



New fault diagnosis approaches for detecting the bearing slight degradation

Saeed Nezamivand Chegini · Mohammad Javad Haghdoust Manjili · Ahmad Bagheri

Received: 23 March 2019 / Accepted: 20 December 2019 / Published online: 11 January 2020
© Springer Nature B.V. 2020

Abstract In this paper, two new methods for detecting the bearing's degradation starting points are presented based on the vibration signal analysis. In the first method, a new feature extraction technique is suggested based on the envelope harmonic-to-noise ratio (EHNR) and the fast ensemble empirical mode decomposition (FEEMD). Each vibration signal is decomposed into its intrinsic mode functions (IMFs) using the FEEMD algorithm. Also, a novel technique has been introduced based on the autocorrelation function (ACF) of the original signal and its IMFs for selecting the most appropriate IMF and eliminate the noisy components. Then, the EHNR of the most sensitive IMF is computed for detecting the early degradation of bearing. In the second method, a new adaptive feature is defined using the ACF of the raw signal and the energy-entropy vector. At first, a novel indicator called the periodicity intensity factor (PIF) is introduced using the energy of the ACF of the raw signal and its maximum points. In the next step, the energy-entropy variations of the PIF factor are investigated for recognizing the fault starting point in

bearings. In this work, the vibration signals of the run-to-failure experiment are used to appraise the presented techniques. The results indicate that the proposed approaches are able to detect the exact moment of the defect occurrence. Also, comparing the results of this paper with other techniques presented recently indicates the superiority of the proposed approaches.

Keywords Incipient fault detection · Bearing · Empirical mode decomposition (EMD) · Envelope harmonic-to-noise ratio (EHNR) · Auto-correlation function (ACF) · Periodicity intensity factor (PIF)

1 Introduction

Rotary machines play a very important role in many production lines. The failure of these machines will result in a huge loss for industrial sets. The rolling bearing is one of the main components of the rotating machinery. In general, their failure leads to full paralysis of the machines [1, 2]. Therefore, the bearing fault detection is one of the main tasks of engineers. The subject of designing a condition monitoring approach becomes important when its main purpose is to detect the degradation starting point. In recent years, many studies were carried out to monitor the condition of the bearings. The analysis of vibration signals and the vibration data-based feature extraction

S. N. Chegini (✉) · A. Bagheri
Department of Dynamics, Control, and Vibrations,
Faculty of Mechanical Engineering, University of Guilan,
Rasht, Iran
e-mail: saeed.nezamivand@gmail.com

M. J. H. Manjili
Department of Mechanical Engineering, Ahrar Institute of
Technology and Higher Education, Rasht, Iran

are the approaches that were used to identify the bearing damages in different working conditions [3].

In recent decades, different signal processing methods were presented to develop the fault diagnosis techniques. The feature extraction methods are utilized to extract the fault signatures and are categorized into three main groups: time domain, frequency domain, and time–frequency domain analysis techniques. In the classical feature extraction methods in the time domain, the features such as standard deviation, skewness, kurtosis, crest factor, shape factor, and clearance factor are extracted from the vibration signals. These features are only able to detect the presence of defects and they cannot be used to diagnose the type of faults [4]. On the other hand, the diagnosis of the fault type is possible by analyzing the vibration data in the frequency domain. When a fault is created in one of the bearing components, the alternative impulses are appeared in the vibration signals by the collision of defect location with the other surfaces of bearings. Each element of the bearing has a specific characteristic frequency. These frequencies are the ball spin frequency, ball fault frequency, ball pass frequency outer race and ball pass frequency inner race. Therefore, if the spectrum of a vibration signal includes one of these frequencies and its harmonics, the type of defect can be identified [5]. The analysis of the spectrum of acquired data is one of the useful tools for identifying the location and type of fault. This approach has been used in many studies [6–8]. The third group of signal processing methods is the time–frequency domain techniques. In recent years, various approaches such as wavelet transform (WT) [9], empirical mode decomposition (EMD) [10], ensemble EMD (EEMD) [11], empirical wavelet transform (EWT) [12] and variational mode decomposition (VMD) [13] were proposed for extracting the time–frequency features. These techniques and their improved versions were used in various researches to develop the sagacious fault diagnosis methods, denoise the noisy vibration signal and extract the fault - susceptible feature.

In papers such as [14–16], researchers developed the bearings and gears fault detection methods based on improving the implementation steps of the EMD and EEMD techniques. Wang et al. [17] proposed a self-adaptive filter using the EEMD method for eliminating the noise from the vibration signals acquired from a damaged locomotive bearing. For

this purpose, an adaptive relationship was suggested for computing the number of the sifting process based on the number of signal IMFs. The results of their work demonstrated that the fault characteristics can easily be seen in the frequency spectrum of the denoised signal. Wei et al. [18] proposed a new signal processing for detecting the bearing fault. They extracted the time and frequency statistical features from the vibration signal using the wavelet packet transform (WPT) and EEMD methods. Then, the authors presented a novel optimal feature selection method based on the adaptive feature selection technique and affinity propagation clustering method. Abdelkader et al. [19] proposed a new strategy for diagnosing the characteristics of the bearing faults using the improvement of the EMD-based denoising method, the kurtosis value, and the envelope spectrum. In their work, the vibration signals were decomposed into several IMFs. Then, the trip point and the singular IMFs were selected based on the energy of all IMFs. Finally, the soft thresholding and the optimized threshold were applied to denoise the singular IMFs. Cao et al. [20] investigated the wheel-bearing fault diagnosis of trains using the EWT approach. In this work, the EWT technique for different case studies, and also for the compound fault of the faulty outer race and faulty rolling elements has been applied. Kedadouche et al. [21] combined the EWT method and the operational modal analysis (OMA) for improving the bearing fault diagnosis. In this study, the kurtosis parameter has been performed for identifying the appropriate modes and extracting the defect frequencies. Zhang et al. [22] designed a novel technique for recognizing the bearing fault of the multistage centrifugal pumps based on the VMD method. Cocconcilli et al. [23] presented a procedure for the condition monitoring of the ball bearing in direct-drive motors with the non-stationary condition. They analyzed the vibration signals to highlight the presence of damage impacts in the time–frequency domain. Bellini et al. [24] suggested a fault detection technique for bearings damage based on the statistical analysis of vibration and current signals. The authors used the spectral kurtosis and the energy of the signal to identify the spreading bandwidth related to generalized roughness and introduce the diagnostic index, respectively. Montechiesi et al. [25] introduced a bearing faults recognition approach using the mechanisms of the immune system. Their proposed

algorithm is based on the Euclidean Distance Minimization in the evaluation of the binding between antigens.

In all of the studies mentioned above, different fault types with various sizes were artificially created on each of the components of rotating machinery such as bearings and gears. In fact, the researchers investigated the validation of their proposed methods with fully aware of the status of their case studies. On the other hand, in practical applications, the mechanism of creating faults is not consciously. In other words, according to the working conditions of the bearing, a particular defect may be created by one of its components and the bearing loses its efficiency by expanding the defect severity. Therefore, the recognition of the presence and type of fault in the initial moments can provide sufficient opportunity for the operator to take the appropriate actions. Recently, limited studies have been conducted on the detection of early degradation. In these works, the researchers utilized the run-to-failure data set provided by the center of Intelligent Maintenance System (IMS) at NASA website [26] for evaluating their suggested approaches. Qiu et al. [27] applied the wavelet filter to extract the weak signature of the mechanical impulse-like defect signals. They used the minimal Shannon entropy and the signature value decomposition (SVD) for finding the optimal values of the Morlet wavelet factor and the scale of the wavelet transform, respectively. Yu [28] proposed a new feature selection technique based on the locality preserving propagation (LPP) for exploiting the most informative fault signatures from the original high dimensional feature set. Then, a new on-line bearing performance degradation evaluation was implemented using a combination of the squared prediction error (SPE) statistic and the Exponential Weight Moving Average (EWMA) statistic. In other work [29], the author applied the dynamic principal component analysis (DPCA) for extracting the most appropriate information from the raw signals. In the next stage, the obtained useful features were used as input of the Hidden Markov Model (HMM) for monitoring the bearing in the test-to-failure experiment. Fernandez-Francos et al. [30] used the one-class ν -SVM to recognize the degradation starting point of bearing based on the healthy vibration data in the run-to-failure test. Ben Ali et al. [31] performed the EMD method to process the non-stationary bearing signals. The authors applied the

combination of the EMD-energy entropy and statistical features to construct the feature matrix as the input of the artificial neural network (ANN) classifier. They suggested a health index (HI) for predicting the degradation point. Xu et al. [32] indicated that the root-mean-square (RMS) and kurtosis are not suitable parameters for appearing the periodicity property of impulses produced by the bearing defects. Therefore, they introduced an effective feature called envelope harmonic-to-noise ratio (EHNR) for monitoring the incipient bearing faults. Jia et al. [33] investigated the capabilities of the minimum entropy deconvolution (MED) and the Convolutional Sparse Filter (CSF) in extracting the impulsive signature. They illustrated the performance of the CSF and the MED techniques for incipient fault detection. Hasani et al. [34] introduced an unsupervised feature extraction by using the auto-encoder correlation (AEC) to diagnose the condition of bearing during the test-to-failure experiment. Li et al. [35] designed a novel weak fault feature extraction technique using the intrinsic character-scale decomposition (ICD) and the turntable Q-factor wavelet transform (TQWT) to estimate the moment of the bearing fault occurrence. Jiang et al. [36] improved the VMD approach by performing the ability of EMD technique in the VMD decomposition process to monitor the weak transient impulses of faulty bearing in the early stage. Dybala [37] presented a new bearing diagnostics approach in the early stage based on the amplitude level-based decomposition of the vibration data. The author proposed a new effective feature from a low-energy component by using the power spectra of the empirically identified local amplitude. Lv et al. [38] utilized a novel strategy based on the complete EEMD with adaptive noise (CEEMDAN) and improved multivariate multi-scale sample entropy (MMSE) for detecting the incipient fault. Qian et al. [39] used the combination of the recurrence quantification analysis (RQA) with Kalman filter for diagnosing the bearing degradation phase. They applied the RQA and Kalman filter to extract the novel feature from the vibration data and identify the bearing conditions, respectively.

In this paper, two novel methods are proposed for diagnosing the degradation starting point. Then, similar to the articles described in this section, the test-to-failure data set provided by IMS is used to check the accuracy of the proposed approaches. The first technique is a combination of the EHNR-based

feature extraction and the EEMD decomposition method. In this method, each vibration signal is decomposed into its constituent components through EEMD. All of these components do not contain the fault-related information and may even be impregnated by noise. Therefore, a new technique based on the correlation coefficient between the auto-correlation coefficient of the original signal and its components has been provided for selecting the most suitable component. Then, the EHNR of the most informative IMF is calculated for appearing the characteristics of the periodic impulses and the incipient faults signatures. In the second proposed technique, a new effective index is defined based on the energy of the auto-correlation of the raw signal and the energy-entropy feature. The results show that the proposed approaches are effective in identifying the fault starting point and are superior to classical techniques such as EHNR and other early fault detection methods.

The remainder of this article is compiled as follows: A summary of the bearings concepts, their characteristic frequencies and the experimental setup utilized in this paper are described in Sect. 2. In Sect. 3, the EEMD algorithm and EHNR are explained. In Sect. 4, both proposed methods and their results are presented. Comparison of suggested approaches with other methods is presented in Sect. 5. Finally, this work has been concluded in Sect. 6.

2 Bearing fault detection in run-to-failure experiment

So far, most researches have been devoted to the bearing fault diagnosis with the artificial defects [5–22]. On the other hand, a few papers studied the condition monitoring of the bearings based on the run-to-failure vibration signals. Although the researchers in most experiments investigated the identification of the small-sized defects, these defects have been artificially created. Hence, their results are not suitable for detecting the fault starting point. In the run-to-failure tests, it is possible to study the early fault detection in the bearings. In other words, in the run-to-failure experiments, fault detection in the earliest stage is one of the most important tasks. Importance of this issue can be discussed as follow:

- The fault diagnosis at a preliminary level allows applying the necessary actions by operators.
- The incipient fault detection in a particular component of the bearing prevents the defect expansion to the other components and as a result, it makes economic sense.
- The fault diagnostic at the early moments provides the possibility of predicting the remaining useful life (RUL) of the bearing. Therefore, maintenance can be performed at a lower cost.

2.1 The bearing fault characteristic frequencies

The vibration signals of the faulty bearings usually include fault-induced periodic impulses which can be used as an indicator to diagnose the bearing conditions. These series of impulses are produced by interactions between the fault location and the rolling elements of bearing. The periodicity of the produced impulses depends on the defect location in the bearing, i.e., the inner race, outer race, and rolling element. Therefore, each type of fault is corresponding to a specific frequency which called the characteristic fault frequency. The presence of a specific fault characteristic frequency along with its harmonics in the envelope spectrum of the vibration signal will be an indication of a particular fault type. The commonly used equations for calculating these frequencies have been presented in Table 1 [30]. In this table, F_o is the ball pass frequency of the outer race, F_i is the ball pass frequency of the inner race, BSF is the ball spin frequency, FTF is the fundamental train frequency, F_s is the rotational frequency, N_b is the number of rolling elements, d is the rolling element diameter, D_m is the pitch diameter and α is the contact angle which is the angle of the load from the radial plane. In these equations, it is assumed that there is no slip between the components of the bearing. Nevertheless, practically there is some slip. Therefore, the angle α is changed with the position of each rolling element in the bearing. As a result, in the practical condition, the characteristic frequency value generally deviates about 1–2% from the theoretical value. Table 2 indicates the list of frequencies corresponding to different fault types that are emerged in the envelope spectrum [30].

