



An improved envelope detection method using particle swarm optimisation for rolling element bearing fault diagnosis

Sunil Tyagi, S.K. Panigrahi *

Department of Mechanical Engineering, Defense Institute of Advance Technology, Girinagar, Pune 411025, India

ARTICLE INFO

Article history:

Received 25 February 2017

Received in revised form 15 May 2017

Accepted 17 May 2017

Available online 22 May 2017

Keywords:

Envelope Detection (ED)

Particle Swarm Optimisation (PSO)

Hilbert transform

Rolling element bearing

Fault diagnosis

High Frequency Resonant Technique (HFRT)

ABSTRACT

Traditionally Envelope Detection (ED) is implemented for detection of rolling element bearing faults by extracting the envelope of band-passed vibration signal and thereafter taking its Fourier transform. The performance of ED is highly sensitive to the envelope window (i.e. central frequency and bandwidth of the passband). This paper employs Particle Swarm Optimisation (PSO) to select the most optimum envelope window to band pass the vibration signals emanating from rotating driveline that was run in normal and with faults induced rolling element bearings. The envelopes of band-passed signals were extracted with the help of Hilbert Transform. The performance of ED whose envelope window was optimised by PSO to identify various commonly occurring bearing faults such as bearing with Outer Race Fault (ORF), Inner Race Fault (IRF) and Rolling Element Fault (REF) were checked under varying load conditions. The performance of 'ED enhanced by PSO' was also checked with increase in the severity of defect. It was shown that the improved ED method is successfully able to identify all types of bearing faults under different load conditions. It was shown that the by selecting envelope window by PSO makes ED especially useful to identify bearing faults at the incipient stage of defect. It was also shown by presenting comparative performance that by optimising the envelope window by PSO the performance of ED gets significantly enhanced in comparison to the traditional ED method for bearing fault diagnosis.

© 2017 Society for Computational Design and Engineering. Publishing Services by Elsevier. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Rolling element bearings or ball/roller bearings are most common machine component of any rotating machinery used in industry. Rolling element bearings endure heavy loads under industrial operating conditions and structural faults, such as wear, pitting, or spall may occur after a long period of running (Randall & Antoni, 2011). A defect in bearing if remain undetected would cause its break down and may ultimately cause catastrophic failure of the machinery. Therefore early detection of bearing defects is most important for effective condition monitoring of industrial plant.

Changes in vibration signals are often an indication that fault is developing in machinery. Measuring the vibration of machine with the help of accelerometer mounted on the bearing housing and monitoring changes in characteristic defect frequency is an established method for machinery fault assessment. However, it is difficult to identify the bearing defect frequency in vibration spectrum

as common bearing defects are pits or spalls located at outer race, inner race and on the rolling element. These defects generate a series of impacts as rolling element passes over the defect due to the metal to metal contact. The resultant vibrations in time domain are sharp peaky bursts. It is difficult to identify the defect frequency in spectrum as these impact vibrations distributes their energy over wide range of frequencies; the bearing's characteristic defect frequency contains low energy (Norton & Karczub, 2003) and hence it get masked by broadband noise and other low frequency effects.

Simple frequency domain or time domain analysis of raw vibration signal is unable to detect the ball bearing faults (Mathew, 1984). Various researchers have tried to solve this problem using wide range of techniques such as spectral kurtosis (Antoni, 2007), improving the SNR by use of adaptive noise cancelling technique (Chaturvedi & Thomas, 1982). Wavelets have also been used for bearing fault diagnosis due to its ability of generating multi resolution analysis (Peng & Chu, 2004; Peter & Wang, 2013; Prabhakar, Mohanty, & Sekhar, 2002). However, Envelope Detection (ED) also known as High Frequency Resonant Technique (HFRT) has been used most successfully for bearing fault detection over the years because the envelope signals provide much more diagnostic insight than the raw signals (Tandon & Choudhury, 1999). A reviews of Envelope Detection have been presented by

Peer review under responsibility of Society for Computational Design and Engineering.

* Corresponding author.

E-mail address: panigrahisk@diat.ac.in (S.K. Panigrahi).

Nomenclature

φ	contact angle	n	Number of balls
$a(t)$	envelope signal	P_d	Pitch diameter
B_d	ball diameter	p_k^g	position of the particle with best global fitness at current move k
BPFI	Ball Pass Frequency Inner Race	p^i	best position of particle i in current and all previous moves
BPFO	Ball Pass Frequency Outer Race	rand	uniformly distributed rand variable between 0 and 1
BPFR	Pass Frequency Rolling Element	v_k	position of particle i in the design space at time k
$c1$	self confidence factor	w	inertia factor
$c2$	swarm confidence factor	x_k	position of particle i in the design space at time k
ED	Envelope Detection	$x(t)$	acquired vibration signal
f	fitness function	$\hat{x}(t)$	Hilbert transform of $x(t)$
g	constraint function	$\tilde{x}(t)$	analytic signal
$H(\omega)$	frequency response of Hilbert transform		
HFRT	High Frequency Resonant Technique		

Randall and Antoni (2011) and McFadden and Smith (1984). Every time a rotating elements passes over the defect on mating element, a pulse of short duration is generated that excites the resonances. This excitation happens periodically at the characteristic frequency that is related to the location of the defect. The resultant vibration is and amplitude modulated signal with resonant frequency as the carrier frequency whose amplitude is modulated at the characteristic defect frequency. By demodulating at one of these resonances, a signal indicative of the bearing condition can be obtained.

The ED technique involved bandpassing the vibration signal while keeping the central frequency of passband at one of the resonance and using the filter of bandwidth equal to 3 to 4 times the highest characteristic defect frequency of the bearing (McFadden and Smith (1984). Although the ED has been proved to be an effective technique for detection of bearing faults but its performance degrade if the selected central frequency and bandwidth of band-pass filter is not optimum. It has been shown by Bechhoefer and Menon (2009) that instinctively placing the envelope band at the resonant frequency is fraught with danger of misidentification of bearing fault and may lead to catastrophic failure. A selective envelope detection method that places the envelope window outside of the structural resonance has been reported in Currech (2000), Barkov, Barkova, and Mitchell (1995). A spectral kurtosis method for selection of optimum envelope window was presented in Bechhoefer, Menon, and Kingsley (2011).

Particle Swarm Optimization (PSO) is a technique used to explore the search space of a given problem to find the settings or parameters required to maximize a particular objective (Blondin, 2009). Working of PSO is inspired by the collaborative behavior of bird flock or school of fish (Poli, Kennedy, & Blackwell, 2007). Ever since its introduction PSO has been enormously successful. In past two decades many successful applications of PSO have been reported from disparate and wide ranging fields such as Biological and pharmaceutical applications (Selvan et al., 2006), bandwidth and channel allocation in communication system (Zhang, Hu, & Zhu, 2010), clustering and classification (Huang, 2011), control (Gaing, 2004), power system optimization (AlRashidi & El-Hawary, 2009). Recently various researchers have presented hybrid methods where PSO was combined with other soft computing techniques for diagnosis of machinery faults (Liu, Cao, Chen, He, & Shen, 2013; Quan, Yongqian, & Yongping, 2012; Sun, Huang, & Huang, 2013; Xiang, Huang, Huang, & Hu, 2006). Inspiration behind this paper is to present a Particle Swarm Optimization based method that will enable selection of optimum envelope window. That will enable diagnosis of bearing fault in the spectrum of the enveloped vibration signal, with high success. It is shown in this paper how PSO optimizes the performance of ED

by selecting the most optimum envelope window (i.e. the central frequency and bandwidth of band-pass filter) to band-pass the vibration signal prior to extraction of its envelope.

Experimental setup included an electric motor rotating a drive-line that was connected to dynamometer and rested on ball bearings. Faults of varying size were artificially introduced in outer race, inner race and rolling element of bearings and vibration data were collected at various loads. It has been shown that the ED improved by PSO is able to perform better than the traditional ED method and it was also shown that the improved ED method is able to detect all three types the bearing faults at various stages of defect.

2. Defective bearing and its vibration signal

Rolling element bearings produces vibrations and noise even if they are in good condition. As the bearing rotates the position of balls/rollers of bearing change their position relative to the load. This change in relative position of load and balls (with rotation of the bearing) causes vibrations due to change in total stiffness of bearing assembly (Liu & Shao, 2015). Apart from the inherent bearing vibrations due to changing stiffness; emergence of localised discrete defect such as pit, spall or crack on outer race, inner race or on the rolling element causes an increase in overall vibration levels. As and when rolling element (ball/roller) passed over localized discrete defect on outer or inner race of bearing, there is an abrupt change in the contact stresses at the edge. This generates a high-level short duration (impulsive) force. This impulsive force excites the bearings structural resonance resulting in high frequency damped oscillations (refer Fig. 1) that are periodically repeated at the frequency of rolling element passing over the defect.

