Research Article

Roller Bearing Monitoring by New Subspace-Based Damage Indicator

G. Gautier, R. Serra, and J.-M. Mencik

INSA Centre Val de Loire, Laboratoire de Mécanique et Rhéologie EA 2640, 3 rue de la Chocolaterie,
CS 23410, 41034 Blois Cedex, France

Correspondence should be addressed to R. Serra; roger.serra@insa-cvl.fr

A frequency-band subspace-based damage identification method for fault diagnosis in roller bearings is presented. Subspace-based damage indicators are obtained by filtering the vibration data in the frequency range where damage is likely to occur, that is, around the bearing characteristic frequencies. The proposed method is validated by considering simulated data of a damaged bearing. Also, an experimental case is considered which focuses on collecting the vibration data issued from a run-to-failure test. It is shown that the proposed method can detect bearing defects and, as such, it appears to be an efficient tool for diagnosis purpose.

1. Introduction

Successfully implementing a condition monitoring procedure allows a mechanical system to operate at full capacity without the need of shutting down the process for periodic inspections. In this context, vibration-based structural health monitoring (SHM) techniques [1] can be considered which involve measuring the vibration signals of a structure in the time domain, as well as proposing damage indicators and efficient statistical analysis for determining its current state of structural health [2].

Vibration-based damage detection methods have gained a large popularity over the last two decades, especially for rotating machines [3]. Rolling element bearings are essential components of rotating machines, which constitute the primary cause of breakdowns [4]. When a rolling element moves through a faulty surface, an impact is induced which in turn excites the resonance frequency of the bearing system. As the bearing rotates, these impulses occur periodically at the characteristic frequency of the defect. This explains why the detection of faults in bearings is commonly achieved by identifying those bearing characteristic frequencies (BCFs) [5]. Notice however that these impulses are very weak at the early stage of fault generation and are thus usually overwhelmed by measurement noises and other vibration sources such as rotor unbalance [6] or gear mesh, which induce difficulties for detecting these emerging defects.

For the purpose of weak signature enhancement, a variety of signal processing techniques have been proposed. They mainly include scalar indicators [7], high frequency resonance techniques [8], spectral kurtosis analysis [9, 10], wavelet analysis [11, 12], empirical mode decomposition [13, 14], or PeakVue analysis [15].

The present work is concerned with the use of subspace identification methods for damage diagnosis [16, 17]. The originality of the present work is to assess the effectiveness and applicability of such methods to identify damage in roller bearing when the output signals are only considered, that is, when the input data are not known a priori. To that aim, vibration response data are periodically collected and are used to obtain a so-called observability matrix which is well suited for damage detection [18]. Here, a subspace-based damage detection procedure is proposed which is combined with a pass-band filtering approach [19]. Clearly speaking, the output signals are filtered around the bearing characteristic frequencies so as to provide efficient subspace-based damage indicators.

The rest of the paper is organized as follows. The basics of the subspace-based methods, for damage detection, are recalled. Also a new subspace damage indicator is proposed.

A presentation of roller bearings is proposed in Section 3 along with a damage identification strategy combined with the aforementioned pass-band filtering procedure. A theoretical study of a damaged bearing and an experimental validation based on a run-to-failure test are proposed in Sections 4 and 5, respectively. This paper is an extended version with several improvements of the scientific and experimental contents of an earlier conference paper version [20].

2. Subspace-Based Damage Detection Procedure

2.1. Subspace-Based Methods. Consider the following $2n$th order discrete-time linear state space system for $k = 0, 1, \ldots, N - 1$ [21]:

$$
\begin{aligned}
\mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{w}_k, \\
\mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{v}_k,
\end{aligned}
\tag{1}
$$

where $\mathbf{x}_k$ is a $2n \times 1$ state vector defined at the discrete time t_k ($k = 1, 2, \ldots, N$); $\mathbf{y}_k$ is a $l \times 1$ vector of output data; also, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ are system matrices of respective sizes $2n \times 2n$, $2n \times m$, $l \times 2n$, and $l \times m$. Besides, $\mathbf{v}_k$ and $\mathbf{w}_k$ are $2n \times 1$ and $l \times 1$ vectors of process and measurement noises, respectively, which are supposed to be white Gaussian with zero-mean distributions and joint covariance matrix [22]

$$
\mathbf{E}\left[\begin{pmatrix} \mathbf{w}_p \\ \mathbf{v}_p \end{pmatrix} \begin{pmatrix} \mathbf{w}_q^T & \mathbf{v}_q^T \end{pmatrix} \right] = \begin{pmatrix} \mathbf{Q} & \mathbf{S} \\ \mathbf{S}^T & \mathbf{R} \end{pmatrix} \delta_{pq},
\tag{2}
$$

where $\mathbf{E}[\cdot]$ is the expectation operator, while δ_{pq} is the Kronecker delta.

Some notations used for the subspace algorithms are introduced here. A block Hankel matrix of output data is first considered, which can be partitioned into past (p) and future (f) outputs as follows:

$$
\mathbf{Y}_{0|2i-1} \overset{\text{def}}{=} \begin{bmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \cdots & \mathbf{y}_{j-1} \\ \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_j \\ \vdots & \vdots & & \vdots \\ \mathbf{y}_{i-1} & \mathbf{y}_i & \cdots & \mathbf{y}_{i+j-2} \\ \hline \mathbf{y}_i & \mathbf{y}_{i+1} & \cdots & \mathbf{y}_{i+j-1} \\ \mathbf{y}_{i+1} & \mathbf{y}_{i+2} & \cdots & \mathbf{y}_{i+j} \\ \vdots & \vdots & & \vdots \\ \mathbf{y}_{2i-1} & \mathbf{y}_{2i} & \cdots & \mathbf{y}_{2i+j-2} \end{bmatrix}
$$

$$
\overset{\text{def}}{=} \begin{bmatrix} \mathbf{Y}_{0|i-1} \\ \hline \mathbf{Y}_{i|2i-1} \end{bmatrix}
$$

$$
\overset{\text{def}}{=} \begin{bmatrix} \mathbf{Y}_p \\ \hline \mathbf{Y}_f \end{bmatrix}.
\tag{3}
$$

Here, the number i of block rows of $\mathbf{Y}_p$ and $\mathbf{Y}_f$ should be greater than the order of the system—that is, $2n$—in order to identify the system; also, the number j of columns is usually chosen so that it is equal to $N - 2i + 1$, where N is the number of time samples. In the same way as (3), block Hankel matrices of measurement noises can be defined as follows:

$$
\begin{aligned}
\mathbf{V}_{0|2i-1} &\overset{\text{def}}{=} \begin{bmatrix} \mathbf{V}_p \\ \hline \mathbf{V}_f \end{bmatrix}, \\
\mathbf{W}_{0|2i-1} &\overset{\text{def}}{=} \begin{bmatrix} \mathbf{W}_p \\ \hline \mathbf{W}_f \end{bmatrix}.
\end{aligned}
\tag{4}
$$

Also, $2n \times j$ matrices of state sequences $\mathbf{X}_p$ and $\mathbf{X}_f$ can be defined as follows:

$$
\begin{aligned}
\mathbf{X}_p &\overset{\text{def}}{=} \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \cdots & \mathbf{x}_{j-1} \end{bmatrix}, \\
\mathbf{X}_f &\overset{\text{def}}{=} \begin{bmatrix} \mathbf{x}_i & \mathbf{x}_{i+1} & \cdots & \mathbf{x}_{i+j-1} \end{bmatrix}.
\end{aligned}
\tag{5}
$$