Table 1 The equations of bearing characteristic frequencies [30]

Bearing characteristic frequencies	Formula
Ball-pass frequency of outer race	$F_o(Hz) = F_s \left(\frac{N_b}{2}\right) \left(1 - \frac{d}{D_m} \cos\phi\right)$
Ball-pass frequency of inner race	$F_i = F_s \left(\frac{N_b}{2}\right) \left(1 + \frac{d}{D_m} \cos\phi\right)$
Fundamental train frequency	$FTF(Hz) = F_s \left(\frac{1}{2}\right) \left(1 - \frac{d}{D_m} \cos\phi\right)$
Ball spin frequency	$BSF(Hz) = F_s \left(\frac{D_m}{2d}\right) \left(1 - \frac{d^2}{D_m^2} \cos^2\phi\right)$

Table 2 The frequencies appeared in the envelope spectrum correspond to different fault types [30]

Bearing fault	Frequencies appeared in the power spectrum	Description
Rolling elements	$2 \times BSF, F_o, F_i$	Modulated by $2 \times BSF$ or FTF
Outer race	F_o	Harmonics may be found
Inner race	F_i	Harmonics may be found Inner race faults are typically modulated by F_s

2.2 Experimental setup

In this paper, it is studied a run-to-failure lifetime testing provided by the center of Intelligent Maintenance Systems (IMS) of the University of Cincinnati [26]. The data package includes three data sets. Each test is a run-to-failure experience of four bearings. So, a total of twelve bearings were used but only four bearings were reached failure with the known defects.

Bearing test rig and schematic of the sensor placements is shown in Fig. 1. According to this illustration, PCB 353B33 High sensitivity Quartz ICP accelerometers installed on the bearing housing in X

and Y axes (shown in Fig. 1a) were used to acquire each vibration signal. The vibration data were collected with the sampling rate of 20 kHz by a National Instruments DAQCard-6062E data acquisition card. The details of case studies used in presented work are summarized in Table 3. In each case study, the number of samples, the experimental full lifetime, the damaged bearing number and the type of fault are reported. Each dataset consists of individual files that are 1-s vibration signal snapshots recorded every 10 min and each file contains 20,480 points. Rexnord ZA-2115 double row bearings were used in this run-to-failure test. The bearings contain 16 rollers in each

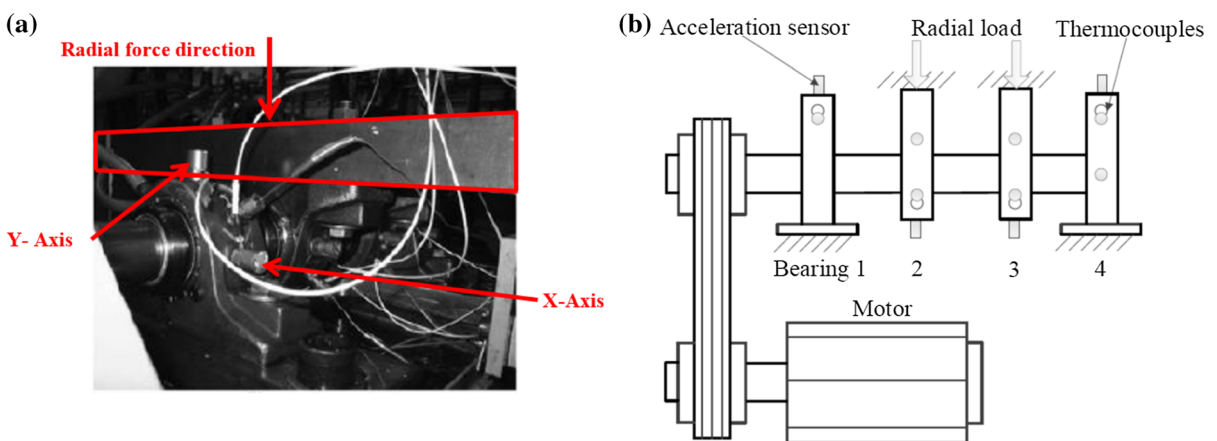


Fig. 1 Bearings run-to-failure test rig [27]. **a** Test system, **b** System structure

Table 3 Summary information of case studies used in the presented work

Case study	Number of samples	Experiment full lifetime (day)	Number of defected bearing (defect type)
Case 1	2156	35	Bearing 3 (inner race)
Case 2	984	8	Bearing 1 (outer race)
Case 3	6324	46	Bearing 3 (outer race)

row, the pitch diameter of 7.150 cm, a roller diameter of 0.840 cm and a tapering contact angle of 15.17 degree. In each case study described in Table 3, the outer ring is stationary and the inner ring rotates with the shaft speed. A constant radial load of 6000 lbs (about 26,690 N) was performed to the shaft and each bearing in Y-Axis by a spring mechanism, and the rotating speed of the shaft was kept constant at 2000 rpm. All of the bearings were force lubricated. A magnetic plug collected debris from the oil circulation. An electronic switch was utilized to stop the test when the debris exceeds a certain level and cause a switch to turn off [27].

Figure 2a–c demonstrate the vibration signals of the faulty bearings presented in Table 3. According to the bearing geometry, the rotational speed and the frequency characteristic equations presented in Table 1, the theoretical value of F_o , F_i , BSF , FTF

are 236.4 Hz, 296.9 Hz, 139.9 Hz, and 14.04 Hz, respectively.

3 Methods

3.1 Ensemble empirical mode decomposition method

Ensemble empirical mode decomposition (EEMD) [11] is an adaptive signal processing method for decomposing the complex signal. This method can decompose a nonlinear signal into a series of functions called intrinsic mode functions (IMFs) and a residue via an iterative procedure named the sifting process. The main idea of the EEMD method is to solve the problems of the empirical mode decomposition (EMD) such as the mode mixing phenomenon and the end effects. In the EEMD technique, the original signal with added white noise is repeatedly decomposed into a series of IMFs by applying the original EMD process. Then the final EEMD decomposition results are calculated by the ensemble average of the extracted IMFs. The EEMD algorithm can be described as follows:

In the first step, the Gaussian white noise with different amplitudes is added to the original signal and ensembles of signals are produced:

$$x_i = x + amp \cdot w_i, \quad i = 1, 2, \dots, n \quad (1)$$

where w_i is the Gaussian white noise (with zero-mean and unit variance), the amp is the amplitude of the added noise, and n is the number of ensembles. The amplitude is commonly computed relative to the standard deviation of the original signal. In the next step, the signal $x_i (i = 1, \dots, n)$ is decomposed by the traditional EMD method into its IMFs. It is assumed that c_{ik} is the k th IMF produced by the i th realization. In the last step, the ensemble mean of the

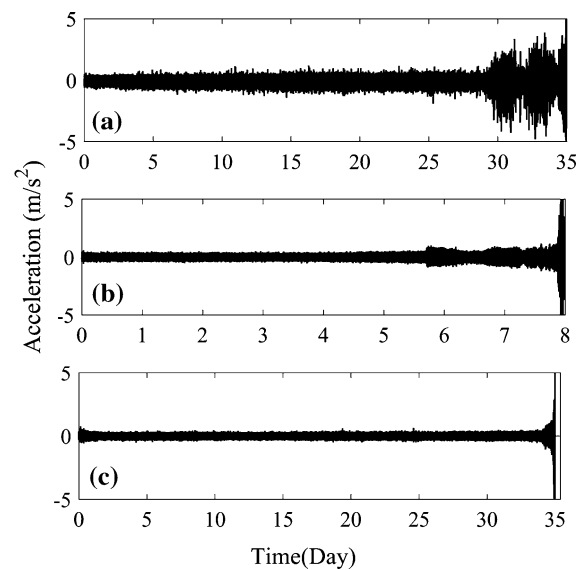


Fig. 2 The vibration signals corresponding to the damaged bearings introduced in Table 3 [27]. **a** Case 1, **b** Case 2 and **c** Case 3

corresponding IMFs is computed based on the following equation:

$$\bar{c}_k = \frac{\sum_{i=1}^n c_{ik}}{n}, \quad k = 1, 2, \dots, I \tag{2}$$

where \bar{c}_k and I are the k th IMF of the original signal x and the minimum number of IMFs among all the trials, respectively.

In the EEMD method, determining the amplitude of added noise and the ensemble number are two vital parameters in the decomposition process. Wu et al. [11] indicated that the EEMD with an ensemble number of a few hundred and the added noise amplitude of 0.2 standard deviation of the original signal will lead to a very good result. On the other hand, Wang et al. [40] proposed an optimized program (FEEMD) to increase the computation speed of EEMD up to about 1000 times faster. In the presented paper, the FEEMD approach is used to process the vibration signals. The original vibration signal corresponding to

Case 2 and its first six IMFs and residual obtained by FEEMD are shown in Fig. 3.

3.2 Envelope harmonic-to-noise ratio for early fault detection

Xu et al. [32] proposed a new method based on envelope harmonic-noise ratio (EHNR) for detecting the periodicity of the fault-induced impulses. According to [32], the implementation steps of the EHNR method are as follows:

1. Calculate the Hilbert transform of the signal based on the following equation:

$$\hat{x}(t) = H\{x(t)\} = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t - \tau} d\tau \tag{3}$$

2. Compute the direct envelope of the signal and remove the direct-current (DC) component from it:

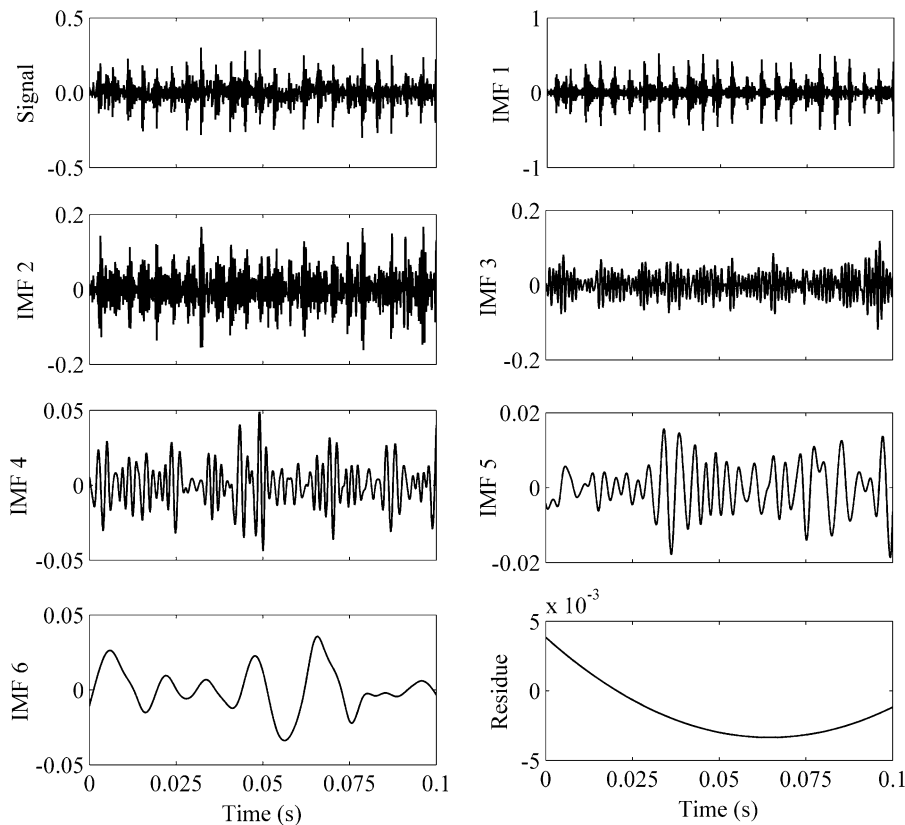


Fig. 3 Signal vibration and its first 6 IMFs and residue obtained via FEEMD

$$Env'_x(t) = \sqrt{\hat{x}(t)^2 + x(t)^2} \quad (4)$$

$$Env_x(t) = Env'_x(t) - \text{mean}(Env'_x(t)) \quad (5)$$

3. Calculate the autocorrelation of $Env_x(t)$:

$$r_{env_x}(\tau) = \int Env_x(t)Env_x(t + \tau)d\tau \quad (6)$$

where τ is the lag in the autocorrelation function (ACF). ACF is a powerful tool for finding the periodic events such as the fault-related impulses.

4. Find the maximum position of the autocorrelation function of the original signal in the lag domain. Then EHNR is defined as follows:

$$EHNR = \frac{r_{env_x}(\tau_{max})}{r_{env_x}(0) - r_{env_x}(\tau_{max})} \quad (7)$$

where τ_{max} is the maximum location of the autocorrelation of $Env_x(t)$. $r_{env_x}(\tau_{max})$ is the amplitude of the autocorrelation spectrum at $\tau = \tau_{max}$ or the energy of harmonics and $r_{env_x}(0)$ is the total energy of the envelope.

In [32], the EHNR results were presented only to identify the moment of the occurrence of the fault for the second case study and no interpretation was provided for analyzing the bearing conditions. Therefore, in this work, the steps of the condition monitoring process of Case 2 are investigated for further explanation. For this purpose, the results of the original EHNR method for vibration signals of Case 2 are shown in Fig. 4. As shown in this figure, the health condition monitoring process clearly can be divided into five segments. These five parts are discussed as follows:

1. Healthy phase: This section takes about 90 h. In this phase, the bearing works in perfectly healthy conditions. The changes of the EHNR values in the healthy phase are very insignificant.
2. Fault occurrence phase: This phase takes about 28 h and is started when a small fault has been created in the bearing. Since the fault-related key characteristics in this stage often are masked by the heavy noise, identifying the start time of this phase is very difficult.
3. Initial defect propagation phase: In this phase, the increasing the defect size is more accelerated and the fault-related characteristics become obvious.

Therefore, this phase is relatively easy to recognize.

