Journal of Sound and Vibration 329 (2010) 5209–5221

Contents lists available at ScienceDirect

Journal of Sound and Vibration

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

ARX model-based gearbox fault detection and localization under varying load conditions

Ming Yang, Viliam Makis*

Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Road, Toronto, ON, Canada M5S 3G8

ARTICLE INFO

Article history:
Received 10 September 2009
Received in revised form
29 June 2010
Accepted 1 July 2010
Handling Editor: K. Shin

ABSTRACT

The development of the fault detection schemes for gearbox systems has received considerable attention in recent years. Both time series modeling and feature extraction based on wavelet methods have been considered, mostly under constant load. Constant load assumption implies that changes in vibration data are caused only by deterioration of the gearbox. However, most real gearbox systems operate under varying load and speed which affect the vibration signature of the system and in general make it difficult to recognize the occurrence of an impending fault.

This paper presents a novel approach to detect and localize the gear failure occurrence for a gearbox operating under varying load conditions. First, residual signal is calculated using an autoregressive model with exogenous variables (ARX) fitted to the time-synchronously averaged (TSA) vibration data and filtered TSA envelopes when the gearbox operated under various load conditions in the healthy state. The gear of interest is divided into several sections so that each section includes the same number of adjacent teeth. Then, the fault detection and localization indicator is calculated by applying *F*-test to the residual signal of the ARX model. The proposed fault detection scheme indicates not only when the gear fault occurs, but also in which section of the gear. Finally, the performance of the fault detection scheme is checked using full lifetime vibration data obtained from the gearbox operating from a new condition to a breakdown under varying load.

1. Introduction

Gears are the most efficient and compact devices used to transmit torques and change the angular velocities. They are widely applied in many machines, such as mining machines, automobiles, helicopters, and aircraft turbine engines. Gearbox vibration data carries a lot of useful information and it has been very popular for condition monitoring and early fault detection of gearboxes. The condition monitoring and fault detection schemes improve gear transmission systems reliability and reduce their failure occurrence. The development of fault detection and diagnostic schemes for gear transmission systems has been an active area of research in recent years, due to the need for many manufacturing companies to reduce unplanned production capacity loss caused by gear transmission systems failure and to improve equipment reliability through condition monitoring and failure prevention. Vibration data are mostly collected by accelerometers mounted on the shell of a gearbox. However, the accelerometers collect not only gear vibration signals, but

* Corresponding author. Tel.: +1 4169784184; fax: +1 4169783453.
E-mail addresses: yangming@mie.utoronto.com (M. Yang), makis@mie.utoronto.ca (V. Makis).

0022-460X/$ - see front matter © 2010 Published by Elsevier Ltd.
doi:10.1016/j.jsv.2010.07.001

also the vibration signals associated with all transmission components (such as rotating shafts, bearings, and motor rotation), which makes it very difficult to recognize the occurrence of an incipient fault.

In recent articles, advanced non-parametric approaches have been considered such as wavelets in Wang et al. [1] and Wang and McFadden [2], short-time Fourier transform in Wang and McFadden [3], and multiresolution Fourier transform in Kar and Mohanty [4] Some papers have applied parametric time-series models for gearbox fault detection, mostly assuming constant load. In Kay [5], the author showed that time series models are appropriate for representing spectra with sharp peaks (but not deep valleys) and are particularly useful for modeling sinusoidal data, which is the case of the gear meshing signals. For example, in Wang and Wong [6], an autoregressive (AR) model was considered for gear fault detection fitted to the vibration signal of the gear of interest in its healthy state. The model was used as a linear prediction error filter to process the future state signal from the same gear. The gear state was assessed using the error signal obtained as the difference between the filtered and unfiltered signals. Zhan and Makis [7] presented an adaptive Kalman filter-based AR model with varying coefficients fitted to the gear motion residual signal in the healthy state of the gear of interest considering several load conditions. The compromised model order was determined using several criteria. Liu and Makis [8] proposed a gear failure diagnosis method based on vector autoregressive (VAR) modeling of the vibration signals, dimensionality reduction applying dynamic principal component analysis and condition monitoring using a multivariate Q control chart. Rofe [9] considered an autoregressive moving-average (ARMA) model as a prediction error filter to detect gear fault accounting for load variation, and the variance and kurtosis of ARMA model residuals were used to indicate the presence or absence of a fault. However, different models were fitted for different loads, i.e., each model was built under the assumption that the load was a constant.

If the load is assumed to be a constant, the vibration signals caused by a fluctuating load are not interpreted correctly. The fluctuating load condition dramatically affects the vibration signature of the system, so that the model developed under the constant load assumption cannot recognize whether the vibration signature changes are caused by the load variation or by a failure occurrence. It should be noted that most real production systems operate under varying load. The previous AR, ARMA and VAR models which were fitted to vibration data did not consider load variation. The proper approach would be to consider load as additional information in a time series model, so that an autoregressive model with an exogenous input (ARX) or a multivariate vector ARX model would be the appropriate choice. To our knowledge, this novel approach has not been considered in the vibration literature.

In this paper, an ARX model is proposed to consider varying load as the input, so that the model and the developed fault detection scheme can be used in real situations. This paper is organized as follows. Section 2 briefly describes the experimental gear rig and the data acquisition system used in this study. In Section 3, a time-synchronous averaging (TSA) algorithm is applied to average out the contributions of both the noise and signals whose periods are not equal to the shaft rotation period of the gear of interest in a gearbox. Filtered TSA envelopes are then extracted from the TSA vibration data for each data file under varying load condition. In Section 4, using the TSA signals as output and the filtered TSA envelopes as input variable, an ARX model is identified and fitted to the data. In Section 5, F-test applied to the ARX model residuals is performed and the corresponding p-value is proposed as a quantitative indicator of fault advancement over a full lifetime of a gearbox. The conclusions and future research are summarized in Section 6.

2. Experimental set-up and data collection

The vibration data used in this research have been obtained from the Applied Research Laboratory of the Pennsylvania State University. The data were collected using the Mechanical Diagnostics Test Bed (as shown in Fig. 1), which was designed to provide experimental vibration data for research on diagnostics of a commercial transmission. The gearbox was driven at a set input speed using a 22.38 kW, 1750 rev/min drive motor, and the torque was applied by a 55.95 kW absorption motor (MDTB [10]). A vector unit capable of controlling the current output of the absorption motor accomplished 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

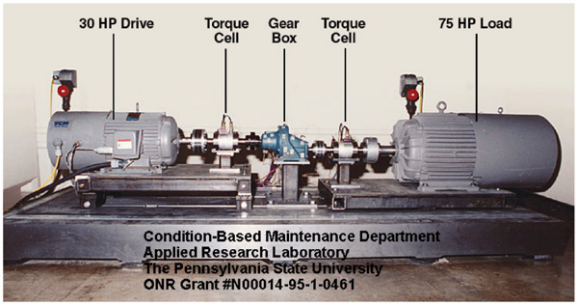

Fig. 1. Mechanical diagnostic test rig.

1.2:1 to 6:1. The gearboxes are nominally in the 3.675–14.7 kW range. The system is defined to provide the maximum versatility to speed and torque settings.

The vibration data of test-run #14 is used in this paper. In this test, the gearbox contained a 70-tooth driven gear and a 21-tooth pinion gear. The nominal output of the gearbox was 62.66 N m. The vibration data were collected by accelerometer A02 mounted on the shell of the gearbox as shown in Fig. 2.

