Journal of Sound and Vibration 327 (2009) 413–423

Contents lists available at ScienceDirect

Journal of Sound and Vibration

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

Autoregressive model-based gear shaft fault diagnosis using the Kolmogorov–Smirnov test

Xiyang Wang [a], Viliam Makis [b,*]

[a] Department of Mechanical Engineering, Nanchang Hangkong University, 330062 Nanchang, Jiangxi, People's Republic of China
[b] Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Road, Toronto, Ont., Canada M5S 3G8

A R T I C L E   I N F O

Article history:
Received 9 August 2007
Received in revised form
3 July 2009
Accepted 6 July 2009
Handling Editor: C.L. Morfey
Available online 26 July 2009

Keywords:
Gearbox vibration
Fault diagnosis
Autoregressive model
Kolmogorov–Smirnov test
Gear shaft fault
Prediction error signal

A B S T R A C T

Vibration behavior induced by gear shaft crack is different from that induced by gear tooth crack. Hence, a fault indicator used to detect tooth damage may not be effective for monitoring shaft condition. This paper proposes an autoregressive model-based technique to detect the occurrence and advancement of gear shaft cracks. An autoregressive model is fitted to the time synchronously averaged signal of the gear shaft in its healthy state. The order of the autoregressive model is selected using Akaike information criterion and the coefficient estimates are obtained by solving the Yule–Walker equations with the Levinson–Durbin recursion algorithm. The established autoregressive model is then used as a linear prediction filter to process the future signal. The Kolmogorov–Smirnov test is applied on line for the prediction of error signals. The calculated distance is used as a fault indicator and its capability to diagnose shaft crack effectively is demonstrated using a full lifetime gear shaft vibration data history. The other frequently used statistical measures such as kurtosis and variance are also calculated and the results are compared with the Kolmogorov–Smirnov test.

1. Introduction

Gearboxes are widely used to transmit power and change speed and/or direction in many mechanical systems. Failures of the gears may have serious consequences, in some cases claiming lives. Development of gear failure diagnostic techniques based on the analysis of vibration signals has been an active area of research for more than two decades [1].

Although most faults in gearboxes are tooth-related problems such as pitting, spalls and tooth cracks, gear shaft cracks, eccentric problem or misalignment can also take place in some cases and lead to a complete failure of the gear. For example, cracks may often occur at the root of the gear in an integrated gear shaft because of working stress concentration. Up to now, detection and diagnostic techniques for gear faults have focused on the tooth fault, especially fatigue crack due to cyclic loading. Vibration behavior induced by shaft cracks is different from that induced by tooth cracks; fault indicator used for detection of tooth fault may not be effective to detect shaft crack. There has been extensive research on the vibrational behavior of cracked shafts and crack identification in rotating shafts. Pennacchi et al. [2] used a model-based transverse crack identification method in the frequency domain that is suitable for industrial machines, and validated their method by experimental results obtained on a large test rig. Sekhar [3] used a model-based approach in the time domain that has been developed by Bach and Markert [4]. But the method was tested using only numerical results obtained from very simple rotor models, which are unable to represent the behavior of real machines. Dilena and Morassi [5] used

* Corresponding author. Tel.: +1 416 978 4184; fax: +1 416 978 3453.
E-mail addresses: wxywzh@yahoo.com.cn (X. Wang), makis@mie.utoronto.ca (V. Makis).

0022-460X/$ - see front matter
doi:10.1016/j.jsv.2009.07.004

resonances in the frequency response curves of non-rotating beams, and Gounaris and Papadopoulos [6] used the axial vibrations of cracked shafts to detect position and depth of cracks. However, all the above papers have focused on the crack identification in a non-gear shaft, specifically in a rotor shaft. As summarized by Hamidi et al. [7], several publications have proposed a number of techniques such as the use of natural frequencies, mode shapes and frequency response functions for damage detection of rotor shafts. For a gear shaft, the tooth-meshing vibration signal is an important part of vibration signal obtained from a gearbox, which will mask the signal induced by shaft cracks. Therefore, the method of crack detection in a gear shaft is different from that in a non-gear shaft. Little research has been done to develop gear shaft crack detection methods based on vibration signal analysis.

In the literature, there are many signal-processing tools for vibration data, such as power spectrum, time-domain averaging, de-noising, demodulation, time series modeling, time–frequency distribution, wavelet transform, neural network, calculation of higher-order statistics, etc. Several fault indicators have been proposed using these tools to detect early tooth damage. Recently, the time series modeling (autoregressive (AR) model, moving average model or autoregressive moving average model—ARMA), known as high-resolution parametric spectrum analysis method, has been applied to vibration signal analysis of rotating machinery. Both accuracy and resolution can be significantly improved using this method. Wang and Wong [8] demonstrated a successful application of an autoregressive model-based filter to isolate the impulse-like effect of localized cracks in a gear tooth. In their paper, kurtosis of AR filter prediction error (residual signal) was used to indicate the severity of gear crack. The AR filter technique was proved to be superior when compared with the traditionally used residual analysis technique based on subtracting a regular gear meshing signal represented by the tooth-meshing harmonics and immediately adjacent sidebands from the spectrum of time synchronously averaged signal. Endo and Randall [1] proposed the use of the minimum entropy deconvolution technique to enhance the ability of AR model-based filtering technique to detect localized faults in gears. Kurtosis was also used as a fault measure in their paper. Zhan et al. [9] used a noise-adaptive Kalman filter to fit a time-varying AR model to the gear motion residual signal, and proposed several statistical measures to diagnose gear tooth faults under varying load conditions. All these papers have focused on the problem of gear tooth fault detection. Little research has been done to develop methods for detecting gear shaft cracks.