4. Healing Phase: When the edges of a crack or a small defect area created in the bearing components have been smoothed due to the continuous contact of the damage location with the rolling elements of bearings, the healing phenomenon is utilized [32]. In this phase, the EHNR amplitude is reduced and the fault-related characteristics are hidden.
5. Severe degradation phase: The bearing fault grows more rapidly, and at the end of this step, the bearing will not have enough efficiency to continue the work. At this stage, maintenance is not appropriate and can cause serious damage to the entire mechanical system.

The results of the original EHNR for Case 1 and Case 3 are presented in Figs. 5 and 6, respectively. Xu et al. [32] assumed that the bearings work in the normal condition at hours 1–80 and they considered the mean of EHNR plus four times the standard deviation of the EHNR in the normal region as a criterion for distinguishing the instant of the fault occurrence. Similarly, in this paper, this criterion is used for early fault identification. As shown in Fig. 5, the first faulty sample determined by the original EHNR method for Case 1 is the 2130th sample. Another noteworthy point in Fig. 5 is the absence of false alarms in the EHNR curve of Case 1. The results of the original EHNR of Case 3 are different from those of the EHNR of Case 1. According to Fig. 6, in the EHNR curve of Case 1, the first, 22th, 25th and 6162th samples are nominated for identifying the incipient fault moment. Since the three samples 1, 22 and 25 are related to the initial hours of the bearing operation and at these moments, the bearing works in the healthy condition, these samples cannot be considered as the bearing degradation starting points. Therefore, the sample corresponding to the fault occurrence time reported by the original EHNR for Case 3 is the 6162th measurement. Three measurements 1, 22 and 25 are the false alarms that confuse the operator in identifying the fault occurrence time. The existence of these false warnings indicates the weakness of the original EHNR method in the early defect diagnosis.

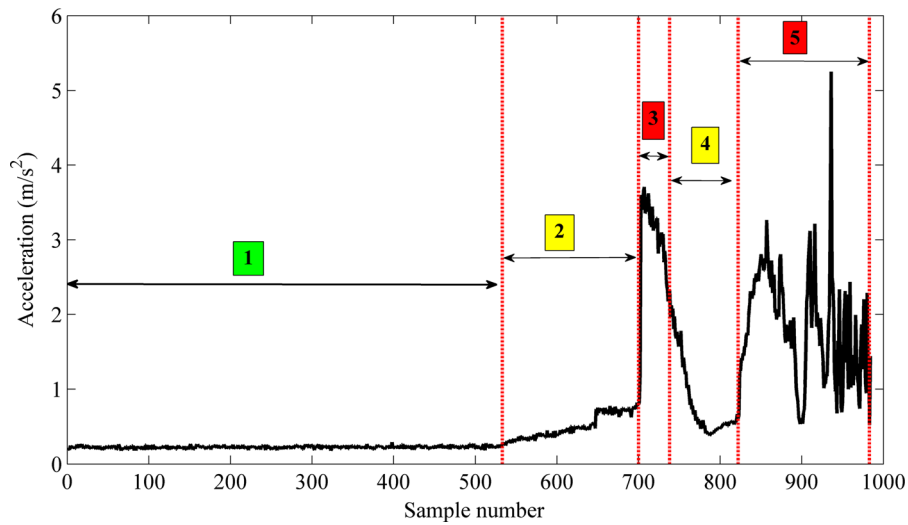


Fig. 4 The results of the original EHNHR method for Case 2

4 Proposed methods

4.1 Proposed method 1: FEEMD–EHNHR

4.1.1 New approach for selecting the most sensitive IMF

The selection of the most informative IMFs for detecting the defects in the early stage is a fundamental problem in the bearing fault feature extraction. The selection of the most sensitive IMFs is equivalent to choosing the IMFs containing the fault-related information and removing the noisy IMFs. In the previous studies, such as [41, 42], the researchers applied the first few IMFs of each signal as the meaningful components for extracting the features from the vibration data and forming the feature vector. The presence of possible noise and unrelated information in these components as well as the high computational cost due to the use of several IMFs are the shortcomings of the approach employed in those researches. In this section, the cross-correlation coefficient between the auto-correlation of each IMF and the original signal is utilized as a criterion for selecting IMF which includes the most dominant fault information. According to this idea and the algorithm proposed in [43], in this work, a new method is proposed for finding the most suitable component. The steps of this technique are described below:

1. Calculate the cross-correlation between the auto-correlation of the original signal and the auto-correlation of the obtained IMFs denoted as $T_i^{(k)}$ ($i = 1, 2, \dots, n$), where $T_i^{(k)}$ is the correlation between the autocorrelation of the k th sample ($k = 1, 2, \dots, m$) and the autocorrelation of the i th IMF of the mentioned sample. n and m are the number of IMFs obtained in the decomposition process and the total samples, respectively.
2. Calculate the cross-correlation between the auto-correlation of the original signal in the healthy region and the obtained IMFs denoted as $H_i^{(k)}$ ($i = 1, \dots, n$), where $H_i^{(k)}$ is the correlation between the autocorrelation of the k th healthy sample ($k = 1, 2, \dots, h$) and the autocorrelation of the i th IMF of the mentioned sample. n and h are the number of IMFs obtained in the decomposition process and the number of healthy samples, respectively.
3. Calculate the mean value of $T_i^{(k)}$ and $H_i^{(k)}$:

$$\alpha_i = \frac{1}{m} \sum_{k=1}^m T_i^{(k)} \quad (k = 1, 2, \dots, m) \quad (8)$$

$$\beta_i = \frac{1}{h} \sum_{k=1}^h H_i^{(k)} \quad (k = 1, 2, \dots, h) \quad (9)$$

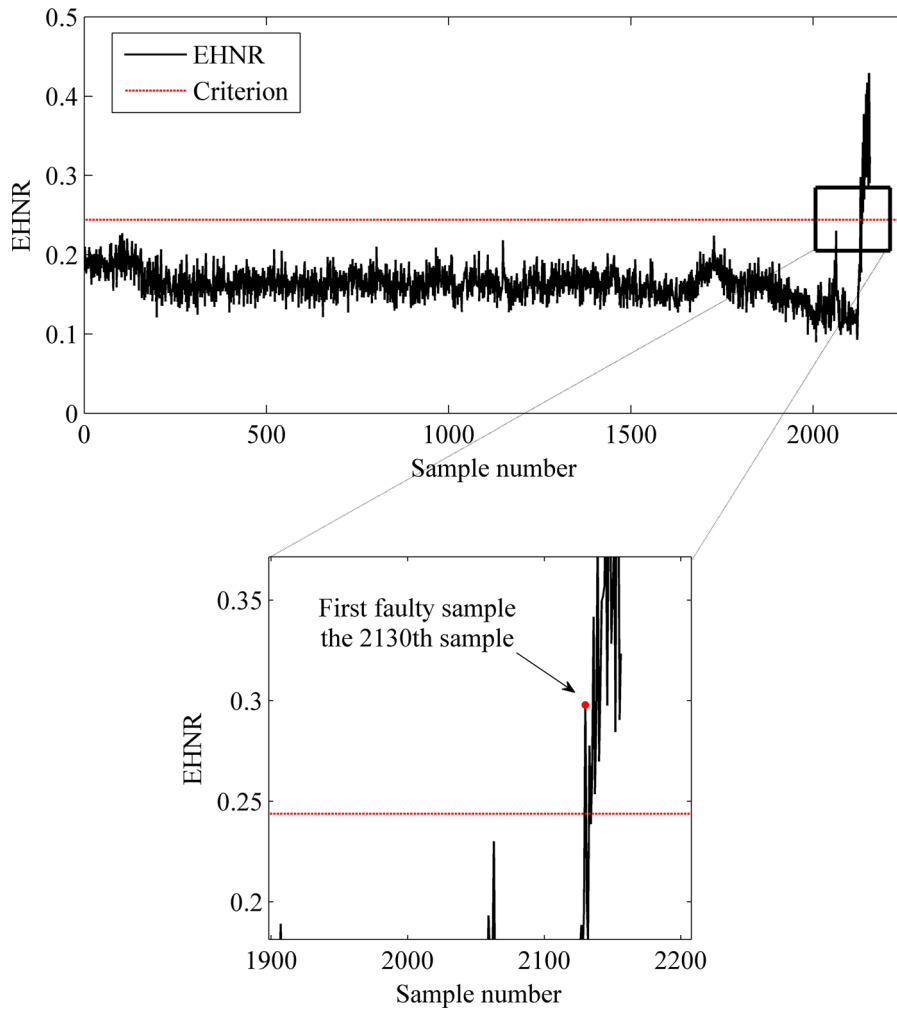


Fig. 5 The results of the original EHNR for Case 1

where α_i is the mean value of $T_i^{(k)}$ for total samples, β_i is the mean value of $H_i^{(k)}$ for normal samples.

4. Compute the absolute difference between two parameters α_i and β_i :

$$\gamma_i = |\alpha_i - \beta_i|, \quad i = 1, 2, \dots, \acute{n} \tag{10}$$

where γ_i called the fault-related coefficient and \acute{n} is the minimum number of the IMFs obtained from the considered samples.

5. Introduce the sensitivity factor λ_i according to the following equation:

$$\lambda_i = \frac{\gamma_i - \min(\gamma)}{\max(\gamma) - \min(\gamma)}, \quad \gamma = \{\gamma_n\}, \quad i = 1, 2, \dots, \acute{n} \tag{11}$$

6. Sort all the IMFs in terms of their sensitivity factors in decreasing order to get the following series:

$$\{\gamma'_i\}, \quad i = 1, 2, \dots, \acute{n} \text{ and } \acute{\lambda}_1 > \acute{\lambda}_2, \dots, \acute{\lambda}_i, \dots, \acute{\lambda}_{\acute{n}-1} > \acute{\lambda}_{\acute{n}} \tag{12}$$

7. Calculate the difference of the sensitivity factor for every two consecutive IMF and find the index i corresponding to the maximum value of d_i as the final sensitive IMF:

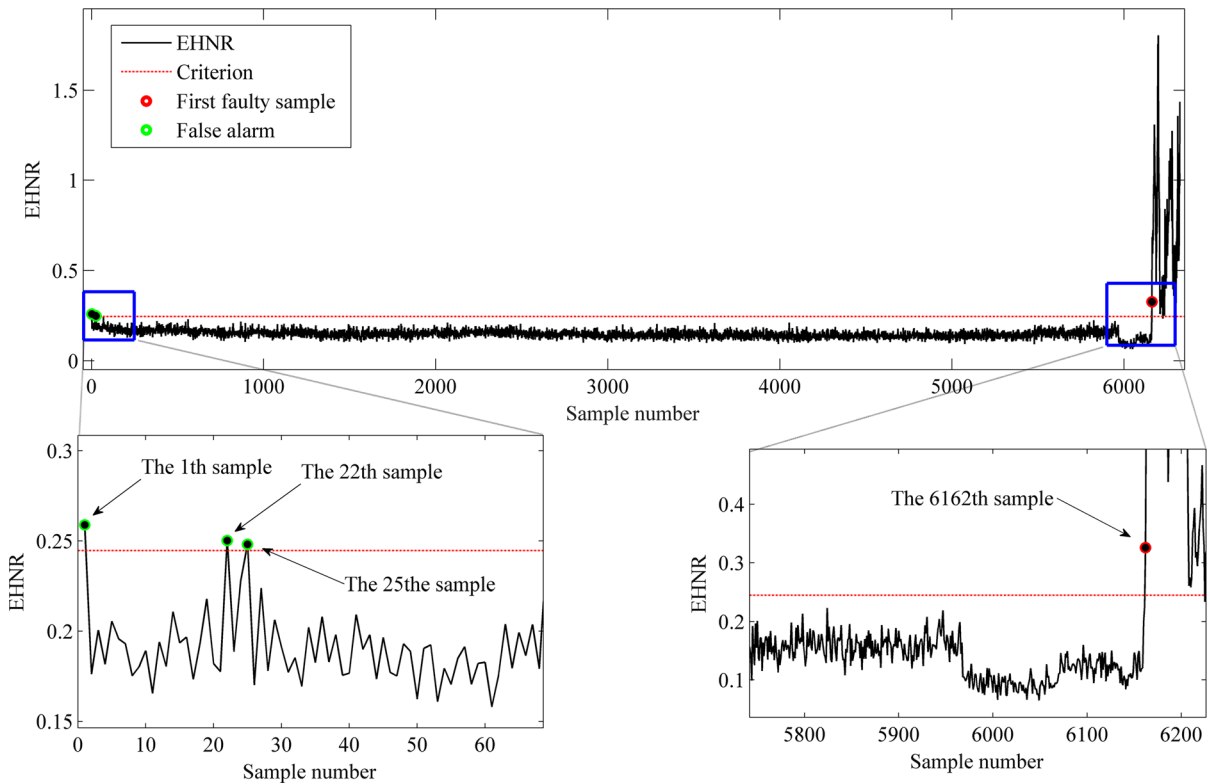


Fig. 6 The results of the original EHNr for Case 3

$$d_i = \hat{\lambda}_i - \hat{\lambda}_{i+1}, \quad i = 1, 2, \dots, \hat{n} \quad (13)$$

The results obtained by the new proposed approach for case studies 1, 2 and 3 are presented in Fig. 7.