Each data file obtained from accelerometer A02 includes 200,000 samples, which were collected in 10 s windows with 20 kHz sampling frequency and triggered by accelerometer rms thresholds. The resolution of analog-to-digital converter was 16-bit, which assured that the accuracy of the accelerometers was preserved. The motor velocity was acquired using the drive motor speed vector feedback V01. The gearbox output load was acquired from the absorption motor torque vector feedback V05. The sampling frequencies of V01 and V05 were 1 kHz. After the vibration data collection, V01 and V05 started collecting load data. In each data file, there were 10,000 samples, which were collected during 10 s periods with 1 kHz sampling frequency. Hence, the vibration data were not synchronized with the V01 and V05 data. The whole test took 116 h, and there were 338 data files in total. The gearbox was run at 100% (62.66 N m) output torque for 96 h. Then, the gearbox was run under varying load from file 194 to 338. The output load ramped up from 50% (31.33 N m) to 300% (187.98 N m) in 8 min, and then ramped down back to 50% again in 8 min. Fig. 3 shows the mean value of the drive motor speed and gearbox output torque.

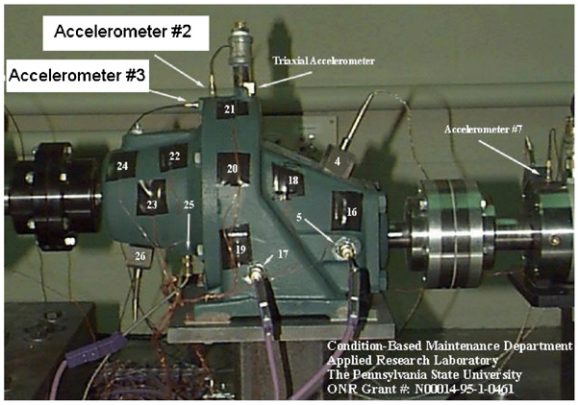

Fig. 2. Locations of accelerometers.

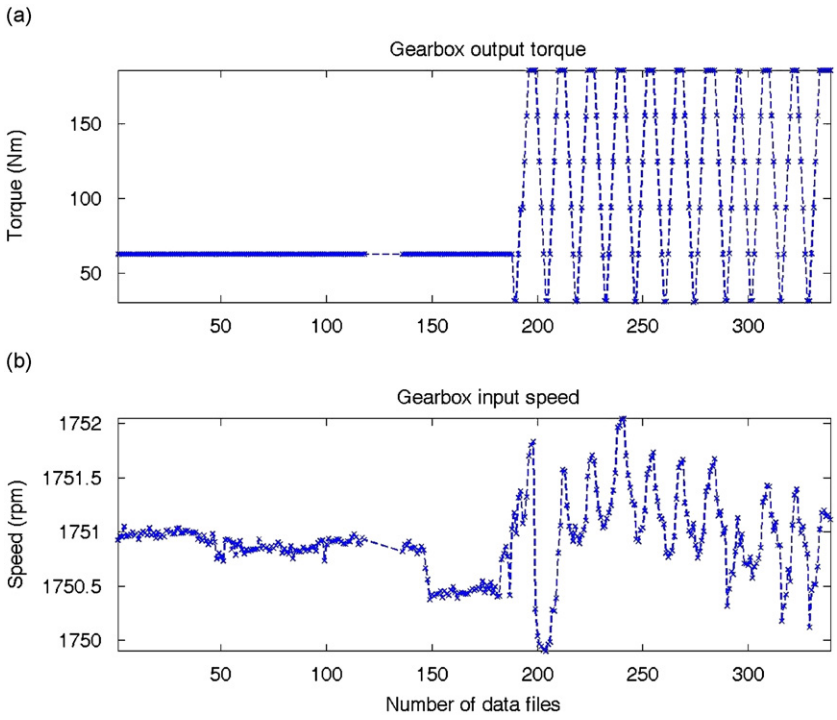

Fig. 3. (a) Gearbox output torque and (b) gearbox input speed.

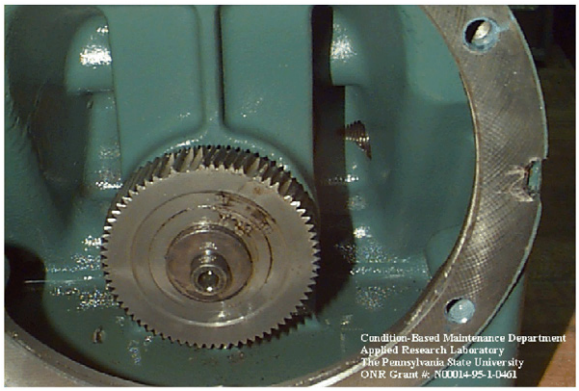

Fig. 4. Broken teeth in test-run#14.

The gearbox input speed fluctuated in the range less than 0.06%, so the speed can be consider as a constant. After 20 h, the test-run #14 was shut down with five fully broken and two partially broken teeth on the driven gear which occurred as shown in Fig. 4.

In this paper, we consider the data files obtained under varying load conditions, so the files from 194 to 338 were analyzed except file 212 that was reported unreliable in MDTB [10] due to some accelerometer problems. All the data files were separated into two groups. The files 194–246 were used to build the ARX model and the files 247–338 were used to detect the gearbox failure.

We chose six data files to demonstrate our procedures. The file 283 was selected, since it was the first file showing that the gear fault occurred. The results obtained for file 283 were compared with the result obtained for file 254, both of these files were collected when the gearbox was ramped up to 300% output load. Also, we chose files 259 and 288 for a comparison, since for both of them the output load was ramped down to 100%. When the gearbox runs at a low output load condition, the detection of gear failure is much more difficult than when the gearbox runs at a high output load condition. For the low load condition, we compared the results using the data files 261 and 290, both were collected at 50% output torque. Based on the inspection performed after gearbox failure, the data files 254, 259, and 261 were collected when the gearbox was in healthy state, and files 283, 288, and 290 were collected when the gearbox was in failure state.

3. Time-synchronous averaging algorithm and filtered envelope extraction

The filtered TSA envelope is composed of the information of load variation and gear shaft vibration due to shaft imbalance. This envelope is considered as an input for the ARX model. A method how to extract the filtered TSA envelope from the collected vibration data is presented in this section. First, the TSA signal is obtained and then the envelope carrying load information is extracted from the TSA signal. Finally, a low pass filter is applied to remove the gear fault signature from the TSA envelope.

Vibration data obtained from the gearbox represent a complex combination of information plus noise produced by the background. Time-domain synchronous averaging algorithm (TSA) providing an average time signal of one individual gear over a large number of cycles has been commonly used in the detection of gear faults (see e.g. [11,12,13]). TSA reduces the contributions of both noise and signals whose periods are not equal to the given gear shaft rotation period by synchronizing the sampling of the vibration signal with the shaft rotation of a particular gear and evaluating the ensemble average over many revolutions with the start of each revolution at the same angular position. Suppose there are n data points in each vibration data file collected from a gearbox operating under a constant speed. $\{\mathbf{V}(k)\}$, $k=1,2,...,n$, is used to denote this discrete time vibration data. Since the gearbox is running under a constant speed, the number of sampling points which correspond to one complete revolution of the gear of interest is described as

$$K=\left\lceil \frac{f_s}{f_m/N_t} \right\rceil \tag{1}$$

where f_s is the sampling frequency, f_m is the fundamental meshing frequency of the gear of interest, N_t is the number of teeth of the given gear, and $\lceil\bullet\rceil$is the ceil function returning the closest higher integer. The number of cycles to be averaged can be obtained from $M_0=\lfloor N/K \rfloor$, where N is total number of sampling points in each data file and $\lfloor\bullet\rfloor$ is the floor function returning the closest lower integer. In the time domain, the conventional TSA signal of $\{\mathbf{V}(k)\}$ is defined as

$$\mathbf{V}^{\mathrm{TSA}}(k)=\frac{1}{M_0}\sum_{i=0}^{M_0-1}\mathbf{V}(k+iK), \quad k=1,2,...,K \tag{2}$$