The bearing characteristic defect frequencies are calculated from geometric dimensions of the bearing, number of balls and shaft frequency as given below:

Ball Pass Frequency Outer Race

$$BPFO = \frac{n}{2} \frac{RPM}{60} \left(1 - \frac{B_d}{P_d} \cos \emptyset \right)$$

Ball Pass Frequency Inner Race

$$BPFI = \frac{n}{2} \frac{RPM}{60} \left(1 + \frac{B_d}{P_d} \cos \emptyset \right)$$

Pass Frequency Rolling Element

$$BPFR = \frac{P_d}{2B_d} \frac{RPM}{60} \left[1 - \left(\frac{B_d}{P_d} \right)^2 \cos^2 \emptyset \right]$$

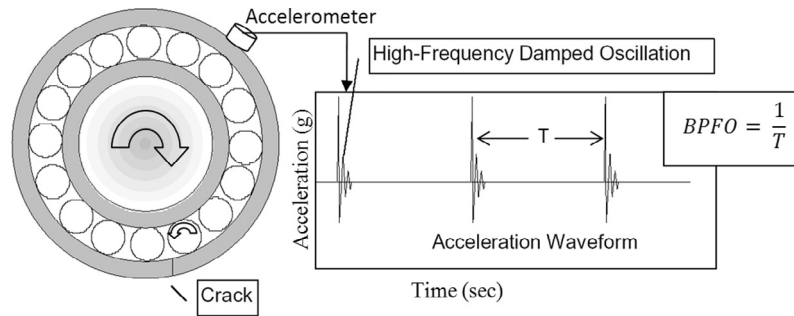


Fig. 1. Vibration signal from bearing with outer race fault.

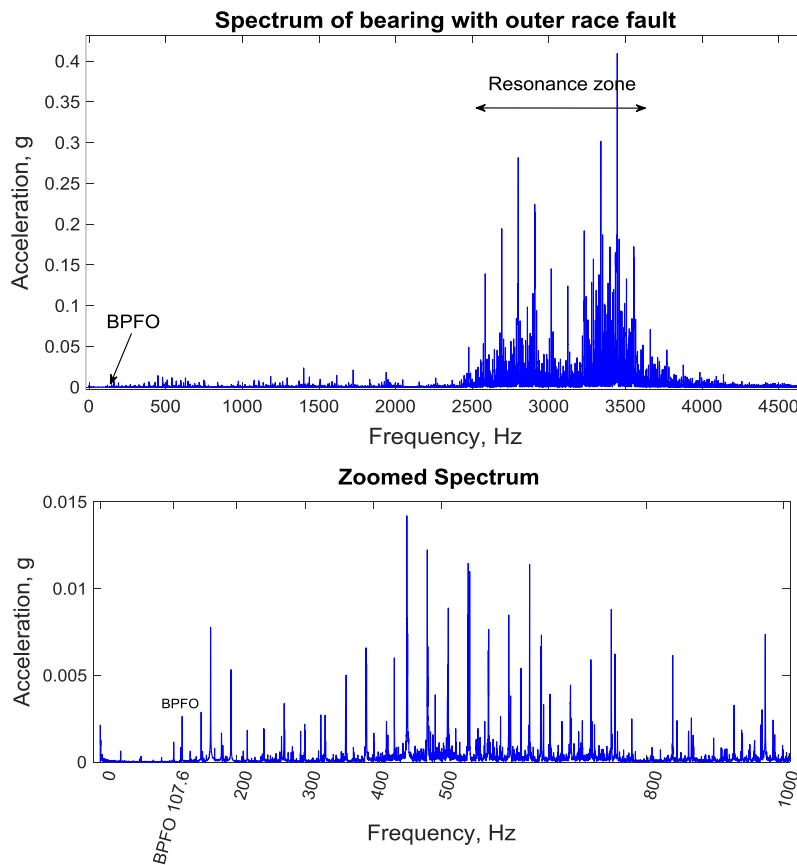


Fig. 2. Spectrum of bearing with outer race fault.

where B_d is Ball diameter, P_d is Pitch diameter, n is number of balls and ϕ is the contact angle (Randall & Antoni, 2011).

As illustrated in Fig. 2 the bearing's characteristic defect frequencies are not evident in the frequency spectrum. The spectrum of defective bearing typically shows energy concentration and large peaks in high frequency range (>2 kHz) as the impacts excite various structural resonances of the bearing.

3. Envelope detection (ED)

To solve this problem of non-detection of bearing characteristic defect frequencies in direct spectrum of vibration signal; envelope detection (also known as High Frequency Resonant Technique [HFRT]) have been employed for bearing fault detection (Barkov et al., 1995; Bechhoefer & Menon, 2009; Bechhoefer et al., 2011; Courrech, 2000; McFadden & Smith, 1984; Randall & Antoni,

2011; Wang, Miao, Fan, & Huang, 2009). The main premise behind ED is that the impulsive force due to passage of rolling elements over the defect excites the bearings structural resonance. The resonances are thus amplitude modulated at the characteristic defect frequency and by demodulating one of these resonances, envelope signal indicative of the bearing condition can be recovered.

There are three main steps involved in implementation of ED. In first step vibration signal is band-passed which removes the high frequency random noise as well as the 'large' low-frequency components and only the burst of high frequency vibrations remains as shown in Fig. 3(a) and (b). Next, an "envelope" of band passed signal is traced around the bursts in the time waveform (Fig. 3(c)) by using the Hilbert Transform. In last step, the envelope signal is converted into frequency domain (Fig. 3(d)) for analysis. In the conventional ED techniques the central frequency of band pass filter is selected to coincide with the resonance in frequency spectrum and choose the passband equal to 2–4 times the highest character-

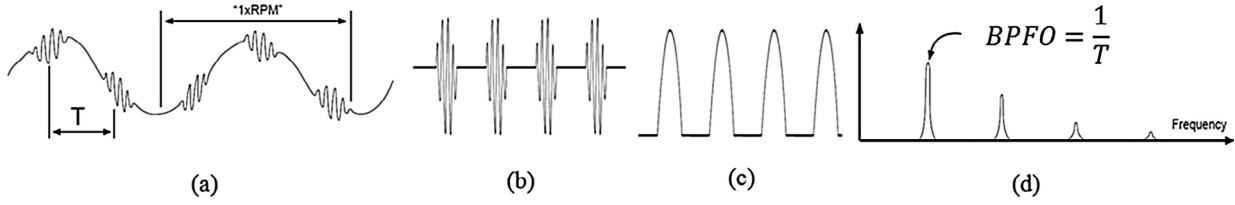


Fig. 3. Envelope detection process. (a) Unfiltered time domain signal, (b) band passed signal, (c) envelope of band passed signal, (d) spectrum of envelope.

istic defect frequency (McFadden & Smith, 1984; Norton & Karczub, 2003; Tandon & Choudhury, 1999).

ED is a good technique but traditional method of automatically placing the envelope window (i.e. selected central frequency and bandwidth of bandpass filter) at the resonant peak in vibration spectrum frequency can lead to misidentification. A case of catastrophic failure of an utility helicopter oil cooler fan bearing was reported in Bechhoefer and Menon (2009). The failure was caused by due to placement of envelope analysis window at resonance that was not excited by bearing faults. Apart from possible misidentification problem another significant aspect is degradation in performance the ED system if the envelope window is not selected properly.

Various researchers have improvised the conventional ED by hybridizing it with SVM (Yang, Yu, & Cheng, 2007) or with Wavelets (Fan & Zuo, 2006; Wang et al., 2009). A selective envelope detection method that places the envelope window outside of the structural resonance has also been reported in Courrech (2000) and Barkov et al. (1995). Bechhoefer et al. (2011) presented a method to select the bearing envelope analysis window using spectral kurtosis technique. Tang, Wang, and He (2016) used cuckoo search algorithm to maximize the maximum correlated kurtosis deconvolution method to extract each fault signal from the compound fault signals.

Work presented in this paper started with the hypothesis that possible reason why sometimes the performance of ED is not adequate is because the conventional ED relies on preselecting the central frequency and bandwidth of band-pass filter. The resonance frequency that is selected to match the central frequency of band-pass filter may not have been excited by bearing fault. Resonance could get excited by impulses from other machine elements, such as the meshing of gears, may also excite the resonance. This defeats the main purpose of the ED.

