Available online at www.sciencedirect.com
ScienceDirect
Journal of Sound and Vibration 317 (2008) 112–126
JOURNAL OF SOUND AND VIBRATION
www.elsevier.com/locate/jsvi

Bearing condition diagnosis and prognosis using applied nonlinear dynamical analysis of machine vibration signal

S. Janjarasjitt[a,b,*], H. Ocak[c], K.A. Loparo[b]

[a]*Department of Electrical and Electronic Engineering, Ubon Ratchathani University, Warinchamrab, Ubon Ratchathani 34190, Thailand*
[b]*Department of Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, OH 44106, USA*
[c]*Mechatronics Engineering Department, Kocaeli University, 41040 Izmit, Turkey*

Received 31 July 2007; received in revised form 25 February 2008; accepted 27 February 2008
Handling Editor: S. Bolton
Available online 11 April 2008

Abstract

In this paper, we introduce a modified form of the correlation integral developed by Grassberger and Procaccia referred to as the partial correlation integral, which can be computed in real time. The partial correlation integral algorithm is then used to analyze machine vibration data obtained throughout a life test of a rolling element bearing. From the experimental results, the dimensional exponent (an approximation of the correlation dimension) as computed from the partial correlation integral algorithm tends to increase as time progresses and the useful remaining life of the bearing is decreasing. The dimensional exponents of a healthy bearing and a bearing close to failure are statistically different. We also propose a computational scheme for bearing condition monitoring (diagnosis and prognosis) using the dimensional exponent integrated with a surrogate data testing technique. As a result, we can characterize the condition of the bearing from the results of the surrogate data test and furthermore, we provide some preliminary evidence that the dimensional exponent can be used to predict the failure of rolling element bearings in rotating machinery from real-time vibration data.

1. Introduction

Motor systems converting almost 60 percent of the electricity produced in the US into other forms of energy to provide power to other equipment are very important in today's modern society [1]. Bearings are one of the most important and frequently encountered elements in motor systems and play an important role in the proper operation of these systems. Even though a bearing is a very inexpensive element, its failure can interrupt the production in a plant causing unscheduled downtime and production losses. A bearing failure has the potential to also damage machinery causing soaring machinery repair and/or replacement costs. As a result, condition monitoring of bearings has been the subject of extensive research for the last 25 years. Condition monitoring can help prevent catastrophic failures in critical rotating machinery such as large

*Corresponding author. Tel.: +66 4535 3332; fax: +66 4535 3333.
E-mail addresses: suparerk.janjarasjitt@case.edu (S. Janjarasjitt), ocak@case.edu (H. Ocak), kenneth.loparo@case.edu (K.A. Loparo).

0022-460X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsv.2008.02.051

motors and pumps in power generation plants, by determining optimal maintenance schedules for the plant and thus avoiding unscheduled downtime and machine failure costs.

A variety of analysis methods exist for condition monitoring of bearings including: vibration analysis, oil analysis, infrared thermography and motor current signature analysis. Of all these methods vibration analysis, see for example Ref. [2] and the references therein, is the most commonly used method and the one that provides the most information from the data acquired. Investigating the statistical parameters derived from the time domain signals is the simplest vibration analysis technique. It has been shown that vibration signals from good and defective bearings manifest statistically different behaviors in the time domain, such as different peak, root-mean square (rms), crest factor, kurtosis, skewness values [3,4] and cyclic spectral analysis [5]. Rolling element bearings have time-varying stiffness characteristics [6] and periodic impulses are generated as the rollers pass over bearing defects with a frequency that is characteristic of the defective element. High frequency resonance analysis (also known as envelope analysis) [7], bicoherence analysis [8] and other frequency domain methods have been used for the identification of these major frequency components in the vibration spectrum for detecting localized defects. Advanced techniques such as time-frequency analysis and the wavelet transform [9–11], neural networks [12,13] and recently hidden Markov models [14,15] have also been employed in machinery vibration analysis. Nevertheless, especially as a machine fails, the vibration characteristics of the machine are changing and nonlinear dynamical analysis techniques may be the most suitable methods for various aspects of machinery analysis including fault detection and diagnosis.

In fact, machine processes often contain complex, nonstationary, noisy and nonlinear characteristics [16]. Such characteristics may range from quasi-periodic to completely irregular behaviors [16]. Therefore, nonlinear dynamical analysis techniques such as the correlation dimension, Lyapunov exponents and Kolmogorov entropy can be used to investigate the nonlinear dynamic behavior and complex generated by these machines over their operating life time. The correlation integral and dimension developed by Grassberger and Procaccia [17] are two of the most popular nonlinear techniques in the field of nonlinear time series analysis [18]. The correlation dimension is derived from computation of the correlation integral and specifies the active degrees of freedom or the complexity of the system operating in a nonlinear regime on the attractor.

In this paper, we present a modified computational algorithm for the correlation integral developed by Grassberger and Procaccia, called the *partial* correlation integral [19]. The partial correlation integral algorithm aims to improve two major drawbacks of the Grassberger–Procaccia algorithm: the extensive computational requirements and the sensitivity of the results to nonstationary characteristics of the time series data. The partial correlation integral algorithm can be used to calculate the dimensional exponent, which is an estimate of the correlation dimension [19].

The partial correlation integral is applied to bearing vibration data to examine the nonlinear dynamics of a bearing system throughout the life of a bearing. The experimental and computational results exhibit that there is a change of the dimensional exponent of the bearing vibration data as the life of the bearing is consumed. The dimensional exponent tends to increase as time progresses and the bearing heads toward failure. This implies that the bearing vibration time series is becoming more complex (possibly stochastic) as the remaining life of the bearing decreases. The dimensional exponent specifies the margin between a healthy and failing bearing and as such, the dimensional exponent can be used to monitor the health and condition of the bearing from the vibration data.

2. Nonlinear dynamical analysis

2.1. Attractors of a dynamical system

Grassberger and Procaccia [17] introduced the correlation integral and the concept of correlation dimension D_2 to quantify the active degrees of freedom or the complexity of a dynamical system on an attractor. Let $s[n]$ be a real-valued observation sequence of a dynamical system where $n = 0, 1, \ldots, N-1$. This observable time series cannot generally provide complete state information for the dynamical system, however a more complete description can be obtained by unfolding the attractor into a higher dimensional (embedding) space $\mathscr{R}^m$ using the time-delay embedding scheme (Takens' reconstruction [20]) as developed in Ref. [21].

The m-dimensional embedding vector $\mathbf{s} \in \mathscr{R}^m$ generated from the time series s is given by [22]

$$\mathbf{s}_n = (s[n] \;\; s[n+\tau] \;\; \cdots \;\; s[n+(m-1)\tau])^{\mathrm{T}}, \tag{1}$$

where $n = 0, 1, \ldots, N_e = N - (m-1)\tau$. Here m and τ are the embedding dimension and the time delay, respectively, and $\cdot^{\mathrm{T}}$ denotes vector transpose. The time-delay embedding method unfolds the attractor into the m-dimensional embedding space and provides a more comprehensive representation of the dynamics of the system.

