Available online at www.sciencedirect.com
ScienceDirect
Mechanical Systems and Signal Processing 21 (2007) 906–919
Mechanical Systems and Signal Processing
www.elsevier.com/locate/jnlabr/ymssp

Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter

H. Endo*, R.B. Randall

School of Mechanical and Manufacturing Engineering, The University of New South Wales, Sydney 2052, Australia

Received 7 October 2005; received in revised form 31 January 2006; accepted 5 February 2006
Available online 29 March 2006

Abstract

This paper proposes the use of the minimum entropy deconvolution (MED) technique to enhance the ability of the existing autoregressive (AR) model based filtering technique to detect localised faults in gears. The AR filter technique has been proven superior for detecting localised gear tooth faults than the traditionally used residual analysis technique. The AR filter technique is based on subtracting a regular gearmesh signal, as represented by the toothmesh harmonics and immediately adjacent sidebands, from the spectrum of a signal from one gear obtained by the synchronous signal averaging technique (SSAT). The existing AR filter technique performs well but is based on autocorrelation measurements and is thus insensitive to phase relationships which can be used to differentiate noise from impulses. The MED technique can make a use of the phase information by means of the higher-order statistical (HOS) characteristics of the signal, in particular the kurtosis, to enhance the ability to detect emerging gear tooth faults. The experimental results presented in this paper validate the superior performance of the combined AR and MED filtering techniques in detecting spalls and tooth fillet cracks in gears.

Keywords: Gear tooth fault diagnosis; Gear tooth fillet cracks; Gear tooth spalls

1. Introduction

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. Failure of the gears in industrial machines and aerial vehicles results in expensive penalties, in some cases claiming lives [1]. Non-intrusive machine diagnostic methods; such as techniques based on vibration signal analysis play an important role in ensuring safe and reliable operation in critical applications.

Development of detection and diagnostic techniques for localised gear tooth faults in helicopter gearboxes is an active research area that has been driven by the interests of military organizations (US Army [2,3], DSTO [1]), large scientific bodies such as NASA [2,3], and by large civil operators such as those servicing oil rigs in

*Corresponding author. Tel.: +61 421 881 010.
E-mail addresses: h.endo@student.unsw.edu.au (H. Endo), b.randall@unsw.edu.au (R.B. Randall).

0888-3270/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ymssp.2006.02.005

the North Sea [4,5], where health and usage monitoring systems (HUMS) are mandatory. Of course, methods developed for helicopter gearboxes will also be applicable to gearboxes in general.

Gear diagnostic techniques developed for vibration-based condition monitoring involve processing and conditioning of the signals through different filters. Wang and Wong demonstrated a successful application of an autoregressive (AR) model-based filter to isolate the impulse-like effect of localised cracks in a gear tooth [6,7]. The AR filter technique was proven superior when compared to the traditionally used residual analysis technique based on subtracting a regular gearmesh signal, as represented by the toothmesh harmonics and immediately adjacent sidebands, from the spectrum of a signal from one gear obtained by the synchronous signal averaging technique (SSAT).

This paper proposes an enhancement for the AR-based linear prediction fault detection technique. By its inherent theory the performance of the AR filter presented in [6,7] is limited by the assumptions in the computation of the AR model: The AR processes are described as recursive events driven by Gaussian white noise and the AR model is estimated from the autocorrelation function of the output signal by solving the Yule–Walker equations (i.e. by using second-order statistical properties).

In reality, the residual of the linear prediction filter in the general case is not Gaussian white noise. It comprises the non-predictable part of the signal and is a mixture of impulses and coloured noise. The AR model has no zeros, and therefore cannot represent any maximum phase properties (all-pass filters) and since it is based on autocorrelation measurements it is insensitive to phase relationships (which differentiate white noise from impulses).

The existing AR filter technique performs well; however its ability to detect an emerging gear tooth fault can be further enhanced by using higher-order statistical properties to separate the impulsive characteristics of localised faults from the noise.

2. Gearmesh signal with the effect of a localised gear tooth fault

The vibration signal '$\boldsymbol{x}$' of a gearbox can be described as a convolution between the impulse response function (IRF) of the transmission path '$\boldsymbol{h}$' and the combined effect of an anomaly caused by a localised gear tooth fault (fault impulses) '$\boldsymbol{w}$', the deterministic signals '$\boldsymbol{d}$' inherent in operating gears and the noise '$\boldsymbol{e}_0$' (illustrated in Fig. 1 and Eq. (1)):

$$\boldsymbol{x} = (\boldsymbol{d} + \boldsymbol{w} + \boldsymbol{e}_0) * \boldsymbol{h}, \tag{1}$$

where $*$ represents convolution.

A gear tooth defect such as a tooth root crack or a spall causes localised impulsive changes '$\boldsymbol{w}$' when the tooth engages in a mesh. The objective of many residual signal methods, including the AR filter technique, is to isolate the effect of '$\boldsymbol{w}$' from the signal mixture to detect and monitor the emergence of gear tooth faults.

A particular challenge in the detection of gear tooth faults in their early stages is that the effect of the impulses caused by the faults is relatively small compared to the gearmesh component '$\boldsymbol{d}$', as well as the noise '$\boldsymbol{e}_0$', and its presence can be masked by these dominant components.

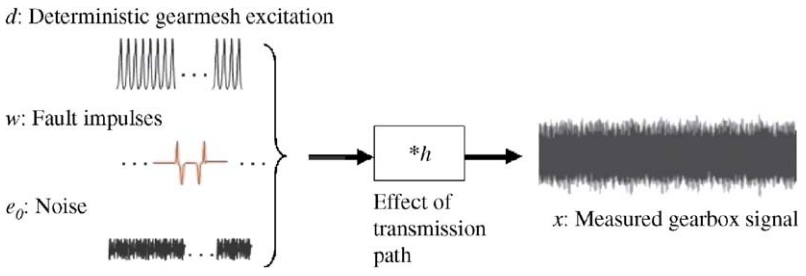


Fig. 1. Vibration signal from a gearbox.

3. Review: AR model-based gear fault diagnosis

In the very first step, the random noise contaminating the signal was minimised by the SSAT. SSAT utilises the property of the random distribution that its expected mean value is zero. The vibration signal measured from the gearbox is averaged over one rotation of the gear. The value of the random noise (and non-related periodicities) will be progressively reduced to zero as the signal is being averaged while the deterministic synchronous contribution from the gear remains constant [8].

Prior to taking the averages, the measured vibration signal is mapped into the angular domain by using the pulses from a shaft encoder as references. The encoders used in the study presented here give 3600 pulses per rotation, divided electronically to 900 pulses per rotation. The benefit of the mapping is that it removes the effect of non-synchronous frequency modulation in the time domain signal caused by the shaft speed variation which occurs in the practical environment. The mapping ensures the deterministic property of the gearmeshing contribution.

