FAULT SEVERITY TRENDING IN ROLLING ELEMENT BEARINGS

Win Sum Siew, Wade A. Smith*, Zhongxiao Peng and Robert B. Randall

School of Mechanical and Manufacturing Engineering
UNSW Australia, Sydney NSW 2052, Australia
*Email: wade.smith@unsw.edu.au

Abstract

Fatigue-induced fault propagation is the most common cause of failure in rolling element bearings. As most classic diagnostic indicators do not trend monotonically with wear development, the ability to assess fault severity remains limited at present. This paper aims to expand current understanding and establish a correlation between the vibration signal and the actual extent of fault propagation. To achieve this, an extensive test program was undertaken using a laboratory test rig fitted with bearings initially seeded with small faults on the inner race. Vibration data from the rig was collected and analysed systematically at regular time intervals. Concurrently, the defect size and raceway topography were examined using laser scanning microscopy. The results from both sources were then combined to provide valuable insight into the effectiveness of common vibration indicators for fault trending. Through a detailed comparison of these indicators, a method of tracking fault severity is suggested which will aid greatly in the prognostics of rolling element bearings.

1. Introduction

Rolling element bearings (REBs) have widespread usage in industry. Due to the harsh operating conditions, they are often prone to potential failure. As a matter of fact, their failure has been cited as one of the most frequent reasons for machine breakdown, particularly in rotating machines [1]. Without proper maintenance and upkeep, they could experience premature failure which would lead to hefty repair costs as well as unwanted down time. On occasion, failure of bearings could result in complete destruction of the entire machinery, such as an air crash accident studied by Salam et al. [2].

To date, condition-based maintenance has been widely accepted by industry as an effective maintenance approach. It relies heavily on the ability to keep track of the current machine condition through the use of condition monitoring techniques, a prevalent one being vibration analysis. Typically, condition-based maintenance can be divided into three distinct phases, namely fault detection, fault diagnosis and fault prognosis. While the first two stages are widely researched, the last stage is not as well-understood. Essentially, prognosis is concerned with the determination of the future condition of the component based on past and currently acquired condition monitoring data. This could be accomplished by trending some parameters which are sensitive to the current condition of the machine. Successful trending of such fault severity metrics would allow the accurate prognosis of the machine which enables a reasonable assessment of remaining useful life (RUL). This brings significant economic advantage in most industrial applications since it avoids the occurrence of unexpected failure, minimises downtime and allows the machine to be operated until repair is truly necessary.

Having said that, at present, the correlation between vibration health indicators and the actual extent of fault severity is not well-established, particularly in the case of REBs. This is due to the fact that most classic diagnostic indicators do not trend monotonically with fault development. Randall [3] has also noted that in some situations, the trend could take the form of stepwise changes instead of uniform variation. Apart from that, in cases where the fault extent directly affects the rate of fault deterioration, the rate of fault development could increase significantly with time, subsequently leading to rapid changes in the trending parameter within a short time frame. These phenomena further exacerbate the difficulty in tracking the fault severity, causing significant challenges in the prognosis of REBs.

2. Trending Parameters

Parameters chosen as the health indicator must be able to reflect the current condition of the component reasonably well. As pointed out by Shakya et al. [4], an ideal trending parameter should exhibit three main attributes: robustness, sensitivity and early detectivity. Robustness refers to the ability of the parameter to register a consistent value for constant load, speed and fault size. Sensitivity refers to the ability of the parameter to indicate a discernible change for an increase in fault severity. Early detectivity, on the other hand, means that the parameter needs to undergo reasonable change upon a defect initiation so that any defect can be diagnosed in its early stage. Similar to these three attributes, Zhang et al. [5] also proposed a measurement of trending parameter suitability that encompasses three aspects: correlation, monotonicity and robustness. It can therefore be concluded that a good health indicator should be robust, sensitive, tolerant to outliers and vary monotonically with defect development. Trending of a parameter with such desirable traits will enable the tracking of the fault evolution.

Generally, trending parameters can be classified into two broad categories:

1. Time domain indicators, which involve the direct examination of the vibration signal and computation of its statistical features
2. Spectral indicators, which involve examination of the signal's variation in the frequency domain, for example a change in amplitude of particular frequency components or bands.