2.2. Correlation integral and dimension

The correlation integral $C(r)$ [17] of the time series s is defined by

$$C(r) = \lim_{N_c \to \infty} \frac{1}{N_c} H_c(r), \tag{2}$$

where N_c denotes the total number of pairs of points $(\|\mathbf{s}_i, \mathbf{s}_j\|)$, $i,j \in (0, 1, \ldots, N_e - 1)$ on the reconstructed attractor. The distribution of distances $H_c(r)$ specifies the total number of the pairs of embedding vectors $(\mathbf{s}_i, \mathbf{s}_j)$ such that the distance between the embedding vectors $\mathbf{s}_i$ and $\mathbf{s}_j$ is less than the specified distance r, i.e.,

$$H_c(r) = 2 \sum_{i=0}^{N_e-1} \sum_{j=i+1}^{N_e-1} \Theta(r - \|\mathbf{s}_i - \mathbf{s}_j\|), \tag{3}$$

where the Heaviside function $\Theta(n) = 1$ if $n \geqslant 0$; 0 otherwise. Practically, the correlation integral $C(r)$ measures the probability that points on the attractor have pairwise distances less than or equal to the corresponding distance r.

A revised algorithm [22,23] corrects for autocorrelation effects by modifying the distribution of distances $H_c(r)$ according to:

$$H_c(r) = 2 \sum_{i=0}^{N_e-1} \sum_{j=i+w}^{N_e-1} \Theta(r - \|\mathbf{s}_i - \mathbf{s}_j\|), \tag{4}$$

where $N_c = (N_e - w + 1)(N_e - w)/2$ and the revised correlation integral calculation given in Eq. (4) is the same as the Grassberger–Procaccia algorithm if $w = 1$.

The correlation integral $C(r)$ behaves as a power of v for small r [17], i.e.,

$$C(r) \propto r^v. \tag{5}$$

The exponent v for small r defines the correlation dimension D_2 and can be calculated by

$$v = \lim_{r \to 0} \frac{\log(C(r))}{\log(r)}. \tag{6}$$

The computational complexity of a direct implementation of the Grassberger–Procaccia algorithm is $\mathcal{O}(N_e^2)$ if the length N_e is large and dependence on the distance vector is neglected.

2.3. Surrogate data testing

The method of surrogate data [24] is one of the most popular tests that have been developed to examine the evidence of nonlinearity in a time series. The use of surrogate data is important when nonlinear time series analysis techniques, such as the correlation integral, are being used to analyze data. Nonlinear time series analysis techniques can often misinterpret linear correlations for nonlinear behavior [25] and the purpose of the surrogate data analysis is to validate certain aspects of the computational analysis of the time series data. In particular when nonlinear time series analysis methods such as correlation dimension are being used, the objective is to investigate if the computational method is truly picking up nonlinear features in the data, or if standard more conventional linear time series analysis methods could accomplish the same results.

In our application, the surrogate data are artificial data that are computationally generated from the original bearing vibration data. The surrogate data sets are generated using the method of iteratively refined surrogates, which is one of the computational approaches for generating a constrained surrogate data realization. Using this approach, the surrogate data resemble a linear Gaussian random process and have the same power spectrum and amplitude distribution as the original bearing vibration data. The comparison of the dimensional exponents from the original bearing vibration data and the corresponding surrogate data will provide evidence for the existence of nonlinearity in the original bearing vibration data. This existence of nonlinearity is investigated by setting up a hypothesis testing problem, with the null hypothesis being that the dimensional exponents from the original and surrogate data are drawn from the same distribution [26]. If the null hypothesis of the surrogate data testing can be rejected, this means that the nonlinear time series measures of the original time series and its corresponding surrogate time series are statistically different with a desired level of significance. This implies that the original time series are not generated from a linear Gaussian random process and that the nonlinear time series measures computed from the original time series can quantify nonlinear characteristics in the data that are statistically different from linear stochastic characteristics.

There are a number of methods of surrogate data generation that use either traditional bootstrap methods or a constrained realizations approach. One of the surrogate data generation methods using constrained realizations is referred to as the iteratively refined surrogate data method [27]. The surrogate data are generated by randomizing the time points of the data with iterative corrections to minimize deviations in the power spectrum and amplitude distribution of the surrogate data as compared to the original time series [25]. As a result, surrogate data generated by the method of iteratively refined surrogates have approximately the same power spectrum and amplitude distribution as the original time series. The method of iteratively refined surrogates provides a substantial improvement and more accuracy over other Fourier-based (frequency domain) surrogate generation methods that use other constrained realizations approaches [27].

3. The partial correlation integral

Because the attractor of a stable dynamical system is confined to a bounded subset of the phase space, a portion of the attractor referred to as a partial-attractor can also provide a reasonable description of the dynamical system. The Grassberger–Procaccia algorithm basically measures all possible pairwise distances of the points on the attractor while the partial correlation integral algorithm measures only pairwise distances of the points on the partial-attractor. The partial correlation integral is defined by [19]

$$C_\rho(r) = \frac{1}{N_\rho} H_\rho(r), \tag{7}$$

where N_ρ denotes the total number of pairwise distances to be used in the calculation. The distribution of distances $H_\rho(r)$ is given by

$$H_\rho(r) = \sum_{i=0}^{N_e} \sum_{j=i+w}^{i-1+(w+\rho)} \Theta(r - \|\mathbf{s}_i - \mathbf{s}_j\|), \tag{8}$$

where the integral limit of the partial correlation integral ρ, usually specified as a multiple of the embedding dimension m, is an arbitrary positive integer, $N_e = N_e - (w + \rho)$ and $N_\rho = \rho(N_e + 1)$ for computation of the partial correlation integral.

The partial correlation integral $C_\rho(r)$ provides a good approximation of Grassberger–Procaccia's correlation integral $C(r)$ and also behaves as a power law for small r. The exponent v_ρ, defined as the dimensional exponent, is also a good approximation to the correlation dimension D_2. The partial correlation integral algorithm is less computationally complex than the Grassberger–Procaccia algorithm. The computational complexity of the partial correlation integral algorithm is linear in the length N_e, i.e., $\mathcal{O}(N_e)$, if the length N_e is large and the dependency on the distance vector $\mathbf{r}$ is neglected, and is thus suitable for real-time implementation.

4. Bearing condition monitoring: diagnosis and prognosis

The partial correlation integral algorithm in combination with surrogate data testing provide a computational framework for bearing condition monitoring. The proposed technique does not require any information about the type of bearing, operating conditions such as speed and load, fault type, etc., and hence has the potential of a bearing diagnosis and prognosis method that could be applied in practice. A diagram of the proposed bearing condition diagnosis and prognosis scheme is illustrated in Fig. 1.