In statistics, Kolmogorov–Smirnov test (K–S test) is used to determine whether two underlying probability distributions differ, or whether an underlying probability distribution differs from a hypothesized distribution. Recently, the K–S test has been found to be an extremely powerful tool in the condition monitoring of rotating machinery. The K–S test-based signal-processing technique compares two signals and tests the hypothesis that the two signals have the same probability distribution. Using this technique, it is possible to determine whether the two signals are similar or not. Therefore, by comparing a given vibration signature to a number of template signatures (i.e. signatures from known gear conditions), it is possible to determine which is the most likely condition of the gear under analysis. Andrade et al. [10] applied this technique to the specific problem of fatigue crack detection. They showed that this test not only successfully identifies the presence of fatigue cracks but also gives an indication regarding the crack advancement. The signal considered in [10] was the time synchronously averaged signal of the gear of interest. Kar and Mohanty [11] discussed the applicability of the K–S test to bearing fault diagnosis. Vibration signatures of good and faulty bearings were compared statistically using the K–S test. It was shown that the K–S test has many advantages when compared with the conventional statistics such as mean, kurtosis, skewness and variance. Zhan et al. [9] also applied the K–S goodness-of-fit test to AR model prediction errors to assess the gear fault advancement under varying load conditions. The K–S goodness-of-fit test compares an empirical distribution function with the distribution function of the hypothesized distribution, which was set to be a normal distribution in [12].

This paper applies the K–S test to the prediction error signal obtained from the AR model filter to indicate the occurrence of gear shaft fault and to estimate the fault severity. The K–S test statistical distances, as well as the variance of prediction error signal, are used as fault indicators. Kurtosis, which is usually used in detecting tooth crack, is shown to be ineffective in this case.

2. Vibration signals induced by gear meshing

Time synchronous averaging (TSA) technique is widely accepted as a powerful tool in the detection and diagnosis of gear faults. By synchronizing the sampling of the vibration signal with the rotation of a particular gear and evaluating the ensemble average over many revolutions with the start of each frame at the same angular position, a TSA signal is obtained, which contains only the components that are synchronous with the revolution of the gear in question. This method reduces considerately the effects of all other sources, including other gears, and noise. The TSA signal can be expressed as [13]

$$g(t) = \sum_{m=0}^{M} X_m(1 + a_m(t))\cos(2\pi m f_x t + \phi_m + b_m(t)), \quad (1)$$

where M is the number of tooth-meshing harmonics, f_x the tooth-meshing frequency, X_m and ϕ_m are, respectively, the amplitude and the phase of the mth meshing harmonic, while the modulation effects concerning the same harmonic are given by the amplitude modulation (AM) function, $1+a_m(t)$, and the phase modulation (PM) function, $b_m(t)$. These modulation functions are periodic with the considered gear rotation frequency. It should be noted that the vibration signal fluctuation induced by shaft crack or by tooth defect will not be removed by TSA technique.

For healthy gears, the TSA signal is normally composed of gear meshing waveforms modulated by some low shaft order (first- and/or second-order) functions. The spectrum of the TSA signal is therefore dominated by sharp spectral peaks at the fundamental gear meshing frequency and its harmonics accompanied by some low-order modulation sidebands. The modulation effects are generally believed to be caused by geometric and/or assembly errors of the gears [8].

When a localized gear tooth fault occurs, say a tooth crack is present, some form of impulsive signal at a comparatively low energy level will infiltrate the TSA signal. This will produce additional amplitude modulation and phase modulation, giving rise to high-order modulation sidebands in the frequency domain. Structural resonances may be excited by impacts produced by the gear fault. The spectrum of the faulty-state TSA signal will still be dominated by the sharp peaks at the gear meshing harmonics with the modulation sidebands stretching to high shaft orders. It is believed that these faults, at least in their early stages, do not affect the periodic nature of the signal, because they contact periodically with exactly the same matching surfaces. Rather, only the strength and shape of the modulations are affected in a deterministic fashion [12]. However, a gear shaft crack will induce whole shape fluctuation in a TSA signal, not a localized impulsive change. Actually, the signal waveform will be modulated by the shaft rotation frequency vibration signal due to shaft eccentricity or misalignment caused by shaft crack. The whole signature curve will also deviate from its centerline. Two typical waveforms of TSA signals induced by gear tooth crack and gear shaft crack are shown in Figs. 1(a) and (b), respectively. The waveform in Fig. 1(a) was calculated using the gearbox Test-run #14 data obtained from an experiment performed on the mechanical diagnostics test bed (MDTB) in the Applied Research Laboratory at the Pennsylvania State University [14,15]. The waveform in Fig. 1(b) was obtained from the gearbox Test-run #13 data, which are used in the subsequent sections where the experiment is described in more detail.

The effect of impulses caused by the faults is relatively small compared with the gear mesh components. The objective of many residual signal methods, including AR filter techniques described in the following sections, is to isolate the impulsive signal from the signal mixture to monitor and detect the emergence of gear fault.

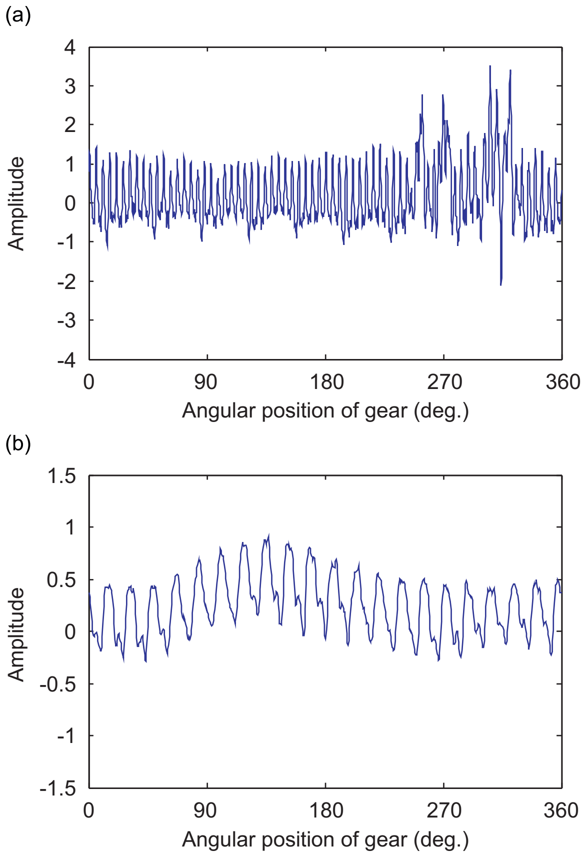

Fig. 1. Vibration behaviors induced by gear tooth fault and by gear shaft fault: (a) gear teeth crack (71 teeth, 4 teeth cracked) and (b) gear shaft crack (21 teeth).