Coefficients of the AR filter are adaptively derived from the vibration signals of undamaged gears. The Akaike information criterion (AIC) was used to ensure optimum adaptation of the AR coefficients to the undamaged gear signal [6]: see Eq. (2). The AIC reflects the effect of spectral variance due to increase in model order k and the prediction error σ computed in estimating the AR coefficients.

$$AIC(k) = N(\ln(\sigma_k^2)) + 2k. \tag{2}$$

Wang and Wong [6] emphasise the importance of selecting the right model order for the AR filter. Excessively large filter order causes spurious spectral peaks and statistical instabilities whereas too small an order causes poor prediction and incomplete peak fit in the spectrum.

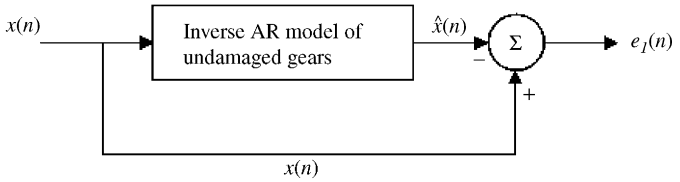
The AR filter well predicts the deterministic pattern of the gearmesh signal '$\boldsymbol{d}$' but is not capable of adapting to the sudden impulses '$\boldsymbol{w}$' caused by a localised gear tooth fault. As a consequence, the fault impulse will be left in the residual of the AR filter: obtained by taking the difference between the actual signal and the AR prediction. The definition of an AR process is given in Eq. (3) and the concept of the AR filter is illustrated in Fig. 2:

$$\hat{x}(n) = \sum_{k=1}^{p} a(k)x(n-k), \tag{3}$$

where $\hat{x}(n)$ is the nth data point predicted by the AR filter, $a(k)$ the kth coefficient in the AR model, p the order of the AR model.

The most common method of determining the coefficients of the AR model is by using the autocorrelation (second-order statistical characteristic) of the signal. The AR process satisfies the Yule–Walker equations (Eqs. (4) and (5)):

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



where: $x(n)$: Input signal with effect of gear fault
$\hat{x}(n)$: AR prediction of undamaged gear signal
$e_l(n)$: Prediction Error (AR residual)

Fig. 2. AR filter.

In matrix form:

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

The coefficients of the AR model $a[l]$ can be determined by using p samples of $\hat{r}_{xx}[k]$ and solving the Yule–Walker equations [9,10]. The values of $\hat{r}_{xx}[k]$ are the autocorrelation sequence obtained by using the biased autocorrelation estimate defined as

$$\hat{r}_{xx}[k] = \frac{1}{N} \sum_{n=0}^{N-1} x[n]x[x-k], \quad 0 \leqslant k \leqslant p-1, \tag{6}$$

where N is the number of samples in the signal.

The solution of the Yule–Walker equations can be computed by different methods. Two of the most commonly accepted methods are the Levinson–Durbin recursion (LDR) algorithm and Burg's (maximum entropy) method (BM). Direct application of the Gaussian elimination method (GEM) is also possible, but generally considered computationally less efficient than the LDR and BM. For more details on LDR please refer to Refs. [6,9,10].

3.1. Limitations of the AR filtering technique

In signal processing, the transfer function of a linear system $H(z)$ (Eq. (7)) can be expressed by using autoregressive moving average models (ARMA). A general definition of the ARMA is given in Eq. (8).

The concept of an ARMA model used in system identification techniques can be applied to effectively model the structure of gearmesh signals. By assuming the residual signal to be a mixture of white noise and impulses, the AR part of the ARMA model represents the deterministic characteristic of the gearmesh signal and the MA part represents the stochastic characteristics of the input signal (coloured noise: filtered white noise and impulses) in the ARMA model [9,10]:

$$H(z) = \frac{B(z)}{A(z)} = \frac{1 + \sum_{k=1}^{-k} b[k]z^{-k}}{1 + \sum_{k=1}^{p} a[k]z^{-k}}, \tag{7}$$

$$x[n] = \underbrace{\sum_{k=1}^{p} a[k] \cdot x[n-k]}_{\text{AR}} + \underbrace{\sum_{k=0}^{q} b[k] \cdot e[n-k]}_{\text{MA}} + e[n] \tag{8}$$

where, $b[k]$ is the kth coefficient of the MA model, $\boldsymbol{e}$ the noise series (note: white noise), q the order of the MA model.

A typical spectrum of gear-excited signals is dominated by the harmonics of the gearmesh frequencies and appropriately represented by using an all-pole expression, i.e. AR model.

Note: a system modelled by one type of model: AR, MA or ARMA can be represented more or less efficiently by any of the models. The advantage of the AR model is that unlike an MA model (and the MA part of ARMA), its parameters can be determined by solving a linear set of equations. Generally it is more efficient in modelling a filter of equivalent performance (for lightly damped systems with sharp peaks): an AR model requires far fewer coefficients than the corresponding MA model [9,10].

The AR filter removes the deterministic (AR) components of the signal and leaves the residual (MA filtered white noise and impulsive effect of a gear tooth fault of the input signal). Higher-order statistical (HOS) properties must be used to define and differentiate the contribution of the different types of signals mixed in the stochastic MA part.

The main objective of this stage of the signal processing is to enhance the clarity of the impulses caused by the gear tooth faults, thus enabling detection of a fault at an earlier stage of its development.

The minimum phase condition assumed in the AR model limits the ability to achieve complete extraction of the impulsive nature of the fault signal. It is worthwhile examining this point prior to proceeding further: The spectrum amplitudes of both noise and impulses are uniform and can be described as broadband excitation. The information that differentiates noise from impulses is contained in the phase of the signal. As the AR filter is phase blind this property in the output residual signal is left untouched.

4. Minimum entropy deconvolution technique

The minimum entropy deconvolution technique (MED) was originally presented by Wiggins [11] and has shown its effectiveness in deconvolving the impulsive sources from a mixture of signals [11–14].

The MED technique is a type of system identification method that was originally developed to aid extraction of reflectivity information in seismic data. The reflectivity information can be used to identify and locate layers of subterranean minerals.

The MED searches for an optimum set of filter coefficients that recover the output signal (of an inverse filter) with the maximum value of kurtosis (Eq. (10)). Kurtosis is an indicator that reflects the "peakiness" of a signal, and therefore the property of impulses.

(Note: This explains the name "minimum entropy" method, because higher entropy corresponds to a tendency to become more Gaussian, i.e. less structure and more disorder, and therefore minimizing the entropy enhances the structured information in the signal). This property of the MED filter was thought to be useful in enhancing the definition of the gear fault impulses.