In the next section, the TSA signal is going to be processed by an ARX model with an optimal model order b, s, and r, which requires that the number of TSA signals is $K+c$, where $c=\max(r,b+s)$, to obtain K residuals corresponding to a

complete revolution of the gear of interest. Thus, the equation to calculate the TSA signal is modified as follows:

$$\mathbf{V}^{\mathrm{TSA}}(k)=\frac{1}{M}\sum_{i=0}^{M-1}\mathbf{V}(k+iK),\quad k=1,2,\ldots,K+c \tag{3}$$

where

$$M=\lfloor N/K\rfloor-\lceil c/K\rceil$$

The TSA signals contain not only gear fault signatures, but also gear meshing frequencies, gear shaft imbalance signals, and load variation signatures. An ARX model is introduced to partially filter out the gear meshing frequencies, shaft imbalance signals, and load variation signatures from the TSA signals using a filtered TSA envelope as an external input. We will now present a method how to extract a filtered TSA envelope. First, we localize peaks of the meshing signal of each gear. The upper envelope is made up of all the local maximum values. We filter out the peaks caused by the local gear fault, and fit a piecewise linear function l_u to the modified upper envelope. Using the same method, the other piecewise function l_d is fitted to the lower envelope using the local minimum value of each gear meshing signal. The estimated load signals $\{\hat{l}(k)\}$ for each data file are given as

$$\hat{l}(k)=\frac{l_u(k)-l_d(k)}{2} \tag{4}$$

When a local gear fault occurs, a peak may be found in the estimated load signals and this kind of peak can be considered as an impulse component when compared with estimated load signals $\{\hat{l}(k)\}$. The filtered TSA envelope $\{\mathbf{x}(k)\}$ is extracted from the estimated load signals $\{\hat{l}(k)\}$ by averaging out the impact this impulse component caused by the local gear fault may have. So, the filtered TSA envelope $\{\mathbf{x}(k)\}$ contain the information related to varying load and shaft imbalance, and this information is not affected by local gear faults.

Then, we applied previously described TSA technique to the vibration data collected by accelerometer A02. Each data file contained the data collected during 10 s. As shown in Fig. 5, the horizontal axis is the time and the vertical axis is the amplitude of the acceleration.

Based on these figures, it is difficult to obtain any information of gear state, so we apply the TSA algorithm and extract filtered TSA envelopes, which are shown in Fig. 6.

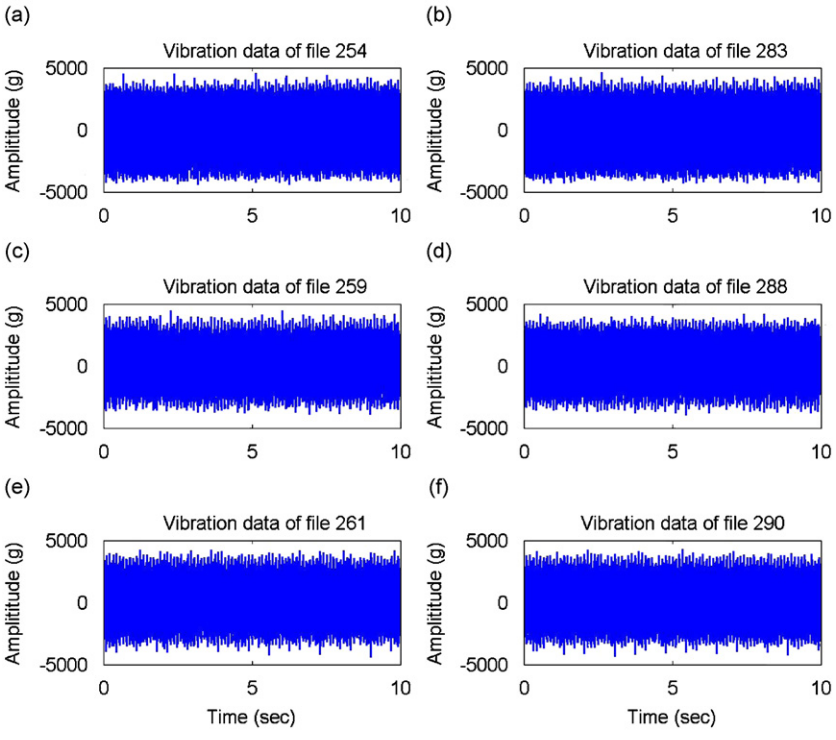

Fig. 5. Vibration data: (a) healthy state 300% output torque, (b) faulty state at 300% output torque, (c) healthy state at 100% output torque, (d) faulty state at 100% output torque, (e) healthy state at 50% output torque, and (f) faulty state at 50% output torque.

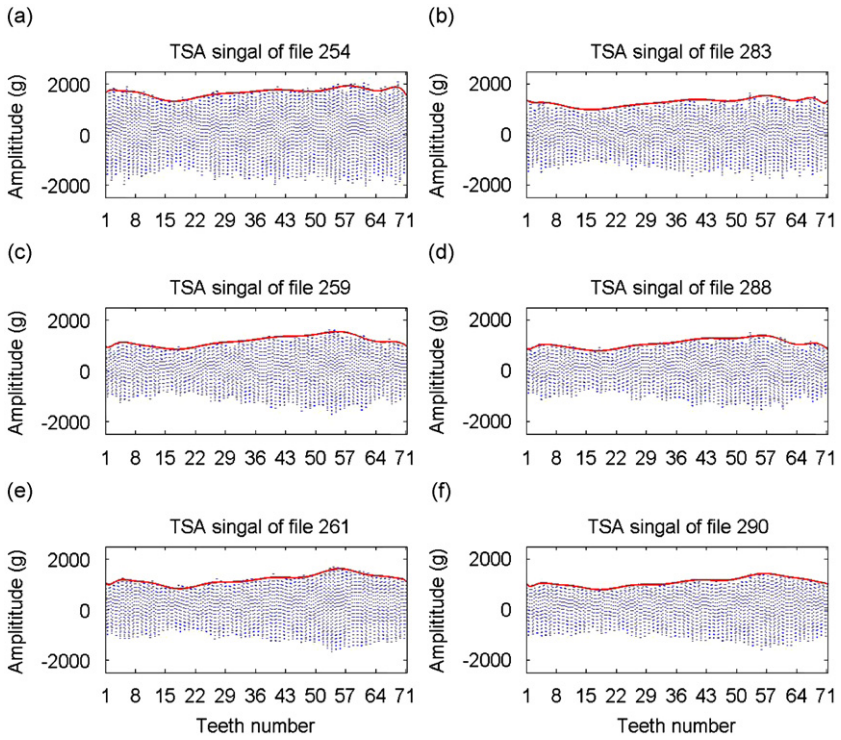

Fig. 6. TSA signals (dotted line) and filtered TSA envelop (solid line): (a) healthy state at 300% output torque, (b) faulty state at 300% output torque; (c) healthy state at 100% output torque, (d) faulty state at 100% output torque; (e) healthy state at 50% output torque, and (f) faulty state at 50% output torque.