3. Theoretical background

3.1. AR model-based diagnosis

An AR model is simply a linear regression of the current value of the series against one or more prior values of the series. The AR model of an order p is defined as

$$x[n] = \sum_{k=1}^{p} a[k]x[n-k] + e[n], \tag{2}$$

where $a[k]$ are the AR coefficients and $e[n]$ is a Gaussian white-noise series with zero mean and variance σ^2. A typical spectrum of a gear-excited signal is dominated by the harmonics of the gear mesh frequencies and appropriately represented by using an all-pole expression, i.e. AR model. In fact, the spectrum of the autoregressive moving average process can be represented purely in terms of AR coefficients without computing the moving average model coefficients. The advantage of the AR model is that unlike a moving average model (and the moving average part of ARMA), its parameters can be determined by solving a linear set of equations. Generally, an AR model requires far fewer coefficients than the corresponding moving average model, and hence, it is more efficient in modeling a filter of equivalent performance (for lightly damped systems with sharp peaks).

The selection of the AR model order is a crucial step because spurious spectral peaks and general statistical instability will be caused if the order is too large, whereas too small an order will lead to smoothed spectral peaks [12]. The conventional model order selection procedure based on the Akaike information criterion (AIC) is usually used to ensure optimum adaptation of the AR coefficients to the undamaged gear shaft signal. The AIC, which reflects the effect of spectral variance due to increase in model order p and prediction errors computed in estimating the AR coefficients, is given by

$$AIC(p) = N\ln(\sigma^2) + 2p. \tag{3}$$

The most common method of determining the coefficients of the AR model is by using autocorrelation (second-order statistical characteristic) of the signal and solving the Yule–Walker equations

$$r_{xx}[k] + \sum_{l=1}^{p} a[l]r_{xx}[k-l] = \left|b[0]^2\right|\delta[k], \quad k \geq 0. \tag{4}$$

In matrix form:

$$\begin{bmatrix} r_{xx}[0] & r_{xx}[-1] & \cdots & r_{xx}[p-1] \\ r_{xx}[1] & r_{xx}[0] & \cdots & r_{xx}[-p+2] \\ \cdots & \cdots & \cdots & \cdots \\ r_{xx}[p-1] & r_{xx}[p-2] & \cdots & r_{xx}[0] \end{bmatrix} \begin{bmatrix} a[1] \\ a[2] \\ \cdots \\ a[p] \end{bmatrix} = -\begin{bmatrix} r_{xx}[1] \\ r_{xx}[2] \\ \cdots \\ r_{xx}[p] \end{bmatrix}. \tag{5}$$

The solution of the Yule–Walker equations can be obtained by several methods. Two of the most commonly accepted methods are the Levinson–Durbin recursion (LDR) algorithm or Burg's (maximum entropy) method (BM). The former one is adopted in this paper.

The first step of AR model-based diagnosis is model construction. The optimal AR model order p and its coefficients $a[k]$ can be estimated from the healthy-state TSA signal of the gear of interest at a given loading condition using the AIC and Yule–Walker equations.

The second step is fault diagnosis. The $AR[p]$ is used as a linear prediction filter to pass the future-state TSA signal under the same load conditions, and the filtered signal is then subtracted from its original version to produce the AR model residual (prediction error) signal. If the gear of interest remains in a healthy condition, its AR residual signal will be randomly distributed. However, when a localized fault is present, the signal region affected by the engagement of the faulty tooth will not be well predicted by the model, and consequently shown as localized changes in the AR residual signal. Statistical measures, such as kurtosis used by many researchers, can then be calculated using AR residual signals to indicate fault advancement. However, as shown in the following sections, such statistical measures are unable to reflect the signature variation induced by shaft crack.

3.2. K–S test

The K–S test considers the null hypothesis that the cumulative distribution function (CDF) of the target distribution, denoted by $F(x)$, is the same as the CDF of a reference distribution, $R(x)$. Hence, it is possible to compare two vibration signatures, and assess if both have the same CDFs [5]. Note that the application of this test for condition monitoring assumes that the fault is strong enough to vary the CDF of the original vibration signature, which is the case of fatigue cracks and many other gear mechanical faults. In this paper, the AR model prediction error signal in healthy state is chosen to represent the reference distribution; the target distribution is represented by the future AR model prediction error signal in an unknown state.

When applying a two-sample K–S test, the distance D between two empirical distribution functions $F_{N_1}(x)$ and $R_{N_2}(x)$is calculated as

$$D = \sup_{-\infty < x < \infty} |F_{N_1}(x) - R_{N_2}(x)|, \tag{6}$$

where N_1 and $N_2$are the number of data points in the first and second sample, respectively. Let

$$N = \frac{N_1 N_2}{N_1 + N_2}. \tag{7}$$

Then (see e.g. Von Mises [16])

$$p(\sqrt{N}D > \lambda) \doteq \sum_{i=1}^{+\infty} (-1)^{i-1} e^{-2i^2\lambda^2}. \tag{8}$$

The null hypothesis that the two distributions are equal is rejected, at significance level α, if $D^* = \sqrt{N}D$ is greater than the corresponding critical value. The approximate p-value $p(D)$, which is termed the similarity probability in the vibration signal-processing literature (e.g. [4,5]), can then be obtained from Eq. (8).

If the two vibration signatures are similar (i.e. have statistically similar CDFs, $D \to 0$), then the similarity probability $p(D)$ tends to 1. On the other hand, if the signatures are different, then the similarity probability $p(D)$ tends to 0. In this paper, the distance value D and corresponding similarity probability $p(D)$ are considered as fault indicators.