Following sections will outline and review the parameters studied in this paper, with emphasis given to parameters that are commonly adopted.

2.1 Root mean square

Root mean square (RMS) value is one of the simplest parameters in the time domain. Referring to Equation (1) below, it is defined as the square root of the mean of the squares of the signal x_i. It describes the power content in the vibration signal.

$$x_{rms} = \sqrt{\frac{1}{N}\left[\sum_{i=1}^{N}(x_i)^2\right]} \tag{1}$$

Mathew and Alfredson [6] noted that this parameter is of limited value in regard to fault detection. In the early stage of fault development, the RMS value may not be affected. Likewise under adverse conditions, there may also be no significant changes in RMS value. Furthermore, as a rolling bearing's overall vibration often increases only in the final stages of fault development, this parameter might offer late warning of fault propagation and failure. Nonetheless, despite these shortcomings, RMS has frequently been chosen as a trending indicator due to its overall simplicity. In this regard, researchers have reported a certain degree of success [7, 8].

2.2 Crest factor

Crest factor (CF) is defined as the ratio of the peak value to RMS value of the vibration signal.

$$x_{cf} = \frac{x_{max}}{x_{rms}} \tag{2}$$

As the bearing deteriorates, the peak value of the waveform increases more rapidly than the RMS value due to the increase in impulsiveness [6]. This subsequently results in an increase in crest factor. Generally, crest factors of higher than 3~3.5 indicate bearing damage [9, 10]. Compared to RMS value, crest factor is capable of providing an earlier warning of bearing failure. Nonetheless, towards the end of bearing life, the crest factor typically decreases due to the increase in RMS value with little change in peak value.

2.3 Kurtosis

First proposed by Dyer and Stewart [11], kurtosis is the fourth moment normalised by the square of the mean square of the vibration signal waveform.

$$x_{kur} = \frac{\sum_{i=1}^{N}(x_i-\bar{x})^4}{{x_{rms}}^4} \tag{3}$$

As its formulation involves a fourth moment, kurtosis is highly sensitive to the impulsiveness of the vibration signal. It gives the difference between the distribution of the sampled values and a normal (Gaussian) distribution. For a bearing in good condition, the value of kurtosis is close to 3. A significant deviation from this value is regarded as an indication of developing fault. As the bearing deteriorates, the kurtosis value initially increases, and it often reduces once the defect is well-advanced. In a study conducted by Bolaers et al. [12], it was shown that kurtosis is a better indicator than crest factor for the detection of an impulsive defect. While studies [13-15] have demonstrated the practicality of kurtosis in bearing fault detection, its application in fault severity trending remains uncertain. In fact, in a study conducted by Lybeck et al. [7], it was found that kurtosis is essentially uncorrelated with spall size and hence will not be a good severity indicator.

2.4 Other parameters

Apart from the three common parameters discussed previously, there are a number of other parameters which could also be used as potential trending indicators. Table 1 provides a summary of these parameters and their formulation. It should be noted that the amplitude of ball pass frequency inner race (BPFI) refers to the component in the envelope spectrum, obtained from the full bandwidth signal after cepstral prewhitening [16].

Table 1. Other potential trending parameters

Parameter	Formulation	References
Skewness	$x_{skew} = \frac{\sum_{i=1}^{N}(x_i-\bar{x})^3}{{x_{rms}}^3}$	[17, 18]
Shape Factor	$x_{SF} = \frac{x_{rms}}{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert}$	[19]
Impulse Factor	$x_{IF} = \frac{x_{peak}}{\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert}$	-
Shannon Entropy	$x_h = -\sum_{i=1}^{N} p_i \log p_i$	[20]
Amplitude of BPFI (envelope spectrum)	-	-

3. Methodology

3.1 Experimental setup

3.1.1 Test Rig

The test rig used in this study was a Bearing Prognostics Simulator provided by SpectraQuest Inc. and is shown in Figure 1.