Fig. 3 illustrates the deconvolution process performed by the MED filtering. The output from the AR filter '$\boldsymbol{e}_1$' can be modelled as in Eq. (9). The signal 'x' represents the fault impulses in original form:

$$\boldsymbol{e}_1 = (\boldsymbol{w} + \boldsymbol{e}_0) * \boldsymbol{h}. \tag{9}$$

The objective of the deconvolution is to find the coefficients of the inverse filter $f[n]$ which achieves '$\boldsymbol{e}_1[n] * f[n] = \delta(n - l_{\mathrm{m}})$' [14]. (Note: The delay of "l_m" is allowed to make the inverse filter causal).

The MED filter was implemented by the objective function method (OFM) given in [13]. The OFM is an optimization process designed to maximise the kurtosis of the MED output (Eq. (10)). OFM achieves this by changing the values of the coefficients in the MED filter. The optimization process finishes when the values of the coefficients converge within the specified tolerance (Eq. (15)). The process in the MED filter is summarised below:

(1) The objective of the process is to find the filter coefficient vector $\boldsymbol{f}$ which optimises the kurtosis of the output signal y:

$$\mathrm{O}_k(f[l]) = \sum_{n=1}^{N} y^4[n] \Bigg/ \left[\sum_{n=1}^{N} y^2[n]\right]^2. \tag{10}$$

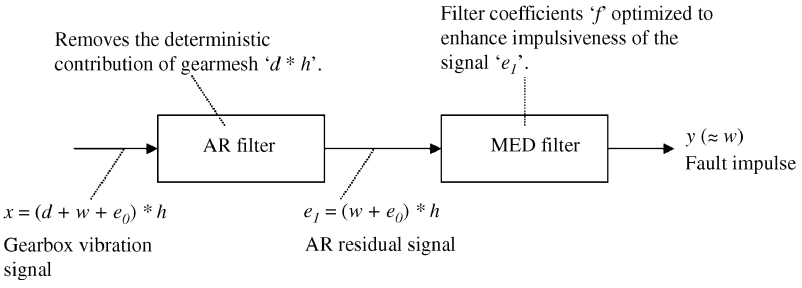


Fig. 3. Inverse filtering (deconvolution) process.

(2) A general expression for inverse filtering is given as a convolution by:

$$y[n] = \sum_{l=1}^{L} f[l] e_1[n-l]. \tag{11}$$

(3) Optimum filter settings occur when the objective function achieves:

$$\partial(\mathrm{O}_4(f[l]))/\partial(f[l]) = 0. \tag{12}$$

(4) By using $\partial y[n]/\partial f[l] = \boldsymbol{e}_1[n-l]$ and Eqs. (11) and (12) can be expressed as

$$\underbrace{\left[\sum_{n=1}^{N} y^2[n] \Big/ \sum_{n=1}^{N} y^4[n]\right] \sum_{n=1}^{N} y^3[n] e_1[n-l]}_{\boldsymbol{b}} = \underbrace{\sum_{p=1}^{L} f[p]}_{\boldsymbol{f}} \underbrace{\sum_{n=1}^{N} e_1[n-l] e_1[n-p]}_{\boldsymbol{A}} \tag{13}$$

Eq. (13) can be written in matrix form:

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{f}. \tag{14}$$

The column vector $\boldsymbol{b}$ is computed from the cross correlation of input and output signals ($\boldsymbol{y}$ and $\boldsymbol{e}_1$) of the inverse filter $\boldsymbol{f}$; $\boldsymbol{A}$ is the Toeplitz autocorrelation matrix of input signal $\boldsymbol{e}_1$ with dimension $(L \times L)$ and $\boldsymbol{f}$ is the column vector of required inverse filter coefficients.

(5) The solution for the optimum inverse filter is found iteratively as Eq. (14) is highly non-linear. The iterative steps involve the following:

Step 1: Compute the autocorrelation matrix $\boldsymbol{A}$. (Note: the matrix $\boldsymbol{A}$ is computed only once at the beginning of the iteration and is reused in the iterative loop).

Step 2: Assume the initial value of the inverse filter coefficients $\boldsymbol{f}^{(0)}$. (Note: a delayed impulse was assumed as the initial vector value).

Step 3: Compute the output signal $\boldsymbol{y}^{(0)}$ using the input signal $\boldsymbol{e}_1^{(0)}$ and initial optimum filter coefficients $\boldsymbol{f}^{(0)}$ (i.e. solve Eq. (11)).

Step 4: Compute the column vector $\boldsymbol{b}^{(1)}$ using $\boldsymbol{y}^{(0)}$ and input signal $\boldsymbol{e}_1^{(0)}$ (i.e. calculate the value in the left-hand side of Eq. (14)).

Step 5: Solve for new filter coefficients $\boldsymbol{f}^{(1)} = \boldsymbol{A}^{-1}\boldsymbol{b}^{(1)}$.

Step 6: Compute the error criterion (change in the filter coefficient values):

$$\boldsymbol{err} = (\boldsymbol{f}^{(1)} - \mu\boldsymbol{f}^{(0)})/\mu\boldsymbol{f}^{(0)}, \tag{15}$$

$$\mu = (E(\boldsymbol{f}^{(0)})^2/(E(\boldsymbol{f}^{(1)})^2)^{1/2}. \tag{16}$$

The iteration finishes when the expected value of error '$E(\boldsymbol{err}) \leqslant tolerance$'; if '$E(\boldsymbol{err}) > tolerance$' then update the filter coefficients: $\boldsymbol{f}^{(0)} = \boldsymbol{f}^{(1)}$and repeat the process from Step 3. A condition was set also to abort the iteration if the value of the $E(\boldsymbol{err})$ was found to be diverging. The effectiveness of the impulse enhancement by the MED technique is demonstrated in the next section.

5. Experimental results: gear fault diagnosis with AR & MED filter technique

Vibration signals measured from the gear test rig (Fig. 4) at the University of New South Wales were used to compare the effectiveness of the AR & MED filter (ARMED) to the original AR filter in detecting gear tooth faults. The values of Kurtosis: a statistical parameter commonly used to compare the peakiness of an impulse, thus clarity of the effect of localised faults, are displayed in Figs. 8 and 10 to allow quantitative comparison of the improvement achieved by the ARMED technique.

Fig. 4. UNSW gear test rig.

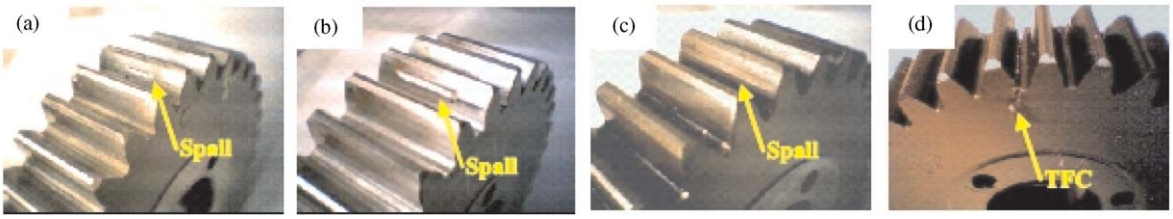


Fig. 5. Seeded gear faults. (a) Sp25%, (b) Sp50%, (c) Sp100% and (d) TFC (plastic gears).