4. Particle Swarm Optimisation (PSO)

Ideally, to select most optimum combination of band-pass filter central frequency and bandwidth full gamut of frequency spread of the vibration spectrum forms the possible solution space. The vast spread of bearing vibration spectrum (typically 0–40 kHz) necessitates use of a suitable optimization technique to select the central frequency and bandwidth that would maximize the performance of ED.

Particle Swam Optimization (PSO) is an evolutionary computation technique that was proposed by Kennedy and Eberhart in 1995 (Eberhart & Kennedy, 1995). PSO is inspired by the social behavior of flock of bird or school of fish. PSO is an algorithm based on the group (swarm) behavior. The algorithm searches for the optimal value by sharing the cognitive and social information among the particles in the solution space.

There are three main steps in the PSO algorithm. First step is to generate particles with their positions and velocities, in second step the particle velocities are updated, and lastly the positions of particles are updated (Hassan, Cohanin, De Weck, & Venter,

2005). The particles are spread all over the search space and they update their position in each iteration based on interaction with all other particle in the swarm. In the first step the initial positions, $\mathbf{x}_k^i$, and velocities, $\mathbf{v}_k^i$, of complete swarm of particles are randomly generated so as to lie within upper and lower limits of the variables values i.e. $\mathbf{x}_{\min}$ and $\mathbf{x}_{\max}$; as given below:

$$\mathbf{x}_0^i = \mathbf{x}_{\min} + \text{rand}(\mathbf{x}_{\max} - \mathbf{x}_{\min})$$

$$\mathbf{v}_0^i = \frac{\text{position}}{\text{time}} = \frac{\mathbf{x}_{\min} + \text{rand}(\mathbf{x}_{\max} - \mathbf{x}_{\min})}{\Delta t}$$

Above equations initializes the swarm by distributing the particle randomly in complete search space of variable subject to upper and lower bounds.

In the second step velocities of particles at $k+1$ th iteration is updated by calculating fitness values of all particles which depends upon current positions of particles in search space at iteration k . Fitness function value is calculated for all particles at each iteration. With this the particle that has best global value in the swarm is identified. Hence, $\mathbf{p}_k^g$ the position of global best particle in iteration k is also identified. Also individual best position of each particle $\mathbf{p}^i$ in current iteration and all previous iterations is determined. The equation (given below) for velocity update of particles uses these two details and adds $\mathbf{p}_k^g$ and $\mathbf{p}^i$ to the velocity vector $\mathbf{v}_k^i$ for each particle in the swarm, to provide a new search direction, $\mathbf{v}_{k+1}^i$, for $k+1$ iteration. New velocity $\mathbf{v}_{k+1}^i$ is influenced by three parameters such as current velocities of particles $\mathbf{v}_k^i$, particle's own memory (i.e. personal best $\mathbf{p}^i$), and swarm influence (i.e. position of global best particle $\mathbf{p}_k^g$). All these three components are summed up as per velocity update equation shown below. The velocity update equation is the weighted sum of current velocity, particle's own memory and swarm influence by multiplying them respectively with inertia weight 'w', cognitive parameter 'c1' and social scaling parameter 'c2':

$$\mathbf{v}_{k+1}^i = \underbrace{\text{current motion}}_{w\mathbf{v}_k^i} + c_1 \underbrace{\text{particle memory influence}}_{\text{rand}(\mathbf{p}^i - \mathbf{x}_k^i)} + c_2 \underbrace{\text{swarm influence}}_{\text{rand}(\mathbf{p}_k^g - \mathbf{x}_k^i)}$$

The factor $w\mathbf{v}_k^i$ in above equation denotes current motion of particle i . It is the velocity of i^{th} particle at iteration k multiplied by inertia weight w . The inertia weight was decreased linearly from about 0.9 to 0.1 during run of algorithm to improve the optimisation performance of algorithm (Eberhart & Shi, 1998).

The factor $\text{rand}(\frac{\mathbf{p}^i - \mathbf{x}_k^i}{\Delta t})$ represents influence of i^{th} particle's own memory. It is consisting of position vector of the difference in particle's previous best position $\mathbf{p}^i$ (in present all past iterations) and $\mathbf{x}_k^i$ that is the position of i^{th} particle in the search space at iteration k . The own memory influence of particle is multiplied with cognitive parameter 'c1' which is varied between 1.5 and 2 (Hassan et al., 2005; Shi & Eberhart, 1998).

The factor $\text{rand}(\frac{\mathbf{p}_k^g - \mathbf{x}_k^i}{\Delta t})$ represents swarm influence on particle i . It is made up of difference in position vector $\mathbf{p}_k^g$ (global best position) and $\mathbf{x}_k^i$ (position of i^{th} particle in the search space at iteration

k). The swarm influence is multiplied with social scaling parameter 'c2' that was varied between 2 and 2.5 for best results (Hassan et al., 2005; Shi & Eberhart, 1998).

In each iteration the position update of particles is the third and last step. New position of particle is calculated using the updated velocity and adding to the current position as shown below:

$$\mathbf{x}_{k+1}^i = \mathbf{x}_k^i + \mathbf{v}_{k+1}^i \Delta t$$

All three steps PSO are repeated iteratively until the stopping criteria is met. The stopping criteria is generally the lower bound of change in value of objective function ε in each iteration that is given by:

$$|f(\mathbf{p}_k^g) - f(\mathbf{p}_{k-q}^g)| \leq \varepsilon, \quad q = 1, 2, \dots, S$$

In the PSO algorithm that is implemented in this paper uses the specified number of moves, $S = 100$ and $\varepsilon = 10^{-5}$.

Optimization by evolutionary computation techniques such as PSO or genetic algorithms (GAs) is quite effective as they ensure convergence to global minima. PSO was selected as optimizer in this paper as PSO has simpler implementation, lower computational complexity, faster convergence rate than GA (Hassan et al., 2005; Kennedy, Kennedy, Eberhart, & Shi, 2001). However the most important consideration to select PSO was that PSO has very few parameters to adjust that require very slight variations (or none at all) and works well in a wide variety of applications (Shi, 2001). In fact PSO has some advantages over GA such as interaction in the group enhances progress toward the solution rather than detracting as in case of GA.

Further, a particle swarm system has memory, which the genetic algorithm has in limited sense when elitism is employed. In particle swarm optimization, individuals who 'fly past' optima are pulled to return toward them as knowledge of good solutions is retained by all particles. Due to these reasons many researchers have used PSO for machinery fault diagnosis (Kanović, Rapaić, & Jeličić, 2011; Samanta & Nataraj, 2009; Yuan & Chu, 2007). PSO has been successfully used specifically for bearing fault diagnosis (Liu et al., 2013; Xu & Chen, 2013; Zhu, Song, & Xue, 2014).

5. Hilbert transform

The Hilbert transform is defined as convolution of time domain signal $x(t)$ with function $1/\pi t$ (Carcattera & Sestieri, 1997). It is time-domain transform that maps a real-valued time domain signal $x(t)$ to another time domain signal $\hat{x}(t)$ as given below:

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

Hilbert Transform in frequency domain act as $\pi/2$ phase shifter. All positive frequencies get a -90° phase shift and all negative frequencies of a signal get a $+90^\circ$ phase shift. The frequency response function of a Hilbert Transformer has the following property:

$$H(\omega) = \begin{cases} -j & \text{for } \omega > 0 \\ j & \text{for } \omega < 0 \end{cases}$$

Because Hilbert transform is a frequency-independent 90° phase shifter, the non-stationary characteristics of the modulating signal are not influenced by taking Hilbert transform of the signal. Demodulation which is extraction of functions of amplitude modulation as well as phase modulation can be achieved by forming a complex valued time-domain signal $\tilde{x}(t)$ that is called as analytic signal. The analytic signal is obtained by summing up the phase shifted and original signals as follows:

$$\tilde{x}(t) = x(t) + j\hat{x}(t) = a(t)e^{j\theta(t)}$$

Envelope and phase can be extracted from this analytic signal as follows:

$$a(t) = \sqrt{x^2(t) + \hat{x}^2(t)}$$

$$\theta(t) = \tan^{-1} \left[\frac{\hat{x}(t)}{x(t)} \right]$$

Analytic signal is a complex time series. Envelope can be obtained from this by taking magnitude of analytic signal added with the original time domain signal $x(t)$. The modulations present in the signal are revealed in the envelope (Fan & Zuo, 2006). The modulations caused by faulty bearings in band-passed signal can also be revealed in this way. The frequency of impacts can thus be obtained by taking FFT of the envelope.