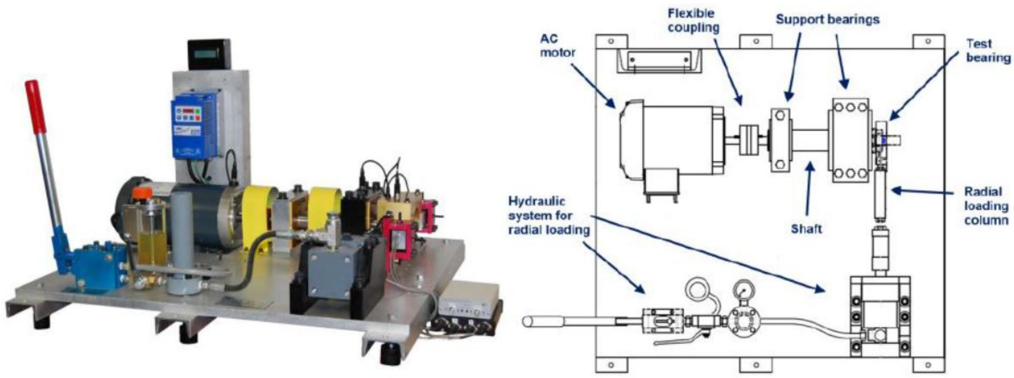

Figure 1. SpectraQuest Bearing Prognostics Simulator. (Left) general view; (Right) floor plan

It can be seen that the shaft is supported by two support bearings, thereby creating a cantilever arrangement for the test bearing. A purely radial force can then be applied through the radial loading column onto the bearing housing. Two PCB 352C04 accelerometers were mounted on the test bearing housing: one horizontally and the other vertically. The rotating speed of the shaft can be adjusted by the motor and its associated control system. A tachometer was also installed to measure and display the actual angular shaft speed.

3.1.2 Bearings and seeded fault

Table 2 summarises the properties of test bearing used in this study.

Table 2. Specifications of test bearing

Model	Nachi 6205-2NSE9
Type	Single-row deep groove ball bearing
Number of balls	9
Ball diameter	7.94 mm
Bearing pitch diameter	39.04 mm
Contact angle	0°

To simulate a localised bearing fault, a single notch was artificially seeded across the entire inner race by using electrical discharge machining. The notch was created such that it is approximately 0.5 mm deep at the middle of the raceway, sufficient to prevent the balls from contacting the bottom of the notch.

In addition to the seeded fault, the bearings tested in this study were also modified in order to facilitate the use of a laser microscope for examination of actual extent of fault propagation. The rivets joining the two cage halves were removed, which allows the cage to be pulled apart and enables complete disassembly of the entire bearing. The reassembly process, on the other hand, was accomplished by using M1.6 screws and nuts to clamp the cage together. This allows the bearing to function normally again.

3.2 Test Details

Two tests were conducted in this paper. In both cases, the inner race was seeded with a notch of 0.4 mm × 0.15 mm (width by depth) and the tests were performed with a nominal shaft speed of 6 Hz. Three sets of data were recorded: a tachometer signal, and horizontal and vertical accelerometer signals. These data were collected at specified time intervals: 30 minutes for Test 1 and 15 minutes for Test 2. However, this sampling interval was reduced upon detection of fault propagation. This is in reflection of the actual industrial practice. The duration of each measurement was 10 s and the sampling frequency used was 131072 Hz. In Test 1, the initial applied radial load was 7 kN and it was later increased to 14 kN. Test 2 was performed by using 14 kN radial load right from the beginning.

In order to obtain actual visual information about the fault severity, the test was paused at predetermined time intervals of 1, 2 or 4 hours. The test bearing was then removed and disassembled. After cleaning, it was subsequently examined under a Keyence laser scanning microscope. The bearing was then re-greased and re-assembled for further testing. For cases where the fault was too extended, a digital vernier caliper was used to measure the fault size. A mean spall size was then calculated by taking the average of 10 measurements at different locations.

4. Results and Discussion

4.1 Test 1

As mentioned previously, Test 1 was initially conducted with an applied load of 7 kN. The bearing was disassembled at a fixed time interval of every four hours and examined under the microscope. Figures 2 and 3 below show the condition of the seeded fault after 20 hours of testing.

