ).
Step 2	Repeat the steps 2–12 until reaching the adequate number of epochs noted N_{ep}
Step 3	For each input vector noted s, run the steps 4–12.
Step 4	Updating of the F_1 layer: (1) Initialization; $u_i = 0$ $w = s_i$ $p_i = 0$ $x_i = s_i/(e + \|s\|)$ $q_i = 0$ $v_i = f(x_i)$ (2) Update; $u_i = v_i/(e + \|v\|)$ $w = s + au_i$ $p_i = u_i + dt_{ji}$ $x_i = w_i/(e + \|w\|)$ $q_i = p_i/(e + \|p\|)$ $v_i = f(x_i) + bf(q_i)$ $where \quad f(x_i) = x_i if \quad x_i \geq \theta$ $else \quad f(x_i) = 0$
Step 5	Send the signal of the P sub-layer to the layer F_2: $y_i = \sum_{i=1}^{n} b_{ij} p_i$
Step 6	While the reset is true, do the steps 7–8
Step 7	Find the winner classifier with the highest activity. Define the index J for the winner neuron where $y_J \geqslant y_j \quad j = 1 \cdots m$

(continued on next page)

Table B (*continued*)

Step 8	Rest test: $u_i = v_i/(e + \|\|v\|\|)$ $p_i = u_i + dt_{Ji}$ $r_i = (u_i + cp_i)/(e + \|\|u\|\| + c\|\|p\|\|)$ ● If $\|\|r\|\| < \rho - e$ then the rest is true and the J neuron is inhibited with $y_J = -1$. Then, run the step 6 ● $\|\|r\|\| \geqslant \rho - e$ then : $w = s + du_i$ $x_i = w_i/(e + \|\|w\|\|)$ $q_i = p_i/(e + \|\|p\|\|)$ The reset is considered as false and go to the step 9 $v_i = f(x_i) + bf(q_i)$
Step 9	Repeat the steps 10–12 until reaching the required number of iterations (N_{it})
Step 10	Update the weights of the winning classifier J: $t_{Ji} = \alpha d u_i + \{1 + \alpha d(d-1)\} t_{Ji}$ $b_{iJ} = \alpha d u_i + \{1 + \alpha d(d-1)\} b_{iJ}$
Step 11	Update the F_1 layer: $u_i = v_i/(e + \|\|v\|\|)$ $w = s + au_i$ $p_i = u_i + dt_{Ji}$ $x_i = w_i/(e + \|\|w\|\|)$ $q_i = p_i/(e + \|\|p\|\|)$ $v_i = f(x_i) + bf(q_i)$
Step 12	Test the required number of iterations (N_{it})
Step 13	Test the required number of epochs (N_{ep})
Step 14	Remove the input vector. Restore all inhibit and inactive F_2 neurons. Return to step 3 with a new input pattern.

The ART2 is able to recognize the underlying similarity of identical patterns superimposed on constant backgrounds having different levels. In other words, two such patterns may appear entirely different, but thanks to the robust ART algorithm, they will be classified as same patterns. The strength of the ART2 algorithm is due to the updating given in step 10. In fact, the old t and b matrices will be updated and presented by a new form (weights) to better encompass and define the different forms of patterns of the same class. The updating process tries to minimize the inter-class variance and to maximize intra-class variance.

ART gets its name from the particular way in which learning and recall interplay in the network. In physics, resonance occurs when a small-amplitude vibration of the proper frequency causes a large-amplitude vibration in an electrical or mechanical system. In an ART network, information in the form of processing-element outputs reverberates back and forth between layers. If the proper patterns develop, a stable oscillation ensues, which is the neural-network equivalent of resonance. During this resonant period, learning or adaptation can occur. Before the network has achieved a resonant state, no learning takes place, because the time required for changes in the processing-element weights is much longer than the time that taken by network to achieve resonance.

A resonant state can be attained in one of two ways. If the network has learned previously to recognize an input vector, then a resonant state will be achieved quickly when that input vector is presented. During resonance, the adaptation process will reinforce the memory of the stored pattern. If the input vector is not immediately recognized, the network will rapidly search through its stored patterns looking for a match. If no match is found, the network will enter a resonant state whereupon the new pattern will be stored for the first time. Thus, the network responds quickly to previously learned data, yet remains able to learn when novel data are presented [52].

References

[1] Barrington-Leigh C, Ouliaris M. The renewable energy landscape in Canada: a spatial analysis. Renew Sustain Energy Rev 2017;75:809–19.

[2] Keller J, Guo Y, Sethuraman L. "Gearbox reliability collaborative investigation of gearbox motion and high-speed-shaft loads", (Technical Report) NREL/TP-5000e65321, national renewable energy laboratory (NREL). CO (US): Golden; 2016 Available at: http://www.nrel.gov/docs/fy16osti/65321.pdf.