Consider now the following $(li) \times 2n$ so-called extended observability matrix $\mathbf{\Gamma}_i$ and $(li) \times (li)$ block Toeplitz matrix $\mathbf{H}_i^w$:

$$
\mathbf{\Gamma}_i \overset{\text{def}}{=} \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{i-1} \end{bmatrix},
$$

$$
\mathbf{H}_i^w \overset{\text{def}}{=} \begin{bmatrix} \mathbf{0} & & & \mathbf{0} \\ \mathbf{C} & \mathbf{0} & & \\ \vdots & \ddots & \ddots & \\ \mathbf{CA}^{i-2} & \cdots & \mathbf{C} & \mathbf{0} \end{bmatrix}.
\tag{6}
$$

By considering (1) and (3)–(6), the following matrix equations can be derived for the past and future parts, respectively [16]:

$$
\begin{aligned}
\mathbf{Y}_p &= \mathbf{\Gamma}_i \mathbf{X}_p + \mathbf{H}_i^w \mathbf{W}_p + \mathbf{V}_p, \\
\mathbf{Y}_f &= \mathbf{\Gamma}_i \mathbf{X}_f + \mathbf{H}_i^w \mathbf{W}_f + \mathbf{V}_f.
\end{aligned}
\tag{7}
$$

The key idea behind subspace methods is to identify the extended observability matrix $\mathbf{\Gamma}_i$, and further the system matrix $\mathbf{A}$. The strategy consists in projecting the row space of the future outputs $\mathbf{Y}_f$ on the row space of the past outputs $\mathbf{Y}_p$, as follows:

$$
\frac{\mathbf{Y}_f}{\mathbf{Y}_p} = \mathbf{Y}_f \mathbf{Y}_p^T \left(\mathbf{Y}_p \mathbf{Y}_p^T \right)^+ \mathbf{Y}_p,
\tag{8}
$$

where $(\cdot)^+$ denotes the Moore-Penrose pseudoinverse. By considering (7), this yields

$$
\frac{\mathbf{Y}_f}{\mathbf{Y}_p} = \frac{\mathbf{\Gamma}_i \mathbf{X}_f}{\mathbf{Y}_p} + \frac{\mathbf{H}_i^w \mathbf{W}_f}{\mathbf{Y}_p} + \frac{\mathbf{V}_f}{\mathbf{Y}_p}.
\tag{9}
$$

The procedure enables one to remove the noise terms; it is understood that the row spaces of $\mathbf{W}_f$ and $\mathbf{V}_f$ are perpendicular to the row space of $\mathbf{Y}_p$ when the number of samples grows to infinity. A proof of this statement lies in the consideration of (2) [22]. Hence, the vector of state sequence $\mathbf{X}_f$ belongs to the joint row space of $\mathbf{Y}_p$; that is,

$$\lim_{N \to \infty} \frac{1}{N} \left(\frac{\mathbf{Y}_f}{\mathbf{Y}_p} \right) = \frac{\mathbf{\Gamma}_i \mathbf{X}_f}{\mathbf{Y}_p}. \tag{10}$$

As a result, one has

$$\frac{\mathbf{Y}_f}{\mathbf{Y}_p} \approx \frac{\mathbf{\Gamma}_i \mathbf{X}_f}{\mathbf{Y}_p}. \tag{11}$$

In other words, there exists some direct connection between the extended observability matrix $\mathbf{\Gamma}_i$ and the matrix $\mathbf{Y}_f$ which is supposed to be known from measurement of the output signals [23]. For practical purpose, a singular value decomposition (SVD) of $\mathbf{Y}_f / \mathbf{Y}_p$ is usually considered as follows:

$$\frac{\mathbf{Y}_f}{\mathbf{Y}_p} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{S}_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix} \simeq \mathbf{U}_1 \mathbf{S}_1 \mathbf{V}_1^T, \tag{12}$$

where $\mathbf{S}_1$ is a $2n \times 2n$ diagonal matrix whose components are the $2n$ largest singular values of $\mathbf{Y}_f / \mathbf{Y}_p$, while $\mathbf{U}_1$ and $\mathbf{V}_1$ are $(li) \times 2n$ and $j \times 2n$ matrices of orthogonal vectors. In other words, it is assumed here that $\mathrm{rank}(\mathbf{Y}_f / \mathbf{Y}_p) = 2n$. Assume that the matrix $\mathbf{\Gamma}_i$ is full column rank; that is, $\mathrm{rank}(\mathbf{\Gamma}_i) = 2n$, which is certainly true because $N \gg 2n$. Thus it turns out that the column space of $\mathbf{\Gamma}_i$ is almost the same as the space spanned by the first $2n$ left singular vectors of the matrix $\mathbf{Y}_f / \mathbf{Y}_p$, that is, those associated with its $2n$ largest singular values.

From (11) and (12), one has $\mathbf{\Gamma}_i \mathbf{X}_f / \mathbf{Y}_p \approx \mathbf{U}_1 \mathbf{S}_1 \mathbf{V}_1^T$. This particularly means that the column space of $\mathbf{\Gamma}_i$ matches that of $\mathbf{U}_1$; that is, $\mathrm{ran}(\mathbf{\Gamma}_i) \approx \mathrm{ran}(\mathbf{U}_1)$. Hence an estimate of the extended observability matrix can be defined as $\mathbf{\Gamma}^e = \mathbf{U}_1$. The determination of the true observability matrix $\mathbf{\Gamma}_i$ follows as

$$\mathbf{\Gamma}_i = \mathbf{\Gamma}^e \mathbf{T}_i, \tag{13}$$

where $\mathbf{T}_i$ is a $2n \times 2n$ full rank matrix, while $\mathbf{\Gamma}^e$ is the so-called experimental extended observability matrix that is built from output measurements.

2.2. Subspace-Based Damage Indicator.

The key idea behind the proposed subspace-based damage detection procedure consists in comparing the extended observability matrix of the safe structure (say, $\mathbf{\Gamma}_i = \mathbf{\Gamma}^s$) with that of the structure in an unknown health condition (say, $\mathbf{\Gamma}_i = \mathbf{\Gamma}^u$). This yields the consideration of an error norm $\|\mathbf{\Gamma}^u - \mathbf{\Gamma}^s\|_F = \|\mathbf{\Gamma}^{eu} \mathbf{T}^u - \mathbf{\Gamma}^{es} \mathbf{T}^s\|_F$, where $\| \cdot \|_F$ is the Frobenius norm, while $\mathbf{\Gamma}^{es} = \mathbf{U}_1^s$ and $\mathbf{\Gamma}^{eu} = \mathbf{U}_1^u$ are experimental extended observability matrices; see (13).