4. Experimental set-up

The vibration data used in this paper were obtained from the mechanical diagnostics test bed in the Applied Research Laboratory at the Pennsylvania State University [14,15]. The MDTB is functionally a motor-drive train-generator test stand. The gearbox is driven at a set input speed using a 30 hp (22.38 kW), 1750 rpm AC drive motor, and the torque is applied by a 75 hp (55.95 kW), 1750 rpm AC absorption motor.

A vector unit capable of controlling the current output of the absorption motor accomplishes the variation of the torque. The MDTB is highly efficient because the electrical power that is generated by the absorber is fed back to the driver motor. The mechanical and electrical losses are sustained by a small fraction of wall power. The MDTB has the capability of testing single and double reduction industrial gearboxes with ratios from about 1.2:1 to 6:1. The gearboxes are nominally in the 5–20 hp (3.73–14.92 kW) range. The system is designed to provide the maximum versatility to speed and torque settings.

Ten accelerometers and an acoustic microphone were placed on the gearbox. Nine of these were of single axis and shear type with a bandwidth of 20 kHz and the tenth accelerometer was a triaxial and shear type with a bandwidth of 8 kHz. The triaxial accelerometer was included to determine whether triaxial data can provide significantly better sensor fusion for gearbox health assessment than the single-axis accelerometers. However, the measurement tradeoff is that the triaxial accelerometer possesses a lower frequency bandwidth than single-axis accelerometers.

Data were collected in a 10 s window at set times that cover 200,000 sampling points in total and triggered by accelerometer RMS thresholds. The time interval between every two adjacent data files is 30 min, and the averaging time for signal synchronous average analysis is 10 s for each data file. The sampling frequency is 20 kHz. The signals of the MDTB accelerometers are all converted to digital data format with the highest resolution to which the accelerometers are accurate. Among all accelerometers located in the MDTB, the single-axis shear piezoelectric accelerometer data A03 for axial direction present the best quality data for state diagnosis of gearbox. Therefore, data recorded by this accelerometer are selected in this study to investigate the proposed AR model-based indicators for gear shaft fault.

Test-run #13 of MDTB was performed with a single reduction helical 1:3.3 ratio gearbox that was run at 100% output torque (62.72 N m) for 96 h, then increased to about 300% torque 188.14 m) until the input shaft broke (the duration of the whole experiment was 15.5 h), decoupling the drive train. The data of Test-run #13 are analyzed in the following sections.

5. Gear shaft fault diagnosis with AR filter technique

5.1. TSA signals and spectrum

It was in the second load phase (i.e. 300% output torque level, from data file number 192 to 231) that the gear shaft ran from healthy state to a completely broken state; so the data files acquired only under this condition were analyzed. The vibration signals measured from the gearbox were synchronously averaged per one rotation of the gear of interest. Because it was the input gear shaft that was finally broken, the gear was chosen to be the pinion gear (21 teeth). The TSA signals and their spectra of the pinion gear shaft for several randomly selected data files (192, 212, 218, 225, 227 and 231) are shown in Fig. 2. Note that each TSA signature includes only one gear revolution. File 192 was the first data file of the second load phase, and the gearbox was in healthy state in the running period of this file. File 231 was the final data file and gear shaft was broken completely at the end of this file. However, the gear shaft states were unknown during the running period of data files 212, 218, 225 and 227.

In the healthy state (Fig. 2(a)), the TSA signatures vary very regularly, oscillating along the centerline. There are 21 signature periods in one revolution, corresponding to 21 teeth. The fundamental frequency is 613 Hz (Table 1), and the harmonic meshing frequencies are $N \times 613$ Hz ($N = 2, 3\ldots$), which are evident in the spectra. In the broken gear shaft state (Fig. 2(k)), there is a very large hump in the whole waveform of TSA signal. Also, the curve deviates far from the centerline, which is induced by the shaft eccentricity due to shaft crack. However, there is no peak impulse in the curve, which is often induced by tooth fault. In the corresponding spectra, there is a considerable increase in the energy level of the sidebands on the left-hand side of the meshing fundamental frequency. The amplitudes of the harmonic frequencies also have an increase, due to the energy variation induced by gear shaft faults. The waveforms in Figs. 2(c)–(i) show the behavior between the two extreme situations.

5.2. Determination of AR model order and coefficients

Data file 192, which was acquired at the beginning of the second load condition, was used as the referential healthy-state signal in order to get AR model order and coefficients. To determine the order of AR model, the AIC given by Eq. (3) was calculated using the TSA signal of one gear rotation obtained from data file 192. The AIC reached its minimum value at the order of 43 (see Fig. 3) and an AR(43) model was therefore selected for diagnosis.