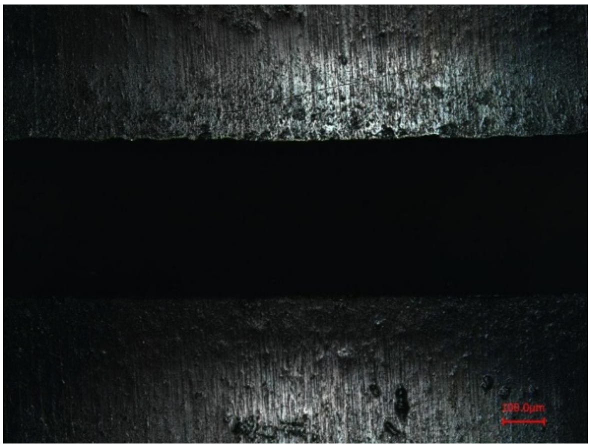

Figure 2. Fault topography after elapsed testing duration of 20 hours

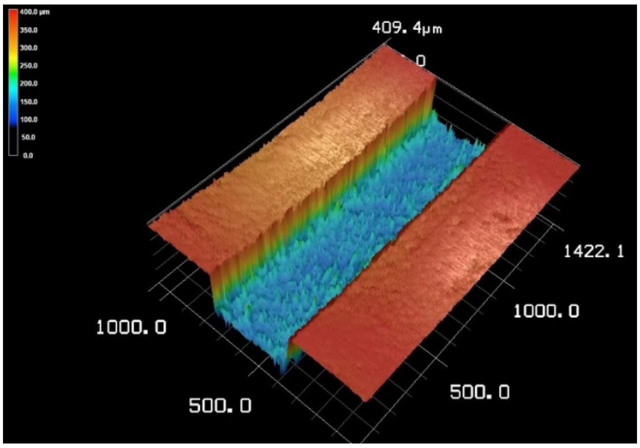

Figure 3. 3D scanning image showing the fault surface topography after test duration of 20 hours

Evident from the figures above, there has been no discernible fault propagation after 20 hours of testing. This was undesirable because wear development is essential in the study of fault severity trending. In order to accelerate the rate of wear development, the experiment was continued by increasing the applied load to 14 kN. This approach proved to be very effective as significant fault propagation was observed upon the disassembly process. Figure 4 shows the extent of deterioration at a cumulative test duration of 24, 26 and 28 hours, respectively.

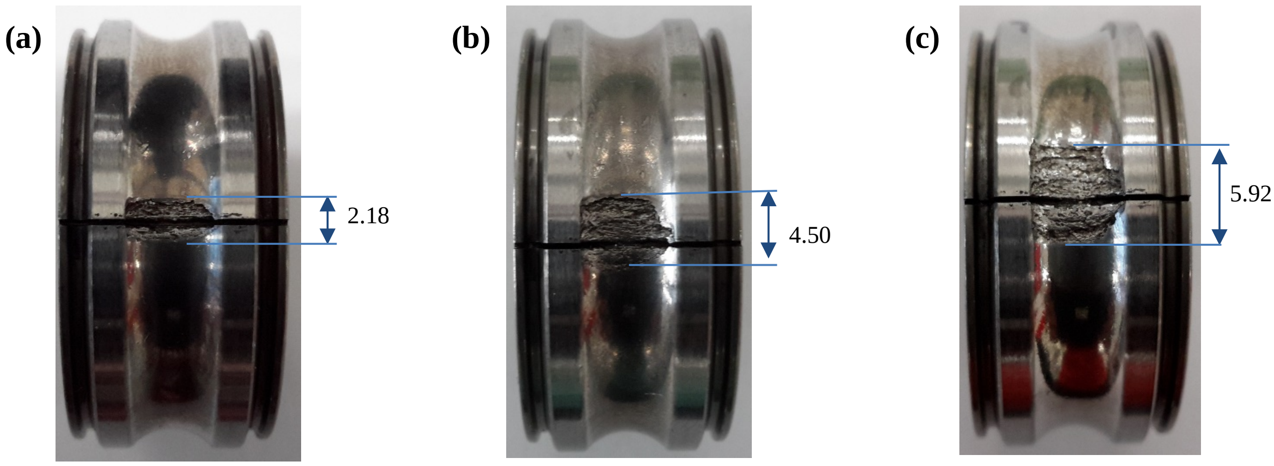

Figure 4. Extent of fault at different elapsed test duration. (a) 24 hours; (b) 26 hours; (c) 28 hours. Indicated are the mean defect sizes in mm.