4. ARX model using gear vibration data under varying load

In this study, an ARX model is utilized to detect and localize gear failure for a gearbox operating under varying load. The residual signal of an ARX model represents the difference from the average tooth-meshing vibration in a healthy state and usually shows evidence of faults earlier and more clearly than the TSA signal. Since the majority of energy in the healthy state of the target gear is concentrated at the meshing fundamental and its harmonics, the ARX model residual signal will be much less sensitive to the load variation and shaft imbalance than the TSA signal. In the next section, we present an approach to identify the model order and to estimate parameters of an ARX model and show numerical results using real data.

4.1. ARX model based on the modified TSA signal and filtered TSA envelope

The TSA signal $\mathbf{V}^{\mathrm{TSA}}$ allows signal frequencies, which are synchronous with the meshing frequency of interest, to be isolated from all other signals. However, $\mathbf{V}^{\mathrm{TSA}}$ still contains a large numbers of components which are not related to the gear fault signature, such as the gear meshing signal $\mathbf{V}_g$, the load variation signature, and the gear shaft vibration signal caused by shaft geometric asymmetry and assembly errors [9]. In gear fault detection and diagnosis, it is necessary to isolate these signals from the TSA signal to obtain a residual signal that is more efficiently indicative of the gear fault.

For healthy gears, the gear meshing signal is modulated by some low shaft order (first and/or second order) functions. The spectrum of the gear meshing signal is therefore dominated by the sharp spectral peaks at the fundamental gear meshing frequency and its harmonics accompanied by some low-order modulation sidebands. The TSA signal of a gear meshing signal can be expressed by the following equation [11]:

$$\mathbf{V}_g(k) = \frac{1}{H}\sum_{h=0}^{H} \overline{A}_h\left[1+\mathbf{a}_h(k)\right]\cos\left\{2\pi f_h k+\beta_h+\boldsymbol{\theta}_h(k)\right\}, \quad k=1,2,\ldots,K+b \tag{5}$$

where $h=0,1,\ldots,H$ is the meshing harmonic number, $\overline{A}_h$ is the amplitude at the hth harmonic frequency (i.e., $f_h = h\times f_s$), $\{\mathbf{a}_h(k)\}$ is the amplitude modulation function, β_h is the initial phase, and $\{\boldsymbol{\theta}_h(k)\}$ is the phase modulation function at the hth harmonic.

Kay [5] proved that the AR model is appropriate to represent spectra of gear meshing signals with sharp peaks. For a gearbox operating under constant load and speed, and the gear shaft imbalances ignored, Wang and Wong [6] successfully

applied an AR model to represent gear meshing signal. The AR model is given by the following equation:

$$\mathbf{V}_g(k) = \boldsymbol{\delta}_r(\mathbf{B})\mathbf{V}_g(k) + \mathbf{n}_g(k) \tag{6}$$

where

$$\boldsymbol{\delta}_r(\mathbf{B}) = \delta_1\mathbf{B} - \cdots \delta_r\mathbf{B}^r \tag{7}$$

$\mathbf{B}$ denote the backshift operator and $\mathbf{n}_g$ are the AR models residuals.

When the load variation and shaft imbalance are present, the filtered TSA envelope $\{\mathbf{x}(k)\}$ can be used to represent the combination of the signature caused by varying load and shaft imbalance. Then, the TSA signal $\mathbf{V}^{\mathrm{TSA}}$ is described as an ARX model with the filtered TSA envelope as an exogenous variable by the following equation:

$$\mathbf{V}^{\mathrm{TSA}}(k) = \mathbf{V}_g(k) + \mathbf{w}_p(\mathbf{B})\mathbf{x}(k) + \mathbf{n}(k) \tag{8}$$

where

$$\mathbf{w}_p(\mathbf{B}) = w_0 + w_1\mathbf{B} + \cdots + w_p\mathbf{B}^p \tag{9}$$

and $\{\mathbf{n}(k)\}$ is an error term. More details about ARX model are introduced in Appendix A. Submitting Eq. (6) into Eq. (8), we can represent the TSA signal as

$$\begin{aligned}\mathbf{V}^{\mathrm{TSA}}(k) &= \boldsymbol{\delta}_r(\mathbf{B})\mathbf{V}_g(k) + \mathbf{n}_g(k) + \mathbf{w}_p(\mathbf{B})\mathbf{x}(k) + \mathbf{n}(k)\\ &= \boldsymbol{\delta}_r(\mathbf{B})\mathbf{V}_g(k) + \mathbf{n}_g(k) + \mathbf{w}_p(\mathbf{B})\mathbf{x}(k) + \mathbf{n}(k) + \boldsymbol{\delta}_r(\mathbf{B})\mathbf{w}_p(\mathbf{B})\mathbf{x}(k) + \boldsymbol{\delta}_r(\mathbf{B})\mathbf{n}_g(k) - \boldsymbol{\delta}_r(\mathbf{B})\mathbf{w}_p(\mathbf{B})\mathbf{x}(k) - \boldsymbol{\delta}_r(\mathbf{B})\mathbf{n}_g(k)\\ &= \boldsymbol{\delta}_r(\mathbf{B})[\mathbf{V}_g(k) + \mathbf{w}_p(\mathbf{B})\mathbf{x}(k) + \mathbf{n}_g(k)] + \mathbf{w}_p(\mathbf{B})[1 - \boldsymbol{\delta}_r(\mathbf{B})]\mathbf{x}(k) + [1 - \boldsymbol{\delta}_r(\mathbf{B})]\mathbf{n}(k) + \mathbf{n}_g(k)\\ &= \boldsymbol{\delta}_r(\mathbf{B})\mathbf{V}^{\mathrm{TSA}}(k) + \mathbf{w}_p(\mathbf{B})[1 - \boldsymbol{\delta}_r(\mathbf{B})]\mathbf{x}(k) + [1 - \boldsymbol{\delta}_r(\mathbf{B})]\mathbf{n}(k) + \mathbf{n}_g(k)\end{aligned} \tag{10}$$

In order to rewrite the above equation as an ARX model, letting $\boldsymbol{\omega}_s(\mathbf{B})\mathbf{B}^b = \mathbf{w}_p(\mathbf{B})[1 - \boldsymbol{\delta}_r(\mathbf{B})]$ and $\mathbf{N}(k) = [1 - \boldsymbol{\delta}_r(\mathbf{B})]\mathbf{n}(k) + \mathbf{n}_g(k)$, we have an ARX model described by the following equation:

$$\mathbf{V}^{\mathrm{TSA}}(k) = \boldsymbol{\delta}_r(\mathbf{B})\mathbf{V}^{\mathrm{TSA}}(k) + \boldsymbol{\omega}_s(\mathbf{B})\mathbf{B}^b\mathbf{x}(k) + \mathbf{N}(k) \tag{11}$$

where $\{\mathbf{N}(k)\}$ is the noise term approximated by the ARX model residuals. For fault detection, only the information embedded in residuals is of interest.

4.2. ARX model order identification and parameters estimation

In this section, we show how to identify an ARX model order and estimate the model parameters. The structure of an ARX model is represented by the equation:

$$\mathbf{y}(k) = \boldsymbol{\delta}_r(\mathbf{B})\mathbf{y}(k) + \boldsymbol{\omega}_s(\mathbf{B})\mathbf{B}^b\mathbf{x}(k) + \mathbf{N}(k) \tag{12}$$

where

$$\omega_s(\mathbf{B}) = \omega_0 + \omega_1\mathbf{B} + \cdots + \omega_s\mathbf{B}^s \tag{13}$$

$$\delta_r(\mathbf{B}) = \delta_1\mathbf{B} + \cdots + \delta_r\mathbf{B}^r \tag{14}$$

$\mathbf{y}$ is the TSA vibration signal $\{\mathbf{V}^{\mathrm{TSA}}(k)\}$, $\mathbf{x}$ is the ARX model exogenous input representing the filtered TSA envelop, and $\{\mathbf{N}(k)\}$ is assumed as a Gaussian white noise with zero mean and variance σ_N^2 when the gearbox is in a healthy state.