6. Experimental setup

Vibration signals used in this research were provided by the CWRU bearing data center (Loparo). The experimental test stand is shown in Fig. 4, the test stand comprised of a 4 hp motor driving a shaft on which a torque transducer and a dynamometer were mounted. The bearing under analysis supported the motor shaft at the drive end. The bearings under analysis were introduced with only one type of fault at only one location. Fault measuring 7 mils, 14 mils and 21 mils in diameters were introduced to the bearings using spark erosion method. The fault depth was kept at 0.011 mils for all type of faults. Bearing type 6205-2RS JEM SKF, deep groove ball bearing was used for analysis and it was installed at the motor drive end. Geometric details and characteristic defect frequencies of bearing are given in Table 1.

Vibration data (acceleration) was collected for 5 s duration at a sampling frequency of 48,000 samples per second. The data was collected by magnetically mounting accelerometers on the bearing housing. Accelerometers were located at the 6 o'clock position on the bearing housing. Outer raceway faults (ORF) are the stationary type of faults and therefore they are sensitive to the relative location of fault and accelerometer. During recording of vibration data for ORF the position of both fault and the accelerometer was kept at 6 o'clock which lies in the load zone of bearing. Vibration data was recorded for four discrete motor loads of 0, 1, 2 and 3 hp. With increase in load the motor speeds reduced from 1797 to 1720 RPM. Three types of faults i.e. Outer raceway faults (ORF), Inner raceway fault and Rolling element fault (REF) were studied. Total 36 cases of various combinations of load, fault type and fault severities were formed. The details of 36 different cases are given in Table 2.

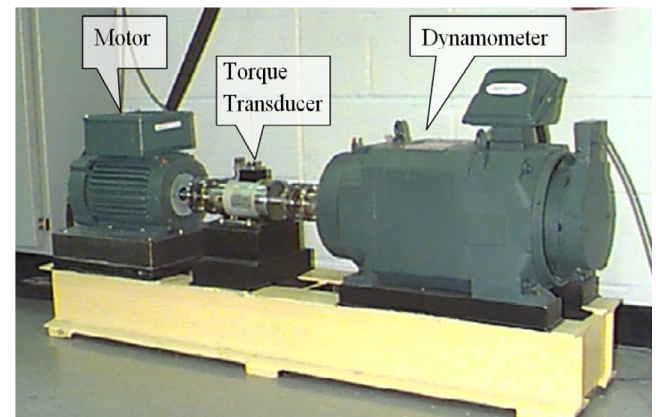


Fig. 4. Experimental setup.

Table 1
Bearing geometric details and characteristic defect frequencies.

Size: (mm)				
Inside Diameter	Outside Diameter	Thickness	Ball Diameter	Pitch Diameter
25	52	15	7.94	39.04
Defect frequencies: (multiple of running speed in Hz)				
Outer race fault (ORF)		Inner race fault (IRF)		Rolling Element fault (REF)
3.5848		5.4152		4.7135

Table 2
Different cases formed by combinations of bearing fault type, size and load.

Case	Fault Type	Fault diameter	Load (hp)	RPM	Defect frequency (Hz)
1	ORF	0.007"	0	1797	107.6
2	ORF	0.007"	1	1772	106.3
3	ORF	0.007"	2	1750	104.9
4	ORF	0.007"	3	1730	103.4
5	ORF	0.014"	0	1797	108.1
6	ORF	0.014"	1	1772	106.0
7	ORF	0.014"	2	1750	105.2
8	ORF	0.014"	3	1730	103.7
9	ORF	0.021"	0	1797	107.6
10	ORF	0.021"	1	1772	106.0
11	ORF	0.021"	2	1750	104.6
12	ORF	0.021"	3	1730	102.9
13	IRF	0.007"	0	1797	162.2
14	IRF	0.007"	1	1772	159.9
15	IRF	0.007"	2	1750	157.9
16	IRF	0.007"	3	1730	156.1
17	IRF	0.014"	0	1797	162.2
18	IRF	0.014"	1	1772	159.9
19	IRF	0.014"	2	1750	157.9
20	IRF	0.014"	3	1730	156.1
21	IRF	0.021"	0	1797	162.2
22	IRF	0.021"	1	1772	159.9
23	IRF	0.021"	2	1750	157.9
24	IRF	0.021"	3	1730	156.1
25	REF	0.007"	0	1797	141.2
26	REF	0.007"	1	1772	139.2
27	REF	0.007"	2	1750	137.5
28	REF	0.007"	3	1730	135.9
29	REF	0.014"	0	1797	141.2
30	REF	0.014"	1	1772	139.2
31	REF	0.014"	2	1750	137.5
32	REF	0.014"	3	1730	135.9
33	REF	0.021"	0	1797	141.2
34	REF	0.021"	1	1772	139.2
35	REF	0.021"	2	1750	137.5
36	REF	0.021"	3	1730	135.9

7. Improving ED by optimisation of envelope window

The first step of conventional ED process is band-passing bearing vibration signal that is obtained in Section 6 above. Conventionally, the envelope window is kept at one of the bearing's resonance. The resonance central frequency of the bandpass filter is tuned to the resonant frequency in the spectrum while keeping the passband equal to thrice the highest defect frequency of bearing. The basic idea to select the resonance is that the impulsive force due to passage of rolling element over the defect excites the bearings structural resonance. The vibration at resonance frequency are thus amplitude modulated and by demodulating the bandpass signal one can obtain a signal that contains impact occurring frequency. The performance of conventional ED degrades significantly if central frequency and band width are not correctly chosen (refer Figs. 6 and 7). Further, the conventional ED works on premise that resonances are caused by bearing faults, which is not always the case. In running machinery resonance captured by accelerometer may not be bearing resonance and could have

been caused by impulse type of force due to a machinery component not related to bearing (Bechhoefer & Menon, 2009).

The amplitude of envelope spectrum is dependent upon the central frequency (F_{cf}) of bandpass filter, bandwidth (W) of the passband by which the signal has been passed prior to extraction of envelope. Since the spectrum is a plot of amplitude vs. frequency; the amplitude of envelope spectrum varies with the frequency. However, the amplitude at the bearing characteristic defect frequency (A_{cdf}) such as BPFO, BPFI and BPFH are the function of the central frequency (F_{cf}) of bandpass filter and its bandwidth (W):

$$A_{cdf} = f(F_{cf}, W)$$

In this paper selection of envelope window is done by using PSO. As peak at defect frequencies in the envelope spectrum varies with variation of central frequency and bandwidth of band-pass filter; the PSO is implemented by maximising the peak of the bearing characteristic defect frequency A_{cdf} in the envelope spectrum. The central frequency (F_{cf}) and bandwidth (W) were kept as indepen-