Table 1
Dimensions of gear tooth faults

	ID	Description of the fault
Spalls	Sp25%	0.5 mm along the tooth profile, 6.25 mm across the tooth face (25% tooth face width), 0.5 mm depth.
	Sp50%	1.0 mm along the tooth profile, 12.5 mm across the tooth face (50% tooth face width), 0.5 mm depth.
	Sp100%	1.5 mm along the tooth profile, 25 mm across the tooth face (100% tooth face width), 0.01 mm depth.
Tooth fillet crack	TFC	Crack length = 3 mm (into the tooth fillet area at an angle of 45° to the axis of symmetry of the tooth).

Four sets of unity ratio spur gears (32×32 teeth): three sets of steel gears with tooth spalls of different sizes and a pair of plastic gears with an artificially cut tooth fillet crack (TFC) were considered. The gears with the seeded faults are shown in Fig. 5 and the dimensions of the faults are given in Table 1. The vibration (acceleration) signals on the gearbox casing were measured while running the gears at a constant speed of 168 rpm.

Signal preconditioning: The vibration signals measured from the gearbox were order tracked, synchronously averaged per one rotation of the gear and low-pass filtered at 2688 Hz (i.e. retaining up to the 30th gear mesh harmonic) before being processed by AR and MED filters. The low-pass filtered signal was down-sampled at frequency slightly above twice the low-pass frequency (5460 Hz) to minimise the number of samples in the signal. The down-sampling removed the excess data points (i.e. the 23,406 samples per gear rotation of the original signal were reduced to 1950 samples).

Optimum AR filter: In a typical gearbox vibration signal, the effect of the gear harmonics dominates the landscape of the spectrum in the lower frequency band. The AIC optimization algorithm of AR filter is designed to attenuate the dominant effect of the gear harmonics in the lower frequency region. It was found by trial and error that the effectiveness of the AR filtering deteriorated when dealing with the higher frequency components because the magnitude of the gear harmonics is very small at the higher frequencies and the signal contains the effects from the other parts of the gearbox system (for example, the rolling element bearings).

The optimum length of the AR filter was determined based on the minimum AIC value. Typically, the order of the AR filter tends to settle around one tooth meshing period (60–70 samples).

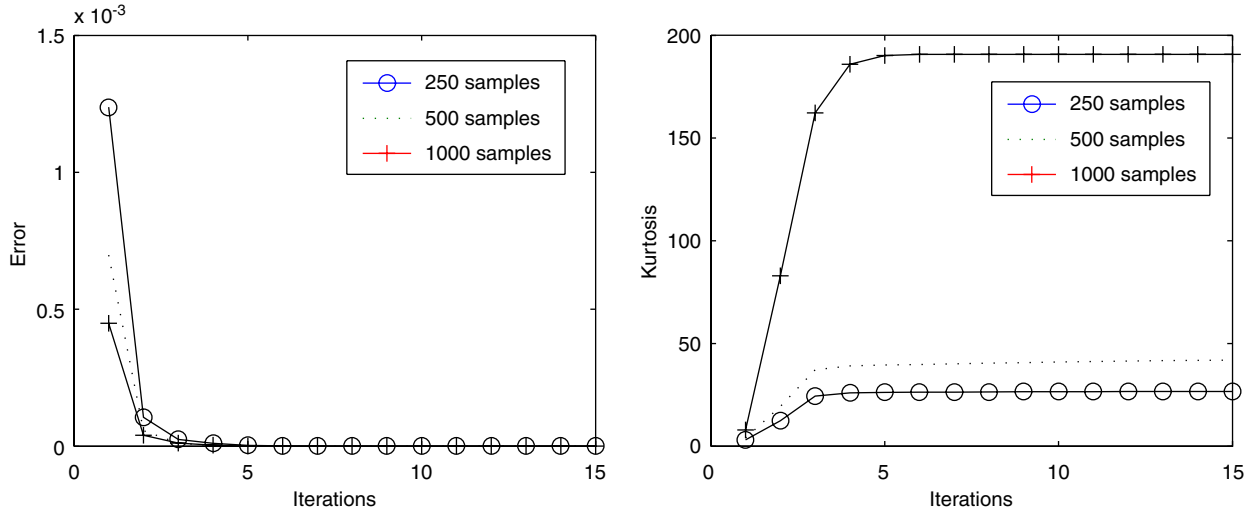


Fig. 6. Convergence of MED algorithm.

Setting for MED filter: Fig. 6(a) and (b) show the convergence characteristics of the '$E(\boldsymbol{err})$' and Kurtosis (of the MED filter outputs) over the number of iterations performed in the MED algorithm. The convergence properties of the MED filters with filter lengths of 250, 500 and 1000 samples were compared. The efficiency of the algorithm is apparent in the steep convergence of the $E(\boldsymbol{err})$ and the kurtosis values shown in Fig. 6(a) and (b). The convergence was achieved within six iterative steps regardless of the lengths of the MED filter.

The length of the MED filter appears to have significant impact on the value of the kurtosis of the MED outputs. The comparison in Fig. 6(b) shows that the longer MED filters tend to converge more rapidly to higher kurtosis values. However, the comparison of the time history signals of the MED outputs in Fig. 7 shows that there is no appreciable benefit from having more than two iterations. Having a greater number of iterations can increase the kurtosis of the MED output; however, the result did not gain any benefit for the purpose of machine diagnosis. Regardless of the large increase in the kurtosis, the clarity of the fault impulses in the actual time signal did not appear to gain much useful improvement.

Another important consideration in performing MED enhancement is associated with its sensitivity to the impulses. It was found by trial and error that the MED filter can erroneously enhance small impulses, unrelated to gear faults, when a suitable convergence condition was not specified.

By trial and error, a consistent diagnostic result was achieved in this experiment when the value of the '*tolerance*' for the '$E(\boldsymbol{err})$' was set to 0.01% (Note: typically this setting resulted in two iterations). This tolerance appears to be sensitive enough to allow effective enhancement of some hidden fault impulses in the AR residual while prohibiting the filter from causing spurious enhancement of the small impulses in undamaged gears. (Note: the comparisons in Figs. 6 and 7 were based the analysis of the signal from the 'Sp25%' case)

5.1. Spall detection

The unprocessed signals and the residuals of the AR filter and the ARMED (signal with the MED enhancement) of gears with spalls are shown in Fig. 8. The signals were measured while running the gears at 168 rpm shaft speed and with loading 67 Nm of torque. The corresponding power spectra of the signals are presented in Fig. 9.