In order to test the normality assumption of $\{\mathbf{N}(k)\}$, we apply the Jarque–Bera test [15,16], that is a goodness-of-fit measure of departure from normality based on the sample kurtosis K_J and skewness S_J with sample size n_J. The test statistic J is defined as

$$J = \frac{n_J}{6}\left[S_J^2 + \frac{(K_J - 3)^2}{4}\right] \tag{15}$$

If the sample size n_J is large (e.g. $n_J > 3$, preferably $n_J > 100$), J is chi-square distribution with two degrees of freedom. Based on the Jarque–Bera test with significant level 0.01, each ARX residual corresponding to from data files 194 to 246 is normal distribution except data file 210. Then we calculate the autocorrelation of all ARX model residual from data files 194 to 246, and Fig. 9 shows that the uncorrelated assumption about $\{\mathbf{N}(k)\}$ is reasonable.

For given model order parameters b, s, and r, the parameters of the ARX model are estimated using the least squares algorithm [17]:

$$\hat{\boldsymbol{\theta}}_m = \left[\sum_{k=1}^{m} \varphi(k)\varphi^T(k)\right]^{-1} \sum_{k=1}^{m} \varphi(k)\mathbf{y}(k) \tag{16}$$

where m is the number of data values, $\hat{\boldsymbol{\theta}}_m$ is the estimated ARX model parameter vector

$$\hat{\boldsymbol{\theta}} = \begin{bmatrix}\hat{\delta}_1 & \ldots & \hat{\delta}_r & \hat{\omega}_0 & \ldots & \hat{\omega}_s\end{bmatrix}^T \tag{17}$$

and φ is the input and output series vector used to build the ARX model,

$$\varphi(k)=\left[-\mathbf{y}(k-1),\ldots,\quad -\mathbf{y}(k-r),\qquad \mathbf{x}(k-b),\ldots,\quad \mathbf{x}(k-b-s),\right]^T \tag{18}$$

The model order parameters b, s, and r are determined using Bayesian information criterion (BIC) [18]

$$\mathrm{BIC}_{b,s,r}=\ln(\hat{\sigma}_N^2)+d\frac{\ln(m)}{m} \tag{19}$$

where $\hat{\sigma}_N^2$ is the maximum likelihood estimate of σ_N^2, and $d=s+r+1+1$ is the number of estimated parameters in the model. BIC imposes a greater penalty for the number of estimated model parameters than does the Akaike information criterion [19]:

$$\mathrm{AIC}_{b,s,r}=\ln(\hat{\sigma}_N^2)+d\frac{2}{m}+\text{constant} \tag{20}$$

Once the ARX model is identified and all its parameters are estimated, the ARX model residuals can be calculated using the following equation:

$$\hat{\mathbf{N}}(k)=\left[1-\hat{\delta}_r(\mathbf{B})\right]\mathbf{y}(k)\text{-}\hat{\omega}_s(\mathbf{B})\mathbf{B}^b\mathbf{x}(k) \tag{21}$$

4.3. Numerical results

In this paper, the filtered TSA envelope and TSA data $\mathbf{V}^{\mathrm{TSA}}$ from files 194 to 246 (shown in Fig. 7) are used to fit an ARX model. Comparing Fig. 7 with Fig. 3, we can find out that the filtered TSA envelope contains the components caused not only by varying load but also by the gear shaft imbalance.

As mentioned above, least squares algorithm is applied to estimate the ARX model parameters. The order of the model is identified by BIC approach that was introduced in Section 4.2. Fig. 8 shows that the minimum value of BIC for our ARX model is obtained for b=0, s=3, and r=66.

The parameters of this ARX model $\mathbf{y}(k)=\boldsymbol{\delta}_{65}(\mathbf{B})\mathbf{y}(k-1)+\boldsymbol{\omega}_3(\mathbf{B})\mathbf{x}(k)+\mathbf{N}(k)$ were obtained as follows:

$$\boldsymbol{\omega}_3(\mathbf{B})=1.082-1.674\mathbf{B}+0.5991\mathbf{B}^2 \tag{22}$$

$$\begin{aligned}\boldsymbol{\delta}_{65}(\mathbf{B})=\;&1.598\mathbf{B}-1.01\mathbf{B}^2+0.2965\mathbf{B}^3+0.1293\mathbf{B}^4+0.008993\mathbf{B}^5-0.1169\mathbf{B}^6\\&-0.011\mathbf{B}^7+0.08425\mathbf{B}^8-0.1159\mathbf{B}^9+0.06916\mathbf{B}^{10}-0.01803\mathbf{B}^{11}-0.1243\mathbf{B}^{12}\\&+0.1699\mathbf{B}^{13}-0.1695\mathbf{B}^{14}+0.04267\mathbf{B}^{15}-0.03125\mathbf{B}^{16}+0.0126\mathbf{B}^{17}-0.0005206\mathbf{B}^{18}\\&-0.136\mathbf{B}^{19}+0.1622\mathbf{B}^{20}+0.1128\mathbf{B}^{21}-0.2211\mathbf{B}^{22}+0.06914\mathbf{B}^{23}+0.09999\mathbf{B}^{24}\\&-0.1493\mathbf{B}^{25}+0.1034\mathbf{B}^{26}+0.001125\mathbf{B}^{27}-0.05813\mathbf{B}^{28}-0.001733\mathbf{B}^{29}+0.01286\mathbf{B}^{30}\\&+0.03604\mathbf{B}^{31}+0.1178\mathbf{B}^{32}+0.01246\mathbf{B}^{33}-0.2306\mathbf{B}^{34}+0.2217\mathbf{B}^{35}-0.1782\mathbf{B}^{36}\\&+0.02661\mathbf{B}^{37}-0.08044\mathbf{B}^{38}+0.123\mathbf{B}^{39}-0.02903\mathbf{B}^{40}-0.02264\mathbf{B}^{41}+0.05132\mathbf{B}^{42}\end{aligned}$$

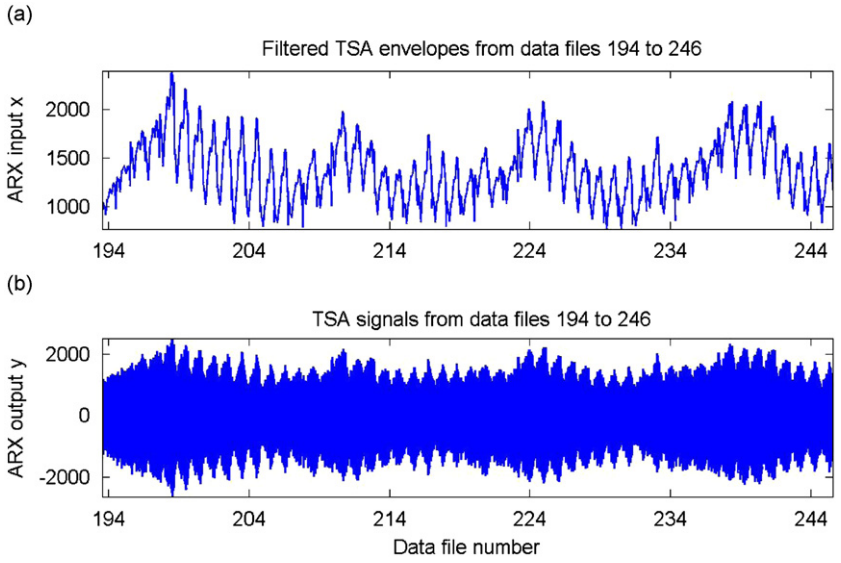

Fig. 7. TSA data for ARX model building: (a) filtered TSA envelopes used as an ARX model input and (b) TSA vibration data used as an ARX model output.