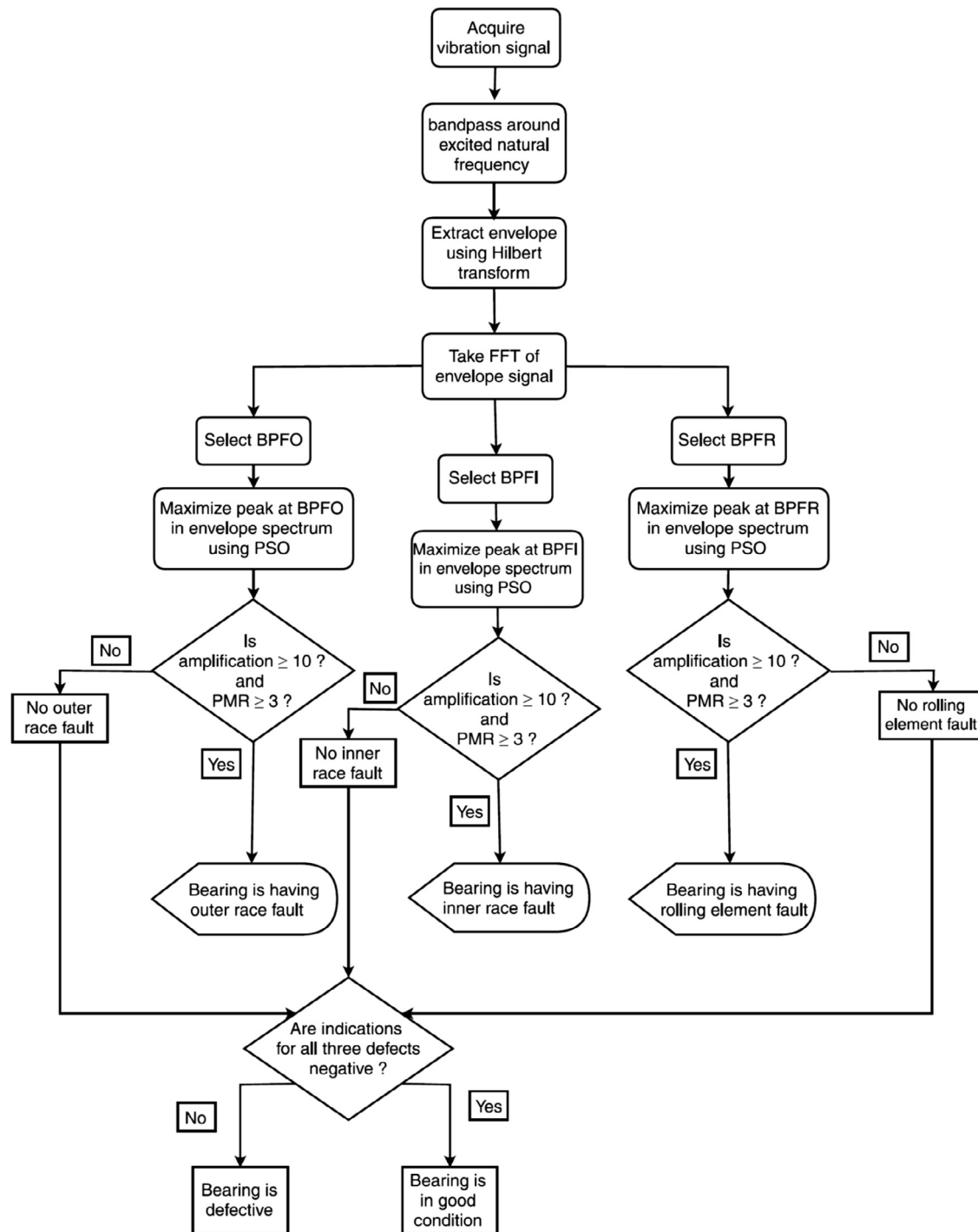


Fig. 5. Improved ED process flow.

dent variables. The central frequency and bandwidth of band-pass filter were optimised by PSO so as to achieve highest peak at the desired bearing characteristic defect frequency in envelope spectrum.

The complete process of improved Envelope Detection is summarised in flow diagram in Fig. 5. The initial three steps of the improved ED process are same as that of traditional ED process. The envelope of vibration signal is extracted by keeping the envelope window at resonance and thereafter spectrum of envelope is taken. In fourth step peaks at corresponding defect frequency is selected for maximising by searching for most optimum central

frequency and bandwidth by using PSO. In fifth step, envelope is obtained after bandpassing the vibration signal through the optimised envelope window that was obtained in previous step. Thereafter, spectrum of envelope is extracted and peaks at defect frequencies are checked for how much amplification and Peak to Mean Ratio (PMR) has been attained (refer details in Section 8). In sixth step the amplification and PMR obtained are checked if they both are found more than 10 and 3 respectively then output that bearing is having Outer race fault/Inner race fault/Rolling Element Fault is given out otherwise it indication that bearing is in good condition is produced.

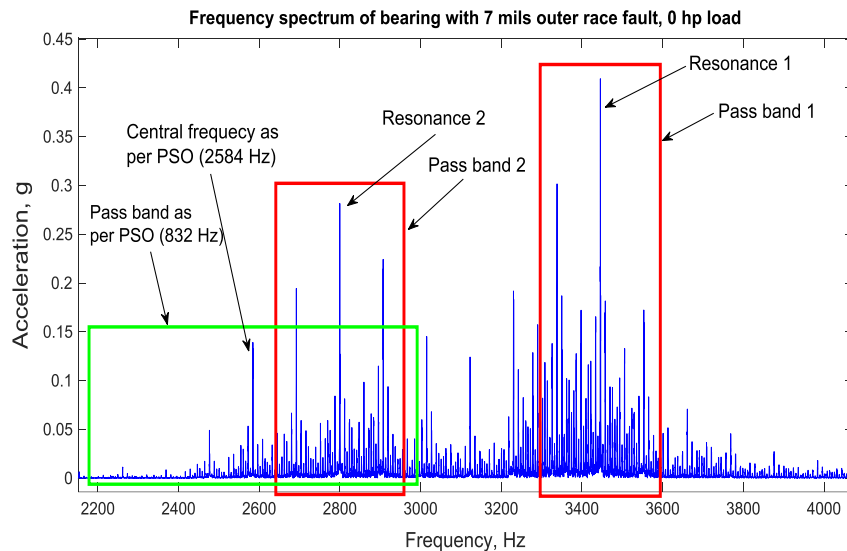


Fig. 6. Envelope window as per conventional and improved ED processes.

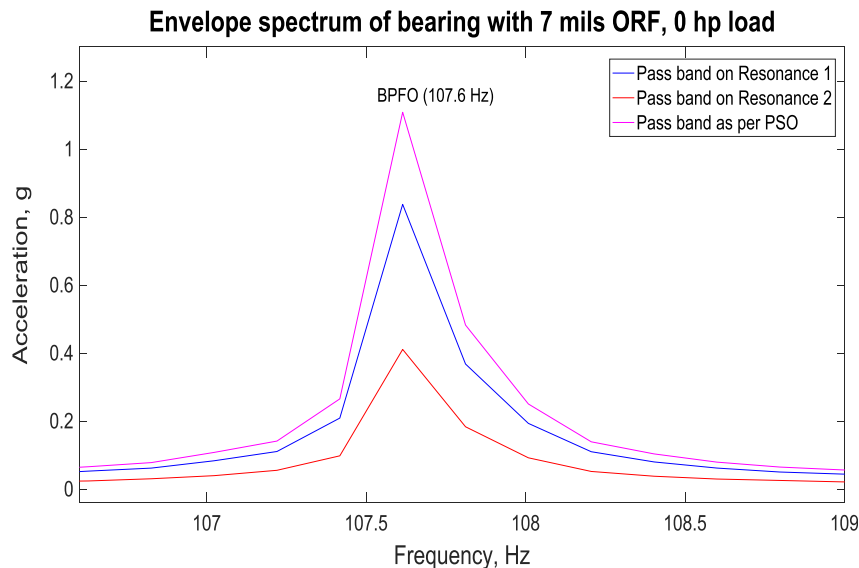


Fig. 7. Comparison of envelope spectrum as per improved and conventional ED.

In the improved ED process the envelope is obtained by making the analytic signal of the bandpassed signal using Hilbert transform as explained in Section 5. Thereafter the envelope signal is converted to frequency domain for analysis. PSO parameters used in the improved ED process are as per Table 3.

The improved ED process has suggested different central frequency and bandwidth than that of conventional ED. Fig. 6 displays the spectrum of vibration signal recorded for case 1 i.e. ORF with 7 mils dia fault size and 0 hp load. Two distinct resonances (at 3446 Hz and 2800 Hz) can be seen in the spectrum shown in Fig. 6. Red rectangles depict the envelope window of the conventional ED process. The PSO improved process has suggested a different envelope window (i.e. central frequency and bandwidth) as shown in green rectangle in Fig. 6. The envelope spectrum obtained from conventional ED process and by the PSO improved ED process is shown in Fig. 7. It can be seen in Fig. 7 that by using the improved ED process the peak (at BPFO) has been significantly increased in magnitude in comparison to traditional ED process. Fig. 8 presents the peak magnitude at bearing defect frequencies

Table 3
PSO parameters.

Parameter	Symbol	Value
Lower bound of variable search (VarMin)	F_{cf}	2000 Hz
Upper bound of variable search (VarMax)	W	20 Hz
Self confidence factor	F_{cf}	25000 Hz
Swarm confidence factor	W	4000 Hz
Inertia factor	$c1$	2
	$c2$	2
	w	0.99 and reducing in each iteration by 0.01 up to 0.4
Swarm size	–	100
Lower bound of change in value of objective function	–	1.0×10^{-6}

in envelope spectrum obtained for ORF, IRF and REF in traditional manner (blue colour) and compared with peak magnitude obtained by using PSO to select optimum pass band. Fig. 8 displays the increase in peak at defect frequencies of envelope spectrum

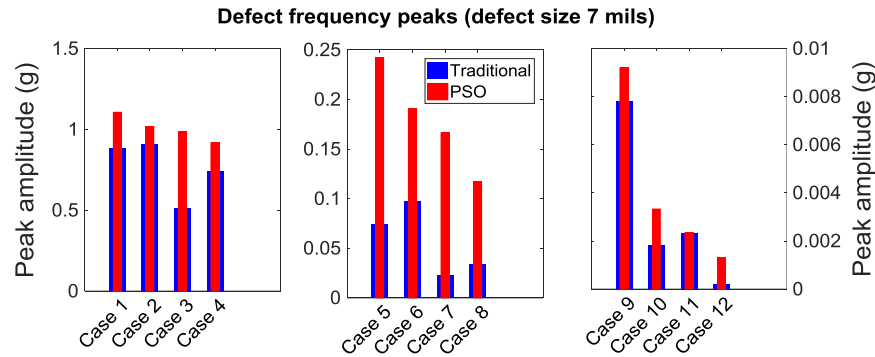


Fig. 8. Comparison of envelope spectrum peaks at defect frequency for traditional ED and ED process improved by PSO.

that is achieved by using the PSO enhanced ED process for twelve cases with defect size $\emptyset$ of 7 mils. It is evident that the 'PSO enhanced ED process' has significantly increased amplitude for 11 out of all 12 cases presented.