The proposed bearing condition diagnosis and prognosis scheme is composed of two main elements: surrogate data testing and bearing condition identification. In the surrogate data testing, the partial correlation integrals are computed for a set of bearing vibration time series data and the corresponding surrogate time series. The dimensional exponents are then estimated and a statistical significance test is used to compare the dimensional exponents of actual bearing vibration data and the surrogate data group. The results of the statistical significance test over a period of time are used for bearing condition identification.

5. Results

The partial correlation integral algorithm is used to analyze bearing vibration data collected throughout the life of a bearing in a laboratory test setup without any maintenance intervention.

5.1. Experimental settings

The bearing test system incorporates two test bearings as shown in Fig. 2. The bearings can run at speeds up to 10 000 rev/min and temperatures up to 1000 F. In this study, two new 6204 2RS1 SKF bearings were mounted on the main drive shaft of the test system. An axial load of 154.22 kg was applied to the bearings while the operating temperature was set to 260 F. The estimated L_{10} life of the bearing is 313 million revolutions which at the operational speed of the system corresponds to about 21.71 days. Two PCP 353C65

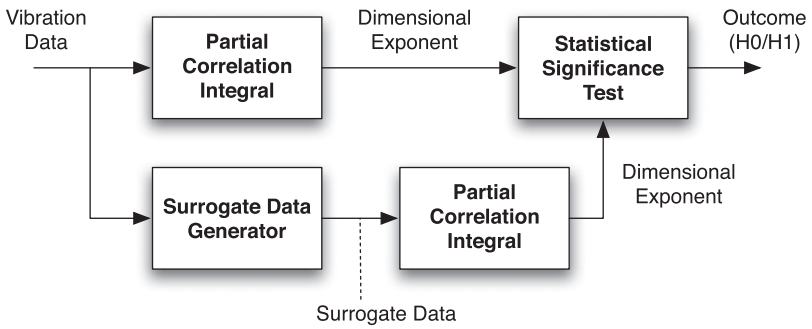

Fig. 1. The proposed bearing condition diagnosis and prognosis scheme.

Fig. 2. The experimental setup of the bearing test system.

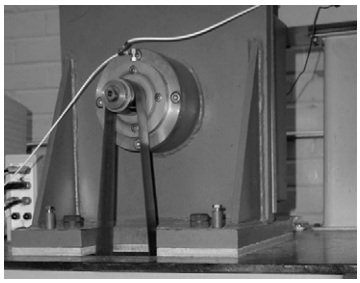

Fig. 3. The location of the PCP 353C65 sensors mounted to the bearing retainer hub.

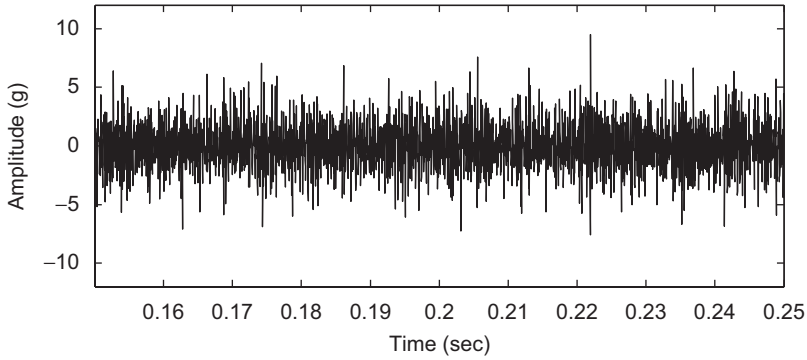

Fig. 4. An exemplary 3000-point bearing vibration data corresponding to the healthy bearing (during the second hour of the experiment).

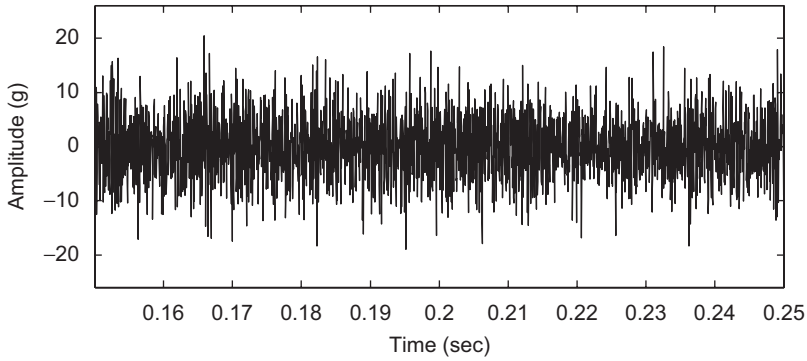

Fig. 5. An exemplary 3000-point bearing vibration data approaching the bearing failure (during the final hour of the experiment).

sensors were mounted to capture bearing vibration data from the test bearing as shown in Fig. 3. The first sensor was stud mounted to the bearing retainer hub and the second sensor was magnetically mounted also on the bearing retainer hub symmetric (with respect to the vertical axis) of the first sensor. Only the bearing vibration data from the stud mounted sensor are used for the analysis given in this paper.

The experimental vibration data were recorded using a National Instruments data acquisition card with 16 bit A/D conversion and a sampling rate of 24 kHz. Every 2 h, 8 s of bearing vibration data were acquired from the beginning of the experiment until the bearing failed. During the experiment, the rotational speed of the drive motor was approximately 3590 rev/min while the bearing inner-race speed was approximately 10 012 rev/min. Examples of a section of the bearing vibration data during the second hour and the final hour of the experiment are illustrated in Figs. 4 and 5, respectively.

5.2. Calculation of the partial correlation integral and dimensional exponents

The partial correlation integral algorithm is applied to epochs of bearing vibration data containing 2400 samples. The time-delay embedding parameters used in the experiment are as follows: the embedding dimension $m = 18$ and the time delay $\tau = 1$. The embedding parameters, m and τ are chosen by using the false nearest neighbor method [28] and the autocorrelation technique [29], respectively. The average number of false neighbors of the bearing vibration data significantly drops to nearly zero around the embedding dimension $m = 18$, while the autocorrelation function of the bearing vibration data drops to $1/e$ of its initial value around $\tau = 1$. The integral limit ρ was set to 600 samples (25 ms). Before the partial correlation integral is computed, each epoch of the vibration data is normalized to have magnitude equal to 1, i.e.,

$$s_n = \frac{s}{\max(s) - \min(s)}, \tag{9}$$

where s and s_n denote the original epoch of vibration data and the normalized epoch of vibration data, respectively.