The effect of the spalls is not visible in the unprocessed signals in Fig. 8(a1)–(a3). The once-per-gear-rotation impulses due to the large spall (Sp100%) became visible in the residuals of the AR filter in the Fig. 8(b1)–(b3); however the symptoms of smaller spalls: Sp25% and Sp50% are not clear (Fig. 8(a1) and (a2)).

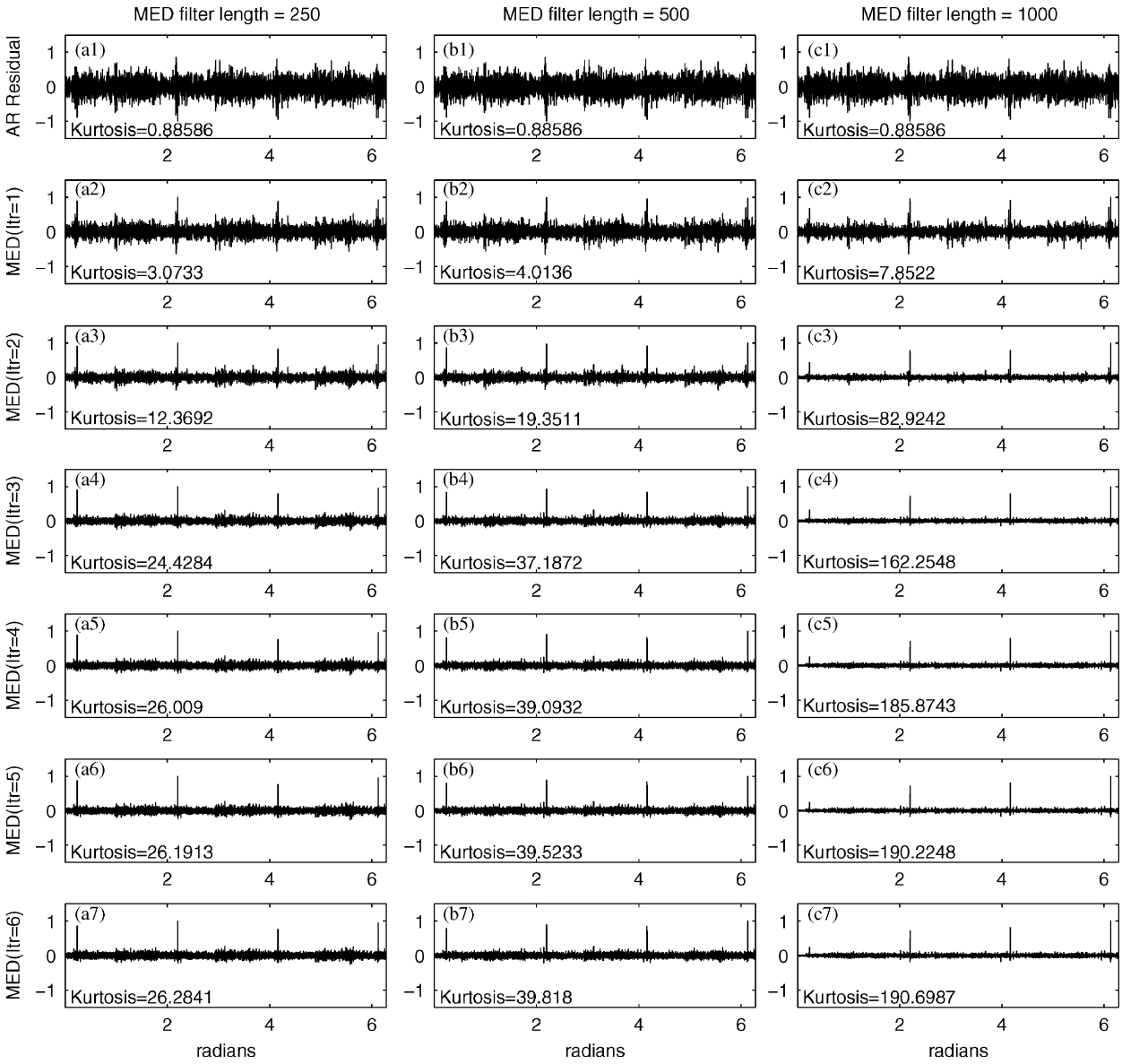


Fig. 7. Effect of MED filter length. (a1), (b1) and (c1): AR residuals. The outputs of the MED with filter length 250samples: (a2)–(a7), 500samples: (b2)–(b7) and 1000samples: (c2)–(c7).

The results of the ARMED shown in Fig. 8(c1)–(c3) clearly demonstrate the effectiveness the MED filters in enhancing the spall impulses. The impulses of the smaller spalls (Sp25% and Sp50%), previously undetected in the AR residuals, were made clearly visible in the ARMED results.

The spectra of the signals in Fig. 8 are presented in Fig. 9. The 'whitening' effect of both AR and MED filters is observed in the flattened amplitude of the power spectra. In general, the spectra of the ARMED results appear flatter than the spectra of AR residuals. The spectra of the AR residual start to roll-off at around 2000 Hz as the AR filter becomes less effective in accurately capturing features in the high-frequency regions: this results in ineffective inverse filtering which causes the roll-off effect. The MED filter appears to compensate the limitation of the AR filter by flattening out the amplitude of the higher-frequency components.

Nevertheless, the main benefit of the MED filter comes from the tuning of phase information. This is evident in the dramatic effect of the MED filter on the signal shapes and their kurtosis when compared to the