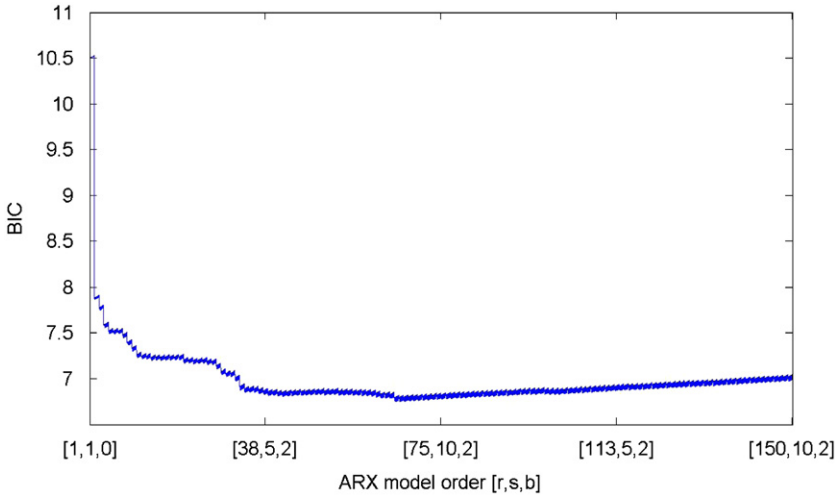

Fig. 8. ARX model order selection using BIC criterion.

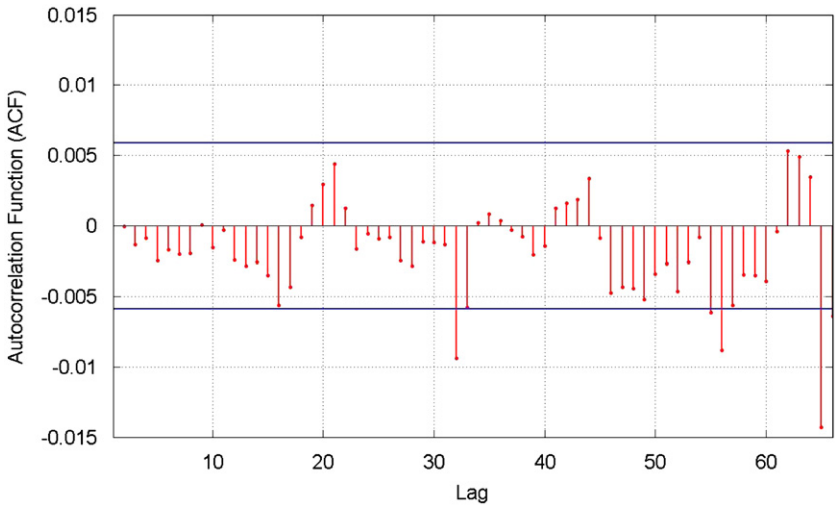

Fig. 9. Autocorrelation function for ARX model residuals.

$$\begin{aligned}&-0.01523\mathbf{B}^{43}+0.04177\mathbf{B}^{44}-0.02584\mathbf{B}^{45}+0.02091\mathbf{B}^{46}+0.02912\mathbf{B}^{47}+0.024\mathbf{B}^{48}\\&-0.03102\mathbf{B}^{49}+0.0716\mathbf{B}^{50}-0.02914\mathbf{B}^{51}+0.02916\mathbf{B}^{52}-0.119\mathbf{B}^{53}+0.06028\mathbf{B}^{54}\\&+0.007412\mathbf{B}^{55}-0.0362\mathbf{B}^{56}+0.02966\mathbf{B}^{57}-0.00834\mathbf{B}^{58}-0.01164\mathbf{B}^{59}-0.03023\mathbf{B}^{60}\\&+0.05247\mathbf{B}^{61}-0.01407\mathbf{B}^{62}+0.04786\mathbf{B}^{63}-0.0418\mathbf{B}^{64}+0.2373\mathbf{B}^{65}-0.196\mathbf{B}^{66}\end{aligned} \tag{23}$$

The residual checking described in Box et al. [14] is used to examine whether the ARX model is correct. Fig. 9 shows the autocorrelations of the ARX model residuals $\{\mathbf{N}(k)\}$ using the data from files194 to 246. All autocorrelations are very close to zero and do not show any marked correlation pattern. Hence, the residual checking indicates that the model is adequate and correct.

Now, we compute the ARX residuals for each data file from 194 to 338. Fig. 10 shows the ARX residuals for data files 254, 283, 259, 288, 261, and 290. The gear failure still cannot be easily detected using the ARX residuals, so the *F*-test is introduced to analyze the ARX model residuals. The results of the *F*-test can tell us not only when the gear failure occurs, but also where the gear failure is localized.

5. Hypothesis testing

The variance of residuals shows a strong relationship with the state of each gear tooth, so variance of the ARX model residuals is investigated in this research. After the ARX model residuals are obtained, an *F*-test is applied to the variance of

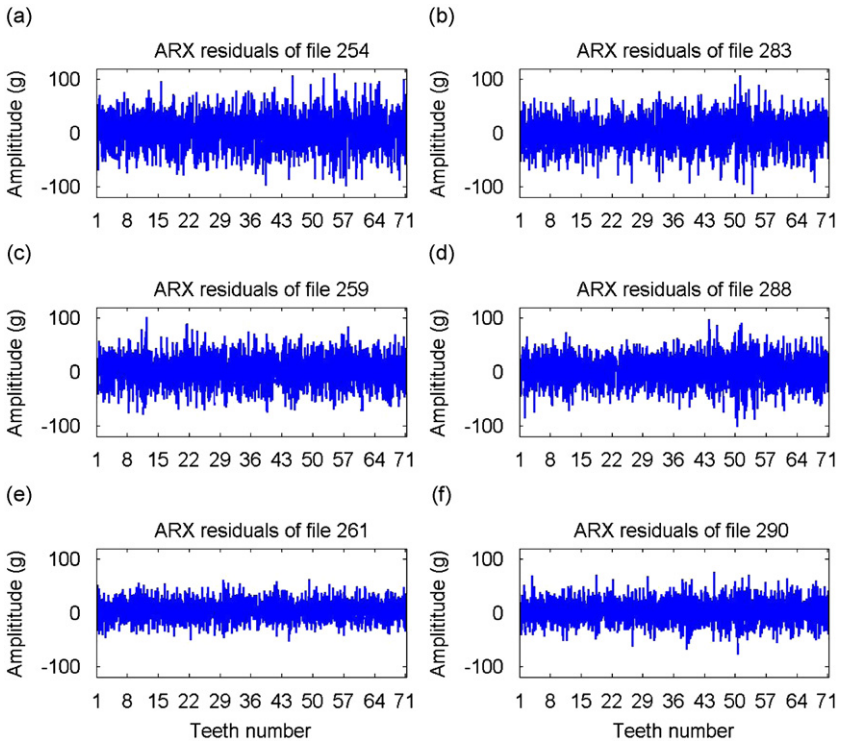

Fig. 10. ARX residuals: (a) healthy state at 300% output torque, (b) faulty state at 300% output torque, (c) healthy state at 100% output torque, (d) faulty state at 100% output torque, (e) healthy state at 50% output torque, and (f) faulty state at 50% output torque.

the residuals. More details about *F*-test are given in Appendix B. The *F*-test result is then used as indicator to decide whether the gear default occurred or not, which reduces the complexity of the decision making based on the gear motion residuals.