Let $\mathbf{r} = \{r[n]\}$, $n = 0, 1, \ldots, N_r$, denote a vector of distances used in the calculation of the partial correlation integral and defined by

$$\mathbf{r} = \{r[n] = r_0 \gamma^{n/N_r}\}_{n=0}^{N_r}, \tag{10}$$

where $\gamma = r_1/r_0$ and the real numbers r_0 and r_1 are the smallest distance and the largest distance in the distance vector $\mathbf{r}$, respectively. In the experiment, the distance $r_0 = 0.24$ and the distance $r_1 = 1.02$, which cover the essential region of the partial correlation integral including the scaling region and the peak of the distribution of distances. The length of the distance vector $\mathbf{r}$ is $N_r = 800$. For each value of the distance vector $r[n]$, the partial correlation integral $\mathbf{C}_\rho[n]$ is calculated corresponding to the distance $r[n]$, as follows:

$$\mathbf{C}_\rho[n] = C_\rho(r[n]). \tag{11}$$

The estimated dimensional exponent is determined from computation of the local slopes of the partial correlation integral. The local slopes are calculated using a linear least-squares regression estimator. The local partial correlation integrals $\tilde{C}_{\rho,i}[n]$ are defined by

$$\tilde{C}_{\rho,i}[n] = \mathbf{C}_\rho[i + n - 1], \tag{12}$$

where $n = 0, 1, \ldots, \tilde{N}_i$ and $i = i_l, i_l + 1, \ldots, i_u$. The lower and upper estimation limits are given by

$$i_l = c_0 + \kappa_l(c_1 - c_0), \tag{13}$$

$$i_u = c_1 - \kappa_u(c_1 - c_0), \tag{14}$$

where $c_0 = \arg_k \min \mathbf{C}_\rho[k] > 0$, $c_1 = \arg_k \mathbf{C}_\rho[k] = \max \mathbf{C}_\rho$, and the coefficients $\kappa_l \in [0, 1]$ and $\kappa_u \in [0, 1]$. In the experiment, $\kappa_l = 0.24$ and $\kappa_u = 0.18$ to focus on the real scaling region. The local distance vectors $\tilde{r}_i[n]$ are further defined by

$$\tilde{r}_i[n] = r[i + n - 1], \tag{15}$$

where $n = 0, 1, \ldots, \tilde{N}_i$. Adopted from the optimal length of the chord estimator given in Ref. [30], the length of the ith local correlation integral is given by

$$\tilde{N}_i = \arg\min_k \frac{\mathbf{C}_\rho[i + k - 1]}{\mathbf{C}_\rho[i]} \geqslant 5. \tag{16}$$

The linear regression model of the ith local partial correlation integral $\tilde{C}_{\rho,i}[\tilde{r}_i]$ is then given by

$$\log(\tilde{C}_{\rho,i}) = \hat{\nu}_{\rho,i}^{\mathrm{LS}} \log(\tilde{r}_i) + \beta_i + \varepsilon_i, \tag{17}$$

where the estimate of the slope of the ith local partial correlation integral, $\hat{\nu}_{\rho,i}^{\mathrm{LS}}$, provides a local estimate of the dimensional exponent. As a result, the local estimate of the dimensional exponent $\hat{\nu}_{\rho,i}^{\mathrm{LS}}$ can be computed

using a linear least-squares regression estimator and the best estimate of the local slope has the minimum rms error, i.e.,

$$\hat{v}_\rho = \left\{ \hat{v}_{\rho,i}^{\mathrm{LS}} : i = \arg\min_i \sqrt{\bar{\varepsilon}_i^2} \right\}. \tag{18}$$

5.3. Dimensional exponents of the bearing vibration data

The average distribution of distances $H_\rho(r)$ of the bearing vibration data during the second hour and during the final hour of the experiment are compared in Fig. 6. Obviously, the distribution of distances $H_\rho(r)$ for these two bearing vibration signals are different. The central distance where the distribution of distances $H_\rho(r)$ is maximal for the bearing vibration data during the second hour is lower than that for the final hour before the bearing failed. In addition, the partial correlation integrals $C_\rho(r)$ of the bearing vibration data during the second hour and during the final hour of the experiment are compared in Fig. 7.

The range of the dimensional exponents of the bearing vibration data throughout the life of the bearing is illustrated in the boxplot shown in Fig. 8. The median and the standard deviations of the dimensional exponents of the bearing vibration data during the entire experiment are shown in Figs. 9 and 10, respectively. Furthermore, Fig. 11 illustrates a third-order polynomial curve fit to the mean of the dimensional exponents of

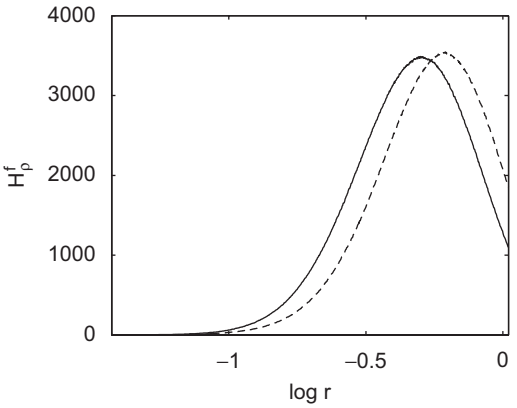

Fig. 6. The average distributions of distances $H_\rho(r)$ of the bearing vibration data during the second hour of the experiment (in solid line) and during the final hour of the experiment before the bearing failure (in dashed line).

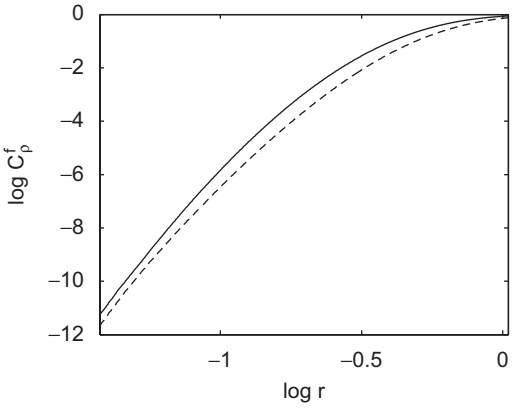

Fig. 7. Comparison between the correlation integral $C_\rho(r)$ of the bearing vibration data corresponding to the healthy bearing (in solid line) and that approaching the bearing failure (in dashed line).

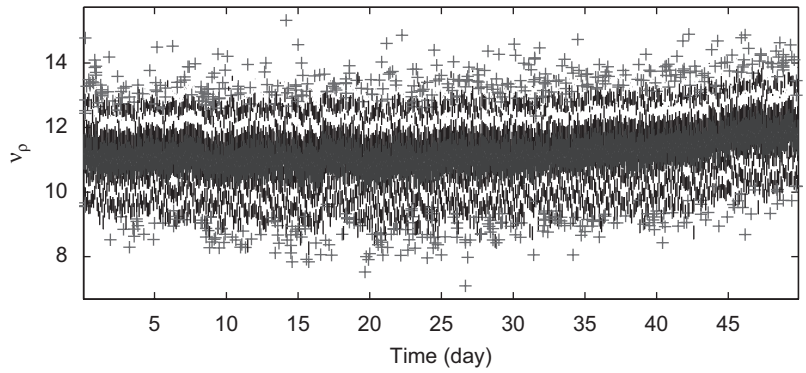

Fig. 8. Boxplot of the dimensional exponent of the bearing vibration data from the beginning to the end of the life testing of the bearing.