8. Performance of improved ED process

Fig. 9 presents plots of spectrums of bearing signals without extracting the envelope along with spectrum of its corresponding envelope. Fig. 9. presents analysis for case 1 (ORF at 0 hp load and defect size 7 mils) and case 25 (REF at 3 hp load defect size 21 mils). It can be seen from Fig. 9 that it is not possible to detect the characteristic defect frequency in spectrum of vibration signal without extracting the envelope. However, in the spectrum of envelope that has been extracted by employing the PSO enhanced 'improved ED' process the bearing characteristic defect frequency is clearly visible. The use of improved ED process has 'magnified' the peak at the characteristic defect frequency and made it possible to detect the bearing condition.

For case 1; the magnitude of peak at bearing characteristic frequency in vibration spectrum (without envelope) is 0.0012 g whereas, in envelope spectrum the magnitude of peak is 1.1081 g. Hence, the use of improved ED process has amplified the peak by 923 times. This amplification at characteristic defect frequency has enabled the detection of bearing characteristic defect frequency in the envelope spectrum.

Fig. 10 presents the amplification achieved in all 36 cases (as per Table 2) that were studied. The improved envelope process has amplified the amplitude of defect frequency peak in the envelope spectrum for all 36 cases. The amplification is achieved for all types of cases irrespective of type of fault, fault size and load. For ORF the amplification varied from minimum 26 in case 6 to max-

imum 923 in case 1. For IRF the amplification varied from minimum 30 in case 18 to maximum of 397 in case 23. For Rolling Element Fault (REF) the amplification achieved varied from minimum 11 in case 32 to 108 in case 33. The amplification achieved for fault on rolling element is least (case Nos. 25 to 36). This could possibly due to nature of ball (rolling element) fault. In case of ball fault the impulsive vibration occur whenever the fault on ball strikes the outer or inner races of the bearing. Due to the spherical shape of the ball; mating of defective pit on ball with both the races on each spin of ball is not certain. Further, the impact due to ball defect takes place alternatively on inner race and outer race. This causes difference in the vibration that is received at the accelerometer, which is quite different due to different system response. The difference in system response for strike that takes place on different races of bearing is due to the different path of travel taken by vibration to reach the vibration pickup.

Amplification alone is not true measure of effectiveness of the defect identification. Peak to Mean Ratio (PMR) which is a measure of how much high the peak at the characteristic defect frequency stands in comparison to its mean peak level in spectrum of envelope is a matter of great relevance. Fig. 11 displays envelope spectrum for ORF at zero load and fault size of 7 mils (case 1). Fig. 11 shows how PMR of 89.7 was achieved in case1. Fig. 12 shows the PMR that was achieved in all 36 cases that were studied in present research. It can be seen from the plot that high PMR 5 to 200 was achieved by outer race fault defect (ORF). PMR was also high (2 to 200). However, the inner race faults (IRF) the PMR achieved was low (less than 5). Eight out of twelve inner race defect cases studied (case No. 13 to 24) had PMR values less than 3. These cases could have been misclassified as good bearing if only the criterion of $PMR > 3$ had been applied. Although, by use of PSO the peak at the BPFI was amplified substantially still the PMR remained low.

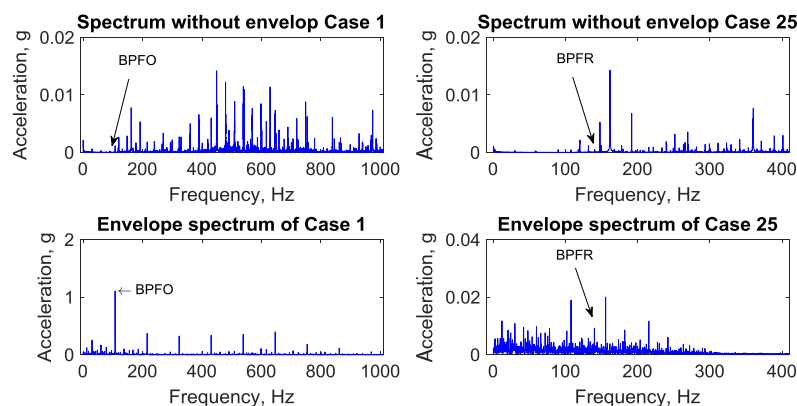


Fig. 9. Performance of improved ED process.

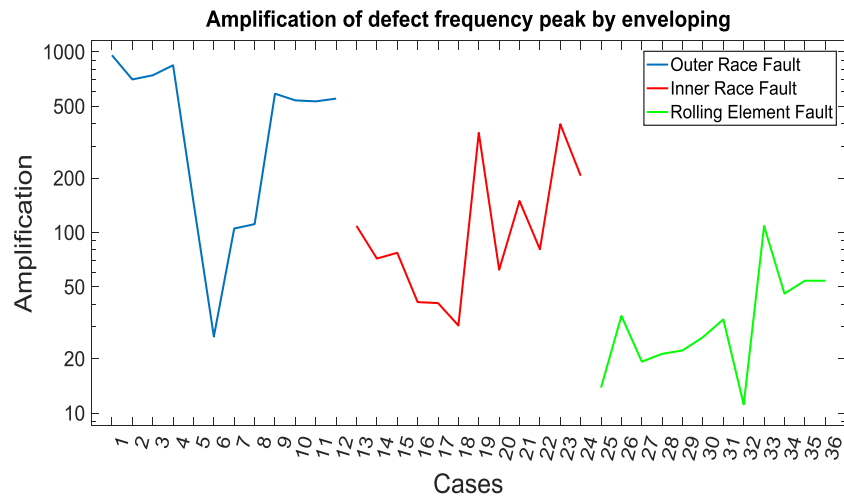


Fig. 10. Amplification of characteristic defect frequency peak by using improved ED.

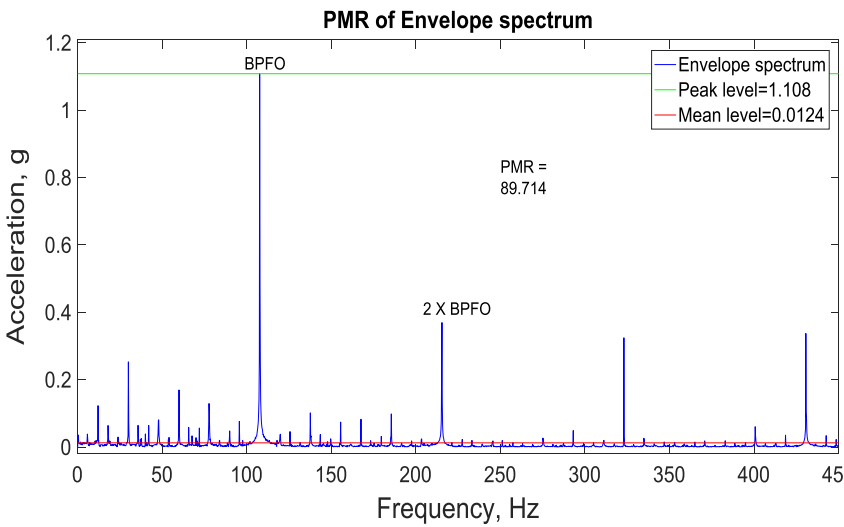


Fig. 11. PMR calculation for case 1.