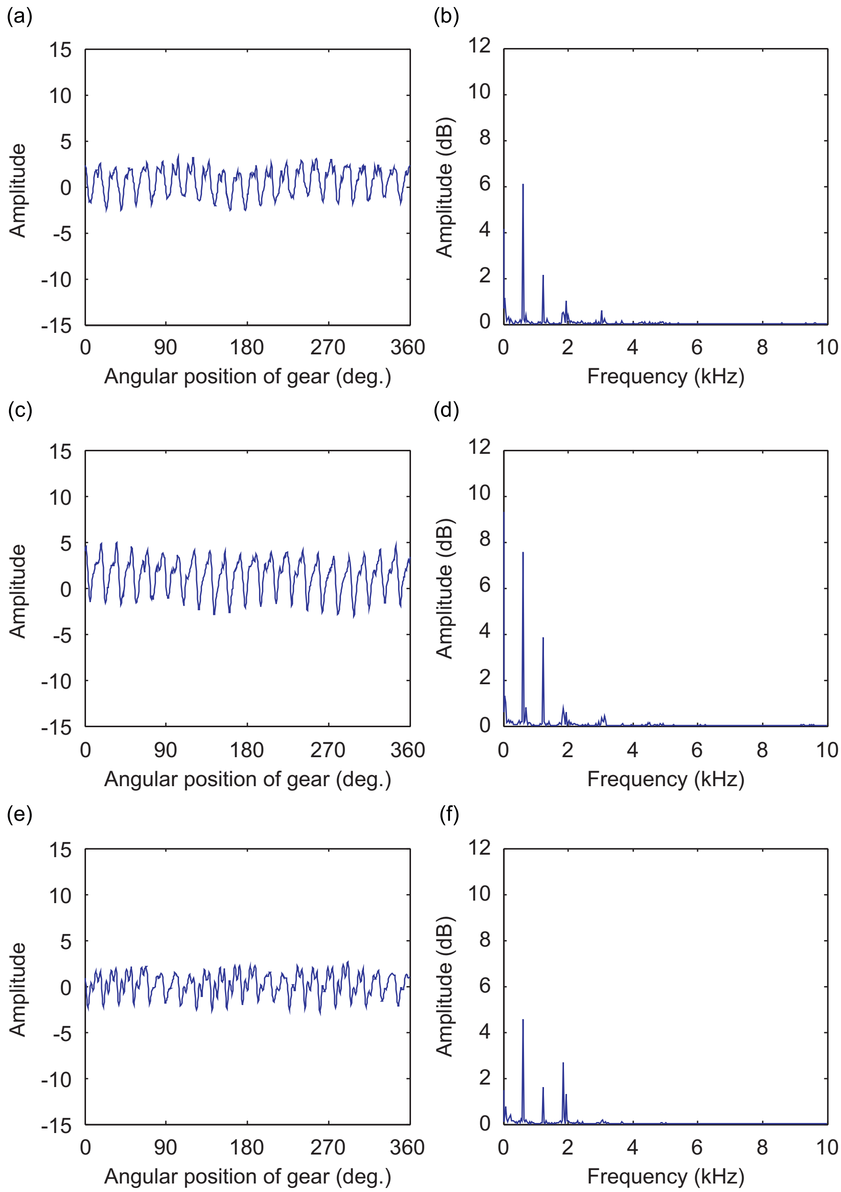

Fig. 2. Time synchronous averaging signals (a), (c), (e), (g), (i) and (k) and the corresponding amplitude spectra (b), (d), (f), (h), (j) and (l). (a) and (b) Signal 192; (c) and (d) signal 212; (e) and (f) signal 218; (g) and (h) signal 225; (i) and (j) signal 227; and (k) and (l) signal 231.

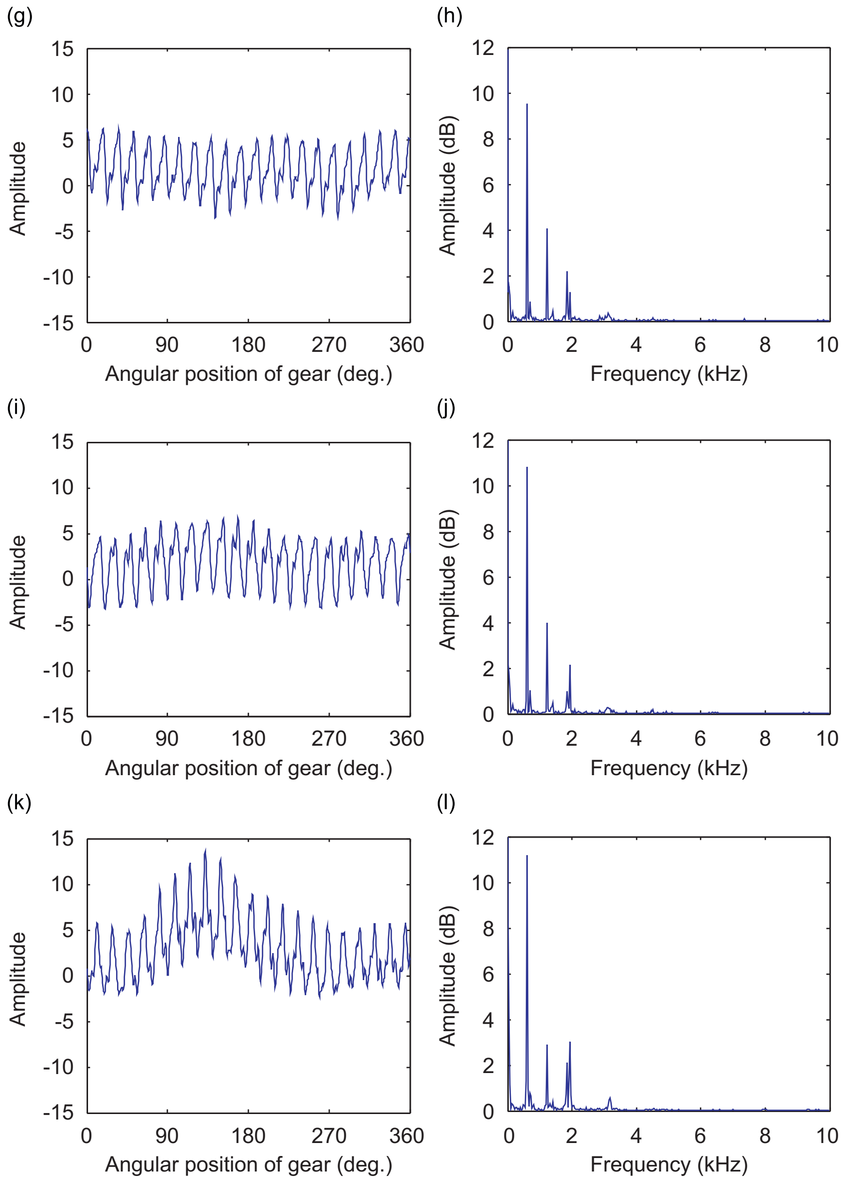

Fig. 2. (*Continued*)

Table 1
Specifications of Test-run #13.