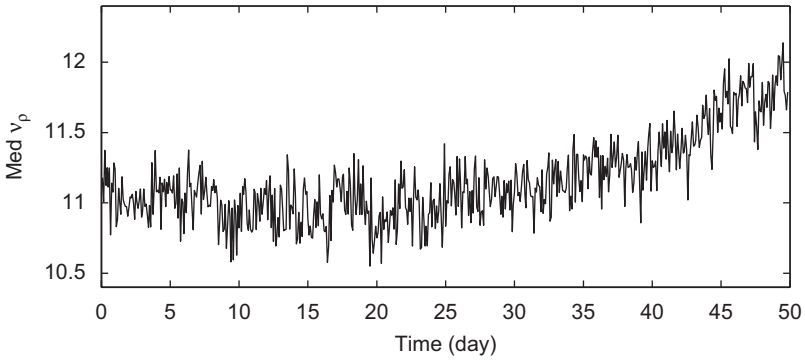

Fig. 9. The median value of the dimensional exponent of the bearing vibration data from the beginning to the end of the life testing of the bearing.

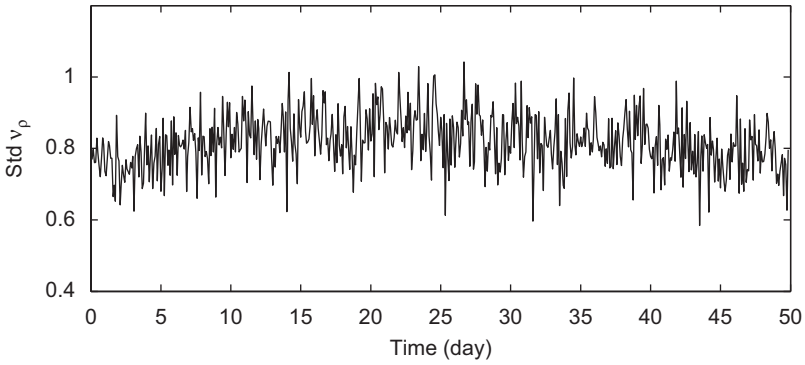

Fig. 10. Standard deviation of the dimensional exponent of the bearing vibration data from the beginning to the end of the life testing of the bearing.

the bearing vibration data. The polynomial fit is given by $f(x) = 8.017 \times 10^{-6}x^3 + 1.929 \times 10^{-4}x^2 + 1.343 \times 10^{-2}x + 11.084$ and the curve fit is shown as a dark line compared to the mean of the dimensional exponents of the bearing vibration data shown in a gray line.

The dimensional exponent of the bearing vibration data slightly decreases during the first quarter of the experiment which corresponds to the natural run-in period of the bearing. The dimensional exponent of the bearing vibration data during the second quarter of the experiment is nearly constant corresponding to normal operation of the bearing. Afterward, the dimensional exponent of the bearing vibration data begins to increase

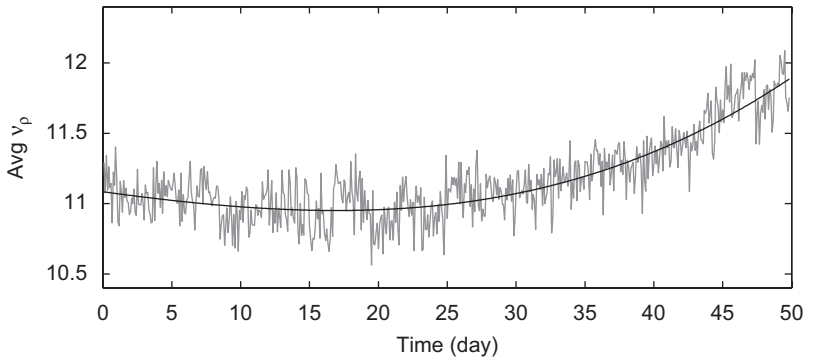

Fig. 11. Third-order polynomial curve fitting of the mean of the dimensional exponent of the bearing vibration data: $f(x) = 8.017 \times 10^{-6}x^3 + 1.929 \times 10^{-4}x^2 + 1.343 \times 10^{-2}x + 11.084$ shown in a dark line compared to the mean of the dimensional exponents of the bearing vibration data from the beginning to the end of the life testing of the bearing shown in a gray line.

during the second half of the experiment until the failure of the bearing. Note that the estimated L_{10} life of the bearing is around 21.71 days and is located in a neighborhood of the valley of the dimensional exponent curve shown in Fig. 11.

Based on the statistical characteristics of the dimensional exponents during the normal operation of the bearing (the average and standard deviation), a threshold for the dimensional exponents can be set at 11.60, which is equal to the average of the dimensional exponents during the initial normal operating and run-in period of the bearing plus one standard deviation. This makes the setting of the threshold bearing specific, but does not require that bearing fault information is available to make a decision related to bearing health. In the current study, the threshold of 11.60 provides failure prognostic information that indicates impending bearing failure approximately 5 days before the bearing catastrophically failed. The dimensional exponent reached 11.60 at the 46th day of the experiment and the bearing failed at day 51.

5.4. Bearing condition monitoring

There have been a variety of schemes that have been proposed for bearing diagnostics and prognostics. The methods can be roughly categorized as data driven or model-based depending on the underlying assumptions and the techniques and methods used. Data driven techniques generally require two steps: feature identification and extraction and decision-making. Time-domain and frequency-domain techniques for feature extraction include: peak vibration, rms vibration, crest factor, kurtosis, skewness, Short Time Fourier Transform (STFT), Wavelets, Cepstrum, Bi-Spectrum, etc. In order to make a decision regarding fault detection, diagnosis or prognosis, it is necessary to have a priori knowledge of the bearing type, operating conditions such as running speed and loading, fault types and the severity of each fault type. The requirement for such information has generally limited the application of most techniques to a laboratory, or experimental setting. Model-based approaches use models for rolling element bearing behavior where fault type and severity information are also used along with information about the relationship between the source of the vibration (the fault) and the observed vibration time series (the transmission path). Mechanistic models for damage propagation have been combined with rolling element bearing models to provide an approach to bearing prognostics. A recent paper by Kurfess et al. [31] provides an overview of the current state-of-the-art in monitoring rolling element bearings and also discusses issues related to sensors, signal processing and diagnostic and prognostic approaches.