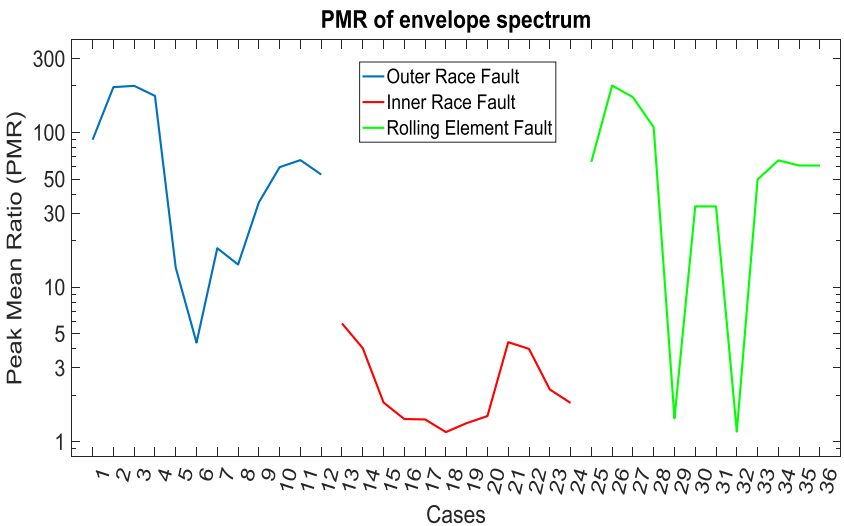


Fig. 12. Case wise PMR of envelope spectrum.

Reason for low PMR for inner race faults is possibly due longer and complicated path of travel of vibrations from defect site till the accelerometer. This causes much diminished peak at defect frequency (BPFI) in the conventional envelope spectrum. The application of PSO enhanced method is unable to enhance the peak to the extent that will make PMR > 3 in most of the cases of inner race faults.

Table 4

Amplification and PMR achieved for bearing in good condition by maximising characteristic defect frequency peaks.

Peak frequency maximised by PSO	Corresponding defect	Load	Amplification	PMR
105.9 Hz	ORF	0 hp	7.9	1.79
162.7 Hz	IRF	0 hp	1.74	2.07
142.1 Hz	REF	0 hp	26.0	1.97
106.7 Hz	ORF	1 hp	7.6	11.58
159.9 Hz	IRF	1 hp	1.18	3.61
140.1 Hz	REF	1 hp	11.6	2.18
104.9 Hz	ORF	2 hp	6.41	5.52
157.5 Hz	IRF	2 hp	0.12	3.00
137.7 Hz	REF	2 hp	18.2	1.53
103.4 Hz	ORF	3 hp	9.4	2.5
156.2 Hz	IRF	3 hp	1.31	1.55
135.5 Hz	REF	3 hp	8.4	2.21

The amplification and PMR achieved in case of bearings in good condition is placed at Table 4. The amplification and PMR for good bearing are calculated for those frequencies which correspond to the defective outer race, inner race and rolling element defects. Although, there were no defect but the characteristic defect frequencies were selected and were maximised by optimising the bandwidth and central frequency of bandpass filter. The process that was followed is depicted in Fig. 5. It can be seen that the amplification of defect frequency that was achieved for good bearing is generally less than 10 which is far less in comparison to that of defective bearings. PMR of the maximised peaks was less than 3 in case of good bearings. There were few exceptions to these values. The bearing condition can be assessed to be in good condition so long as both amplification and PMR are less than 10 and 3 respectively.

9. Effect of fault size and load

How increase in defect severity impacts the performance of the fault detection technique is an important practical consideration. The effect of fault size on performance of detecting bearing condition by using the PSO improved ED process was studied. Cylindrical faults of diameters of 7 mils, 14 mils and 21 mils having 11 mils depth were created in the outer race, inner race and the rolling

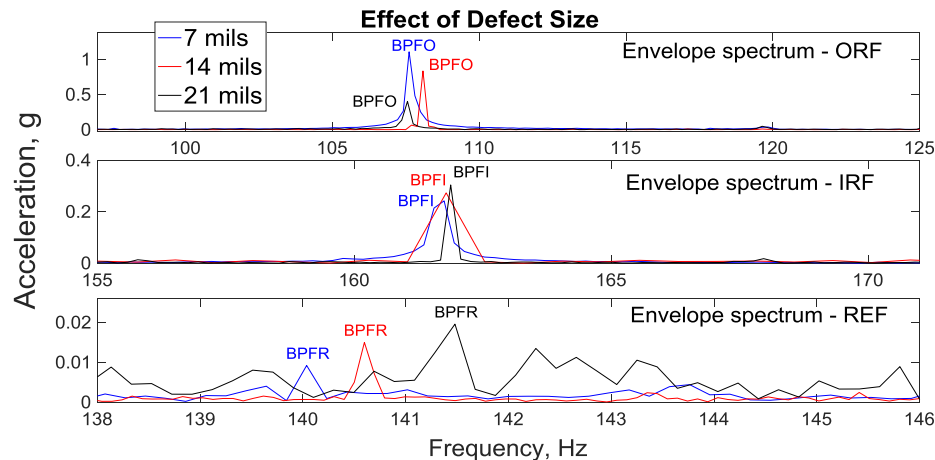


Fig. 13. Envelope spectrum of bearings with varying fault size.

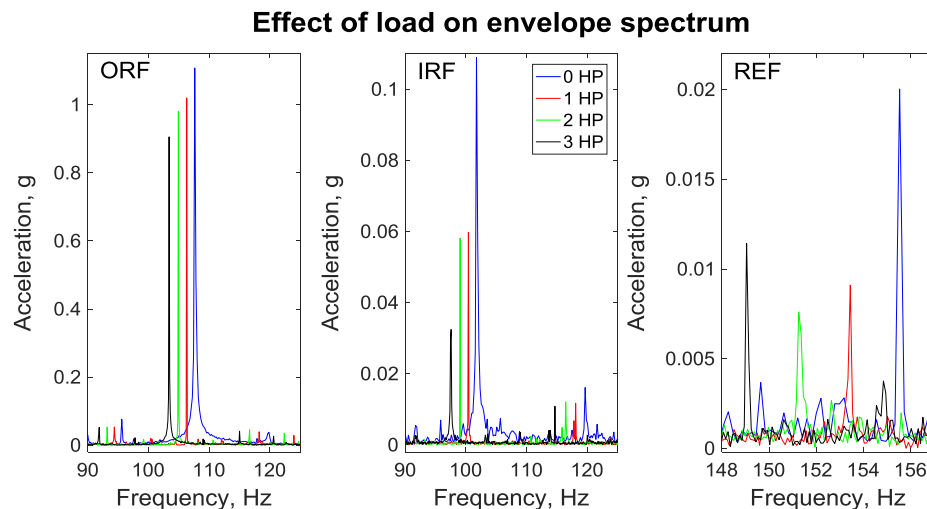


Fig. 14. Envelope spectrum of bearings with varying load.

element of the test bearings using electro-discharge machining. The Fig. 13 presents the envelope spectrum of all three type of bearing faults with varying fault size while keeping the load at zero hp. For the outer race fault (ORF) the increase in defect severity tends to decrease the peak at the defect frequency in envelope spectrum. This is possibly due the fact that with advance in fault severity, the defect frequencies may become submerged in the rising background level of the spectrum (McFadden & Smith, 1984). This may happen due to the reduced severity of impacts which are generated so frequently that the leading edge of the impact is buried in the decay of the previous impact. However, for inner race fault (IRF) and rolling element (REF) fault, with increase in damage the peak in envelope spectrum increased. This could be due to the fact that for IRF and REF vibrations have to travel a longer path to reach accelerometer. Hence higher impact is needed for detecting these faults.

Fig. 14 presents the envelope spectrum of all three type of bearing faults under increasing load from 0 hp to 3 hp while keeping the fault size same at 7 mils. It can be seen that with increase in load the peak at the defect frequency reduces. This could be due to the fact that with advance in fault severity, the defect frequencies may become submerged in the rising background level of the spectrum. As the peak at the defect frequency reduces with increase in load, hence the effectiveness of the improved ED process also reduces.

10. Conclusion

A novel improved envelope detection method that uses particle swarm optimisation for identification of rolling element bearing faults is presented here. The method employs PSO to optimise selection of central frequency and bandwidth of the bandpass filter prior to extraction of envelope. It was shown that by optimising the central frequency and bandwidth of filter with the help of PSO the performance ED is enhanced in comparison to the conventional ED process. It was also shown that by using improved ED process the peak at characteristic defect frequency is substantially amplified in the spectrum of envelope than direct spectrum of vibration signal. It has also been shown that in the envelope spectrum the ratio of peak at characteristic defect frequency and mean spectrum level is quite high which makes identification of bearing fault very effective. Further, it has been shown that as the defect severity increases the identification of bearing faults becomes progressively difficult and the increase in load on bearing also adversely affect the bearing fault identification.