Here, an alternative error norm is considered which involves right-multiplying the residual $\mathbf{\Gamma}^{eu} \mathbf{T}^u - \mathbf{\Gamma}^{es} \mathbf{T}^s$ by $(\mathbf{T}^u)^{-1}$; that is,

$$\|\mathbf{\Gamma}^{eu} - \mathbf{\Gamma}^{es} \mathbf{T}\|_F \quad \text{where } \mathbf{T} = \mathbf{T}^s (\mathbf{T}^u)^{-1}. \tag{14}$$

Besides, an approximate expression of the matrix $\mathbf{T}$ is considered which involves solving the following least squares problem $\mathbf{\Gamma}^u = \mathbf{\Gamma}^s \mathbf{T}$, which leads to $\mathbf{T} = (\mathbf{\Gamma}^s)^+ \mathbf{\Gamma}^u$. The relevance of this approximate expression can be justified in the sense that the error norm $\|\mathbf{\Gamma}^u - \mathbf{\Gamma}^s \mathbf{T}\|_F$ is supposed to be small. Hence, (14) can be rewritten as

$$\left\| \mathbf{\Gamma}^{eu} - \mathbf{\Gamma}^{es} (\mathbf{\Gamma}^{es})^+ \mathbf{\Gamma}^{eu} \right\|_F. \tag{15}$$

Assume that the QR decomposition of $\mathbf{\Gamma}^{es}$ and $\mathbf{\Gamma}^{eu}$ is expressed as

$$\mathbf{\Gamma}^{es} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix} = \mathbf{Q}_1 \mathbf{R}_1,$$

$$\mathbf{\Gamma}^{eu} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{bmatrix} = \mathbf{Q}_1 \mathbf{R}_1 + \mathbf{Q}_2 \mathbf{R}_2, \tag{16}$$

which appears to be consistent with the fact that the error norm (15) is small (indeed, both QR decomposition types are based on the same subspaces $\mathrm{ran}(\mathbf{Q}_1)$, while $\|\mathbf{R}_2\|_F$ is supposed to be small compared to $\|\mathbf{R}_1\|_F$); this yields $\|\mathbf{\Gamma}^{eu} - \mathbf{\Gamma}^{es}(\mathbf{\Gamma}^{es})^+ \mathbf{\Gamma}^{eu}\|_F = \|\mathbf{Q}_2 (\mathbf{Q}_2)^T \mathbf{\Gamma}^{eu}\|_F$. The proof of this result lies in the fact that the pseudoinverse of $\mathbf{Q}_1 \mathbf{R}_1$ is given by $(\mathbf{R}_1)^{-1} (\mathbf{Q}_1)^T$, where $(\mathbf{Q}_1)^T (\mathbf{Q}_1) = \mathbf{I}$. By considering the fact that the Frobenius norm is unitarily invariant, this yields $\|\mathbf{Q}_2 (\mathbf{Q}_2)^T \mathbf{\Gamma}^{eu}\|_F = \|(\mathbf{Q}_2)^T \mathbf{\Gamma}^{eu}\|_F$. Hence, a subspace damage indicator can be expressed as follows:

$$\left\| \mathbf{\Gamma}^{eu} - \mathbf{\Gamma}^{es} \mathbf{T} \right\|_F \approx \left\| (\mathbf{Q}_2)^T \mathbf{\Gamma}^{eu} \right\|_F. \tag{17}$$

It is expected that structural damage is detected by an increase of the subspace damage indicator in (17). The important point is to identify the relevant changes of the damage indicator, that is, those which are not due to noise. To address this issue, it is proposed to normalize the damage indicator through the consideration of n_{ref} reference data sets for the safe structure. Within this framework, a residual covariance matrix Σ can be defined as follows:

$$\Sigma = \frac{1}{n_{\mathrm{ref}} - 1} \sum_{k=1}^{n_{\mathrm{ref}}} \mathbf{r}_k \mathbf{r}_k^T, \tag{18}$$

where $\mathbf{r}_k$ refers to the residual vector for a given data set k $(k = 1, \ldots, n_{\mathrm{ref}})$:

$$\mathbf{r}_k = \mathrm{vec} \left\{ (\mathbf{Q}_2)^T \mathbf{\Gamma}^{eu} \right\}_k, \tag{19}$$

where $\mathrm{vec}\{\cdot\}$ denotes the column stacking operator. As a result, a normalized damage indicator can be defined as follows:

$$\zeta = \mathbf{r}^T \Sigma^{-1} \mathbf{r}, \tag{20}$$

where $\mathbf{r} = \mathrm{vec}\{(\mathbf{Q}_2)^T \mathbf{\Gamma}^{eu}\}$ is the residual vector that concerns any arbitrary data set k of the structure in an unknown health condition.

Define $\bar{\zeta}$ and σ as the mean value and standard deviation of the error indicator ζ over the first n_{ref} reference data sets
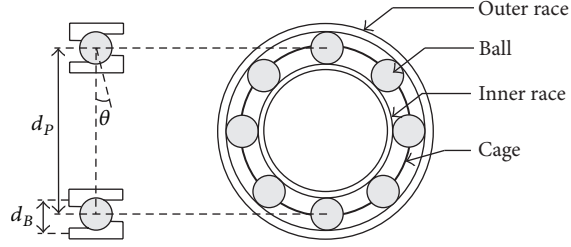
FIGURE 1: Schematic representation of a roller bearing.

for the safe structure. Then, an X-bar control chart [24] can be considered which consists in a centerline (CL) with upper and lower control limits (UCL and LCL) as follows:

$$CL = \overline{\zeta};$$
$$UCL = \overline{\zeta} + 3\sigma; \qquad (21)$$
$$LCL = \overline{\zeta} - 3\sigma.$$

By considering such a control chart, this yields an efficient means to assess the variation of a process and clearly identify damage when the damage indicator ζ exceeds the UCL [25].

3. Roller Bearing Damage Identification

3.1. Introduction. Rolling bearings are mechanical systems whose components—that is, rolling elements, inner raceway, outer raceway, and cage—usually induce a complex vibration behavior [26]. The vibration signature of a defective bearing is characterized by harmonics at particular bearing frequencies [27]. These harmonics can be identified from subspace identification methods but can be blurred by higher energy vibrations which are generated by other components of the same machine. Therefore a signal processing technique is required so as to clearly identify defects in rolling element bearings. This consists in applying a pass-band filter, which focuses on selecting a frequency region of interest.

3.2. Characteristic Frequencies. Generally, rolling bearings consist of two concentric rings, called the inner and outer races, with a set of rolling elements running on their tracks. Standard shapes of rolling elements include the ball, cylindrical roller, tapered roller, needle roller, and symmetrical and unsymmetrical barrel roller (see Figure 1). Typically, the rolling elements in a bearing are guided in a cage that ensures uniform spacing and prevents mutual contact.

There are four basic motions that are used to describe the bearing dynamics whose corresponding frequencies are called the bearing characteristic frequencies (BCFs). These frequencies relate the fundamental train frequency (FTF), the ball passing frequency inner race (BPFI), the ball passing frequency outer race (BPFO), and the ball spin frequency (BSF) and depend on the rotation speed. These are defined as follows [28]:

(i) The fundamental train frequency is related to the motion of the cage:

$$FTF = \frac{f_r}{2}\left(1 - \frac{d_B}{d_P}\cos(\theta)\right), \qquad (22)$$

where N_B, d_P, d_B, θ, and f_r are the number of balls, the pitch diameter, the ball diameter, the contact angle, and the rotation frequency of the bearing.

(ii) The ball passing frequency inner race indicates the rate at which the balls pass a point on the track of the inner race:

$$BPFI = \frac{N_B}{2}f_r\left(1 + \frac{d_B}{d_P}\cos(\theta)\right). \qquad (23)$$

(iii) The ball passing frequency outer race is defined as the rate at which the balls pass a point on the track of the outer race:

$$BPFO = \frac{N_B}{2}f_r\left(1 - \frac{d_B}{d_P}\cos(\theta)\right). \qquad (24)$$

(iv) The ball spin frequency is the rate of rotation of a ball about its own axis in a bearing:

$$BSF = \frac{d_P}{2d_B}f_r\left[1 - \left(\frac{d_B}{d_P}\cos(\theta)\right)^2\right]. \qquad (25)$$

Those four bearing characteristic frequencies can be determined provided that the following assumptions are satisfied [29]:

(1) The balls/rollers have the same diameter.