[3] Chen J, Pan J, Li Z, Zi Y, Chen X. Generator bearing fault diagnosis for wind turbine via empirical wavelet transform using measured vibration signals. Renewable Energy 2016;89:80–92.

[4] Sheng S. "Wind Turbine Gearbox Reliability Database, Condition Monitoring, and O&M Research Update", 19 National Renewable Energy Laboratory (NREL), NREL/PR-5000-63868. Colorado (US): Golden; 2016. p. 1–20.

[5] Ben Ali J, Fnaiech N, Saidi L, Chebel-Morello B, Fnaiech F. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. Appl Acoust 2015;89:16–27.

[6] Hameed Z, Hong Y, Cho Y, Ahn S. C. Song, Condition monitoring and fault detection of wind turbines and related algorithms: a review. Renew Sustain Energy Rev 2009;13:1–39.

[7] Gao Z, Cecati C, Ding S. A survey of fault diagnosis and fault-tolerant techniques -part I: fault diagnosis with model-based and signal-based approaches. IEEE Trans Ind Electron 2015;62:3757–67.

[8] Gao Z, Cecati C, Ding S. 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.

[9] de Azevedo HDM, Araújo AM, Bouchonneau N. A review of wind turbine bearing condition monitoring: State of the art and challenges. Renew Sustain Energy Rev 2016;56:368–79.

[10] Randall RB, Antoni J. Rolling element bearing diagnostic-a review. Mech Syst Sig Process 2011;25:485–520.

[11] Ben Ali J, Chebel-Morello B, Saidi L, Malinowski S, Fnaiech F. Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network. Mech Syst Sig Process 2015;56–57:150–72.

[12] Saidi L, Ben Ali J, Fnaiech F. Bi-spectrum based-EMD applied to the non-stationary vibration signals for bearing faults diagnosis. ISA Trans 2014;53:1650–60.

[13] Borghesani P, Pennacchi P, Chatterton S, Ricci R. The velocity synchronous discrete Fourier transform for order tracking in the field of rotating machinery. Mech Syst Sig Process 2013;44:118–33.

[14] Urbanek J, Barszcz T, Antoni J. A two-step procedure for estimation of instantaneous rotational speed with large fluctuations. Mech Syst Sig Process 2013;38:96–102.

[15] Li Z, Jiang Y, Guo Q, Hu C, Peng Z, Multi-Dimensional Variational Mode Decomposition for Bearing-Crack Detection in Wind Turbines with Large Driving-Speed Variations. Renewable Energy, doi: 10.1016/j.renene.2016.12.013.

[16] Zhou Y, Chen J, Dong GM, Xiao WB, Wang ZY. Application of the horizontal slice of cyclic bispectrum in rolling element bearings diagnosis. Mech Syst Sig Process 2012;26:229–43.

[17] Zhang Y, Randall RB. Rolling element bearing fault diagnosis based on the combination of genetic algorithms and fast kurtogram. Mech Syst Sig Process 2009;23:1509–17.

[18] Bellini A, Filippetti F, Tassoni C, Capolino G-A. Advances in diagnostic techniques for induction machines. IEEE Trans Ind Electron 2008;55:4109–26.