Fault severity trending was then carried out by computing the values of various trending parameters and plotting their evolution with time. Figure 5 shows the result. Note that a step between lines indicates an instance of disassembly.

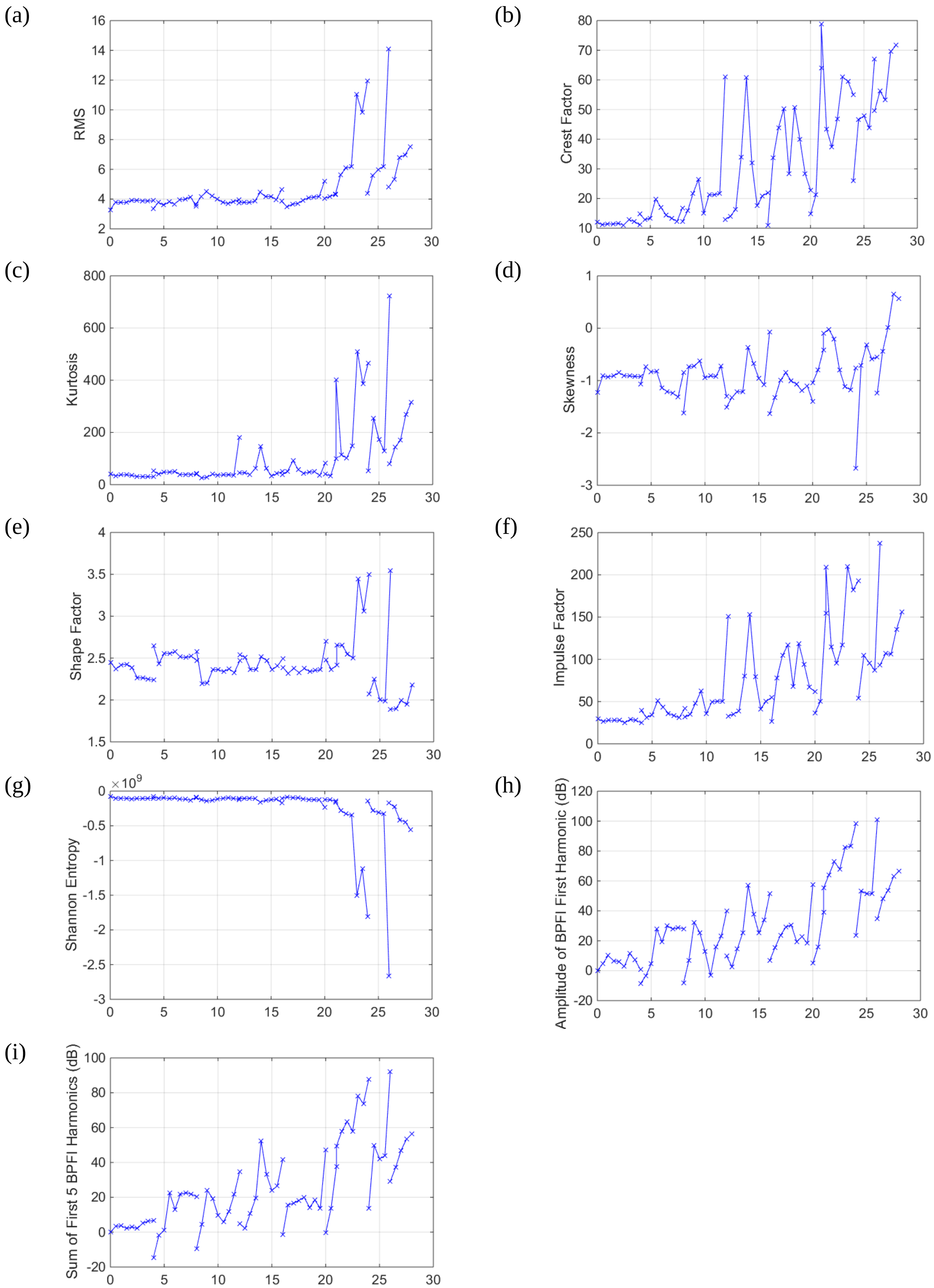

Figure 5. Trending plot of different indicators with time. Horizontal axis represents time in hours. (a) RMS; (b) Crest factor; (c) Kurtosis; (d) Skewness; (e) Shape factor; (f) Impulse factor; (g) Shannon entropy; (h) Amplitude of first BPFI harmonic (envelope spectrum); (i) Sum of first 5 BPFI harmonics (envelope spectrum)