Gearbox ID #	DS3S034013	Number of teeth (drive)	70
Make	Dodge APG	Number of teeth (pinion)	21
Model	R86005	Maximum rated input hp	4.66
Rated input speed (rpm)	1750	Gear ratio	3.333
Maximum rated output torque (lb)	555	Gear mesh frequency (Hz)	613

5.3. *Gear shaft crack detection*

The prediction error filter based on the AR(43) model was applied to the TSA signals shown in Fig. 2. The waveforms of AR model prediction error outputs are shown in Figs. 4(a)–(f).

The prediction error signal of data file 193, during which the gear shaft was in healthy state, was used as the K–S test reference signal. The K–S test was performed to compare the six AR prediction error signals with the reference signal. The

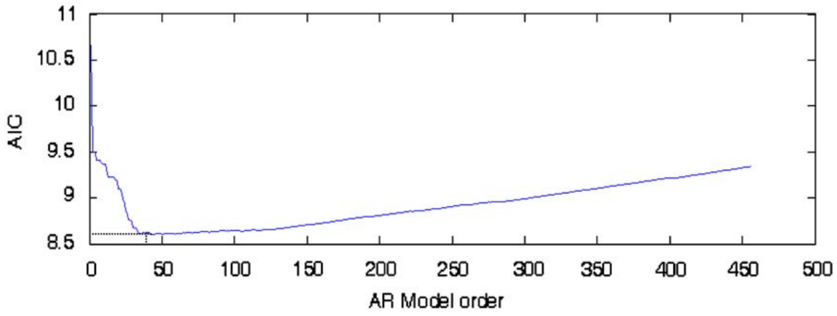

Fig. 3. Optimization of autoregressive model order.

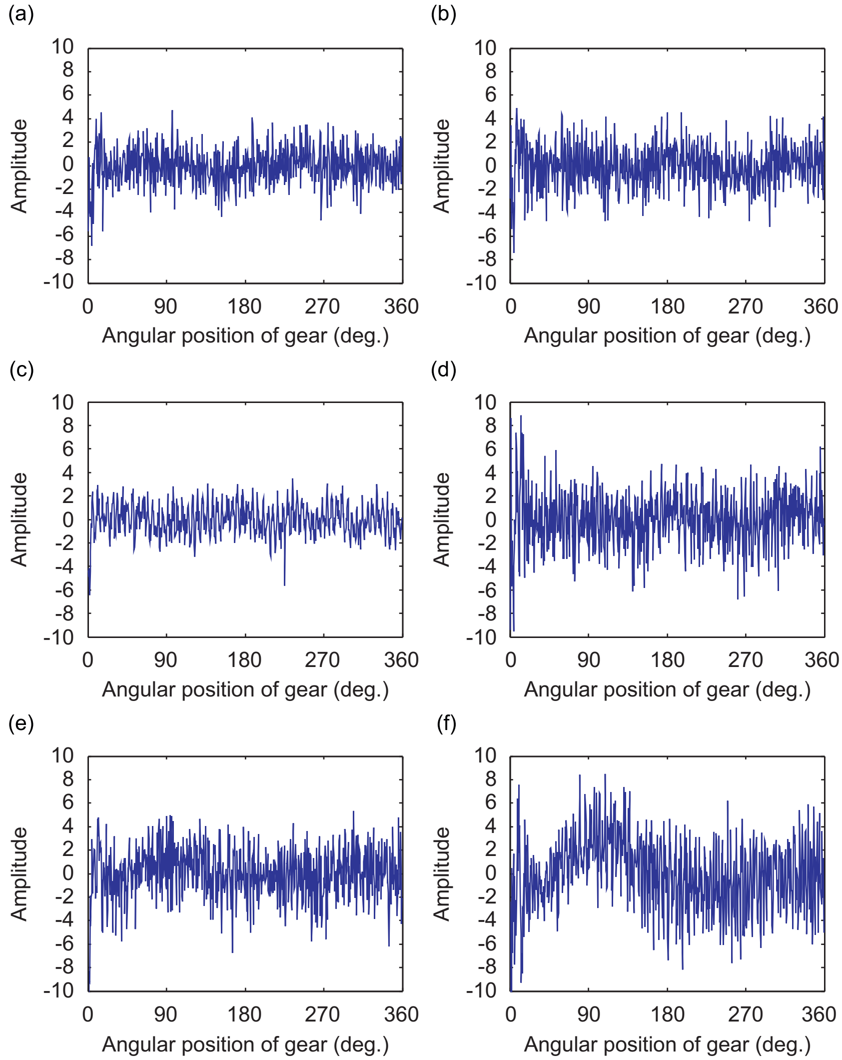

Fig. 4. Autoregressive model prediction errors: (a) data file 192, (b) data file 212, (c) data file 218, (d) data file 225, (e) data file 227 and (f) data file 231.

numbers of points both in the reference and the target prediction error signal equal the number of sampling points per gear rotation, i.e. $N_1 = N_2 = (\text{sampling frequency} \times 60)/\text{gear speed} = 686$, $N = N_1/2 = N_2/2 = 343$. For significance level $\alpha = 0.01$, the critical value for the test statistic D is equal to 0.0875. The graphs of the statistical distance D and the similarity probability $p(D)$, as well as the kurtosis and variance of AR prediction error signals are shown in Fig. 5. Note that in Figs. 5(a) and (b) the distance D and the $p(D)$-value are highly negatively correlated. Obviously, because the $p(D)$-value is an