(2) The interactions between the balls, inner race, and outer race are only due to rolling contacts.

(3) There is no slipping between the shaft and the bearing.

(4) The outer race is fixed, while the inner race is in rotation.

3.3. Damage Identification Strategy. The subspace-based damage detection method can be achieved by filtering the vibration signals of a bearing around the BCFs. Here, a pass-band filtering technique is considered [27], which focuses on the consideration of narrow bands around the BCFs. The method is detailed as follows:

(i) It includes determination of the center frequency of each pass-band filter, defined as a specific BCF.

(ii) It includes determination of the bandwidth, around each BCF. Notice that the bandwidths should not overlap each other, that is, among the set of filters which are considered.

(iii) It includes generation of the filtered data, for each BCF.

(iv) It also includes calculation of the subspace-based damage indicator (Section 2.2) for each BCF.

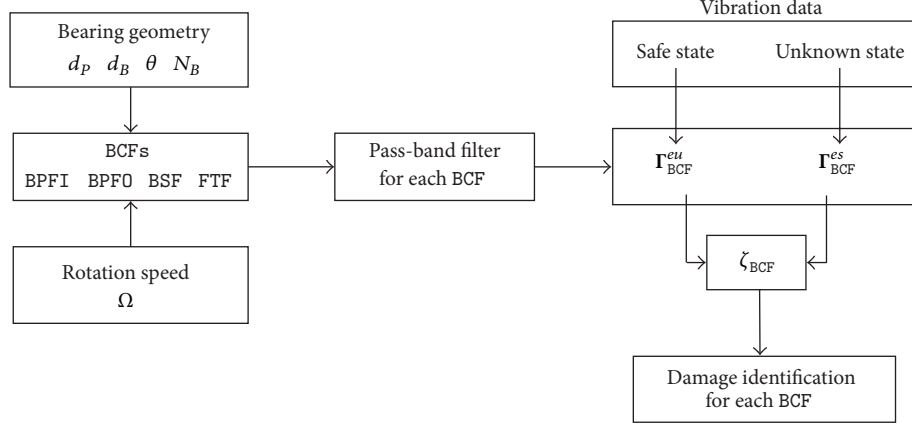The flowchart depicted in Figure 2 summaries the proposed methodology.

FIGURE 2: Flowchart of the subspace-based damage identification strategy.

TABLE 1: Parameters related to the simulated test signal $y_1(t)$.

Parameter	Sinusoidal element (N)	
(i)	1	2
$A_1(i)$ [m]	0.033	0.012
$f_1(i)$ [Hz]	250	40
$\theta_1(i)$ [rad]	0	0.32π

TABLE 2: Parameters related to the simulated test signal $y_2(t)$.

A_2	f_2	C	BCF	θ_2
(m)	(Hz)	(s^{-1})	(Hz)	(rad)
1.10^{-6}	2000	500	150	0

4. Numerical Simulation

4.1. Roller Bearing Damaged Model. A number of well-established models which describe the vibration signals produced by faulty bearings have already been proposed by using finite element model updating [30] or analyzing different physical effects [31–33]. In general, all of these models simulate bearing signals as a series of exponentially decaying high frequency oscillations, which appear repeatedly due to the contact between a fault and the mating surface and low-frequency phenomena which act as amplitude modulators. These signals can be generally decomposed into three components:

$$y(t) = y_1(t) + y_2(t) + n(t), \qquad (26)$$

where

(i) $y_1(t)$ represents the periodic component in the signal, given by

$$y_1(t) = \sum_{i=1}^{N} A_1(i) \sin\left[2\pi f_1(i) t + \theta_1(i)\right], \qquad (27)$$

where $A_1(i)$, $\theta_1(i)$, and $f_1(i)$ are the amplitude, initial phase, and frequency of the ith sinusoidal element, respectively. Here, two frequencies $f_1(1) = 40$ Hz and $f_1(2) = 250$ Hz are chosen to construct the periodic signal $y_1(t)$. The values of the parameters used are displayed in Table 1.

(ii) $y_2(t)$ represents the transient component of the signal. Physically, it reflects the evolution of the structural defects within the bearing, such as spalling on the surface of the bearing raceways or rolling elements. The transient component is modeled as a series of M exponentially attenuated vibrations, given by

$$y_2(t) = \sum_{i=1}^{M} A_2(i) \Theta(t - t_i) e^{-C(t-t_i)} \sin\left[2\pi f_2(t - t_i) + \theta_2\right], \qquad (28)$$

where $A_2(i)$, C, t_i, θ_2, and f_2 are the amplitude, attenuation factor, time-delay, initial phase, and frequency of the ith impact, respectively. $A_2(i)$ is modeled as a random variable that varies between 0 and A_2, while t_i is defined as $t_i = i \times 1/\text{BCF}$, where BCF is the monitored frequency. Also, $\Theta(t - t_i)$ in (28) is the conventional Heaviside function:

$$\Theta(t - t_i) = \begin{cases} 1, & t - t_i \geq 0 \\ 0, & t - t_i < 0. \end{cases} \qquad (29)$$

The values of the parameters used are displayed in Table 2.

(iii) $n(t)$ is a white Gaussian noise, with signal-to-noise ratio (SNR$_{\text{dB}}$) defined as

$$\text{SNR}_{\text{dB}} = 10 \times \log_{10}\left[\left(\frac{A_{\text{signal}}}{A_{\text{noise}}}\right)^2\right], \qquad (30)$$

where A is the root mean square amplitude of the signal and the noise.

The signal depicted in Figure 3 is obtained from the chosen parameters reported in Tables 1 and 2, with SNR$_{\text{dB}}$ = 100 dB and a signal length and sampling frequencies of 0.5 s and 10000 Hz, respectively.

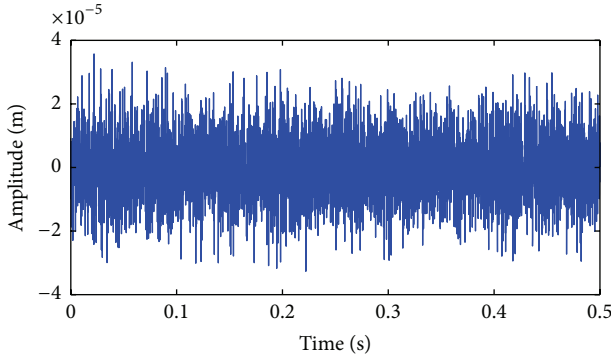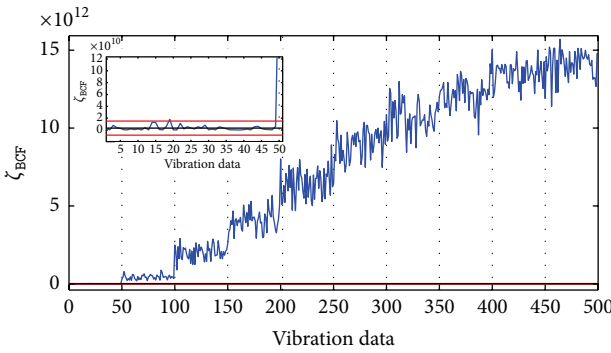Figure 3: Vibration data of the damaged bearing.

Figure 4: Evolution of the subspace-based damage indicator.

Table 3: Geometric parameters of the roller bearing.

Pitch diameter	Roller diameter	Contact angle	Number of rollers
d_P (mm)	d_B (mm)	(degree)	N_B
71.5	8.4	15.17	16

Table 4: Rexnord ZA-2115 BCFs for running speed 2000 RPM.