First, the 70-tooth driven gear is divided into 10 sections so that each section contains 7 adjacent teeth corresponding to 36° out of the total of 360° of one shaft rotation. The ARX model residuals are uncorrelated and independent when the teeth are in the healthy state. Hence, the model residuals would be normally distributed with zero mean. In the healthy state, it is reasonable to assume that all the gears are identical, so we denote the population variance of the ARX residuals of each data file as σ^2, the sample variance of the n_s residuals in each section as S_s^2, and the sample variance of the n_n residuals of the other 63 teeth as S_n^2. Then, *F*-statistic is utilized to test the equality of the population variance $\sigma_s^2$and the no section variance σ_n^2. Hence, we test the hypothesis

$$H_0{:}\sigma_s^2 = \sigma_n^2$$
$$H_1{:}\sigma_s^2 > \sigma_n^2$$

using the following *F*-statistic

$$F_{v_s,v_n} = \frac{S_s^2}{S_n^2}, \tag{24}$$

where $v_s = n_s - 1$ and $v_n = n_n - 1$ are degrees of freedom. Since we have an one sided test, the rejection of the null hypothesis H_0 means that the variance σ_s^2 is greater than σ_n^2. The residuals of the selected section are more variable than the residuals of other sections, which indicate a gear fault occurrence in this section. Fig. 11 shows that there is no rejection of the null hypothesis H_0 for gear teeth 1–35 and 57–70. In the section containing gear teeth 43–49, the H_0 is rejected starting from the data file 283, which indicates that the gear fault occurred. Also, for the sections of 36–42 and 50–56, the hypothesis H_0 is rejected starting from the data files 335 and 338, respectively.

6. Conclusions

In this paper, a gearbox fault detection scheme has been proposed using vibration data collected under varying load condition. A modified TSA algorithm has been introduced to compute a TSA signal for an ARX model building. The ARX model has been fitted using the filtered TSA envelope as an external input and the TSA signal as the output. The parameters of the ARX model have been estimated using the least squares algorithm, and the order of the ARX model has been

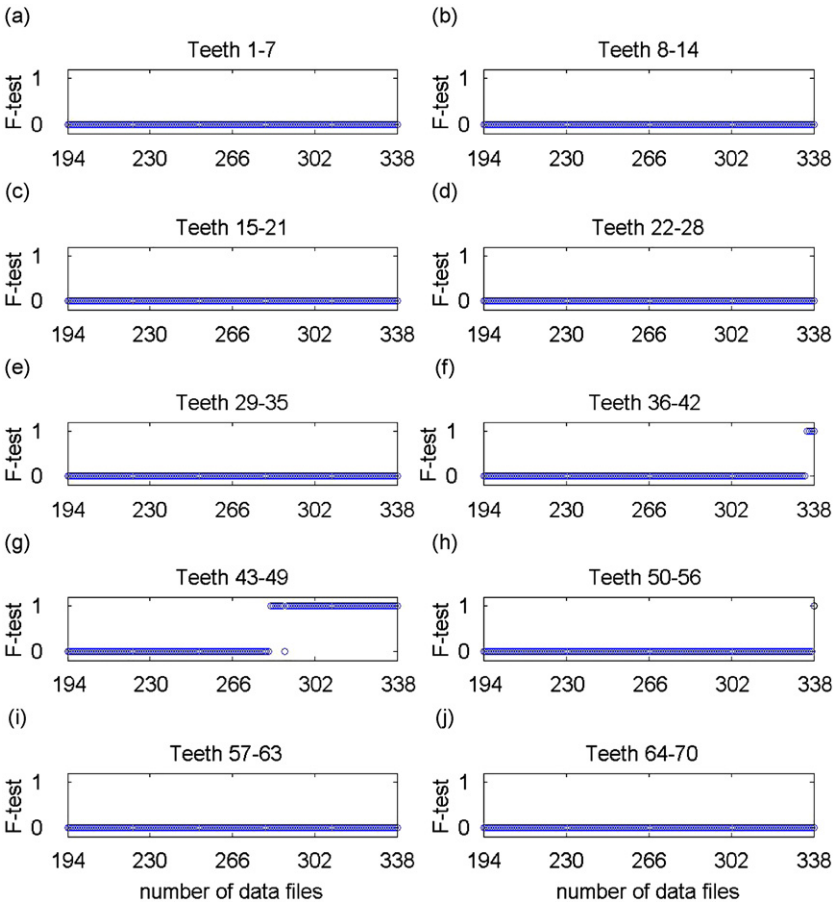

Fig. 11. *F*-test of ARX model residuals (1 indicates rejection of H_0).

determined using BIC criterion. Then, the *F*-test has been applied to the variances of the ARX model residuals for each data file, and the corresponding *p*-value has been proposed as a quantitative indicator of fault advancement over a full lifetime of a gearbox. The method presented in this paper has been compared with the method in Zhan and Makis [21]. The data from the test-run #14 has been analyzed in both papers. Based on the results obtained, we can conclude that the method presented in this paper can detect a gear fault occurrence earlier than their method. Also, our method is much faster and hence, very suitable for on-line gearbox fault detection and damage localization. The results obtained in this paper are very promising and more research should be done in this area.

It is well known that multidimensional vibration data contain more information about gear deterioration than one-dimensional data. Also, in some situations both load and speed vary. Thus, a vector ARX model with possibly two exogenous variables can be considered which should have the potential of further improving performance of the on-line fault detection scheme developed in this paper.

Appendix A. ARX model

ARX model is a time-series model that relates the exogenous input $\mathbf{x}(k)$ and output $\mathbf{y}(k)$ that corrupted by noise $\mathbf{N}(k)$. If we assume that the exogenous input process $\mathbf{x}(k)$ is generated independently of the noise process $\mathbf{N}(k)$ with standard deviation σ_N, we can write the ARX model as [14]

$$\mathbf{y}(k)=\delta_1\mathbf{y}(k-1)+\delta_2\mathbf{y}(k-2)+\cdots+\delta_r\mathbf{y}(k-r)\quad+\omega_0\mathbf{x}(k-b)+\omega_1\mathbf{x}(k-b-1)+\cdots+\omega_s\mathbf{x}(k-b-s)+\mathbf{N}(k) \tag{25}$$

where $b\geq 0$. Let $\mathbf{B}$ denote the backshift operator, which is defined by $\mathbf{B}\mathbf{x}(k)=\mathbf{x}(k-1)$; hence $\mathbf{B}^b\mathbf{x}(k)=\mathbf{x}(k-b)$. The above difference equation can be written as

$$\begin{aligned}\mathbf{y}(k)&=\delta_1\mathbf{B}\mathbf{y}(k)+\delta_2\mathbf{B}^2\mathbf{y}(k)+\cdots+\delta_r\mathbf{B}^r\mathbf{y}(k)+\omega_0\mathbf{x}(k-b)+\omega_1\mathbf{B}\mathbf{x}(k-b)+\cdots+\omega_s\mathbf{B}^s\mathbf{x}(k-b)+\mathbf{N}(k)\\&=(\delta_1\mathbf{B}+\delta_2\mathbf{B}^2+\cdots+\delta_r\mathbf{B}^r)\mathbf{y}(k)+(\omega_0+\omega_1\mathbf{B}+\cdots+\omega_s\mathbf{B}^s)\mathbf{B}^b\mathbf{x}(k-b)+\mathbf{N}(k)\end{aligned} \tag{26}$$

If we define two operators by

$$\boldsymbol{\omega}_s(\mathbf{B}) = \omega_0 + \omega_1\mathbf{B} + \cdots + \omega_s\mathbf{B}^s$$

$$\boldsymbol{\delta}_r(\mathbf{B}) = \delta_1\mathbf{B} + \cdots + \delta_r\mathbf{B}^r$$

The ARX model can be rewritten as

$$\mathbf{y}(k) = \boldsymbol{\delta}_r(\mathbf{B})\mathbf{y}(k) + \boldsymbol{\omega}_s(\mathbf{B})\mathbf{B}^b\mathbf{x}(k) + \mathbf{N}(k) \tag{27}$$

The model order r, s and b are identified based on different criteria, and the unknown parameters δ_1, δ_2, …, δ_r, ω_0, ω_1, …, ω_r, σ_N^2, which are estimated from the observation data.