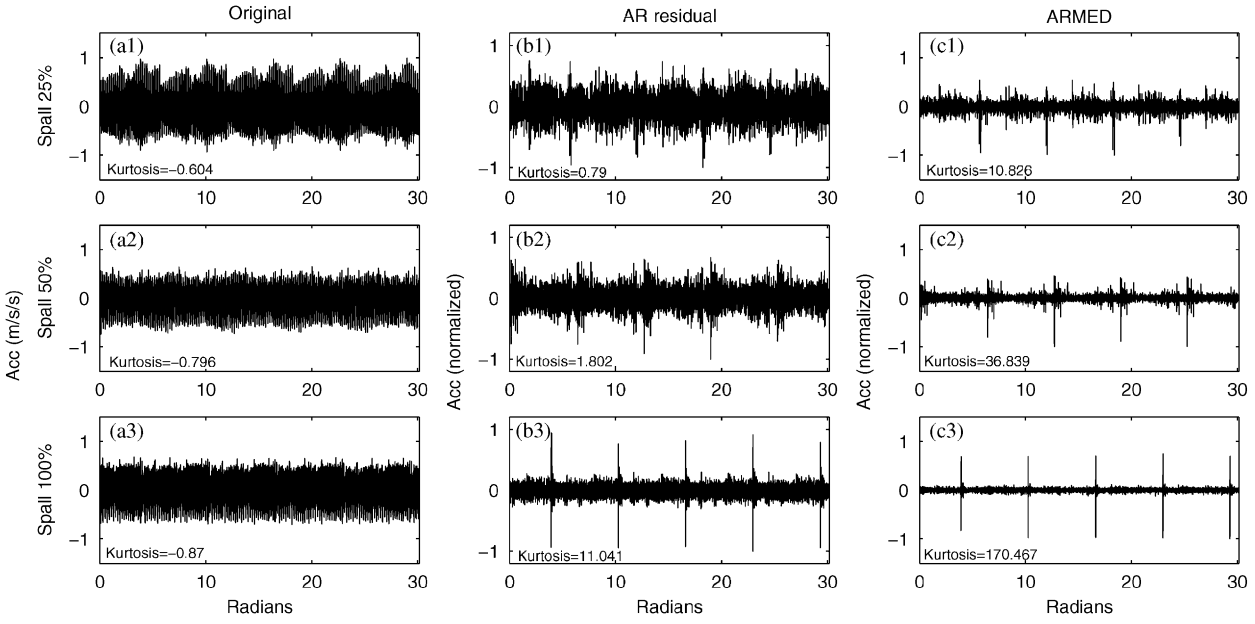


Fig. 8. Comparison of signals (gears with spalls).

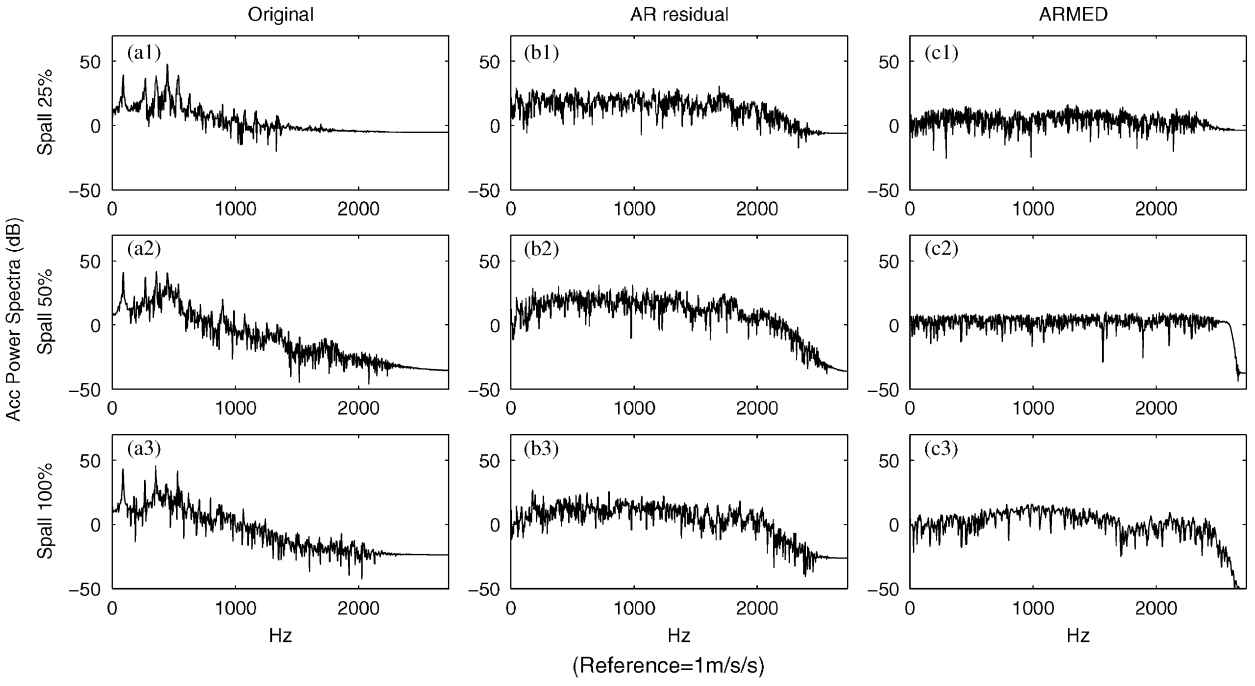


Fig. 9. Comparison of power spectra (gear Spalls). *Note*: In the Figs. 8 and 9, columns (a), (b) and (c) correspond to unprocessed signals, AR residuals and ARMED outputs, respectively. The rows 1, 2 and 3 correspond to the types of damages: Spall 25%, Spall 50% and Spall 100%, respectively.

rather subtle change in the power spectra. The information in the spectra relates only to the magnitude and therefore cannot reflect the effect of phase changes. There are some cases with little discernible difference in spectrum shape (eg Figs. 9(b1, c1) and 11(b2, c2) for a TFC), where there is still a very large difference in the impulsiveness of the time signals.