Bearing characteristic frequencies	Theoretical values (Hz)
FTF	14.78
BSF	280.08
BPFO	236.43
BPFI	296.90

4.2. Damage Identification. To validate the proposed detection procedure, 500 samples of bearing fault data are simulated. Every 50 samples, the amplitude of the fault A_2 in (28) is increased, starting from 10^{-7} m (reference state) until 10^{-6} m (maximum damage).

The BCF subspace damage indicator is obtained from the whole set of simulated data. Here, the number of reference data sets is fixed to $n_{\text{ref}} = 50$ and the filter bandwidth is chosen as [BCF $\pm$ 0.1BCF]. The order of the observability matrices is chosen as $i = 10$.

The results are reported in Figure 4 regarding the damage indicator ζ_{BSF}; see (17). It is shown that the damage is identified as soon as the amplitude of the fault increases from 10^{-7} m to 2×10^{-7} m (i.e., from the undamaged state to the first level of damage). The value of the damage indicator ζ_{BSF} increases in good agreement with the evolution of the amplitude of the fault, which fully gives credit to the proposed methodology.

5. Experimental Validation

5.1. Preliminary Comments. The proposed damage identification procedure is applied to an experimental test rig which hosts four bearings on one shaft as shown in Figure 5. The shaft is driven by an AC motor and is connected to rubber belts. A radial load of 6000 lbs is applied to the shaft and the bearing by means of a spring mechanism. A magnetic plug is

installed in the oil feedback pipe to collect debris, from the oil, as an evidence of bearing degradation.

High sensitive ICP piezoaccelerometers sensors are placed on the bearing housing, in radial horizontal direction as shown in Figure 5. The vibration data are collected every 10 minutes with a sampling rate of 20 kHz. Each data sample contains 20480 points. Data collection is conducted by a NI Labview program and is generated by IMS Center with support from Rexnord Corp. in Milwaukee, WI [34].

The 984 data pieces are recorded during 7 days, by considering sample of about 140 data pieces each day, until a significant amount of metal debris is found on the magnetic plug of the test bearing. The test stops until the accumulated debris adherent to the magnetic plug exceeds a certain threshold. At this time, a visual inspection is made. All failures occurred once the designed lifetime of the bearing is reached which is 100 million revolutions.

Four Rexnord ZA-2115 double row bearings are mounted on the shaft. According to the geometric parameters of the bearing listed in Table 3, the nominal BCFs are calculated for a constant rotation speed of 2000 RPM and listed in Table 4.

5.2. Run-to-Failure Test. From visual inspection, an outer race defect is discovered in bearing 1. A sample of temporal and frequency representations of the data set is shown in Figure 6.

Subspace damage indicators are obtained for each accelerometer and each BCF from the whole set of accelerometers data. Here, the number of reference data sets is $n_{\text{ref}} = 140$ and the filter bandwidth is chosen as [BCF $\pm$ 0.1BCF]. The order of the observability matrices (3) is chosen as $i = 12$ by analyzing the singular values of $\mathbf{\Gamma}_i$ (as shown in Figure 7).

Results of subspace damage indicators ζ_{FTF}, ζ_{BSF}, ζ_{BPFI}, and ζ_{BPFO} are reported in Figures 8, 9, 10, and 11, respectively. A comparison of results for each BCF and each accelerometer position shows that the first bearing is identified as damaged approximately between two and three days before the experiment is stopped. It is seen that the subspace damage indicator ζ_{BPFO} on accelerometer 1 has highest level and earliest sensitivity. According to this result, it is thus expected that the outer ring of the first bearing is the most damaged
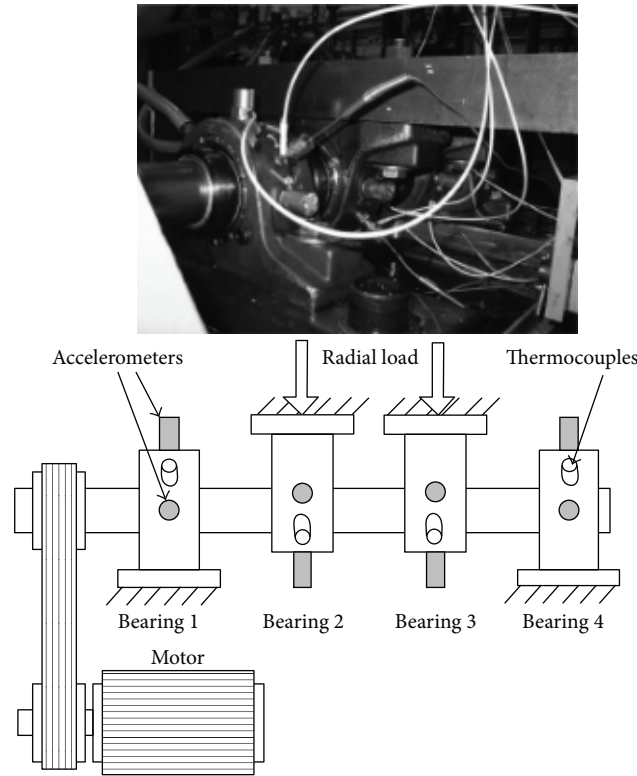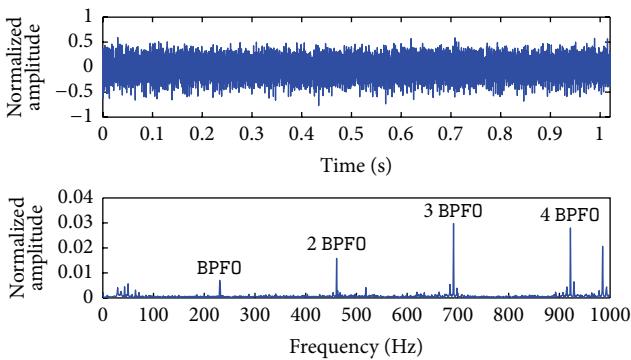
FIGURE 5: Experimental bearing test rig.

FIGURE 6: Normalized amplitude of the acceleration over time and frequency domains for the 900th data sample on accelerometer 1.
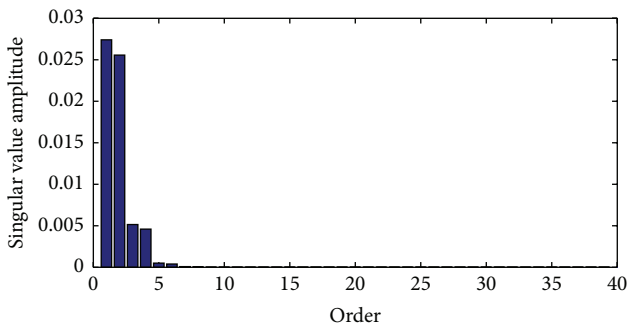

FIGURE 7: Model order selection (here $2n = 6$).

component, which is successfully demonstrated by visual inspection.

It can be shown that ζ_{BPFO} also increases on accelerometers 3 and 4 probably because the dynamic behavior of the damaged bearing in position 1 is also recorded on the whole set of accelerometers or because bearings 3 and 4 start undergoing some deterioration, which seems to be the most probable situation although it is not validated by visual inspection at the end of the test.