Appendix B. *F*-test

The fundamentals of *F*-test are well presented, e.g. in Walpole et al. [20], and the following will provide only a brief summary useful for the analysis in this paper. A continuous random variable z has a *F*-distribution, with degrees of freedom $v_1 > 0$ and $v_2 > 0$ if its probability density function is given by

$$f_{v_1,v_2}(z) = \begin{cases} \dfrac{\boldsymbol{\Gamma}\left[\frac{1}{2}(v_1+v_2)\right](v_1/v_2)^{v_1/2}z^{(v_1-2)/2}}{\boldsymbol{\Gamma}\left(\frac{1}{2}v_1\right)\boldsymbol{\Gamma}\left(\frac{1}{2}v_2\right)(1+v_1/v_2)z^{(v_1+v_2)/2}}, & z>0, \\ 0, & \text{elsewhere} \end{cases} \tag{28}$$

where $\boldsymbol{\Gamma}(\bullet)$ is the Gamma Function defined by

$$\boldsymbol{\Gamma}(a) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx \text{ for } \alpha > 0 \tag{28a}$$

An *F*-distribution is a continuous probability distribution with enormous applications in testing whether the population variances of two samples are equal. Suppose there are two normal populations with population variance σ_1^2 and σ_2^2, and two samples **X** and **Y** are random selected from these two populations, respectively.

	Sample **X**	Sample **Y**
Samples size	n_1	n_2
Samples	$\mathbf{X} = X_1, X_2, \ldots, X_{n_1}$	$\mathbf{Y} = Y_1, Y_2, \ldots, Y_{n_2}$
Samples mean	$\overline{X} = 1/n_1 \sum_{i=1}^{n_1}(X_i)$	$\overline{Y} = 1/n_2 \sum_{i=1}^{n_2}(Y_i)$
Samples variance	$S_1^2 = \sum_{i=1}^{n_1}(X_i-\overline{X})/n_1-1$	$S_2^2 = \sum_{i=1}^{n_2}(Y_i-\overline{Y})/n_2-1$
Population variance	σ_1^2	σ_2^2

Then, the random variable

$$F_{v_1,v_2} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \tag{29}$$

has a *F*-distribution with $v_1 = n_1-1$ and $v_2 = n_2-1$ degrees of freedom. To test the null hypothesis H_0 that $\sigma_1^2 = \sigma_2^2$ against the alternative hypothesis H_1 that $\sigma_1^2 > \sigma_2^2$, the ratio $f = S_1^2/S_2^2$ is a value of the *F*-distribution. Therefore, the critical regions of significance level α corresponding to the alternative hypothesis H_1 is $f > f_{v_1,v_2}(\alpha)$, that is

$$\frac{S_1^2}{S_2^2} > f_{(n_1-1),(n_2-1)}(\alpha) \tag{29a}$$

If the above equation is satisfied, we can say null hypothesis H_0 is rejected at significance level α and conclude that σ_1^2 is greater than σ_2^2 with $100\cdot(1-\alpha)\%$ confidence.

References

[1] D. Wang, Q. Miao, R. Kang, Robust health evaluation of gearbox subject to tooth failure with wavelet decomposition, *Journal of Sound and Vibration* 324 (3–5) (2009) 109–132.
[2] W.J. Wang, P.D. McFadden, Application of wavelets to gearbox vibration signals for fault detection, *Journal of Sound and Vibration* 192 (5) (1996) 109–132.
[3] W.J. Wang, P.D. McFadden, Early detection of gar failure by vibration analysis—I. Calculation of the time–frequency distribution, *Mechanical Systems and Signal Processing* 7 (3) (1994) 289–307.
[4] C. Kar, A.R. Mohanty, Vibration and current transient monitoring for gearbox fault detection using multiresolution Fourier transform, *Journal of Sound and Vibration* 311 (1–2) (2008) 109–132.
[5] S.M. Kay, *Modern Spectral Estimation: Theory and Application*, Prentice-Hall, 1988.

[6] W. Wang, A.K. Wong, Autoregressive model-based gear fault diagnosis, *Journal of Vibration and Acoustics, Transactions of the ASME* 124 (2) (2002) 172–179.
[7] Y. Zhan, V. Makis, A robust diagnostic model for gearboxes subject to vibration monitoring, *Journal of Sound and Vibration* 290 (3–5) (2006) 928–955.
[8] B. Liu, V. Makis, Gearbox failure diagnosis based on vector autoregressive modelling of vibration data and dynamic principal component analysis, *IMA Journal of Management Mathematics* 19 (2008) 39–50.
[9] S. Rofe, Signal processing methods for gearbox fault detection, Defence Science and Technology Organization Technical Report, DSTO-TR-0476, Australia, 1997.
[10] MDTB data (Data CDs: test-runs #14), Condition-based Maintenance Department, Applied Research Laboratory, The Pennsylvania State University, 1998.
[11] P.D. McFadden, Examination of a technique for the early detection of failure in gears by signal processing of the time domain average of the meshing vibration, *Mechanical Systems and Signal Processing* 1 (1987) 173–183.
[12] P.D. McFadden, Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration, *ASME Transactions Journal of Acoustics Stress and Reliability in Design* 108 (1986) 165–170.
[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] E.P.G. Box, M.G. Jenkins, C.G. Reinsel, *Time Series Analysis: Forecasting and Control*, 4th ed., John Wiley & Sons, Inc., 2008.
[15] C.M. Jarque, A.K. Bera, Efficient tests for normality, homoscedasticity and serial independence of regression residuals, *Economics Letters* 6 (3) (1980) 255–259.
[16] C.M. Jarque, A.K. Bera, Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence, *Economics Letters* 7 (4) (1981) 313–318.
[17] L. Ljung, *System Identification: Theory for the User*, 2nd ed, Prentice-Hall, Inc., 1999.
[18] G. Schwarz, Estimating the dimension of a model, *The Annals of Statistics* 6 (1978) 461–464.
[19] H. Akaike, Information theory and an extension of the maximum likelihood principle, in: B.N. Petrov, F. Csaki (Eds.), *2nd International Symposium on Information Theory*, Akademiai Kiado, Budapest, 1973, pp. 267–281.
[20] R.E. Walpole, H.R. Myers, L.S. Myers, K. Ye, *Probability & Statistics For Engineers & Scientists*, 8th ed, Person Education, Inc., 2007.
[21] Y. Zhan, V. Makis, Adaptive state detection of gearboxes under varying load conditions based on parametric modelling, *Mechanical Systems and Signal Processing* 20 (2006) 188–221.