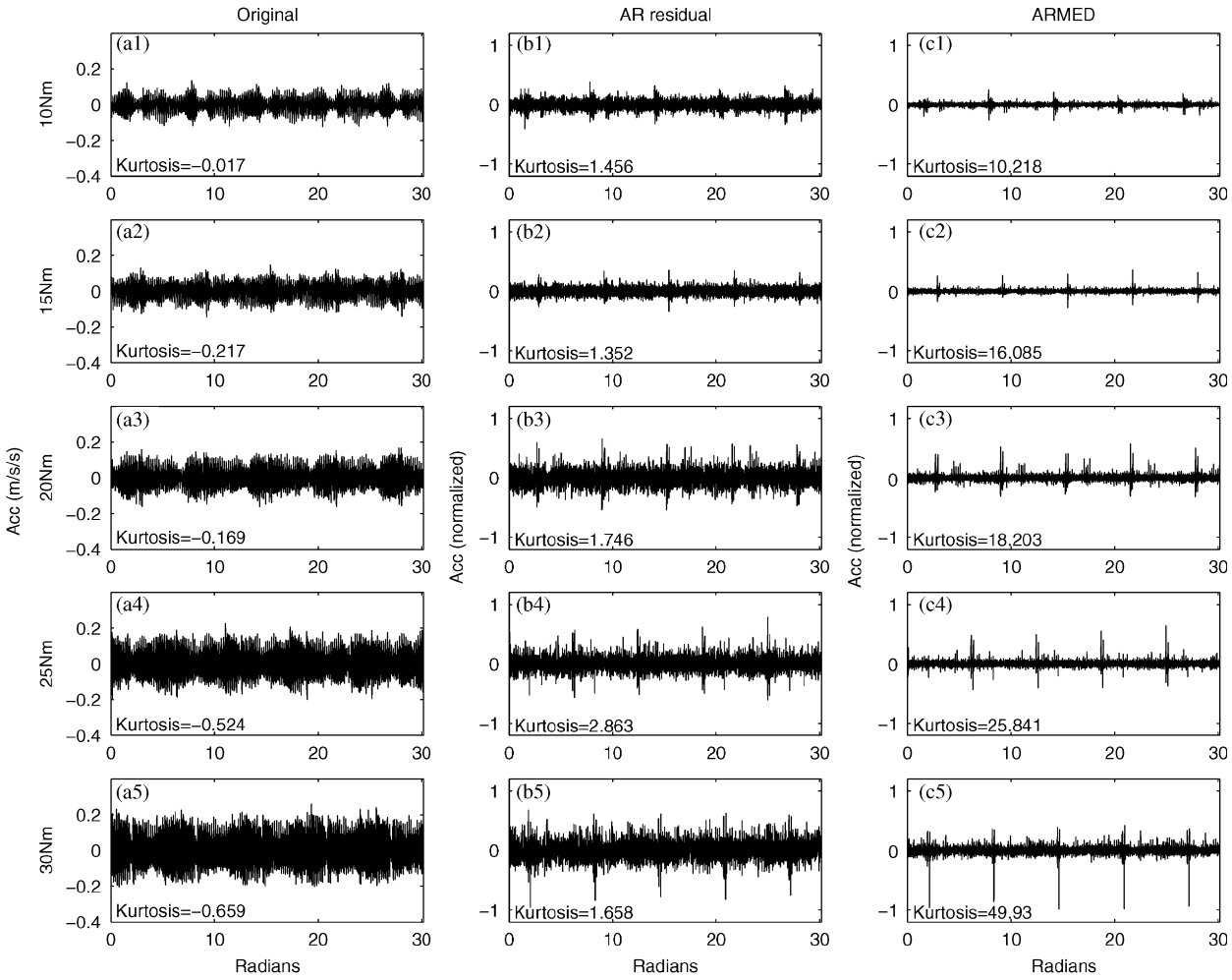


Fig. 10. Comparison of signals (TFC with different loads). *Note*: In the Fig. 10, columns (a), (b) and (c) correspond to unprocessed signals, AR residuals and ARMED outputs, respectively. Rows 1–5 correspond to different levels of loads applied to the gear: 10, 15, 20, 25 and 30 Nm, respectively.

Note: The outputs of the MED can be contaminated by the high-frequency components near the Nyquist frequency when the high-frequency noise is amplified as a result of the inverse filtering. The same low-pass filter as used in the earlier step of the signal processing was reapplied to the output of the MED filter to prevent this contamination. The steep roll-off near the Nyquist frequency of the ARMED spectra is the effect of the supplementary low-pass filtering Fig. 9(c1)–(c3) and Fig. 10(c1)–(c5).

5.2. Tooth fillet cracks

Endo et al explain in [15] that the effect of TFCs on gear dynamics results as a consequence of locally reduced stiffness at the gearmesh: a TFC causes a localised reduction of gearmesh stiffness when the faulty tooth enters the mesh (Note: surface faults such as spalls are almost a purely geometrical effect and do not affect stiffness).

The effect of TFCs is load dependent, i.e. it gives locally increased deflection due to the faulty gear tooth and is therefore directly affected by the amount of loading on the gears. It was anticipated that the effect of TFCs in lightly loaded gears would be more difficult to detect than for gears loaded with greater torque.

The effectiveness of the AR and the ARMED filtering techniques on detecting TFCs was studied by comparing the vibrations of the gears loaded with five incremented levels of torque: 10, 15, 20, 25 and 30 Nm.

A pair of nylon gears was used in the TFC experiment instead of steel gears to compensate for the slender shafts in the gearbox that were designed to have low-frequency resonances within the operational speed range of the gearbox for the purpose of another experiment [16]. Typically gear shafts are designed to have much greater stiffness so as to provide rigid support of the gears. The lack of stiffness in the slender gearshafts allows the centre distance of the gears to change and dominates over the deflection of the gear teeth in the mesh. The excessive shaft flexibility was found to reduce the effect of TFCs that should otherwise be observable in rigidly mounted gears [17]. In this experiment, a pair of nylon gears was used instead of steel gears to make the stiffness at the gear mesh of the same order as that of the shafts.

The unprocessed signals, the residuals of the AR and the ARMED filters of the TFC cases, are shown in Fig. 10 and their power spectra are presented in Fig. 11. The results of both the AR and the ARMED filters show the load-dependent characteristic of the TFCs discussed above. The kurtosis of the fault signal in the ARMED results (Fig. 10(c1)–(c5)) increases with the magnitude of the torque.