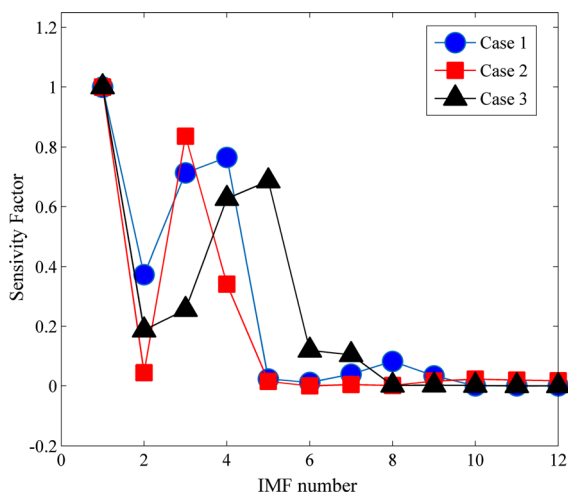


Fig. 7 Sensitivity rank of the IMFs of the vibration signal for all case studies

According to this figure, IMF1 with a significant difference for every three cases has the most sensitive factor. On the other hand, the components IMF4, IMF3, and IMF5 for Case 1, Case 2 and Case 3 are placed in the next ranks, respectively. As can be seen in Fig. 9, IMFs 6–12 in compare with IMF1 have the sensitivity factors with insignificant values and it can be concluded that these components are not useful for diagnosing the bearing fault. In other words, these components are irrelevant IMFs.

For evaluating the IMF selection technique proposed in this work, the original EHNr method is performed to the first three IMFs of the vibration signals of Case 2. The results have been shown in Fig. 8. As shown in Fig. 8a, by calculating the EHNr of the most sensitive IMF i.e., IMF1, the early fault can be identified in the 535th sample without any false alarm. According to Fig. 8b, many false alarms have appeared in the EHNr curve of the IMF2 in the healthy region, and this phenomenon complicates the fault diagnosis procedure. It can be concluded that the existence of noise in IMF2 leads to emerging false

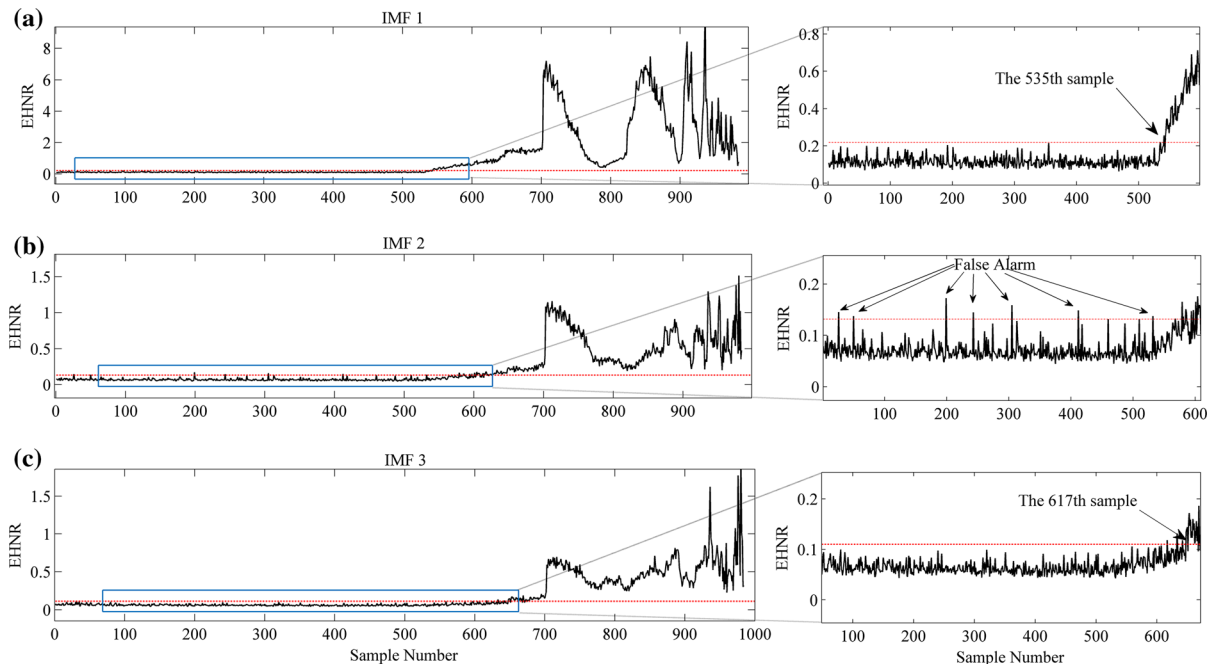


Fig. 8 The results of EHNHR for the three first IMFs of the vibration signal of Case 2

alarms in the EHNHR of IMF2. On the other hand, the investigation of the EHNHR for IMF3 shows that the first indications of the presence of defect are recognizable in the 617th sample (see Fig. 8c). The above results demonstrate that the EHNHR of the first IMF can detect the defect signatures easier than that of IMF3. In other words, IMF1 has the most information about the defect in comparison with other components and gives us a better estimation from the moment of the occurrence of the fault.

4.1.2 Early fault detection using FEEMD–EHNHR

The new hybrid incipient fault detection method, introduced in this subsection, is based on the FEEMD and EHNHR techniques. The implementation steps of the proposed approach called FEEMD-EHNHR are as follows:

1. Decompose the input signal into a series of IMFs by FEEMD.
2. Select the most sensitive IMF using the new approach proposed in Sect. 4.1.1.
3. Compute the EHNHR for the most appropriate IMF chosen in step 2.

4. Calculate the mean of the EHNHR, μ , and the standard deviation of the EHNHR, σ , in the normal working status of bearing. Compute the alarm threshold as $\mu + 4\sigma$ for identifying the early fault.
5. Consider the first intersection of the EHNHR curve with the alarm threshold value computed in the previous step as the moment of the fault occurrence.

The flowchart of the proposed method 1 (FEEMD-EHNHR) has been shown in Fig. 9.

The results of the FEEMD-EHNHR for Case 1, Case 2 and Case 3 are presented in Figs. 10, 8a and 11, respectively. According to Fig. 10, the first faulty sample determined by FEEMD-EHNHR for Case 1 is the 2120th sample. By investigating the results of the FEEMD-EHNHR and EHNHR in Figs. 10 and 5, respectively, it is observed that FEEMD-EHNHR can diagnose the incipient fault ten samples earlier than EHNHR (about 90 min). As illustrated in Fig. 8a, the alarm time recognized by the FEEMD-EHNHR approach for Case 2 is the 89th hour or the 535th sample. As demonstrated in Fig. 11 for Case 3, the 25th and 6079th measurements are reported as the degradation starting point by the FEEMD-EHNHR technique. On the other hand, the 25th sample is corresponding to the

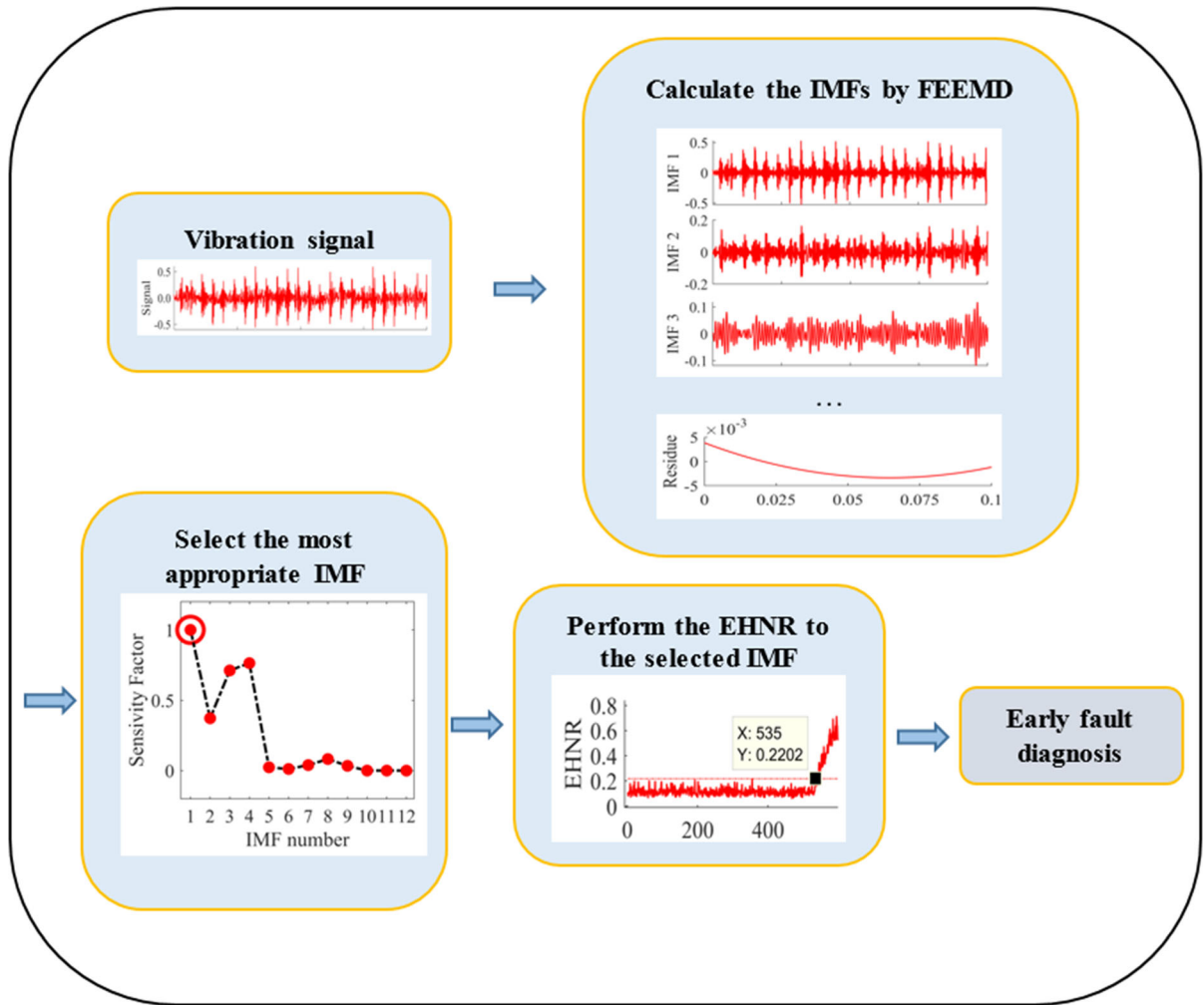


Fig. 9 The flowchart of the proposed method 1 (FEEMD–EHNR)

healthy zone. Therefore, the first faulty sample diagnosed by the proposed method 1 for Case 3 is the 6079th point. As can be seen, the FEEMD-EHNR diagnoses the fault starting point 83 samples (about 14 h) earlier than the original EHNR method (see Fig. 6). Also, the number of incorrect alerts that appeared in the FEEMD-EHNR curve is less than one in the original EHNR curve.

Figure 12 is presented for investigating the capabilities of the EHNR and FEEMD-EHNR methods to analyze the fault occurrence stage. For this purpose, in each of these curves, the points of the beginning and the end of the fault occurrence phase are connected by a line. The slope of these lines for the EHNR and the FEEMD-EHNR methods are 0.003195 and 0.009215, respectively. This result shows that the proposed

technique can indicate the change of the fault characteristics during the fault occurrence very well. Also, the result accuracy of the proposed approach is more than the EHNR method.

4.2 Proposed method 2: a new index based on the energy entropy of auto-correlation function

In this section, a novel health monitoring index or a new feature for the early fault diagnosis is proposed based on the autocorrelation function. The analysis of the auto-correlation is a powerful mathematical tool for finding the intermittent patterns such as the fault-induced impulses. ACF can be calculated by the following equation:

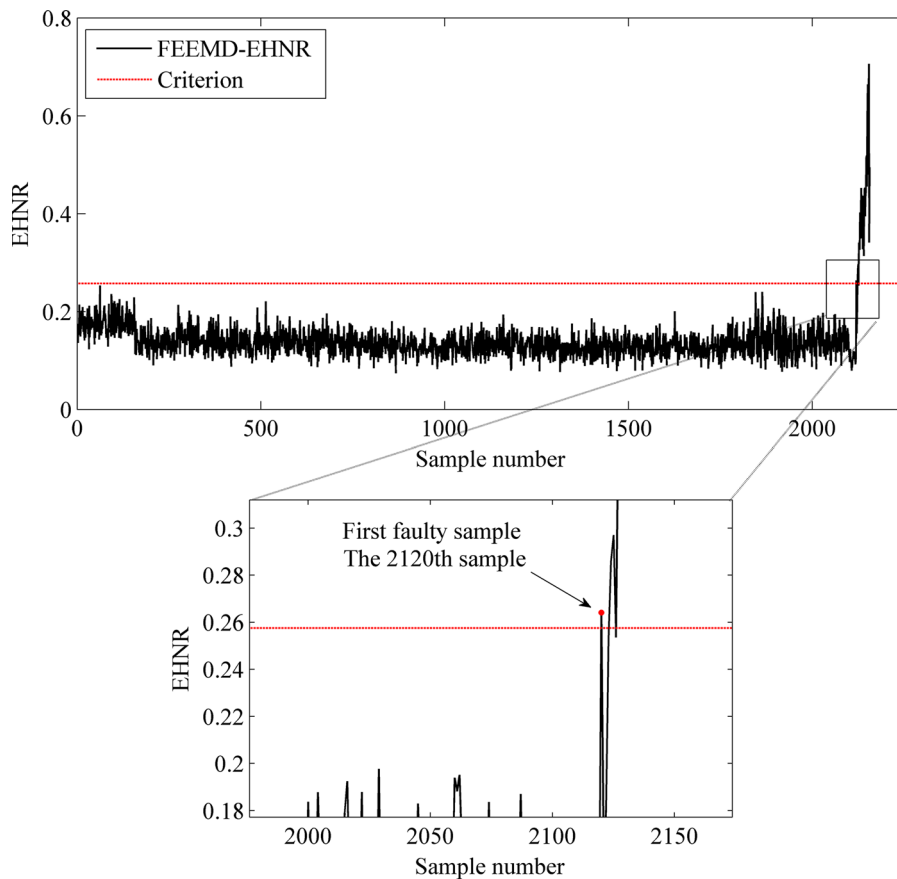


Fig. 10 The results of the FEEMD-EHNR for Case 1

$$\mu_i(t) = \int x_i(t)x_i(t + \tau)d\tau, \quad i = 1, \dots, N \quad (14)$$

where $x_i(t)$ is the vibration signal corresponding to the i -th sample and N is the number of samples.