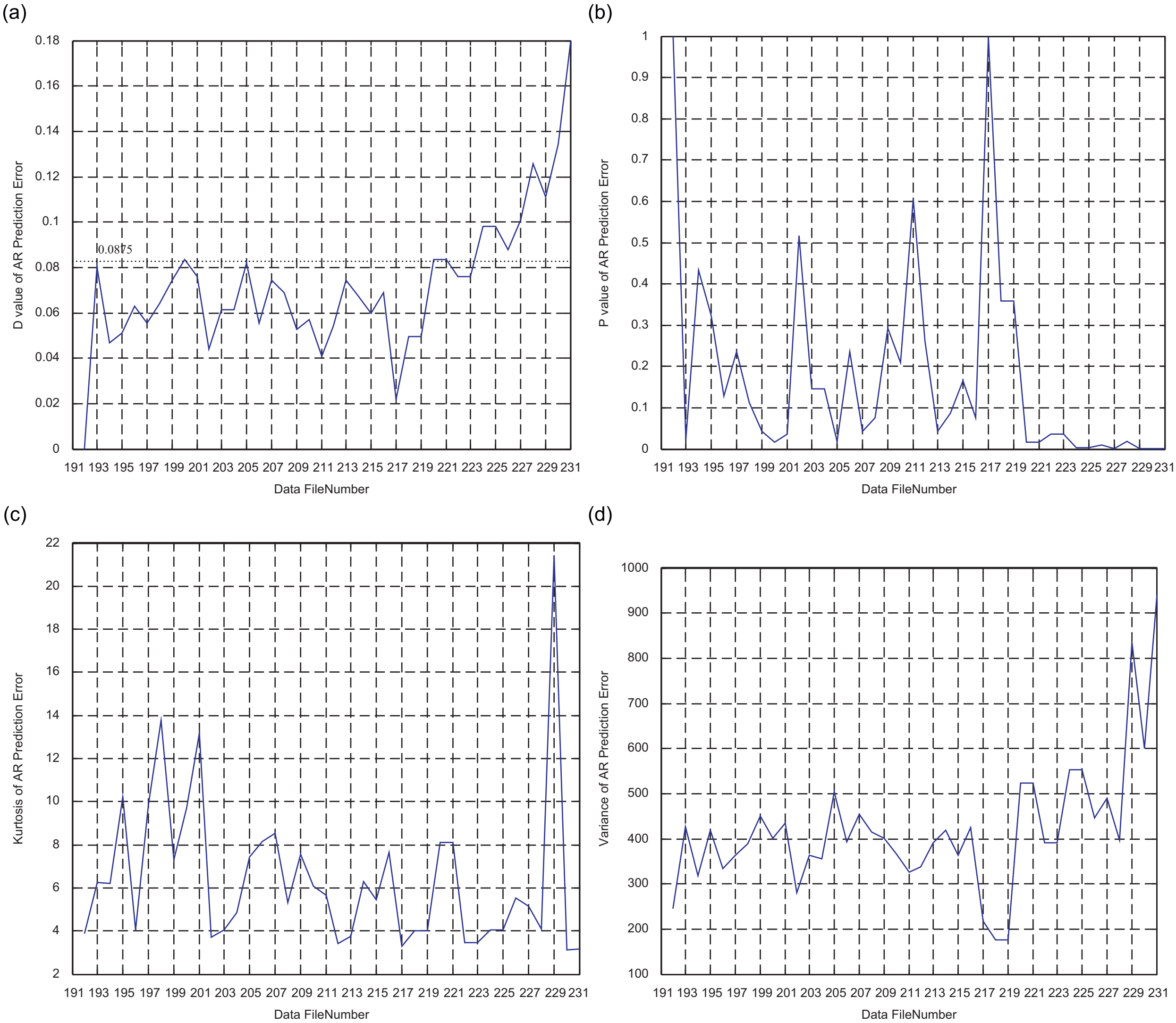

Fig. 5. Statistical measures of autoregressive prediction error over the whole lifetime of gear shaft: (a) K–S test distance, (b) K–S test similarity probability, (c) kurtosis and (d) variance.

approximation of the p-value of the K–S test, the larger the value of the test statistic $D^* = \sqrt{N}D$, the smaller the p-value and vice-versa.

Kurtosis is a statistical parameter commonly used to assess the peakedness of a signal. The kurtosis of a distribution is defined as

$$kurtosis = \frac{E(x-\mu)^4}{\sigma^4}, \tag{9}$$

where μ is the mean of x, σ the standard deviation of x, and $E(Y)$ represents the expected value of quantity Y. However, as shown in the table, kurtosis is not an adequate indicator for gear shaft crack because its value varies considerably when the gear shaft is running from the healthy state at the beginning of the second phase to the completely broken state. The behavior of kurtosis can be understood from the fact that vibration signal fluctuation induced by gear shaft crack does not show a sharp impulse, but a hump in the shape of the waveform. The variance value of AR prediction error signals also changes irregularly.

The following conclusions can be drawn from Fig. 5:

(1) The K–S test statistical distance D between the AR prediction error target signal and the reference signal (Fig. 5(a)) has an evident increase in the data file 220, revealing that the gear shaft crack may have occurred during that time. The average value of D before data file 220 is 0.0621. The D-value of data file 220 is 0.0833, rising about 22.4% above the

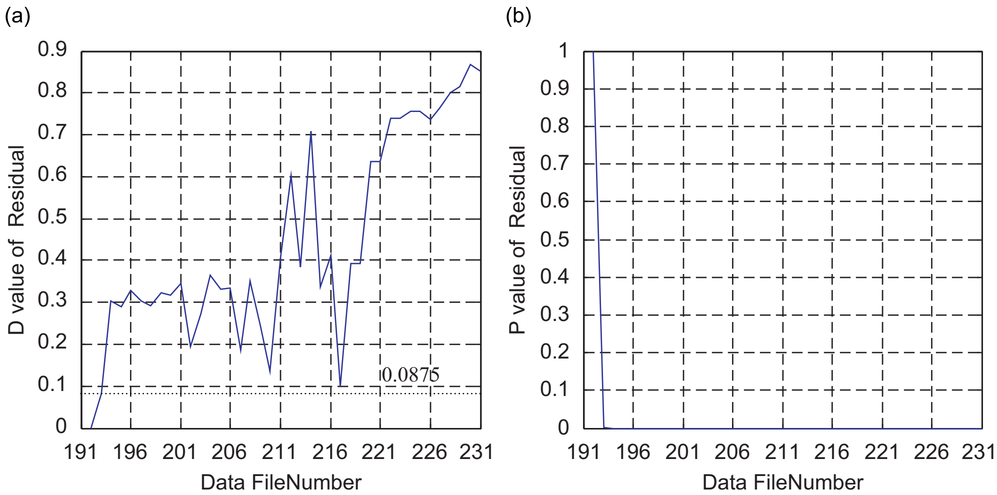

Fig. 6. K–S test of gear motion residual signal over the whole lifetime of gear shaft: (a) K–S test distance and (b) K–S test similarity probability.

average. After data file 220, the *D*-value has a nearly increasing trend, which implies the advancement of the gear shaft crack.