The approach proposed in this paper falls into the general category of data driven methods. However, unlike most data driven methods the proposed detection, diagnosis and prognosis approach is based on the characteristics of a normal bearing and therefore does not require a priori information about the fault type, fault severity or operating conditions. The essence of this technique, similar in spirit to our earlier work on this same data set using hidden Markov modeling (HMM) [14,15], is that characteristics of the normal bearing can be extracted in real time from the vibration data and prognostics is then based on deviations from this normal

trend. The basic assumption in the HMM and dimensional exponent approaches is that monitoring is initiated on a normal bearing. Using vibration data in this way allows the bearing to be used as its own "statistical control", thereby reducing the amount of information that is needed for real-world applications.

We should keep in mind that the dimensional exponent is quantifying the active degrees of freedom of the dynamic system generating the measured bearing vibration signal and from the preliminary experimental results obtained in this study, the characteristics of the dimensional exponents of the bearing vibration data can be divided into three temporal regions. In the first quarter of the experiment, the dimensional exponent slightly decreases which most probably corresponds to the natural run-in period of the bearing. In the second quarter of the experiment, the dimensional exponent is nearly constant corresponding to normal operation of the bearing. In the second half of the experiment, the dimensional exponent begins to increase until the failure of the bearing. Although the dimensional exponent can be used to characterize the condition and health status of the bearing, the interpretation of the dimensional exponent as a diagnostic or prognostic index may depend on the specific bearing, fault type, etc. To make the approach more generally applicable, we use a surrogate data testing technique to help with diagnosing the condition of a bearing with the potential of developing a bearing diagnosis and prognosis method that is independent of the specific bearing, fault type, etc. A diagram of the proposed bearing condition diagnosis and prognosis scheme is illustrated in Fig. 1.

The proposed bearing health monitoring scheme is composed of two main stages: surrogate data testing and bearing condition identification. In the surrogate data testing stage, surrogate data are generated from the original vibration time series using the iteratively refined surrogates method. The partial correlation integrals are then computed for the actual bearing vibration data and the surrogate time series data. Next, the dimensional exponents are estimated from the partial correlation integrals. A statistical significance test between the dimensional exponents of the actual bearing vibration data segment and the surrogate data segment is used to examine the existence of nonlinearity in the vibration time series data. Subsequently, the results of these statistical significance tests over a period of time are used for bearing condition identification.

Fig. 12 shows the dimensional exponents of the surrogate data of the associated bearing vibration data. The mean of the dimensional exponents of the surrogate data is shown in Fig. 13 as a gray line. A third-order polynomial fit to the mean of the dimensional exponents of the surrogate data is $f(x) = 1.591 \times 10^{-5}x^3 - 6.615 \times 10^{-4}x^2 + 8.394 \times 10^{-3}x + 11.339$ and is shown as a black line in Fig. 13.

The dimensional exponents of the surrogate data of the associated bearing vibration data can, on the other hand, be classified according to their characteristics into two regions. The dimensional exponent of the surrogate data is nearly constant for most of the experiment, and then tends to increase when the bearing is close to failure much the same as the dimensional exponents of the original bearing vibration data.

The results of two-tail, paired t-tests between the dimensional exponents of the actual bearing vibration data and the dimensional exponents of the associated surrogate data using a significance level of 0.01 are illustrated in Fig. 14. Here 0 denotes that the null hypothesis cannot be rejected and 1 denotes that the null hypothesis can be rejected. There are a number of 0s at the beginning of the experiment which is during the bearing run-in period. The density of the number of 1s where the null hypothesis can be rejected increases around the middle

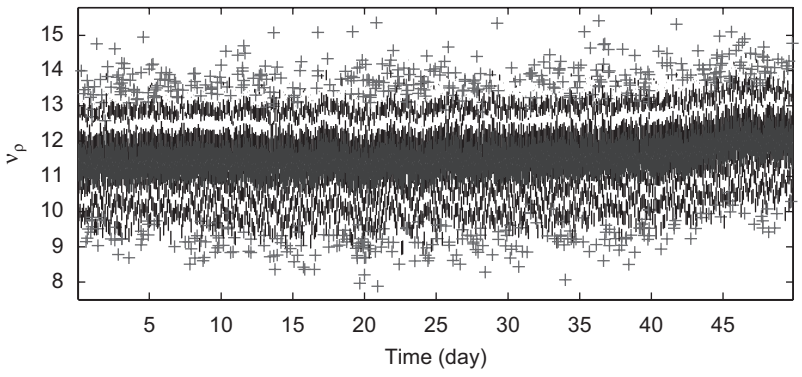

Fig. 12. Boxplot of the dimensional exponent of surrogate data of the bearing vibration data from the beginning to the end of the life testing of the bearing.

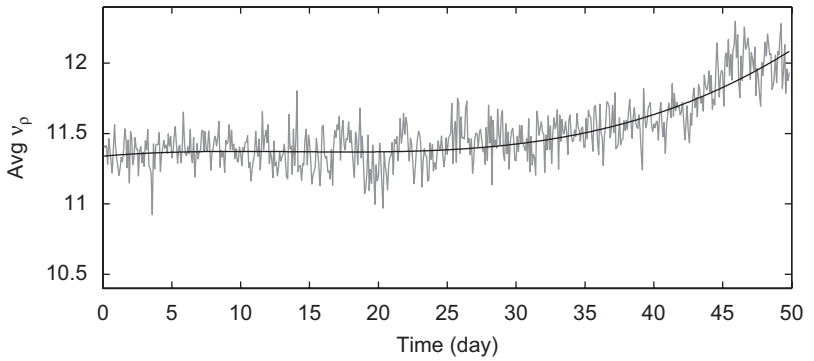

Fig. 13. Third-order polynomial curve fitting of the mean of the dimensional exponent of surrogate data of the bearing vibration data: $f(x) = 1.591 \times 10^{-5}x^3 - 6.615 \times 10^{-4}x^2 + 8.394 \times 10^{-3}x + 11.339$ shown in a dark line compared to the mean of the dimensional exponents of surrogate data of the bearing vibration data from the beginning to the end of the life testing of the bearing shown in a gray line.

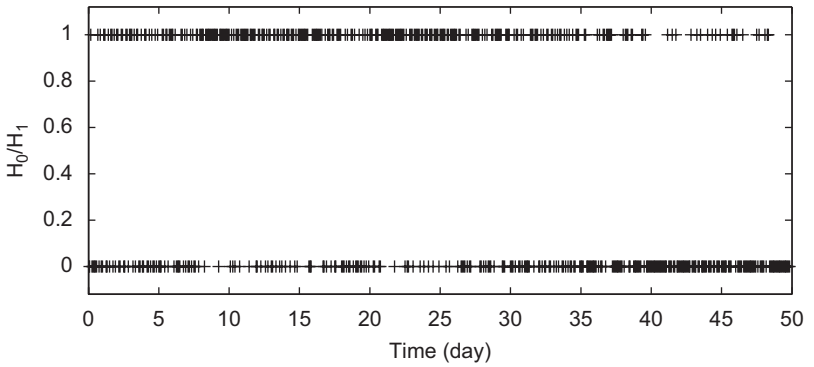

Fig. 14. The results of the surrogate data testing where 0 denotes that the null hypothesis cannot be rejected and 1 denotes that the null hypothesis can be rejected.