6. Conclusion

A damage identification procedure has been proposed which makes use of a subspace method combined with a pass-band data filtering technique. Within this framework, several damage indicators have been considered in the diagnosis of faults in rotating machines. The present methodology has been successfully applied to identify damage in a roller bearing, by considering simulated data. Also, the efficiency of the method has been highlighted regarding an experimental test that consists in monitoring real bearings. It has been shown that a roller bearing defect can be detected at an early stage with accurate precision. Future works may concern the comparison of this new indicator with other methods and extend the diagnosis to other rotating machinery components such as gearbox.
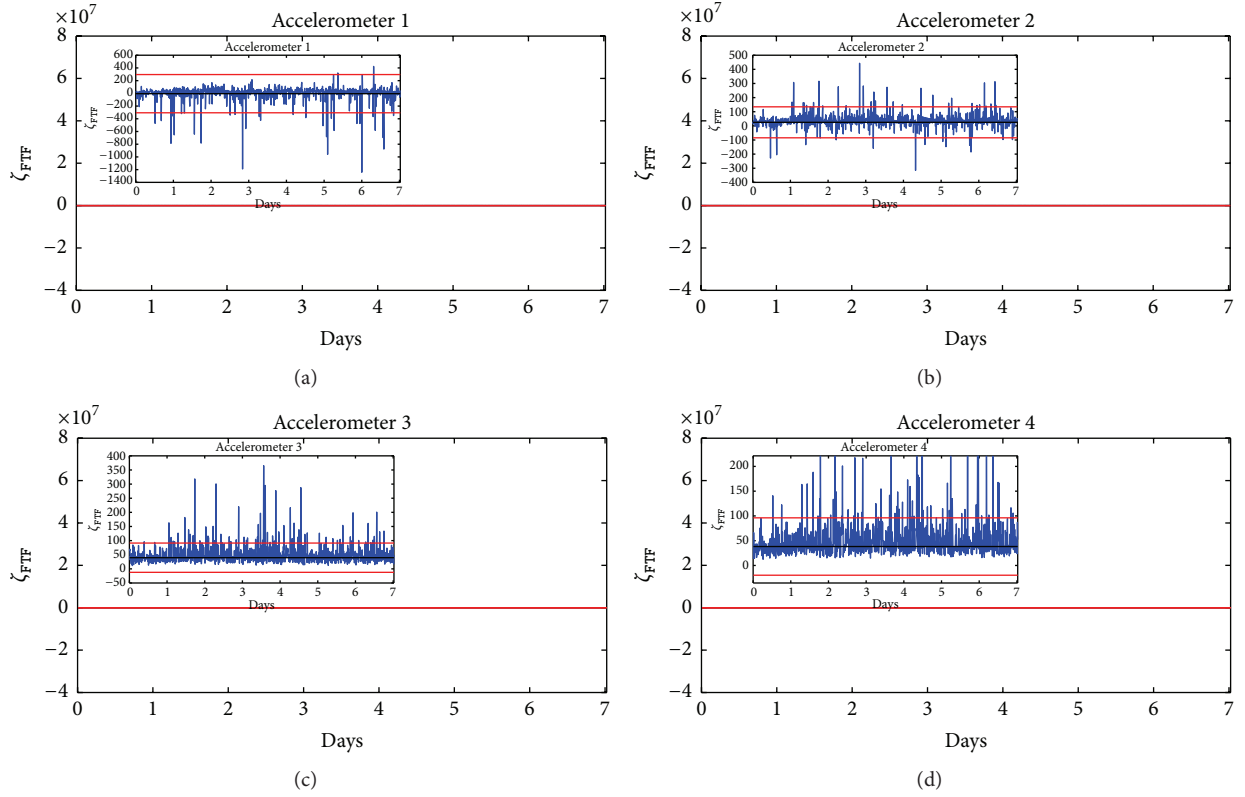
FIGURE 8: Damage indicators associated with the FTF frequency from each accelerometer.
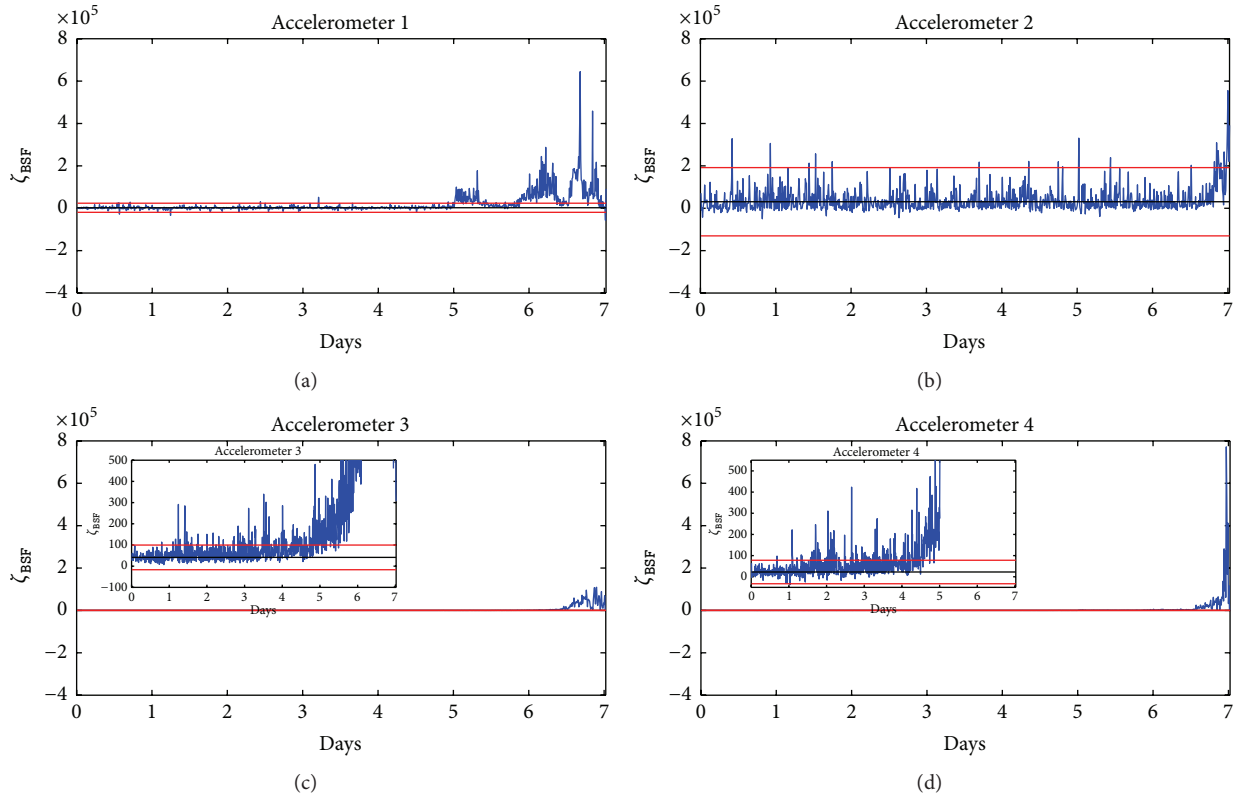