[19] Dubey R, Agrawal D. Bearing fault classification using ANN-based Hilbert footprint analysis. IET Sci Meas Technol 2015;9:1016–22.
[20] Prieto MD, Cirrincione G, Espinosa AG, Ortega JA, Henao H. "Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks". IEEE Trans Ind Electron 2013;60:3398–407.
[21] Cococcioni M, Lazzerini B, Volpi SL. Robust diagnosis of rolling element bearings based on classification techniques. IEEE Trans Ind Electron 2013;60:2256–63.
[22] Saidi L, Ben Ali J, Fnaiech F. Application of higher order spectral features and support vector machines for bearing faults classification. ISA Trans 2015;54:193–206.
[23] Saidi L, Ben Ali J, Bechhoefer E, Benbouzid MEH. Wind turbine high-speed shaft bearings health prognosis through a spectral Kurtosis-derived indices and SVR. Appl Acoust 2017;120:1–8.
[24] Tandon N, Choudhury A. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. Tribol Int 1999;32:469–80.
[25] Lei Y, Lin J, He Z, Zuo MJ. A review on empirical mode decomposition in fault diagnosis of rotating machinery. Mech Syst Sig Process 2013;35:108–26.
[26] Obeid Z, Picot A, Poignant S, Regnier J, Darnis O, Maussion P. Experimental comparison between diagnostic indicators for bearing fault detection in synchronous machine by spectral Kurtosis and energy analysis. 38th Annual Conference of IEEE Industrial Electronics (IECON2012), 2012;3901–3906.
[27] Antoni J. Fast computation of the kurtogram for the detection of transient faults. Mech Syst Sig Process 2007;21:108–24.
[28] Guoji S, Laughlin McS, Yongcheng X, White P. "Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears". Mech Syst Sig Process 2014;43:76–89.
[29] Chua KC, Chandranb V, Acharyaa UR, Lima CM. Application of higher order statistics/spectra in biomedical signals-a review. Med Eng Phys 2010;32:679–89.
[30] Rato RT, Ortigueira MD, Batista AG. On the HHT, its problems, and some solutions. Mech Syst Sig Process 2008;22:1374–94.
[31] Chen GD, Wang ZC. A signal decomposition theorem with Hilbert transform and its application to narrowband time series with closely spaced frequency components. Mech Syst Sig Process 2012;28:258–79.
[32] Rilling G, Flandrin P, Gonçalvés P. On empirical mode decomposition and its algorithms. IEEE–EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP), 2003.
[33] Yao J, Xiang Y, Qian S, Wang S, Wu S. Noise source identification of diesel engine based on variational mode decomposition and robust independent component analysis. Appl Acoust 2017;116:184–94.
[34] Ben Ali J, Sayadi M, Fnaiech F, Morello B, Zerhouni N. Importance of the fourth and fifth intrinsic mode functions for bearing fault diagnosis. in: 14th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), 2013;259–264, .
[35] Dragomiretskiy K, Zosso D. Variational mode decomposition. IEEE Trans Signal Process 2014;62:531–44.
[36] Carpenter GA, Grossberg S. ART2: Self organisation of stable category recognition codes of analog input patterns. IEEE First International Conference on Neural Networks. 1987;727–735.
[37] Carpenter GA, Grossberg S. Fuzzy ARTMAP : a neural network architecture for incremental supervised learning of analog multidimensional maps. IEEE Transact Neural Net 1992;3:698–713.
[38] Grossberg S. Adaptive pattern classification and universal recoding, I: Parallel development and coding of neural feature detectors. Biol Cybern 1976;23:121–34.
[39] Grossberg S. Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, and illusions. Biol Cybern 1976;23:187–202.
[40] Guoji S, McLaughlin S, Yongcheng X, White P. Theoretical and experimental analysis of bispectrum of vibration signals for fault diagnosis of gears. Mech Syst Sig Process 2014;43:76–89.
[41] Frosini L, Bassi E. Stator current and motor efficiency as indicators for different types of bearing faults in induction motors. IEEE Trans Ind Electron 2010;57:244–51.
[42] SKF bearing catalog. Available at: http://www.skf.com/fr/products/bearings-units-housings/roller-bearings/tapered-roller-bearings/single-row-tapered-roller-bearings/single-row/index.html?designation=32222%20J2 (accessed date : 06-03-2017).
[43] Shakya P, Kulkarni MS, Darpe AK. A novel methodology for online detection of bearing health status for naturally progressing defect. J Sound Vib 2014;333:5614–29.
[44] Ben Ali J, Saidi L, Mouelhi A, Chebel-Morello B, Fnaiech F. Linear feature selection and classification using PNN and SFAM neural networks for a nearly online diagnosis of bearing naturally progressing degradations. Eng Appl Artif Intell 2015;42:67–81.
[45] Fernández-Francos D, Martínez-Rego D, Fontenla-Romero O, Alonso-Betanzos A. Automatic bearing fault diagnosis based on one-class v-SVM. Comput Ind Eng 2013;64:357–65.
[46] Jian-Bo Y. Bearing performance degradation assessment using locality preserving projections. Expert Syst Appl 2011;38:7440–50.
[47] Jianbo Y. Bearing performance degradation assessment using locality preserving projections and Gaussian mixture models. Mech Syst Sig Process 2011;25:2573–88.
[48] Liu X, Bo L, Luo H. Bearing faults diagnostics based on hybrid LS-SVM and EMD method. Measurement 2015;59:145–66.
[49] Liu Z, Cao H, Chen X, He Z, Shen Z. Multi fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings. Neurocomputing 2013;99:399–410.
[50] Carpenter GA, Grossberg S. Invariant pattern recognition and recall by an attentive self-organizing ART architecture in a nonstationary world. IEEE First International Conference on Neural Networks, 1987; 737–745.
[51] Carpenter GA, Grossberg S. A massively parallel architecture for a self-organizing neural pattern recognition machine. Computer Vision, Graphics, and Image Processing 1987;37:54–115.
[52] Carpenter GA, Stephen Grossberg. The ART of adaptive pattern recognition by a self-organizing neural network. Computer 1988;21:77–88.
[53] Carpenter GA, Grossberg S, Mehanian C. Invariant recognition of cluttered scenes by a self-organizing ART architecture: CORT-X boundary segmentation. Neural Networks 1989;2:169–81.
[54] Freeman IA, Skapura DM. Neural network, algorithms, applications and programming techniques. Addison Wesley; 1991.