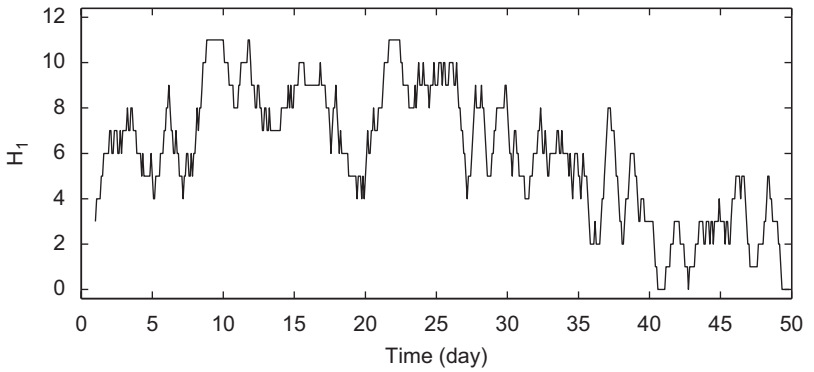

Fig. 15. The number of the null hypotheses of the surrogate data testing that can be rejected during a 1-day period.

of the experiment, indicating that the dimensional exponent is quantifying a nonlinear feature in the vibration time series data. This period is associated with the normal operation of the bearing and the results suggest that the bearing vibration is dominated by nonlinear characteristics during this period. Before the bearing fails, the surrogate data null hypotheses cannot be rejected. Fig. 15 illustrates the total number of the null hypotheses that can be rejected during a 1-day period. From the trend in the dimensional exponents of the vibration data, as the bearing wears the "complexity" (measured by the dimensional exponent) of the vibration signal is

increasing. Surrogate data testing suggests that the bearing vibration time series during the normal period generally contains nonlinear features while the bearing vibration time series behaves more like a linear stochastic process as the bearing tends to failure. From observations of the results of surrogate data testing, all 12 surrogate data null hypotheses cannot be rejected at day 42 of the test. Accordingly, if we use the epochs from the first 1-day period for which the null hypothesis cannot be rejected as an indication of a pending bearing failure, then in this experiment catastrophic failure of the bearing can be predicted about 9 days in advance of the actual failure. This is consistent with our previous HMM results [15] where failure was predicted 6 days (day 44) in advance of a catastrophic failure of the bearing.

6. Conclusion

Bearing vibration data corresponding to the operation of a test bearing in an accelerated life experiment was analyzed using the partial correlation integral and the computation of the dimensional exponent. This is the same data that was used in Ref. [15] and the result reported in this paper compares favorably with these previously published results. The paper also provides further insight into the dynamics of a bearing as it approaches failure. We have shown that the dimensional exponents of the bearing vibration data slightly decrease at the beginning of the experiment during the bearing run-in period. This is consistent with experimental observations reported in the literature. During the next period of time where the bearing is experiencing progressive wear conditions, the dimensional exponents of the bearing vibration data remain eventually constant. Finally, as the bearing progresses toward catastrophic failure at the end of the experiment, the dimensional exponents begin to increase. Increasing dimensional exponents are indicative of an increase in complexity of the vibration time series as quantified by an estimate of the active degrees of freedom of the dynamic system generating these data. Thus, we conclude that the bearing vibration time series becomes more complex as the bearing is closer to failure. This is not surprising and is consistent with what has been observed in experiments as well as real-world applications. This increase in complexity, alone, may not be sufficient for bearing prognostics. The trend of the dimensional exponents may also depend on certain mechanical characteristics of the bearing, the bearing running speed and load, the acquisition of the bearing vibration data such as the sampling rate and resolution, and the characteristics of the mechanical transmission path between the bearing and vibration sensor. As a result, the dimensional exponents computed from vibration data corresponding to different types of bearings in different operating environments may be different making it difficult to use this information directly in a bearing health condition monitoring scheme. To circumvent this difficulty, we proposed the use of surrogate data testing to essentially use the bearing vibration data as its own statistical control. That is, we only need to analyze how the dimensional exponent is different between the original and surrogate data samples as the bearing wears. We observed that the vibration characteristics during normal operating periods tended to include significant nonlinear features, were less complex and had smoother temporal patterns with a reduced number of principal spectral components, while the vibration characteristics as the bearing approached failure tended to be dominated by more stochastic features with increasing complexity that included more irregular patterns with a greater number of principal spectral components. With these observation as a guide for the development of a scheme for tracking bearing health, the comparative analysis methodology developed in this paper that uses the original and surrogate vibration time series data does not depend on the type of bearing. Further empirical and analytical work validated by additional experimental results will be reported on in a subsequent paper.

The method of surrogate data testing, that is commonly used to identify the evidence of nonlinearity in time series data, is used for bearing condition monitoring and prognosis. With the integration of surrogate data testing in the diagnostic process, we can use the results of surrogate data testing to further help characterize the condition of the bearing. We observed that the bearing vibration time series during normal operation generally contained nonlinear features while the bearing vibration time series when the bearing was close to failure behaved more like a linear stochastic process. In nonlinear signal analysis, there are generally two sources of signals that are of interest: linear stochastic signals and nonlinear deterministic signals. Clearly a rolling element bearing is a complex nonlinear system, and we have observed and quantified in our analysis that the vibration characteristics during the normal operating period of a bearing are dominated by nonlinear deterministic signal features and as the bearing wears toward failure, these nonlinear features become

dominated by stochastic signal features. These results are reasonable because as a bearing fails, more and more defects are introduced into the bearing surfaces and these defects are becoming more fully developed. As the rolling elements pass over these defects, they induce a vibration response of the bearing that is modulated by the dynamics of the mechanical transmission path to the sensor. As the defects develop and propagate in a seemingly random pattern, the central limit theorem suggests that as the number of faults grows the response through a linear dynamic transmission path should approach that of a Gaussian random process. This is consistent with what we observed in our experimental and computational work, where we observed increased signal complexity as the bearing wears, and that the increased signal complexity is dominated by stochastic features. Our current work is directed at developing a more rigorous mathematical model that can further substantiate the experimental and computational work presented in this paper.