From Figure 5, it is observed that the skewness and shape factor do not exhibit a trend with the progress of time. Their values fluctuate and do not follow a discernible pattern. As such, they are deemed unsuitable to be used as a severity indicator. Other parameters, to a certain extent, all show some form of variation pattern with time. Due to the similarities in their formulation, a number of parameters indicate trends that are alike. It can be seen from Figure 5 that RMS and Shannon entropy exhibit similar trends; crest factor and impulse factor show a similar pattern. Besides these, an interesting observation is that the trend of amplitude of first BPFI harmonic and the sum of first 5 BPFI harmonics are quite alike. This implies that the first harmonic dominates over the higher harmonics.

As it is known from laser scanning microscopy that there was no fault propagation (albeit the surface roughened) in the initial 20 hours, one would expect the trending parameter to be steady and maintain a constant value in that time interval. Looking at Figure 5, it is seen that RMS and kurtosis exhibit this expected trending behaviour. Crest factor, on the other hand, appeared to undergo significant fluctuation despite the fact that there was no damage development. Thus, it is concluded that crest factor and the similar impulse factor will not be good trending parameters. Apart from that, the plots of BPFI amplitude show a general monotonically increasing trend and they display a significant change after 20 hours. These behaviours show that BPFI amplitude could be a potentially useful severity indicator.

Another interesting feature of the plots in Figure 5 is the apparent drop in value of the parameters after every instance of disassembly (i.e. at 4, 8, 12, 16, 20, 24 and 26 hours). This phenomenon was suspected to be caused by the removal of wear debris during the disassembly process. In the early stages of test, not much wear debris was present and so this phenomenon was not pronounced. However, towards the end of testing, as the wear debris increased in size and amount, this drop in parameter value is quite considerable. This possible causation is further investigated in Test 2.

4.2 Test 2

In contrast to the complete grease removal in Test 1, Test 2 was conducted such that the lubricant and debris condition were maintained after every act of disassembly. Through this configuration, the effect of debris condition on parameter behaviour could be examined.

Due to the fact that Test 2 was performed with an applied load of 14 kN from the very beginning, the rate of fault propagation was quite remarkable. Figure 6 below shows the surface topography of the raceway defect and demonstrates the contrast between the initial condition and the extent of deterioration after just 1 hour of testing.

It can be seen from Figure 6 that there has been considerable fault deterioration. This observed fault propagation is persistent throughout the test. Its progress with time is shown in Figure 7.

Figure 8 shows the trending of different severity indicators with time. Note that only four selected parameters are presented here for the purpose of brevity.