FIGURE 9: Damage indicators associated with the BSF frequency from each accelerometer.
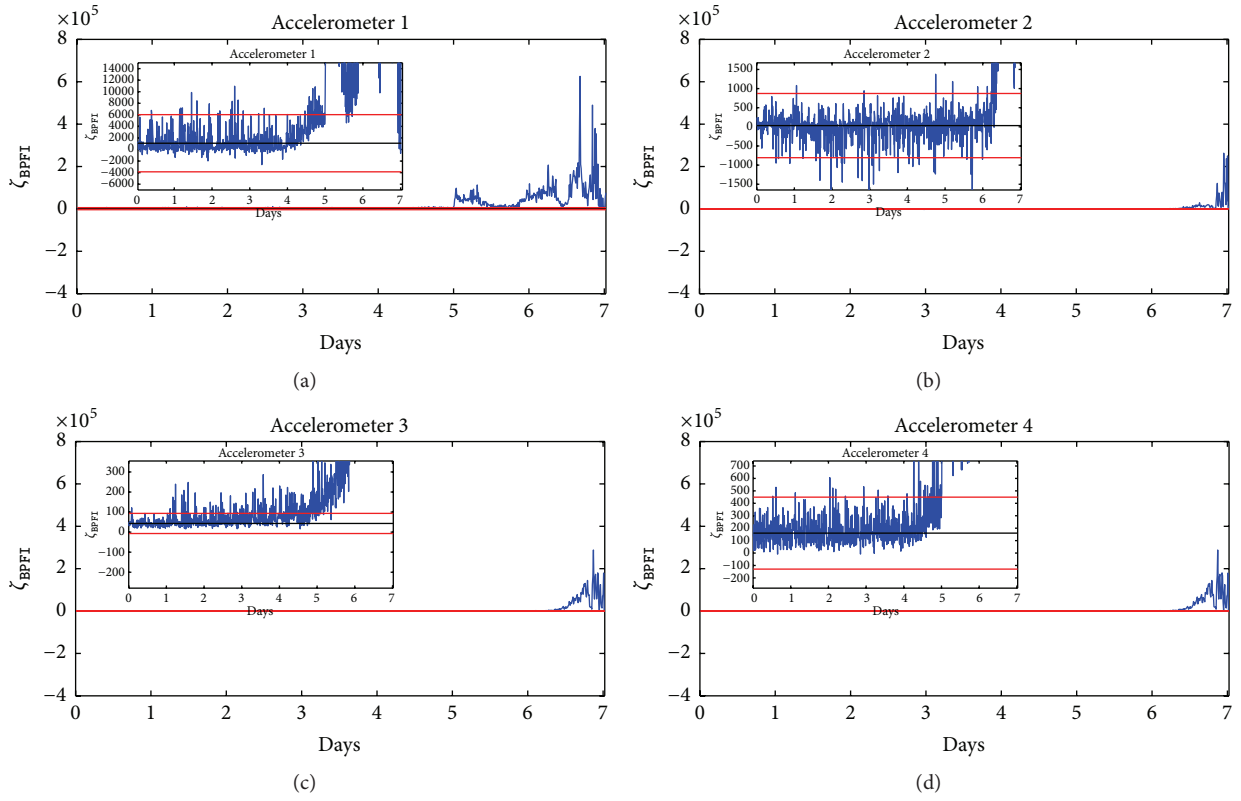
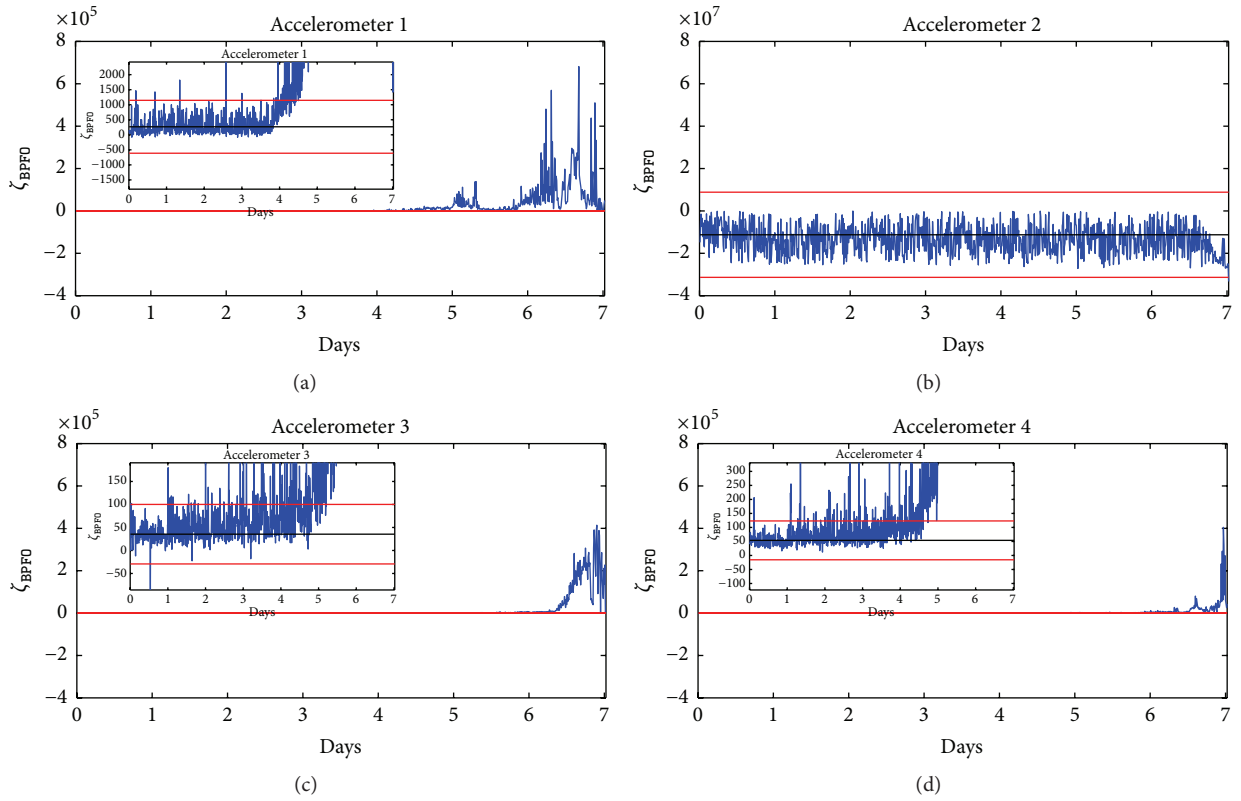Figure 10: Damage indicators associated with the BPFI frequency from each accelerometer.

Figure 11: Damage indicators associated with the BPFO frequency from each accelerometer.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.
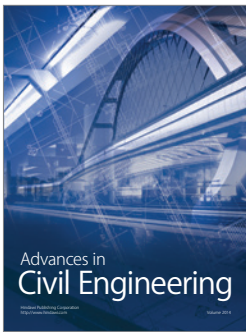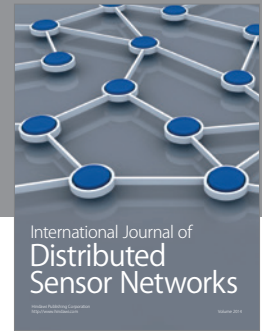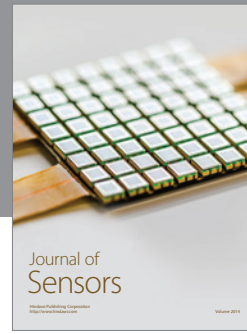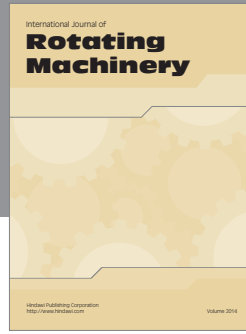
Acknowledgments

References

[1] S. W. Doebling, C. R. Farrar, M. B. Prime, and D. W. Shevitz, *Damage Identification and Health Monitoring of Structural and Mechanical Systems from Changes in Their Vibration Characteristics: A Literature Review*, Los Alamos National Laboratory, Los Alamos, NM, USA, 1996.