According to the ACF definition, when the repetitive impulses appear in a vibration signal, the magnitude of the ACF of the signal is increased. The number of ACF maximum points for the samples of Case 2 is shown in Fig. 13. It can be seen that the number of the maximum points of the ACF corresponding to the faulty samples is more than that of the healthy samples. On the other hand, by increasing the fault size, the value and number of these maximum points have a relatively increasing behavior. As a result, the energy of the maximum points of ACF has the same treatment. The ACF energy of all the samples of Case 2 is illustrated in Fig. 14. As can be seen in this configuration, the total energy of the ACF has a relatively decreasing trend with the occurrence of the

defect and increasing its severity. Therefore, « the energy of the maximum points of the ACF » and « the total energy of the ACF » are the suitable factors for finding the early fault. In this section, a new early fault detection indicator is defined based on these factors.

The feature extraction presented in this section is a combination of the energy-entropy operator and the auto-correlation function and denoted as EEACF. The steps of the proposed algorithm are described as follows:

1. Calculate the autocorrelation of the original signal using Eq. (14).
2. Find the local maxima points of the ACF series obtained in the previous step. These points are denoted as $M_{ij}, j = 1, \dots, m$ where M_{ij} is the j th local maxima detected in the ACF of the i th sample and m is the number of the local maxima found in the ACF series.

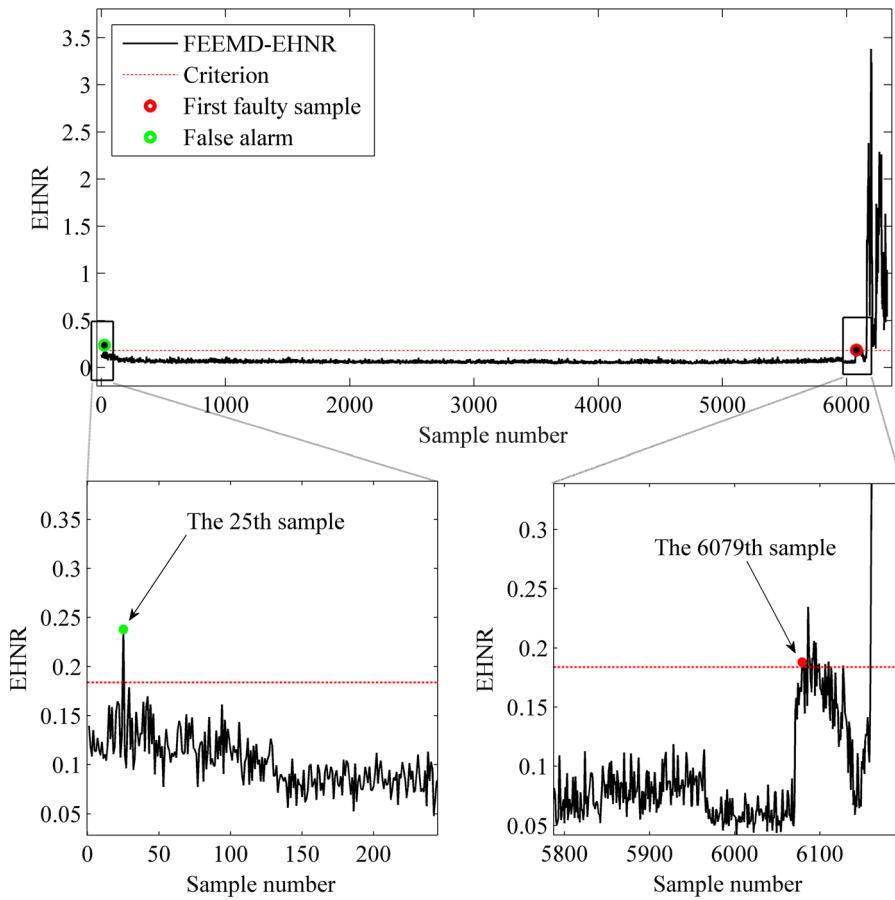


Fig. 11 The results of the FEEMD-EHNR for Case 3

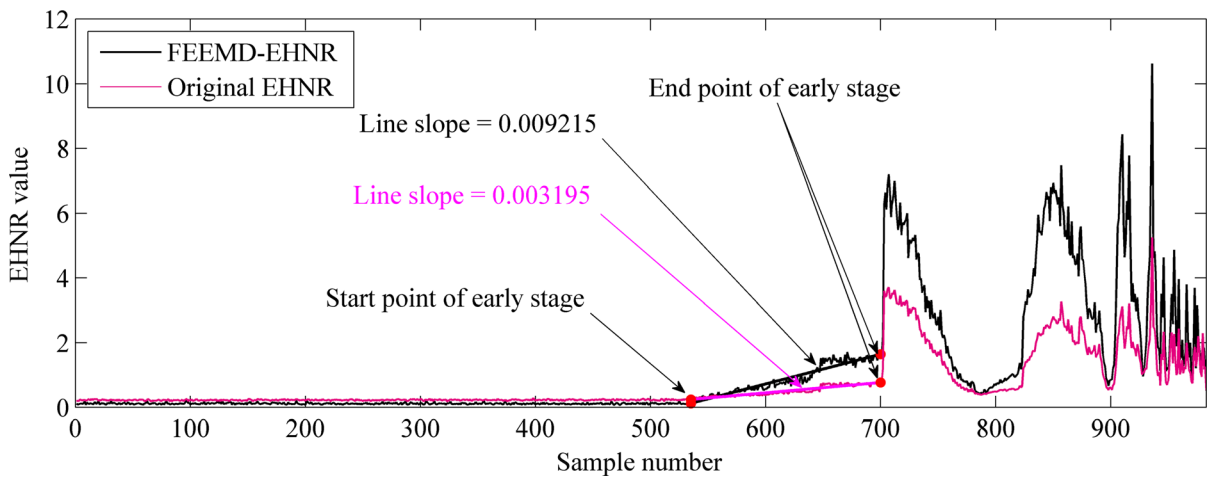


Fig. 12 The comparison of the FEEMD-EHNR and the original EHNR methods for analyzing the early stage for Case 2

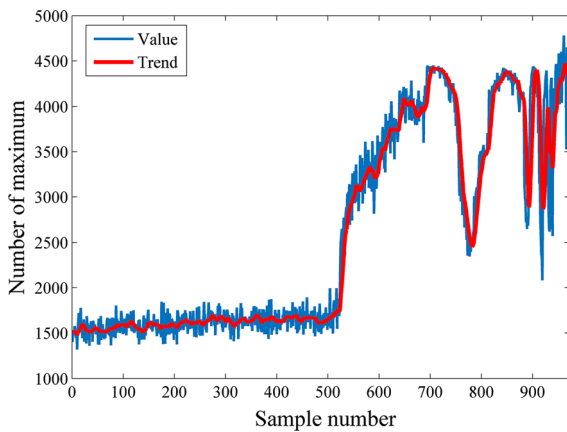


Fig. 13 The number of the maximum points of the ACF series for Case 2

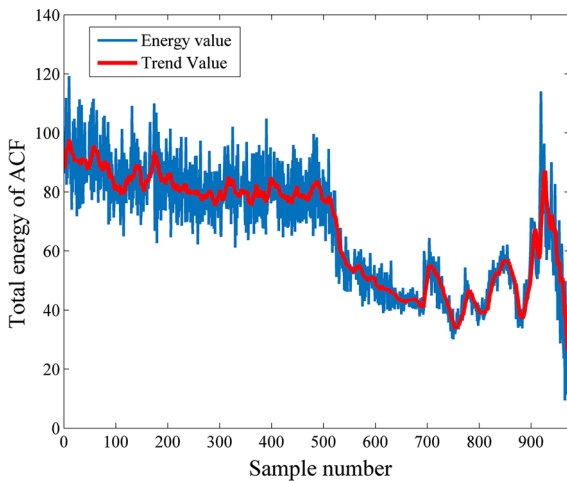


Fig. 14 The total energy of ACF series for Case 2

3. Calculate the following factor called the periodicity intensity factor (PIF):

$$\lambda_i = \frac{\sum_{j=1}^m M_{ij}^2}{\sum_{t=1}^T \mu_i(t)^2}, \quad i = 1, \dots, N \tag{15}$$

where $\sum_{j=1}^m M_{ij}^2$ is the energy of the ACF maximum points, $\sum_{t=1}^T \mu_i(t)^2$ is the total energy of the ACF series, T is the length of the ACF series and N is the total number of samples.

4. Compute the energy-entropy vector of the factor λ_i or EEACF as follows:

$$H_i = -\lambda_i * \log(\lambda_i), \quad i = 1, \dots, N \tag{16}$$

5. Compute the criterion value as $\mu + 4\sigma$ for distinguishing the fault occurrence instance. The parameters μ and σ are the mean and the standard deviation of the EEACF curve, respectively.

The results of the EEACF approach for Case 1, Case 2 and Case 3 have been shown in Figs. 15, 16 and 17, respectively. According to Fig. 15, the EEACF feature is ineffective in determining the moment of fault occurrence of the first case study. As can be seen in Table 3, Case 1 is corresponding to the defective inner race. Stack et al. [44] studied the effect of the fault location on the quality of appearing the fault-related features. They pointed out when a defect is created on the inner race, during each revolution of the shaft, this fault now rotates in and out of the load zone. In this instance, the strong fault signatures produced while the defect is in the load zone are averaged with the weaker signatures acquired while the defect is outside the load zone. This has the effect of attenuating the signatures of the faulty inner race.

As shown in Fig. 16, the alarm time recognized by EEACF for Case 2 is obviously 88 h and 40 min or the 533th sample. According to the results shown in Sects. 3.2 and 4.1.2, the EEACF technique can identify the incipient fault about 20 min, and 1 h and 30 min earlier than the FEEMD-EHNR and the original EHNR, respectively. As can be seen, the EEACF method can diagnose the incipient fault earlier than the other methods. For Case 3, according to Fig. 17, the first faulty sample identified by EEACF is the 6072th measurement. This sample is marked in green in Fig. 17. In this case study, in addition to the faulty sample, an incorrect alert has appeared in the 6037th measurement in the EEACF graph. In Sect. 4.3, the healthy state of the bearing at the 6037th sample will be investigated by analyzing the envelope spectrum of this sample. As mentioned in Sect. 4.1.2 (Fig. 11), a false alarm appeared at the 25th measurement in the FEEMD-EHNR curve. The false alarm displayed in the FEEMD-EHNR corresponds to the early moments of the experiment and the bearing is healthy at this moment. On the other hand, the false alarm appeared in the EEACF curve (the 6037th measurement) is close to the faulty sample **reported by the second proposed method (i.e., the 6072th sample). These results illustrate that even if the 6037th sample is reported as the fault starting point, the prediction error of the EEACF method is insignificant

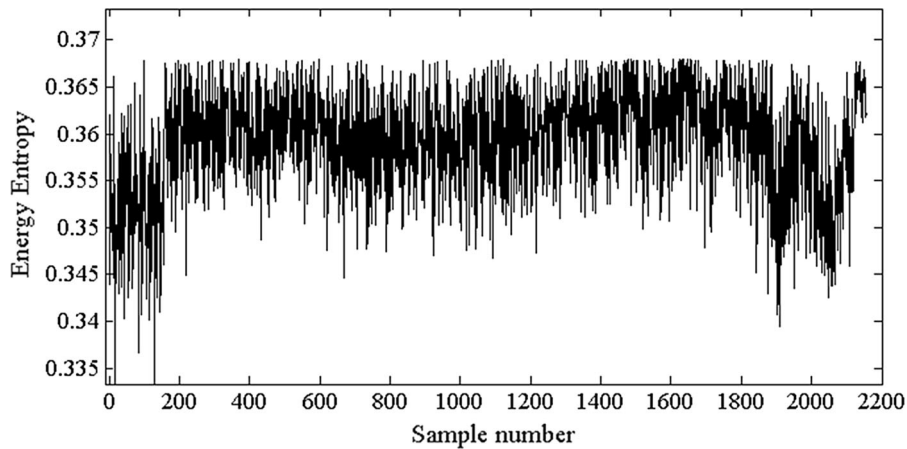


Fig. 15 The results of EEACF for Case 1

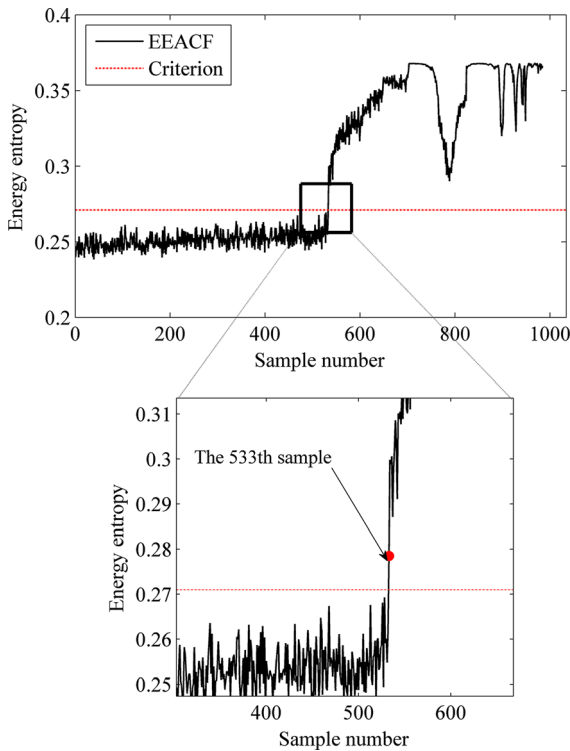


Fig. 16 The results of EEACF for Case 2

compared to the FEEMD-EHNR method. By looking at Figs. 6, 11 and 17, it can be observed that EEACF has been accomplished the early fault detection at about 1 and 15 h earlier than FEEMD-EHNR and EHNR, respectively.