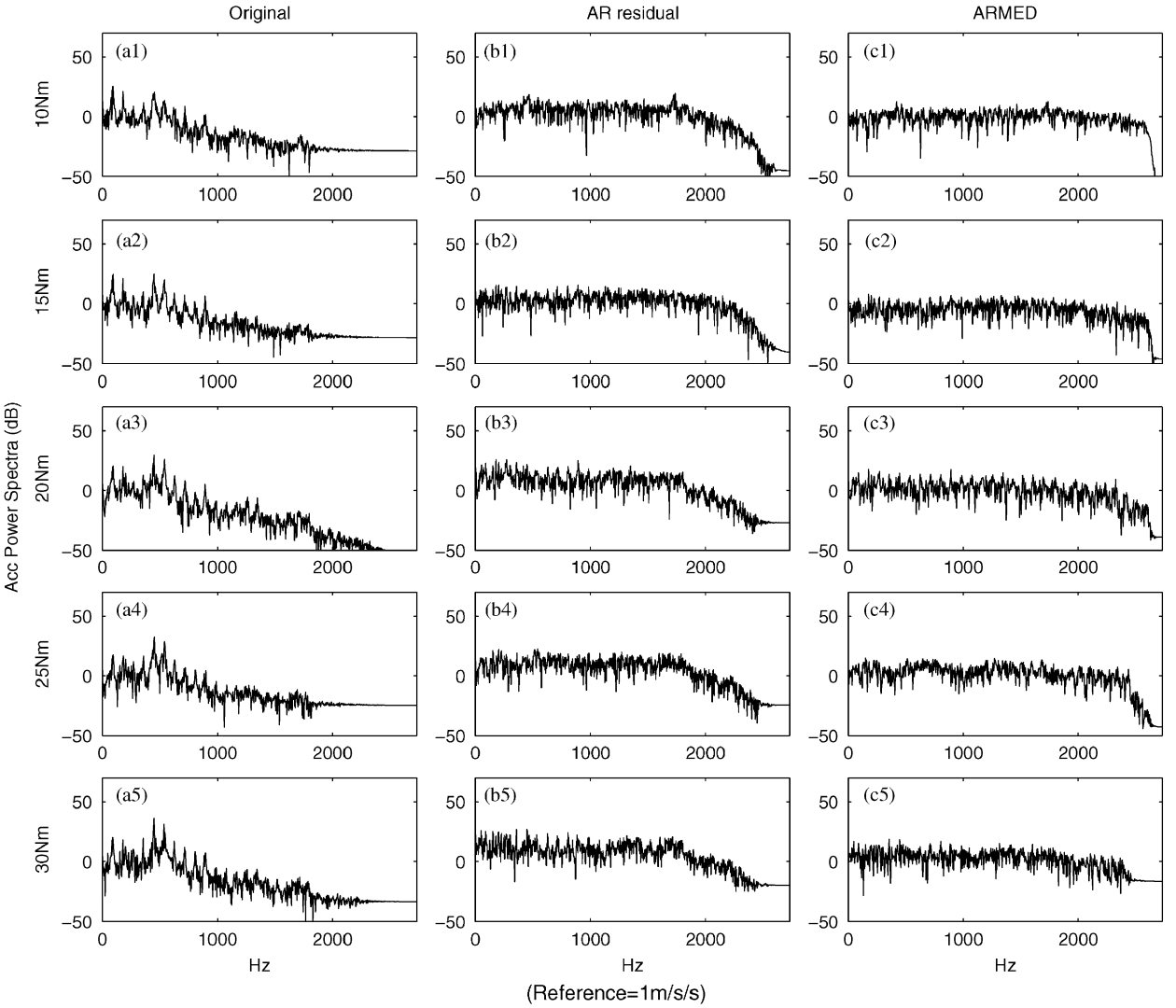


Fig. 11. Comparison of power spectra (TFC with different loads). *Note*: In Fig. 11, columns (a), (b) and (c) correspond to unprocessed signals, AR residuals and ARMED outputs, respectively. Rows 1–5 correspond to different levels of loads applied to the gear: 10, 15, 20, 25 and 30 Nm, respectively.

The effects of the TFC which are undetected in the unprocessed signals (Fig. 10(a1)–(a5)) are made visible in the residuals of the AR (Fig. 10(b1)–(b5)) and the presence of the fault impulses is further enhanced in the ARMED outputs (Fig. 10(c1)–(c5)). While the effect of the TFCs can be recognised in the outputs of both AR and ARMED, the clarity of the fault impulses is significantly better in the ARMED results. This is reflected also in the kurtosis of the signals.

6. Conclusion

The ability of the MED filter to enhance the kurtosis (peakiness) of the impulse-like effects of localised gear tooth faults has been studied for both spalls and TFCs. The MED filter has been shown to complement the existing AR filtering technique to improve its ability to detect gear tooth faults.

Experiments with gears with seeded faults showed promising results. The combined AR and MED (ARMED) filtering technique showed significant improvement in detecting spalls and TFCs. Some of the faults not clearly identified by using only the original AR filter were successfully identified in the outputs of the ARMED filters.

In future work, it will be interesting to study the effectiveness of the ARMED technique in diagnosing more complex gearing systems; such as a helicopter gearbox. Also, the ARMED technique will be applied to the diagnosis of rolling element bearings (REBs) where many faults are known to manifest themselves as impulses.

References

[1] A.K. Wong, Vibration-based Helicopter Health Monitoring—An Overview of the Research Program in DSTO, DSTO-HUMS2001, 2001.
[2] H.J. Decker, Gear crack detection Using tooth analysis, NASA-TM2002-211491, 2002a.
[3] H.J. Decker, Crack detection of aerospace quality gears, NASA-TM2002-211492, 2002b.
[4] B.D. Larder, Helicopter HUM/FDR: Benefits and developments, American Helicopter Society 55th Annual Forum, Montreal, Quebec, Canada, 25–27 May 1999.
[5] J. McColl, Overview of Transmissions HUM performance in UK, North Sea Helicopter Operations, Institution of Mechanical Engineers, Seminar S553, 1997.
[6] W. Wang, A.K. Wong, Model-based Gear Diagnostic Techniques, DSTO, TR- 1079, Airframes and Engine Division, Aeronautical and Maritime Research Laboratory, 2000.
[7] W. Wang, A.K. Wong, Autoregressive model-based gear fault diagnosis, Transaction of ASME, Journal of Vibration and Acoustics 124 (2002) 172–179.
[8] P.D. MacFadden, 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 (2) (1987) 173–183.
[9] S.L. Marple, Digital Spectral Analysis: With Applications, Prentice-Hall, Englewood Cliffs, NJ, ISBN 0132141493, 1987.
[10] S.M. Kay, Modern Spectral Estimation, Prentice-Hall, Englewood Cliffs, NJ, ISBN 013598582X, 1988.
[11] R.A. Wiggins, Minimum Entropy Deconvolution, Geoexploration, vol. 16, Elsevier Scientific Publishing, Amsterdam, 1978. pp. 21–35.
[12] M. Boumahdi, J. Lacoume, Blind identification using the kurtosis: results of field data processing, IEEE Transactions of Signal Processing 0-7803-2431-5/95 (1995) 1960–1983.
[13] J.Y. Lee, A.K. Nandi, Extraction of impacting signals using blind deconvolution, Journal of Sound and Vibration 232 (5) (1999) 945–962.
[14] A. K. Nandi, D. Mampel, B. Roscher, Blind deconvolution of ultrasonic signals in non-destructive testing applications, IEEE Transactions of Signal Processing 45(5) (1997).
[15] H. Endo, R. B. Randall, C. Gosselin, Differential diagnosis of spall vs. cracks in the gear tooth fillet region, Journal of Failure Analysis and Prevention 4(5) (2004).
[16] P. Sweeney, Gear transmission error measurement and analysis, Ph.D. dissertation, University of New South Wales, 1994.
[17] H. Endo, Simulation of gear faults and its application to the development of a differential diagnostic technique, Ph.D dissertation, University of New South Wales, 2005.

Further reading

[18] C.A. Cadzow, Blind deconvolution via cumulant extrema, IEEE Signal Processing Magazine, May 1996.
[19] O. Shalvi, E. Weinstein, New criteria for blind deconvolution of nonminimum phase systems (Channels), IEEE Transactions on Information Theory 36(2) (1990).

[20] M. Boumahdi, J. Lacoume, Blind identification using the kurtosis: results of field data processing, IEEE 0-7803-2431-5/95, 1995.
[21] W. Wang, Identification of gear mesh signal by kurtosis maximisation and its application to CH46 Helicopter gearbox data, 11th IEEE Workshop on Statistical Signal Processing, Singapore, August 6–8, 2001.
[22] R.B. Randall, Frequency Analysis, third ed, Bruel & Kjaer, Copenhagen, 1998.
[23] S. W. Smith, The Scientist and Engineer's Guide to Digital Signal Processing, second ed., California Technical Publishing, 0-9660176-6-8, 1999.