[2] A.-M. Yan and J.-C. Golinval, "Null subspace-based damage detection of structures using vibration measurements," *Mechanical Systems and Signal Processing*, vol. 20, no. 3, pp. 611–626, 2006.

[3] H. Sohn, C. Farrar, F. Hemez et al., "A review of structural health monitoring literature: 1996–2001," Tech. Rep. LA-13976-MS, Los Alamos National Laboratory, Los Alamos, NM, USA, 2004.

[4] I. S. Bozchalooi and M. Liang, "A joint resonance frequency estimation and in-band noise reduction method for enhancing the detectability of bearing fault signals," *Mechanical Systems and Signal Processing*, vol. 22, no. 4, pp. 915–933, 2008.

[5] H. Saruhan, S. Saridemir, A. Çiçek, and I. Uygur, "Vibration analysis of rolling element bearings defects," *Journal of Applied Research and Technology*, vol. 12, no. 3, pp. 384–395, 2014.

[6] Z. Kiral and H. Karagülle, "Vibration analysis of rolling element bearings with various defects under the action of an unbalanced force," *Mechanical Systems and Signal Processing*, vol. 20, no. 8, pp. 1967–1991, 2006.

[7] J. P. Dron, F. Bolaers, and L. Rasolofondraibe, "Improvement of the sensitivity of the scalar indicators (crest factor, kurtosis) using a de-noising method by spectral subtraction: application to the detection of defects in ball bearings," *Journal of Sound and Vibration*, vol. 270, no. 1-2, pp. 61–73, 2004.

[8] P. D. McFadden and J. D. Smith, "Vibration monitoring of rolling element bearings by the high-frequency resonance technique a review," *Tribology International*, vol. 17, no. 1, pp. 3–10, 1984.

[9] N. Sawalhi, R. B. Randall, and H. Endo, "The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2616–2633, 2007.

[10] J. Antoni, "The spectral kurtosis: a useful tool for characterising non-stationary signals," *Mechanical Systems and Signal Processing*, vol. 20, no. 2, pp. 282–307, 2006.

[11] P. W. Tse, Y. H. Peng, and R. Yam, "Wavelet analysis and envelope detection for rolling element bearing fault diagnosis—their effectiveness and flexibilities," *Transactions of the ASME—Journal of Vibration and Acoustics*, vol. 123, no. 3, pp. 303–310, 2001.

[12] C. T. Yiakopoulos and I. A. Antoniadis, "Wavelet based demodulation of vibration signals generated by defects in rolling element bearings," *Shock and Vibration*, vol. 9, no. 6, pp. 293–306, 2002.

[13] D. Yu, J. Cheng, and Y. Yang, "Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings," *Mechanical Systems and Signal Processing*, vol. 19, no. 2, pp. 259–270, 2005.

[14] J. Cheng, D. Yu, J. Tang, and Y. Yang, "Application of SVM and SVD technique based on EMD to the fault diagnosis of the rotating machinery," *Shock and Vibration*, vol. 16, no. 1, pp. 89–98, 2009.

[15] T. R. Lin, E. Kim, and A. C. C. Tan, "A practical signal processing approach for condition monitoring of low speed machinery using Peak-Hold-Down-Sample algorithm," *Mechanical Systems and Signal Processing*, vol. 36, no. 2, pp. 256–270, 2013.

[16] M. Verhaegen, "Identification of the deterministic part of MIMO state space models given in innovations form from input-output data," *Automatica*, vol. 30, no. 1, pp. 61–74, 1994.

[17] P. Van Overschee and B. De Moor, "N4SID: subspace algorithms for the identification of combined deterministic-stochastic systems," *Automatica*, vol. 30, no. 1, pp. 75–93, 1994.

[18] N. Viet-Ha and J.-C. Golinval, "Localization and quantification of damage in beam-like structures using sensitivities of principal component analysis results," *Mechanical Systems and Signal Processing*, vol. 24, no. 6, pp. 1831–1843, 2010.

[19] S. Y. Nalakath, K. Saeed, and K.-A. Alnefaie, "Detection of a Notch type damage using subspace identification and artificial neural networks," *International Journal of Emerging Technology and Advanced Engineering*, vol. 4, no. 8, pp. 406–411, 2014.

[20] G. Gautier, R. Serra, and J.-M. Mencik, "Subspace-based damage identification of roller bearing," *Matec Web of Conferences*, vol. 20, 2015.

[21] J. Lardies and T. Minh-Ngi, "Modal parameter identification of stay cables from output-only measurements," *Mechanical Systems and Signal Processing*, vol. 25, no. 1, pp. 133–150, 2011.

[22] M. Viberg, B. Wahlberg, and B. Ottersten, "Analysis of state space system identification methods based on instrumental variables and subspace fitting," *Automatica*, vol. 33, no. 9, pp. 1603–1616, 1997.

[23] G. Gautier, J.-M. Mencik, and R. Serra, "A finite element-based subspace fitting approach for structure identification and damage localization," *Mechanical Systems and Signal Processing*, vol. 58-59, pp. 143–159, 2015.

[24] H. Sohn and C. R. Farrar, "Damage diagnosis using time series analysis of vibration signals," *Smart Materials & Structures*, vol. 10, no. 3, article 446, 2001.

[25] A.-M. Yan, G. Kerschen, P. De Boe, and J.-C. Golinval, "Structural damage diagnosis under varying environmental conditions—part I: a linear analysis," *Mechanical Systems and Signal Processing*, vol. 19, no. 4, pp. 847–864, 2005.

[26] R. Tiwari, A. W. Lees, and M. I. Friswell, "Identification of dynamic bearing parameters: a review," *The Shock and Vibration Digest*, vol. 36, no. 2, pp. 99–124, 2004.

[27] A.-G. Hassaan, "Frequency spectrum filtering for machinery fault diagnostics," *International Journal of Scientific & Technology Research*, vol. 3, no. 8, pp. 200–203, 2014.

[28] B. P. Graney and K. Starry, "Rolling element bearing analysis," *Materials Evaluation*, vol. 70, no. 1, pp. 78–85, 2012.

[29] A. Selvaraj, R. Marappan, and P. Govindarajan, "Effect of slip in performance of rolling element bearings," SAE Technical Paper 2004-28-0043, 2004.

[30] G. Gautier, R. Serra, and J.-M. Mencik, "Vibratory diagnosis by finite element model updating and operational modal analysis," *Mechanics & Industry*, vol. 14, no. 2, pp. 145–149, 2013.

[31] R. B. Randall, J. Antoni, and S. Chobsaard, "Comparison of cyclostationary and envelope analysis in the diagnostics of rolling element bearings," in *Proceedings of the IEEE Interntional Conference on Acoustics, Speech, and Signal Processing (ICASSP '00)*, vol. 6, pp. 3882–3885, June 2000.

[32] J. Antoni, F. Bonnardot, A. Raad, and M. El Badaoui, "Cyclostationary modelling of rotating machine vibration signals," *Mechanical Systems and Signal Processing*, vol. 18, no. 6, pp. 1285–1314, 2004.

[33] M. Zhao, J. Lin, X. Xu, and X. Li, "Multi-fault detection of rolling element bearings under harsh working condition using IMF-based adaptive envelope order analysis," *Sensors*, vol. 14, no. 11, pp. 20320–20346, 2014.

[34] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *Journal of Sound and Vibration*, vol. 289, no. 4-5, pp. 1066–1090, 2006.