Comparing the results of the EEACF method with the original EHNR and FEEMD-EHNR techniques

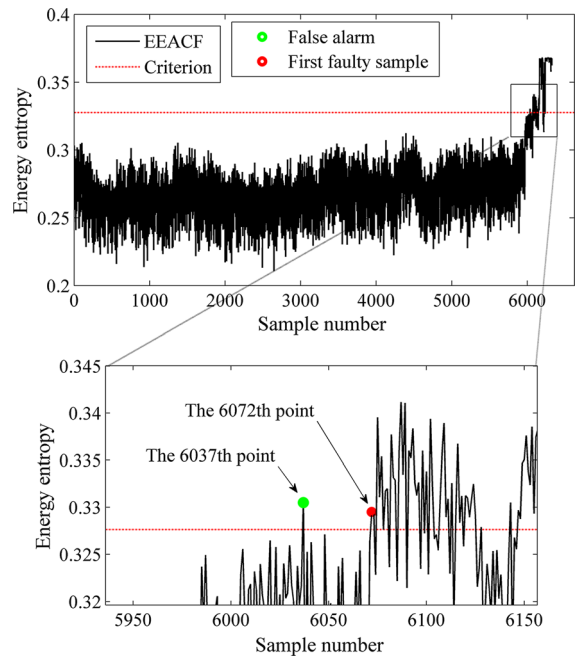


Fig. 17 The results of EEACF for Case 3

indicates that EEACF can detect the first faulty sample earlier than the other methods. It can be concluded that the proposed method has a high sensitivity to the changes caused by the defects with very small sizes. It can be clearly seen that the suggested feature in this paper can monitor the bearing conditions with the huge vibration data and the long run time. These results confirm the capability of the proposed approach.

4.3 Evaluation of the results using the hilbert envelope spectrum

In order to investigate the validation of the obtained results, the envelope spectrum of the vibration signals is used to extract the characteristic frequencies of the bearing fault. For this purpose, three samples of each bearing discussed in Table 3 have been selected. These samples relate to three situations: the last healthy sample, the first faulty sample, and the completely defective sample. The results of the Hilbert envelope spectrum for Case 1, Case 2 and Case 3 are shown in Figs. 18, 19 and 20, respectively. Figure 18a corresponds to the last healthy sample of Case 1, i.e., the 2119th sample. As shown in this figure, none of the characteristic frequencies related to the defected inner ring have appeared in the envelope spectrum. The envelope spectrum of this signal sample indicates that the bearing works in the healthy region. According to Subsection 4.1.2, the first faulty sample of Case 1 found by the proposed FEEMD-EHNR method is the 2120th point. The envelope spectrum of this sample is plotted in Fig. 18b. According to Fig. 18b, the rotational frequency, i.e., F_s , and its harmonics, the characteristic frequency of the inner ring, F_i , and its harmonics and the frequencies modulated with F_s in the envelope spectrum indicate that the 2120th sample corresponds to the moment of the occurrence of a small defect in the inner race. Figure 18c is the spectrum of a sample in the completely defective area. As can be seen, the spectrum amplitude of the characteristic frequencies in the entirely defective region is larger than the corresponding values of these frequencies at the moment of the defect occurrence.

The results of the EEACF method show that the last healthy sample and the first defective sample for Case 2 are the 532th and 533th measurements, respectively. In the spectrum of the 532th point in Fig. 19a, it can be seen that the frequency F_o is hardly detectable, and none of the harmonics of this frequency is observed. This result indicates that the bearing is healthy in the 532th specimen. On the other hand, appearing the characteristic frequency of F_o and its harmonics in Fig. 19b implies the fact that the moment of occurring the damage in the bearing corresponds to the 533th sample or 88 h and 40 min. In the completely faulty samples such as the 700th sample, the frequencies of

Fig. 18 The envelope spectrum of the three samples of Case 1, **a** the 2119th sample, **b** the 2120th sample, **c** the 2151th sample

the defective outer ring are much more clearly seen (see Fig. 19c).

The results of the envelope spectrum for Case 3 are presented in Fig. 20. As can be seen in Figs. 20a, b, emerging the characteristic frequency of the bearing and its harmonics in the spectrum of the sample 6072 and the absence of these frequencies in the sample 6071 indicate that the measurements of 6072 and 6071 are the first defective sample and the last healthy sample, respectively. It was seen in Fig. 17 that an alarm emerged at the 6037th sample in the EEACF feature vector calculated for Case 3. It was also claimed that there is no fault in the bearing in this measurement. Here, to prove this allegation, the envelope spectrum of the 6037th measurement is illustrated in Fig. 21. According to this illustration, none of the characteristic frequency and its harmonics have appeared. In other words, the bearing is healthy at this moment.

5 Comparison with other methods

This section is dedicated to comparing our proposed methodologies with previous researches that studied the run-to-failure case. For this purpose, the results reported by the authors in [28–34, 36, 37, 39] have been used. The summary of the fault diagnosis techniques used in these papers is described in Table 4. The results of the proposed methods and other techniques have been presented in Tables 5 and 6 for Case 1 and Case 2, respectively. Recognizing the slight degradation of bearing at the early stage is a criterion for evaluating these approaches. When a fault occurs in the inner ring of the bearing, it is difficult to detect its occurrence time [44]. Therefore, there are few studies that have investigated the detecting this defect type in the run-to-failure working conditions, and most researches are related to the defect diagnosis in the outer ring. On the other hand, the vibration data of the third case study used in this work have recently been updated. In the articles we have studied so far, no results have been provided for newly updated data of

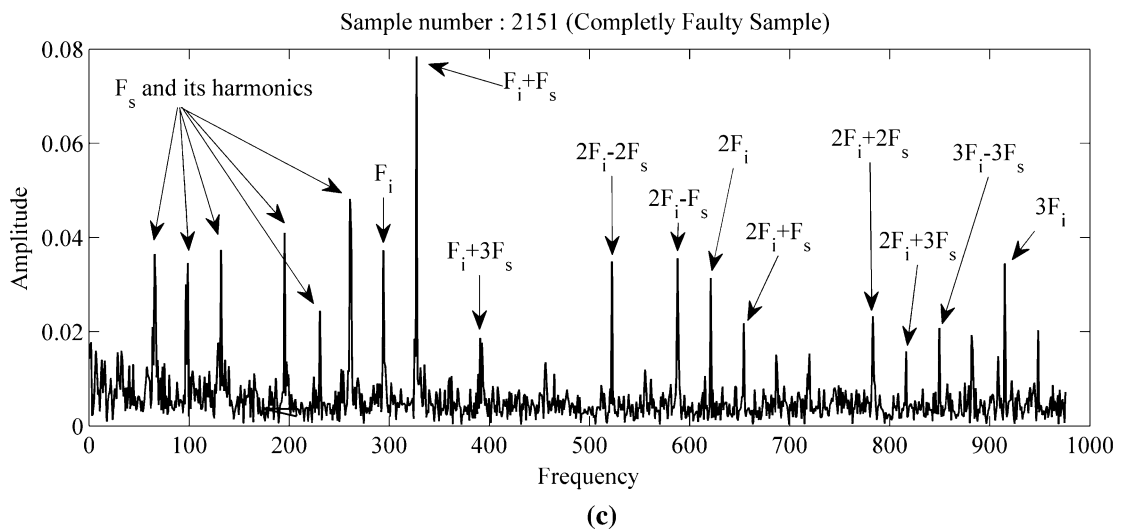
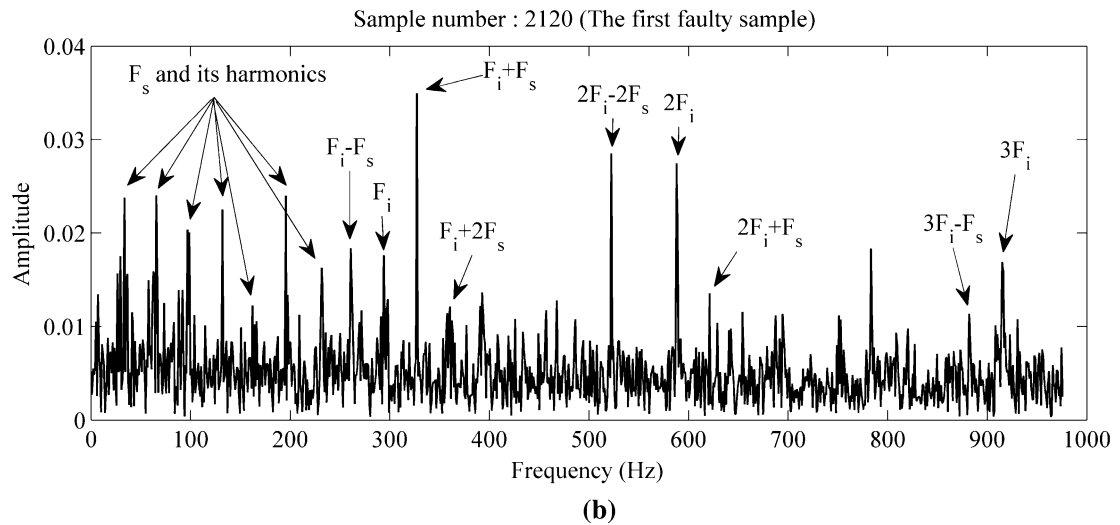
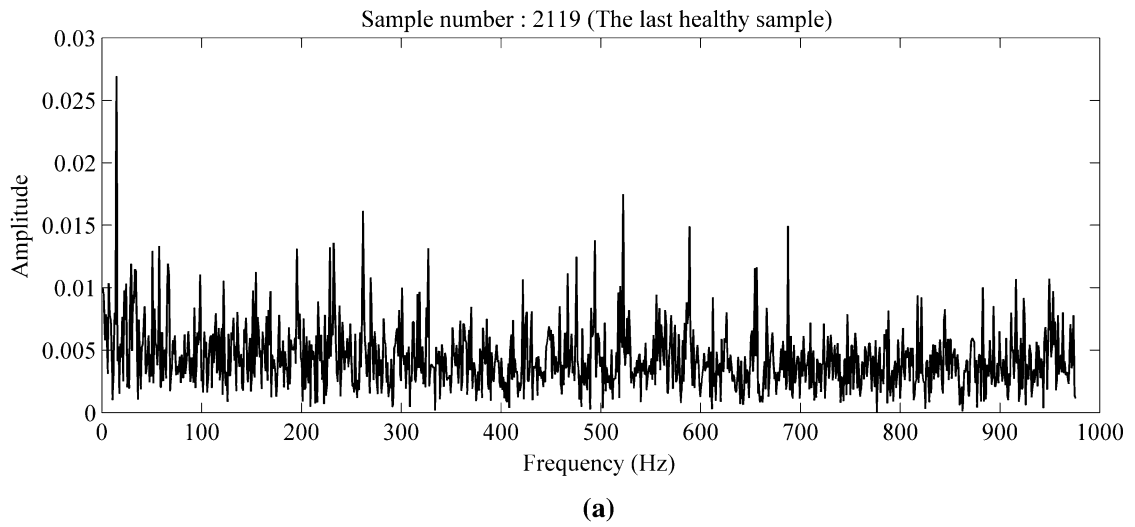
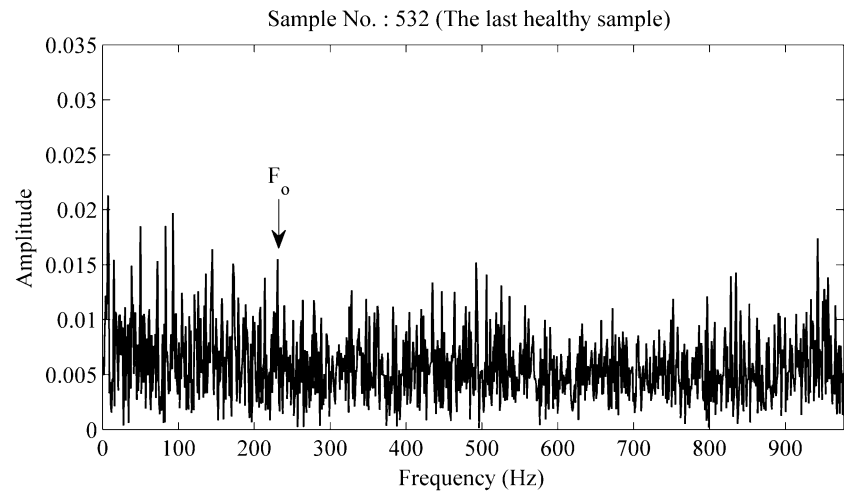
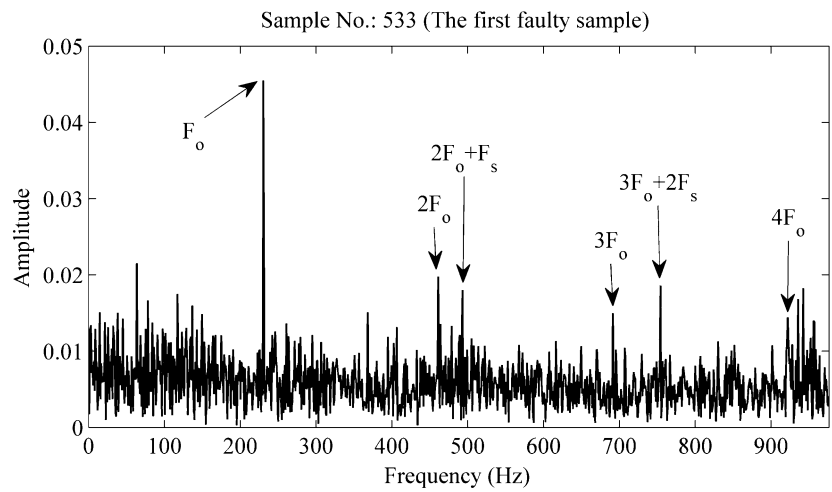