(2) The null hypothesis is rejected for the first time for data file 224 (the *D*-value is 0.0981) and then for all subsequent data files. Furthermore, even the smallest value of *D* after data file 224 is considerably larger than the average value before file 220 (the *D*-value in data file 226 falls to 0.0877, but it still rises 41% above the average). Gear shaft during data file 224 can be diagnosed as being in the cracked state.

(3) The similarity probability $p(D)$ of AR prediction error signal (Fig. 5(b)) decreases dramatically in data file 220 (from 0.9897 in file 194 to 0.0162 in file 220) and remains very low in the subsequent data files, which conforms to the results obtained in (1) and (2).

(4) The kurtosis of AR prediction error signal cannot be used as gear shaft fault indicator, because the value changes considerably over all lifetime of gear shaft (Fig. 5(c)).

(5) The variance of AR prediction error signal also increases in data file 220 (Fig. 5(d)), but it decreases in several subsequent data files and hence, it is unreliable for diagnosis.

It can be concluded that the gear shaft was in a healthy state during data files 212–219; there is an indication of a crack occurrence in data file 220, and the gear shaft can be diagnosed as being in the faulty state at the time of data file 224. The diagnosis indicates the faulty state earlier than the actual shutdown time of gearbox due to shaft cracks.

For a comparison, the K–S test has been applied also to the gear motion residual signal, which has been used traditionally in the literature for fault detection. The gear motion residual signal of data file 193 was used as the reference signal. The results are shown in Fig. 6. The *D*-value varies irregularly, and the null hypothesis was rejected for all the data files after file 193 even when the gear shaft was in a healthy state. Also, the $p(D)$-value falls to 0 after data file 193. Thus, we can conclude that K–S test applied to the traditional gear motion residual signal cannot indicate the occurrence of gear shaft fault.

6. Conclusions

It has been shown that the AR model prediction error signal is capable of revealing the occurrence of gear shaft crack clearly. The K–S test statistical distances *D* between the AR model prediction error target signal and the reference signal as well as the similarity probability $p(D)$ are able to detect the gear shaft crack occurrence, its advancement and the faulty state effectively. Other statistical measures, such as kurtosis and variance of the AR model prediction error signal are unable to reveal the occurrence and advancement of gear shaft cracks.

Acknowledgements

The authors are most grateful to the Applied Research Laboratory at Penn State University and the Department of the Navy, Office of the Chief of Naval Research (ONR), for providing the data used to develop this work. Referees' and editor's comments and suggestions are much appreciated.

References

[1] H. Endo, R.B. Randall, Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter, *Mechanical Systems and Signal Processing* 21 (2007) 906–919.
[2] P. Pennacchi, N. Bachschmid, A. Vania, A model-based identification method of transverse cracks in rotating shafts suitable for industrial machines, *Mechanical Systems and Signal Processing* 20 (2006) 2112–2147.
[3] A.S. Sekhar, Crack identification in a rotor system: a model-based approach, *Journal of Sound and Vibration* 270 (4–5) (2004) 887–902.
[4] H. Bach, R. Markert, Determination of the fault position in rotors for the example of a transverse crack, in: Fu-Kuo Chang (Ed.), *Structural Health Monitoring*, Technomic Publ., Lancaster/Basel, 1997, pp. 325–335.
[5] M. Dilena, A. Morassi, The use of antiresonances for crack detection in beams, *Journal of Sound and Vibration* 276 (1–2) (2004) 195–214.
[6] G.D. Gounaris, C.A. Papadopoulos, Crack identification in rotating shafts by coupled response measurements, *Engineering Fracture Mechanics* 69 (3) (2002) 339–352.
[7] L. Hamidi, J.B. Piand, H. Pastorel, W.M. Mansour, M. Massoud, Modal parameters for cracked rotors—models and comparisons, *Journal of Sound and Vibrations* 175 (1994) 265–278.
[8] W. Wang, A.K. Wong, Autoregressive model-based gear fault diagnosis, *Journal of Vibration and Acoustics* 124 (2002) 172–179.
[9] Y. Zhan, V. Makis, A.K.S. Jardine, Adaptive state detection of gearboxes under varying load conditions based on parametric modeling, *Mechanical Systems and Signal Processing* 20 (2006) 188–221.
[10] F.A. Andrade, I. Esat, M.N.M. Badi, A new approach to time-domain vibration condition monitoring: gear tooth fatigue crack detection and identification by the Kolmogorov–Smirnov, *Journal of Sound and Vibration* 240 (5) (2001) 909–919.
[11] C. Kar, A.R. Mohanty, Application of KS test in ball bearing fault diagnosis, *Journal of Sound and Vibration* 269 (2004) 439–454.
[12] W. Wang, Early detection of gear tooth cracking using the resonance demodulation, *Mechanical Systems and Signal Processing* 15 (2001) 887–903.
[13] G. Dalpiaz, A. Rivola, R. Rubini, Effectiveness and sensitivity of vibration processing techniques for local fault detection in gears, *Mechanical System and Signal Processing* 3 (2000) 387–412.
[14] MDTB data (Data CDs: test-runs #9, #7, #5 and #13), Condition-Based Maintenance Department, Applied Research Laboratory, The Pennsylvania State University, 1998.
[15] C.S. Byington, J.D. Kozlowski, Transitional data for estimation of gearbox remaining useful life, Mechanical Diagnostic Test Bed Data, Condition-Based Maintenance Department, Applied Research Laboratory, The Pennsylvania State University, 2000.
[16] R. Von Mises, *Mathematical Theory of Probability and Statistics*, Academic Press, New York, 1964.