Conflict of interest

We confirm that there is no conflict of interest of any kind regarding this paper.

Acknowledgment

The authors would like to thank Professor K.A. Loparo of Case Western Reserve University for providing free access to the bearing vibration datasets that are available at Bearing Data Center Website.

References

AlRashidi, M. R., & El-Hawary, M. E. (2009). A survey of particle swarm optimization applications in electric power systems. *IEEE Transactions on Evolutionary Computation*, 13, 913–918.

Antoni, J. (2007). Fast computation of the kurtogram for the detection of transient faults. *Mechanical Systems and Signal Processing*, 21, 108–124.

Barkov, A., Barkova, N., & Mitchell, J. S. (1995). Condition assessment and life prediction of rolling element bearings-part 1. *Sound and Vibration*, 29, 10–17.

Bechhoefer, E., Menon, P., & Kingsley, M. (2011). Bearing envelope analysis window selection using spectral kurtosis techniques. In *IEEE conference on prognostics and health management (PHM)*, 2011 (pp. 1–6). IEEE.

Bechhoefer, E., & Menon, P. (2009). Bearing envelope analysis window selection. *Annual Conference of the Prognostics and Health Management Society*, 1–7.

Blondin, J. (2009). Particle swarm optimization: A tutorial. Available from: <<http://CsArmstrongEdu/Saad/csci8100/PsoTutorialPdf>>.

Carcattera, A., & Sestieri, A. (1997). Complex envelope displacement analysis: A quasi-static approach to vibrations. *Journal of Sound and Vibration*, 201, 205–233.

Chaturvedi, G., & Thomas, D. (1982). Bearing fault detection using adaptive noise cancelling. *ASME, Transactions, Journal of Mechanical Design*, 104, 280–289.

Courrech, J. (2000). Envelope analysis for effective rolling-element fault detection-fact or fiction? *Up Time Magazine*, 8, 14–17.

Eberhart, R., & Kennedy, J. (1995). A new optimizer using particle swarm theory. In *Proceedings of the sixth international symposium on micro machine and human science*, 1995. *MHS'95* (pp. 39–43). IEEE.

Eberhart, R., & Shi, Y. (1998). *Comparison between genetic algorithms and particle swarm optimization. Evolutionary programming VII*. Springer.

Fan, X., & Zuo, M. J. (2006). Gearbox fault detection using Hilbert and wavelet packet transform. *Mechanical Systems and Signal Processing*, 20, 966–982.

Gaig, Z.-L. (2004). A particle swarm optimization approach for optimum design of PID controller in AVR system. *IEEE Transactions on Energy Conversion*, 19, 384–391.

Hassan, R., Cohanin, B., De Weck, O., & Venter, G. A. (2005). comparison of particle swarm optimization and the genetic algorithm. In *46th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference* (p. 1897).

Huang, K. Y. (2011). A hybrid particle swarm optimization approach for clustering and classification of datasets. *Knowledge-Based Systems*, 24, 420–426.

Kanović, Z., Rapačić, M. R., & Jeličić, Z. D. (2011). Generalized particle swarm optimization algorithm—Theoretical and empirical analysis with application in fault detection. *Applied Mathematics and Computation*, 217, 10175–10186.

Kennedy, J. F., Kennedy, J., Eberhart, R. C., & Shi, Y. (2001). *Swarm intelligence*. Morgan Kaufmann.

Liu, Z., Cao, H., Chen, X., He, Z., & Shen, Z. (2013). Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings. *Neurocomputing*, 99, 399–410.

Liu, J., & Shao, Y. (2015). A new dynamic model for vibration analysis of a ball bearing due to a localized surface defect considering edge topographies. *Nonlinear Dynamics*, 79, 1329–1351.

Loparo, K. A. Case Western Reserve University Bearing Data Center <<http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>>.

Mathew, J. (1984). The condition monitoring of rolling element bearings using vibration analysis. *Journal of Vibration, Acoustics, Stress, and Reliability in Design*, 106, 447–453.

McFadden, P., & Smith, J. (1984). Vibration monitoring of rolling element bearings by the high-frequency resonance technique—A review. *Tribology International*, 17, 3–10.

Norton, M., & Karczub, D. (2003). *Fundamentals of noise and vibration analysis for engineers*.

Peng, Z., & Chu, F. (2004). Application of the wavelet transform in machine condition monitoring and fault diagnostics: A review with bibliography. *Mechanical Systems and Signal Processing*, 18, 199–221.

Peter, W. T., & Wang, D. (2013). The automatic selection of an optimal wavelet filter and its enhancement by the new sparsogram for bearing fault detection: Part 2 of the two related manuscripts that have a joint title as “Two automatic vibration-based fault diagnostic methods using the novel sparsity measurement—Parts 1 and 2”. *Mechanical Systems and Signal Processing*, 40, 520–544.

Poli, R., Kennedy, J., & Blackwell, T. (2007). Particle swarm optimization. *Swarm Intelligence*, 1, 33–57.

Prabhakar, S., Mohanty, A., & Sekhar, A. (2002). Application of discrete wavelet transform for detection of ball bearing race faults. *Tribology International*, 35, 793–800.

Quan, L., Yongqian, L., & Yongping, Y. (2012). Fault diagnosis method of wind turbine gearbox based on bp neural network trained by particle swarm optimization algorithm. *Acta Energetica Solaris Sinica*, 1, 021.

Randall, R. B., & Antoni, J. (2011). Rolling element bearing diagnostics—A tutorial. *Mechanical Systems and Signal Processing*, 25, 485–520.

Samanta, B., & Nataraj, C. (2009). Use of particle swarm optimization for machinery fault detection. *Engineering Applications of Artificial Intelligence*, 22, 308–316.

Selvan, S. E., Xavier, C. C., Karssemeijer, N., Sequeira, J., Cherian, R. A., & Dhala, B. Y. (2006). Parameter estimation in stochastic mammogram model by heuristic optimization techniques. *IEEE Transactions on Information Technology in Biomedicine*, 10, 685–695.

Shi, Y., & Eberhart, R. (1998). A modified particle swarm optimizer. In *The 1998 IEEE international conference on evolutionary computation proceedings*, 1998. *IEEE World Congress on Computational Intelligence* (pp. 69–73). IEEE.

Shi, Y. et al. (2001). Particle swarm optimization: Developments, applications and resources. In *Proceedings of the 2001 Congress on evolutionary computation* (Vol. 1; pp. 81–86). IEEE.

Sun, H.-C., Huang, C.-M., & Huang, Y.-C. (2013). Fault diagnosis of steam turbine-generator sets using an EPSO-based support vector classifier. *IEEE Transactions on Energy Conversion*, 28, 164–171.

- Tandon, N., & Choudhury, A. (1999). A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology International*, 32, 469–480.
- Tang, G., Wang, X., & He, Y. (2016). Diagnosis of compound faults of rolling bearings through adaptive maximum correlated kurtosis deconvolution. *Journal of Mechanical Science and Technology*, 30, 43–54.
- Wang, D., Miao, Q., Fan, X., & Huang, H.-Z. (2009). Rolling element bearing fault detection using an improved combination of Hilbert and wavelet transforms. *Journal of Mechanical Science and Technology*, 23, 3292–3301.
- Xiang, C., Huang, X., Huang, D., Hu, J. (2006). Wavelets neural network based on particle swarm optimization algorithm for fault diagnosis. In *First international conference on innovative computing, information and control*, 2006. ICICIC'06 (Vol. 3, pp. 320–323). IEEE.
- Xu, H., & Chen, G. (2013). An intelligent fault identification method of rolling bearings based on LSSVM optimized by improved PSO. *Mechanical Systems and Signal Processing*, 35, 167–175.
- Yang, Y., Yu, D., & Cheng, J. (2007). A fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM. *Measurement*, 40, 943–950.
- Yuan, S.-F., & Chu, F.-L. (2007). Fault diagnostics based on particle swarm optimisation and support vector machines. *Mechanical Systems and Signal Processing*, 21, 1787–1798.
- Zhang, B., Hu, K., & Zhu, Y. (2010). Spectrum allocation in cognitive radio networks using swarm intelligence. In *Second international conference on communication software and networks*, 2010. ICCSN'10 (pp. 8–12).
- Zhu, K., Song, X., & Xue, D. (2014). A roller bearing fault diagnosis method based on hierarchical entropy and support vector machine with particle swarm optimization algorithm. *Measurement*, 47, 669–675.