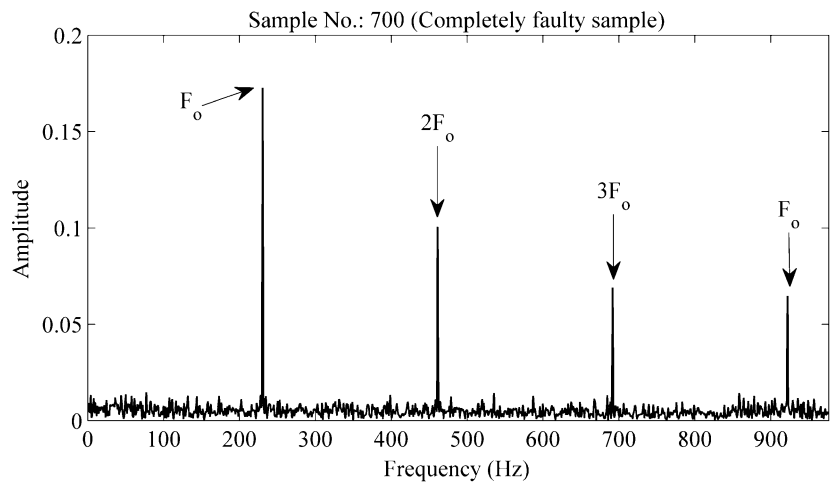
Fig. 19 The envelope spectrum of the three samples of Case 2, **a** the 532th sample, **b** the 533th sample, **c** the 700th sample



(a)



(b)



(c)

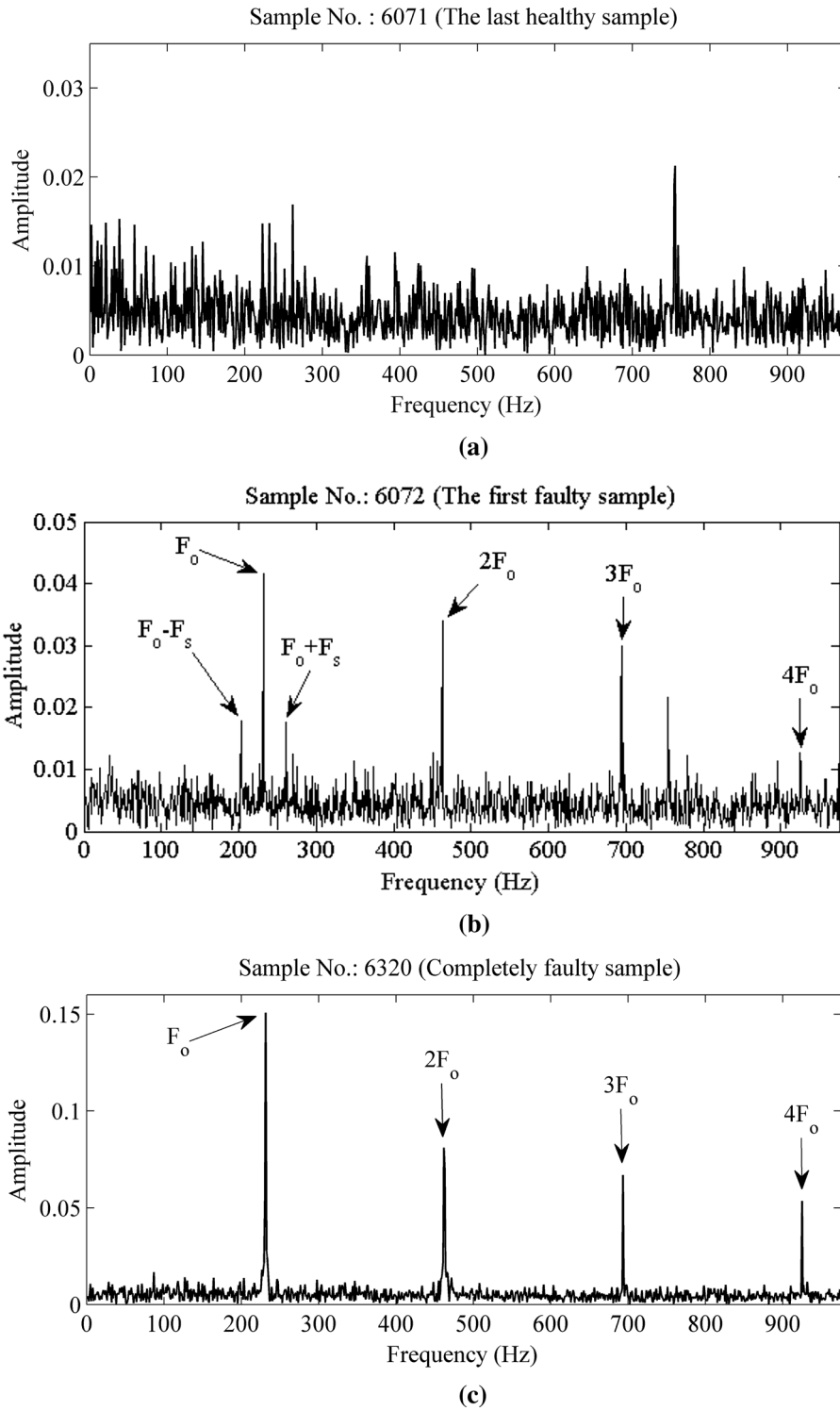


Fig. 20 The envelope spectrum of the three samples of Case 3, **a** the 6071th sample, **b** the 6072th sample, **c** the 6320th sample

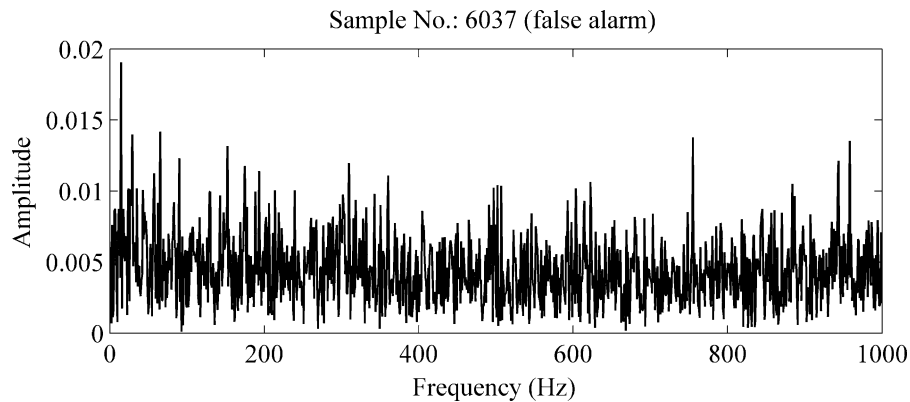


Fig. 21 The envelope spectrum of the 6037th sample appeared as the false alarm for Case 3

Table 4 The description of the proposed methods and the other fault diagnosis method

Method	Description of method
The proposed method 1	Select the most appropriate IMF using the new presented index + early fault detection using EHNR of the selected IMF (FEEMD-EHNR)
The proposed method 2	Diagnose the moment of the occurrence of the incipient fault by the auto-correlation function of the signal and energy entropy (EEACF)
Method 1 [28]	Extract the time and frequency domain features + find the appropriate low-dimensional information using local preserving projection (LPP) + recognize the slight degradation of the bearing using exponential weighted moving average statistic (EWMA)
Method 2 [29]	Apply principle component analysis (PCA) and dynamic PCA (DPCA) to extract the effective features from the vibration signal + distinguish the machine health degradation using hidden Markov model (HMM) and contribution analysis—based method
Method 3 [30]	Discriminant between normal and faulty condition using the combination of one-class ν -SVM and envelope analysis
Method 4 [31]	Drive time–frequency domain features from EMD energy entropy of the first eight IMFs and statistical measurement + classify the bearing defects by the artificial neural network (ANN)
Method 5 [32]	Identify the periodic impulses using the envelope harmonic-to-noise ratio (EHNR)
Method 6 [33]	Use the conventional sparse filter for detecting the impulsive signature
Method 7 [34]	Train a sparse auto-encoder to extract unsupervised features
Method 8 [36]	Detect the incipient fault using improve variational mode decomposition (IVMD)
Method 9 [37]	The amplitude level-based decomposition of the machine vibration signal + spectral analysis of the empirically determined local amplitude of the low-amplitude signal component
Method 10 [39]	Apply the recurrence quantification analysis (RQA) to extract the features from the vibration signal + predict the bearing degradation state using the Kalman filter

Case 3. So, no comparison is presented for Case 3 in this section.

The results reported in Table 5 are relevant to the diagnosis of the early defect created in the inner ring of bearing in accordance with Case 1. In the preliminary review of this table, it seems that the proposed techniques are more incapacitated than the other methods in detecting the degradation starting point.

The schemed approaches in this work determine the moment of the incipient fault at the 2120th measurement or 353 h 20 min. Also, to prove this allegation, the envelope spectrums of some instances that are corresponding to the last healthy sample (the 2119th measurement) and the first defective measurement (the 2120th measurement) were computed and plotted in Fig. 18. Existing the characteristic frequencies in

Table 5 Comparison between the previous researches and the presented works for Case 1

Methods	Time of fault starting point
Method 1 [28]	298 h 20 min
Method 2 [29]	353 h 20 min
Method 4 [31]	260 h 10 min
Method 10 [39]	305 h 30 min
The proposed method 1	353 h 20 min
The proposed method 2	353 h 20 min

Table 6 Comparison between the previous researches and the presented works for Case 2

Methods	Time of fault starting point
Method 2 [29] (HMM-DPCA)	89 h 30 min
Method 2 [29] (HMM-PCA)	89 h 40 min
Method 3 [30]	89 h
Method 4 [31]	92 h 10 min
Method 5 [32]	90 h 30 min
Method 6 [33]	90 h 30 min
Method 7 [34]	91 h 10 min
Method 8 [36]	88 h 30 min
Method 9 [37]	91 h 40 min
Method 10 [39]	90 h 40 min
The proposed method 1	89 h
The proposed method 2	88 h 40 min

the obtained spectrum confirms the accuracy of these results. On the other hand, the authors in articles such as Method 1 [28], Method 4 [31] and Method 10 [39] have not provided any evidence to prove their results regarding the detection of the degradation starting point time for Case 1. Here, in order to demonstrate that the samples reported by these methods are not the first faulty samples, the spectra of their reported samples are shown in Fig. 22. It is necessary to point out that in order to eliminate the effects of noise in the spectrum of these samples, the most sensitive IMF of each signal is determined by using the method presented in Sect. 4.1.1, and then its spectrum is calculated. As shown in Fig. 22, none of the characteristic frequencies of the damaged inner ring of the bearing 3 for Case 1 are clearly recognizable. These

observations indicate that the samples introduced by [28, 31, 39] are in the healthy zone.

The results of the prediction of the degradation starting point in bearing 1 of Case 2 are reported in Table 6. These values have been determined by the different techniques and the presented method. It can be seen that most of these results are close together. On the other hand, the suggested approach 2 and the technique presented in [36] are superior to other methods in the fault starting point detection. Of course, it should be noted that the spectrum provided for Case 2 in Fig. 19 confirms the accuracy of the result reported by the proposed method.

Consequently, the results show that the suggested approaches in this paper are suitable and authoritative in the online bearing fault detection. Also, the results presented for Case 1 indicate that the techniques proposed in this paper are effectively able to identify the moment of degradation occurring in the inner ring of the bearing.