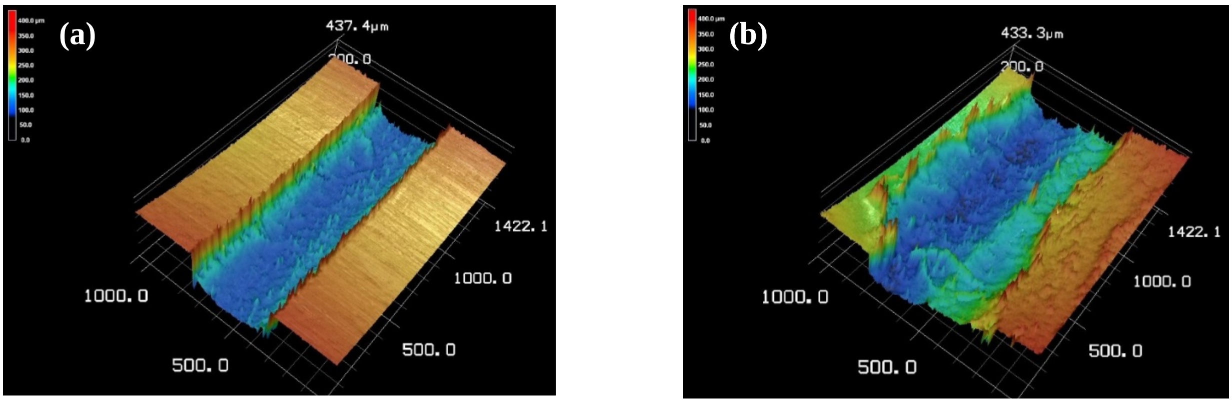

Figure 6. 3D surface topography of the defect. (a) Initial condition; (b) After 1 hour of testing

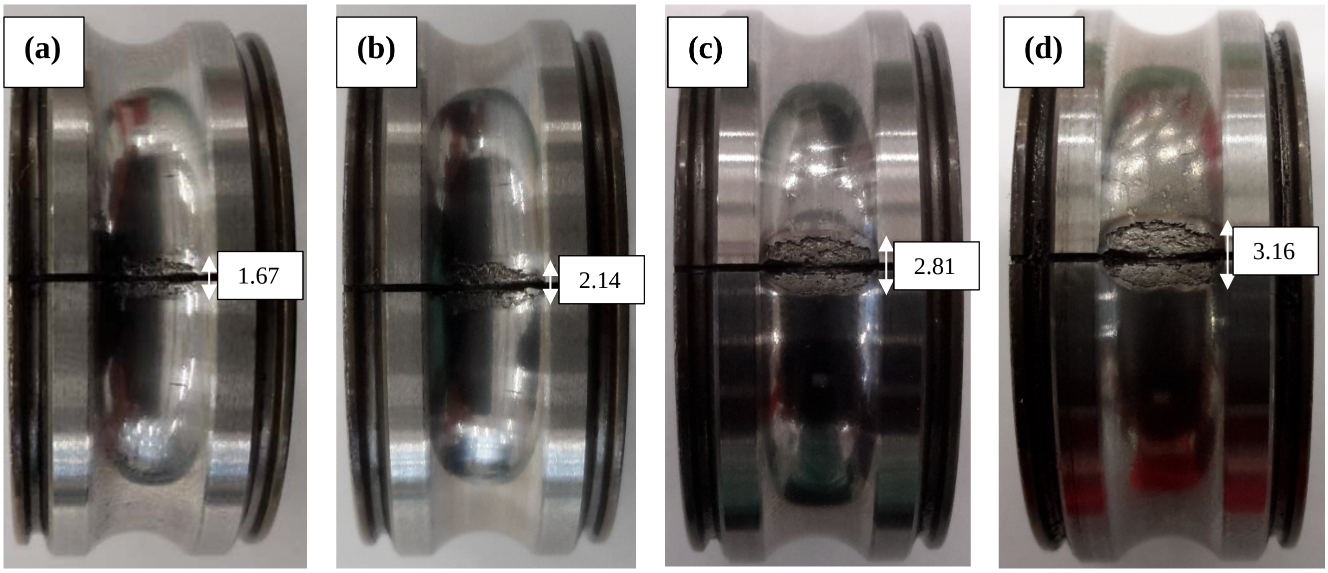

Figure 7. Fault extent after different test duration. (a) 2 hours; (b) 3 hours; (c) 4 hours; (d) 4.25 hours. Indicated are the mean defect sizes in mm.