From this preliminary study, we have identified an event that can be used to forewarn about impending bearing failure about 9 days prior to the catastrophic failure of the test bearing used in our experiment. The warning event is determined by the epochs where the null hypotheses of the surrogate data tests over a 1-day period cannot be rejected. From a previous study using the same set of bearing vibration data used in this experiment, wavelet packet decomposition and hidden Markov modeling (HMM) was used for prognostics [15]. In this prior work it was found that the HMM probabilities decreased as the bearing damage progressed toward bearing failure, and the HMM probabilities dramatically dropped indicating severe damage and imminent bearing failure at about day 44. This provided prognosis of an impending bearing failure at about the point in time where only about 10 percent of the bearing life remained. The results of this paper, based on a dimensional exponent threshold for 11.60, arrived at a similar prognostic indication.

In summary, we have shown that bearing vibration data are becoming increasing more complex, and include the addition of stochastic features, as the bearing tends toward failure when compared to a normal bearing. The partial correlation integral and the dimensional exponent provide useful diagnostic and prognostic information for monitoring the health and condition of rolling element bearings in rotating machinery through complexity analysis of the bearing vibration data. From the preliminary experimental results given in this paper, if a proper threshold of the dimensional exponent is chosen, the dimensional exponent can be used as a prognostic indicator to provide a suitable advanced warning of an impending bearing failure. The addition of surrogate data testing enhances the diagnostic and prognostic utility of the dimensional exponents, and because the bearing is used as its own statistical control the technique should be independent of the type of bearing and fault, and the operating conditions of the bearing. Nevertheless, additional experimental, modeling and computational work is needed to further validate the methodology and to determine if threshold determination is specific to a particular bearing geometry and test setup, or is a general property of the failure mechanisms for rolling element bearings.

References

[1] B. Li, M.-Y. Chow, Y. Tipsuwan, J.C. Hung, Neural-network based motor rolling bearing fault diagnosis, *IEEE Transactions on Industrial Electronics* 47 (2000) 1060–1069.

[2] P.D. Samuel, D.J. Pines, A review of vibration-based techniques for helicopter transmission diagnostics, *Journal of Sound and Vibration* 282 (2005) 475–508.

[3] H.R. Martin, F. Honarvar, Application of statistical moments to bearing failure detection, *Applied Acoustics* 44 (1995) 67–77.

[4] R.B.W. Heng, M.J.M. Nor, Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition, *Applied Acoustics* 53 (1998) 211–226.

[5] J. Antoni, Cyclic spectral analysis of rolling-element bearing signals: facts and fictions, *Journal of Sound and Vibration* 304 (2007) 497–529.

[6] H.-V. Liew, T.C. Lim, Analysis of time-varying rolling element bearing characteristics, *Journal of Sound and Vibration* 283 (2005) 1163–1179.

[7] P.D. McFadden, J.D. Smith, Vibration monitoring of rolling element bearings by high frequency resonance technique—a review, *Tribology International* 17 (1984) 3–10.

[8] C.J. Li, J. Ma, B. Hwang, Bearing localized defect detection by bicoherence analysis of vibrations, *Journal of Engineering for Industry* 117 (1995) 625–629.

[9] H. Qiu, J. Lee, J. Linb, G. Yu, Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics, *Journal of Sound and Vibration* 289 (2006) 1066–1090.

[10] S. Seker, E. Ayaz, Feature extraction related to bearing damage in electric motors by wavelet analysis, *Journal of the Franklin Institute* 340 (2003) 125–134.
[11] C.J. Li, J. Ma, Wavelet decomposition of vibrations for detection of bearing localized defects, *NDT and E International* 30 (1997) 143–149.
[12] H. Taplak, I. Uzmay, S. Yildirim, An artificial neural network application to fault detection of a rotor bearing system, *Industrial Lubrication and Tribology* 58 (2006) 32–44.
[13] B. Samanta, K.R. Al-Balushi, S.A. Al-Araimi, Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection, *Engineering Applications of Artificial Intelligence* 16 (2003) 657–665.
[14] H. Ocak, K.A. Loparo, An hmm based fault detection and diagnosis scheme for rolling element bearings, *Journal of Vibration and Acoustic* 127 (2005) 299–306.
[15] H. Ocak, K.A. Loparo, F.M. Discenzo, Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: a method for bearing prognostics, *Journal of Sound and Vibration* 302 (2007) 951–961.
[16] L.M. Hively, V.A. Protopopescu, Machine failure forewarning via phase-space dissimilarity measures, *Chaos* 14 (2004) 408–419.
[17] P. Grassberger, I. Procaccia, Characterization of strange attractors, *Physical Review Letters* 50 (1983) 346–349.
[18] W.S. Pritchard, D.W. Duke, Measuring chaos in the brain: a tutorial review of EEG dimension estimation, *Brain and Cognition* 27 (1995) 353–397.
[19] S. Janjarasjitt, K.A. Loparo, Partial correlation integral: an efficient computational method for approximating the correlation integral and dimension, Technical Report, EECS Department, Case Western Reserve University, Cleveland, Ohio, 2007.
[20] F. Takens, Detecting strange attractors in turbulence, in: D.A. Rand, L.-S. Young (Eds.), *Dynamical Systems and Turbulence*, in: *Lecture Notes in Mathematics*, Vol. 898, Springer, Berlin, 1981, pp. 366–381.
[21] N.H. Packard, J.P. Crutchfield, J.D. Farmer, R.S. Shaw, Geometry from a time series, *Physical Review Letters* 45 (1980) 712–716.
[22] J. Theiler, Estimating fractal dimension, *Journal of the Optical Society of America A* 7 (1990) 1055–1073.
[23] J. Theiler, Spurious dimension from correlation algorithms applied to limited time-series data, *Physical Review A* 34 (1986) 2427–2432.
[24] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, J.D. Farmer, Testing for nonlinearity in time series: the method of surrogate data, *Physica D* 58 (1992) 77.
[25] T. Schreiber, A. Schmitz, Improved surrogate data for nonlinearity tests, *Physical Review Letters* 77 (1996) 635–638.
[26] T. Schreiber, Constrained randomization of time series data, *Physical Review Letters* 80 (1998) 2105–2108.
[27] T. Schreiber, A. Schmitz, Surrogate time series, Technical Report, University of Wuppertal, 2001.
[28] M.B. Kennel, R. Brown, H.D.I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction, *Physical Review A* 45 (1992) 3403–3411.
[29] A.M. Albano, J. Muench, C. Schwartz, A.I. Mees, P.E. Rapp, Singular-value decomposition and the Grassberger–Procaccia algorithm, *Physical Review A* 38 (1988) 3017–3026.
[30] S. Borovkova, Estimation and prediction for nonlinear time series, Technical Report, University of Groningen, 1998.
[31] T.R. Kurfess, S. Billington, S. Liang, Advanced Diagnostic and Prognostic Techniques for Rolling Element Bearings, in: *Springer Series in Advanced Manufacturing*, Springer London, 2006, pp. 137–165.