6 Conclusion

The main target of this paper was to provide the powerful methods for identifying the starting point of the defects and the assessment of the damage severity in bearings. For this aim, two novel approaches were suggested in this paper. In the first method, the EHNR of the most efficient IMF obtained by the FEEMD algorithm was computed for revealing the signatures of the degradation starting point. In this approach, a new indicator was defined based on the auto-correlation of both the original signal and its IMFs for picking out the most informative IMF. In the second proposed method, a new feature was defined which was a combination of the energy of maximum points of the raw signal autocorrelation and the energy-entropy vector. The vibration signals in the run-to-failure test corresponding to the defective outer race and the defective inner race were used in order to check the ability of the presented works. The results indicated that the presented techniques in this paper can effectively identify faults in their early stages of development. It was found that most of the methods presented in previous studies were ineffective in detecting the fault in the inner race. The results of this paper imply the capability of the presented approaches to identify the degradation moment in

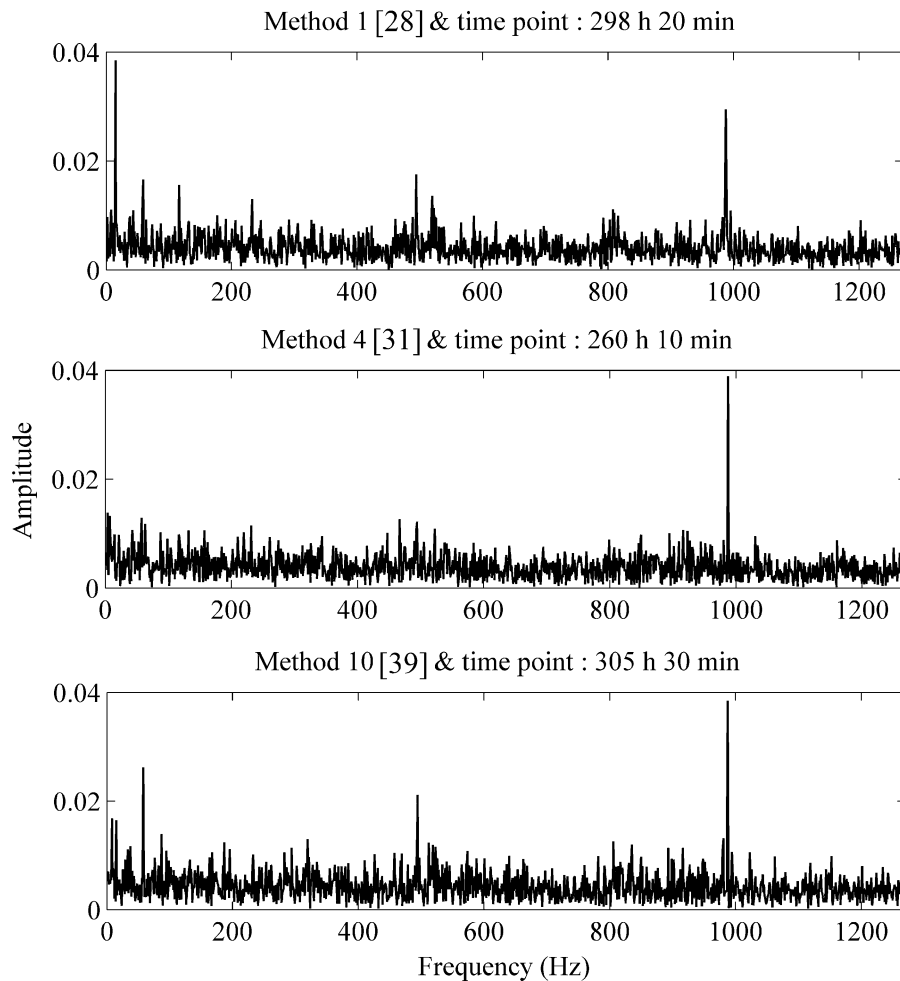


Fig. 22 The envelope spectrum of the measurement reported by other methods as fault starting point for Case 1

the bearing inner race. The envelope spectrums of the last healthy sample and the first faulty sample determined by the proposed techniques proved the accuracy of the results. Also, the results illustrated that the presented methods were more precise than the other methods proposed in the recent years.

The Remaining Useful Life (RUL) can be predicted by determining the moment of degradation on the various components of the bearing. It seems that combination of the indicators introduced in this article and the Prognostics and Health Management (PHM) techniques can be used for diagnosing the bearing statues and estimating its RUL.

Authors' contribution SNC and MJHM designed and coordinated the study. SNC wrote the manuscript. SNC and

AB reviewed the manuscript and contributed to its revision. All the authors discussed the results and gave their final approved for publication.

Funding This study was financially supported by the University of Guilan to S.N.C. The funder had no role in study design, data collection, and analysis, decision to publish or preparation of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

1. Li C, Liang M (2012) Continuous-scale mathematical morphology-based optimal scale band demodulation of

- impulsive feature for bearing defect diagnosis. *J Sound Vib* 331:5864–5879. <https://doi.org/10.1016/j.jsv.2012.07.045>
2. Meng L, Xiang J, Wang Y, Jiang Y, Gao H (2015) A hybrid fault diagnosis method using morphological filter–translation invariant wavelet and improved ensemble empirical mode decomposition. *Mech Syst Signal Process* 50:101–115. <https://doi.org/10.1016/j.ymssp.2014.06.004>
 3. Liu X, Bo L, He X, Veidt M (2012) Application of correlation matching for automatic bearing fault diagnosis. *J Sound Vib* 331:5838–5852. <https://doi.org/10.1016/j.jsv.2012.07.022>
 4. Jiang F, Zhu Z, Li W, Chen G, Zhou G (2013) Robust condition monitoring and fault diagnosis of rolling element bearings using improved EEMD and statistical features. *Meas Sci Technol* 25:025003. <https://doi.org/10.1088/0957-0233/25/2/025003>
 5. McFadden PD, Smith JD (1984) Vibration monitoring of rolling element bearings by the high-frequency resonance technique—a review. *Tribol Int* 17:3–10. [https://doi.org/10.1016/0301-679X\(84\)90076-8](https://doi.org/10.1016/0301-679X(84)90076-8)
 6. Pan J, Chen J, Zi Y, Li Y, He Z (2016) Mono-component feature extraction for mechanical fault diagnosis using modified empirical wavelet transform via data-driven adaptive Fourier spectrum segment. *Mech Syst Signal Process* 72:160–183. <https://doi.org/10.1016/j.ymssp.2015.10.017>
 7. Li Z, Yan X, Tian Z, Yuan C, Peng Z, Li L (2013) Blind vibration component separation and nonlinear feature extraction applied to the nonstationary vibration signals for the gearbox multi-fault diagnosis. *Measurement* 46:259–271. <https://doi.org/10.1016/j.measurement.2012.06.013>
 8. Yuan J, He Z, Ni J, Brzezinski AJ, Zi Y (2013) Ensemble noise-reconstructed empirical mode decomposition for mechanical fault detection. *J Vib Acoust* 135:021011
 9. Mallat SG (1989) A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans Pattern Anal Mach Intell* 11:674–693. <https://doi.org/10.1109/34.192463>
 10. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen NC, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc R Soc Lond A Math Phys Sci* 454:903–995. <https://doi.org/10.1098/rspa.1998.0193>
 11. Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv Adapt Data Ana* 1:1–41. <https://doi.org/10.1142/S1793536909000047>
 12. Gilles J (2013) Empirical wavelet transform. *IEEE Trans Signal Processing* 61:3999–4010. <https://doi.org/10.1109/TSP.2013.2265222>
 13. Dragomiretskiy K, Zosso D (2014) Variational mode decomposition. *IEEE Trans Signal Process* 62:531–544. <https://doi.org/10.1109/TSP.2013.2288675>
 14. Li Y, Xu M, Wei Y, Huang W (2015) An improvement EMD method based on the optimized rational Hermite interpolation approach and its application to gear fault diagnosis. *Measurement* 63:330–345. <https://doi.org/10.1016/j.measurement.2014.12.021>
 15. Xue X, Zhou J, Xu Y, Zhu W, Li C (2015) An adaptively fast ensemble empirical mode decomposition method and its applications to rolling element bearing fault diagnosis. *Mech Syst Signal Process* 62:444–459. <https://doi.org/10.1016/j.ymssp.2015.03.002>
 16. Guo T, Deng Z (2017) An improved EMD method based on the multi-objective optimization and its application to fault feature extraction of rolling bearing. *Appl Acoust* 127:46–62. <https://doi.org/10.1016/j.apacoust.2017.05.018>
 17. Wang CS, Sha CY, Su M, Hu YK (2017) An algorithm to remove noise from locomotive bearing vibration signal based on self-adaptive EEMD filter. *J Cent South Univ* 24:478–488. <https://doi.org/10.1007/s11771-017-3450-8>
 18. Wei Z, Wang Y, He S, Bao J (2017) A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection. *Knowl Based Syst* 116:1–12. <https://doi.org/10.1016/j.knsys.2016.10.022>
 19. Abdelkader R, Kaddour A, Derouiche Z (2018) Enhancement of rolling bearing fault diagnosis based on improvement of empirical mode decomposition denoising method. *Int J Adv Manuf Technol* 97:1–19. <https://doi.org/10.1007/s00170-018-2167-7>
 20. Cao H, Fan F, Zhou K, He Z (2016) Wheel-bearing fault diagnosis of trains using empirical wavelet transform. *Measurement* 82:439–449. <https://doi.org/10.1016/j.measurement.2016.01.023>
 21. Kedadouche M, Liu Z, Vu VH (2016) A new approach based on OMA-empirical wavelet transforms for bearing fault diagnosis. *Measurement* 90:292–308. <https://doi.org/10.1016/j.measurement.2016.04.069>
 22. Zhang M, Jiang Z, Feng K (2017) Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump. *Mech Syst Signal Process* 93:460–493. <https://doi.org/10.1016/j.ymssp.2017.02.013>
 23. Cocconcelli M, Zimroz R, Rubini R, Bartelmus W (2012) STFT based approach for ball bearing fault detection in a varying speed motor. *Condition Monitoring of Machinery in Non-Stationary Operations* 41–50. Springer, Berlin. https://doi.org/10.1007/978-3-642-28768-8_5
 24. Bellini A, Cocconcelli M, Immovilli F, Rubini R (2008) Diagnosis of mechanical faults by spectral kurtosis energy. In: 2008 34th Annual conference of IEEE industrial electronics. IEEE, pp 3079–3083. <http://doi.org/10.1109/IECON.2008.4758452>
 25. Montechiesi L, Cocconcelli M, Rubini R (2016) Artificial immune system via Euclidean Distance Minimization for anomaly detection in bearings. *Mech Syst Signal Process* 76:380–393. <https://doi.org/10.1016/j.ymssp.2015.04.017>
 26. Lee J, Qiu H, Yu G, Lin J (2007) Rexnord Technical Services, ‘Bearing Data Set’, IMS, University of Cincinnati, NASA Ames Prognostics Data Repository. <https://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/#bearing>
 27. Qiu H, Lee J, Lin J, Yu G (2006) Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *J Sound Vib* 289:1066–1090. <https://doi.org/10.1016/j.jsv.2005.03.007>
 28. Yu JB (2011) Bearing performance degradation assessment using locality preserving projections. *Expert Syst Appl* 38:7440–7450. <https://doi.org/10.1016/j.eswa.2010.12.079>

29. Yu J (2012) Health condition monitoring of machines based on hidden Markov model and contribution analysis. *IEEE Trans Instrum Meas* 61:2200–2211. <https://doi.org/10.1109/TIM.2012.2184015>
30. Fernández-Francos D, Martínez-Rego D, Fontenla-Romero O, Alonso-Betanzos A (2013) Automatic bearing fault diagnosis based on one-class v-SVM. *Comput Ind Eng* 64:357–365. <https://doi.org/10.1016/j.cie.2012.10.013>
31. Ali JB, Fnaiech N, Saidi L, Chebel-Morello B, Fnaiech F (2015) Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. *Appl Acoust* 89:16–27. <https://doi.org/10.1016/j.apacoust.2014.08.016>
32. Xu X, Zhao M, Lin J, Lei Y (2016) Envelope harmonic-to-noise ratio for periodic impulses detection and its application to bearing diagnosis. *Measurement* 91:385–397. <https://doi.org/10.1016/j.measurement.2016.05.073>
33. Jia X, Zhao M, Di Y, Jin C, Lee J (2017) Investigation on the kurtosis filter and the derivation of convolutional sparse filter for impulsive signature enhancement. *J Sound Vib* 386:433–448. <https://doi.org/10.1016/j.jsv.2016.10.005>
34. Hasani RM, Wang G, Grosu R (2017) An Automated Auto-encoder Correlation-based Health-Monitoring and Prognostic Method for Machine Bearings. *arXiv preprint arXiv 1703.06272*
35. Li Y, Liang X, Xu M, Huang W (2017) Early fault feature extraction of rolling bearing based on ICD and tunable Q-factor wavelet transform. *Mech Syst Signal Process* 86:204–223. <https://doi.org/10.1016/j.ymssp.2016.10.013>
36. Jiang F, Zhu Z, Li W (2018) An improved VMD with empirical mode decomposition and its application in incipient fault detection of rolling bearing. *IEEE ACCESS* 6:44483–44493. <https://doi.org/10.1109/ACCESS.2018.2851374>
37. Dybała J (2018) Diagnosing of rolling-element bearings using amplitude level-based decomposition of machine vibration signal. *Measurement* 126:143–155. <https://doi.org/10.1016/j.measurement.2018.05.031>
38. Lv Y, Yuan R, Wang T, Li H, Song G (2018) Health degradation monitoring and early fault diagnosis of a rolling bearing based on CEEMDAN and improved MMSE. *MTRL* 11:1009. <https://doi.org/10.3390/mtrl11061009>
39. Qian Y, Yan R, Hu S (2014) Bearing degradation evaluation using recurrence quantification analysis and Kalman filter. *IEEE Trans Instrum Meas* 63:2599–2610. <https://doi.org/10.1109/TIM.2014.2313034>
40. Wang J, Xu G, Zhang Q, Liang L (2009) Application of improved morphological filter to the extraction of impulsive attenuation signals. *Mech Syst Signal Process* 23:236–245. <https://doi.org/10.1016/j.ymssp.2008.03.012>
41. Zhang J, Ma W, Lin J, Ma L, Jia X (2015) Fault diagnosis approach for rotating machinery based on dynamic model and computational intelligence. *Measurement* 59:73–87. <https://doi.org/10.1016/j.measurement.2014.09.045>
42. Tabrizi A, Garibaldi L, Fasana A, Marchesiello S (2015) Early damage detection of roller bearings using wavelet packet decomposition, ensemble empirical mode decomposition and support vector machine. *Meccanica* 50:865–874. <https://doi.org/10.1007/s11012-014-9968-z>
43. Lei Y, Zuo MJ (2009) Fault diagnosis of rotating machinery using an improved HHT based on EEMD and sensitive IMFs. *Meas Sci Technol* 20:125701. <https://doi.org/10.1088/0957-0233/20/12/125701>
44. Stack JR, Habetler TG, Harley RG (2006) Fault-signature modeling and detection of inner-race bearing faults. *IEEE Trans Ind Appl* 42(1):61–68. <https://doi.org/10.1109/TIA.2005.861365>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.