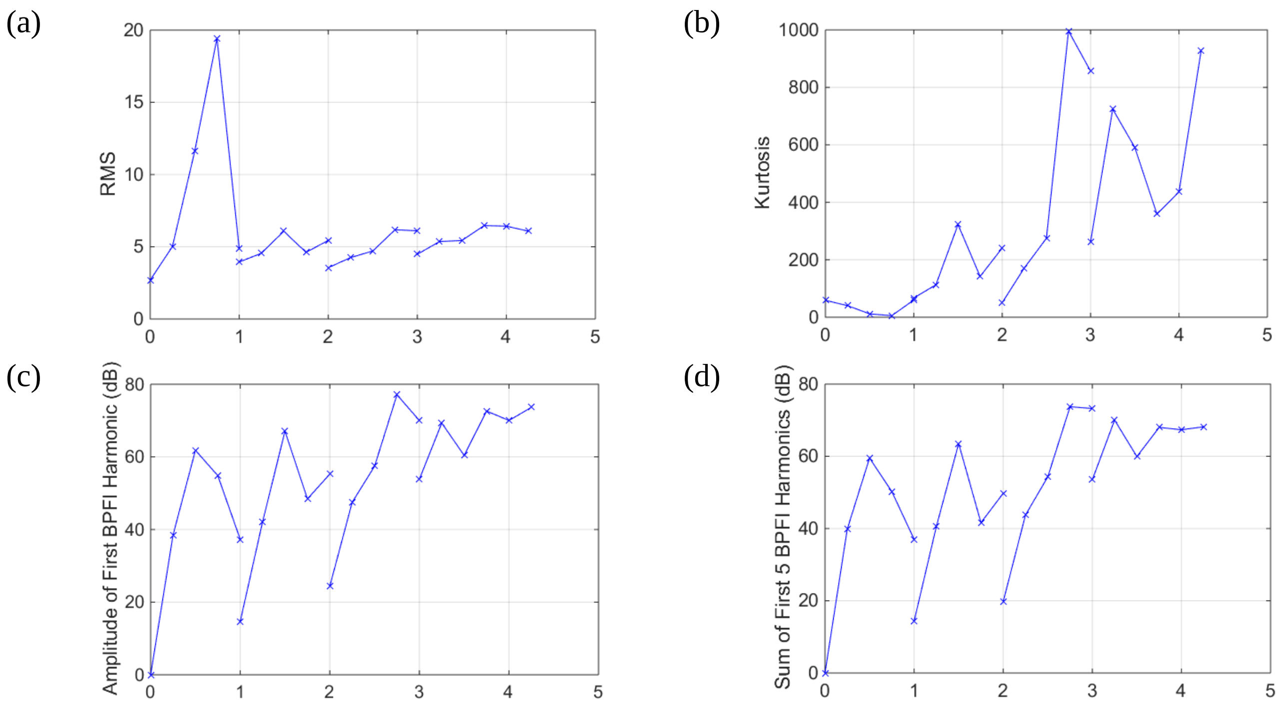

Figure 8. Trending plot of selected indicators with time. Horizontal axis represents time in hours. (a) RMS; (b) Kurtosis; (c) Amplitude of first BPFI harmonic; (d) Sum of first 5 BPFI harmonics

From Figure 8, it is observed that the performance of RMS is not as good as that of Test 1. It appears to be relatively constant even though the fault was propagating (as revealed in Figure 7). Nonetheless, it should be noted that the two anomaly points (0.5 and 0.75 hour) in the RMS plot were caused by accidental frictional rubbing between the components in the test rig and should be disregarded.

On the other hand, kurtosis and BPFI amplitude exhibit an observable pattern in which a general increase in value is seen as the time progresses. This implies that they could potentially be used as a trending indicator which will perform reasonably well. In addition, it can be seen from Figure 8(c) and (d) that the trending plots of BPFI amplitude show identical patterns, whether it be the amplitude of the first harmonic or the sum of the first 5 harmonics. This observation conforms to that of Test 1, and collectively they imply that the amplitude of the first BPFI harmonic is a sufficient representation of the harmonic family. On a side note, consistent with the result in Test 1, skewness and shape factor exhibit no trend whatsoever again with the fault progression in this test.

Lastly, it is also observed that the drop in value after instance of disassembly seems to have slightly diminished in value. This is particularly noticeable in the RMS plot and is probably a direct consequence of retaining wear debris.

4.3 Correlation with spall size

Combining the results (vibration data and actual defect size) obtained in Tests 1 and 2, the relationship between the trending parameter and the spall size was studied. As the faults in this study have been allowed to degrade in a more natural manner (compared to just having a pre-seeded large fault), it is believed the results here will be a closer reflection of reality. Figure 9 shows the variation of four severity indicators with spall size. In each plot, a linear least-squares regression line is fitted to the points in order to assess the goodness of fit.

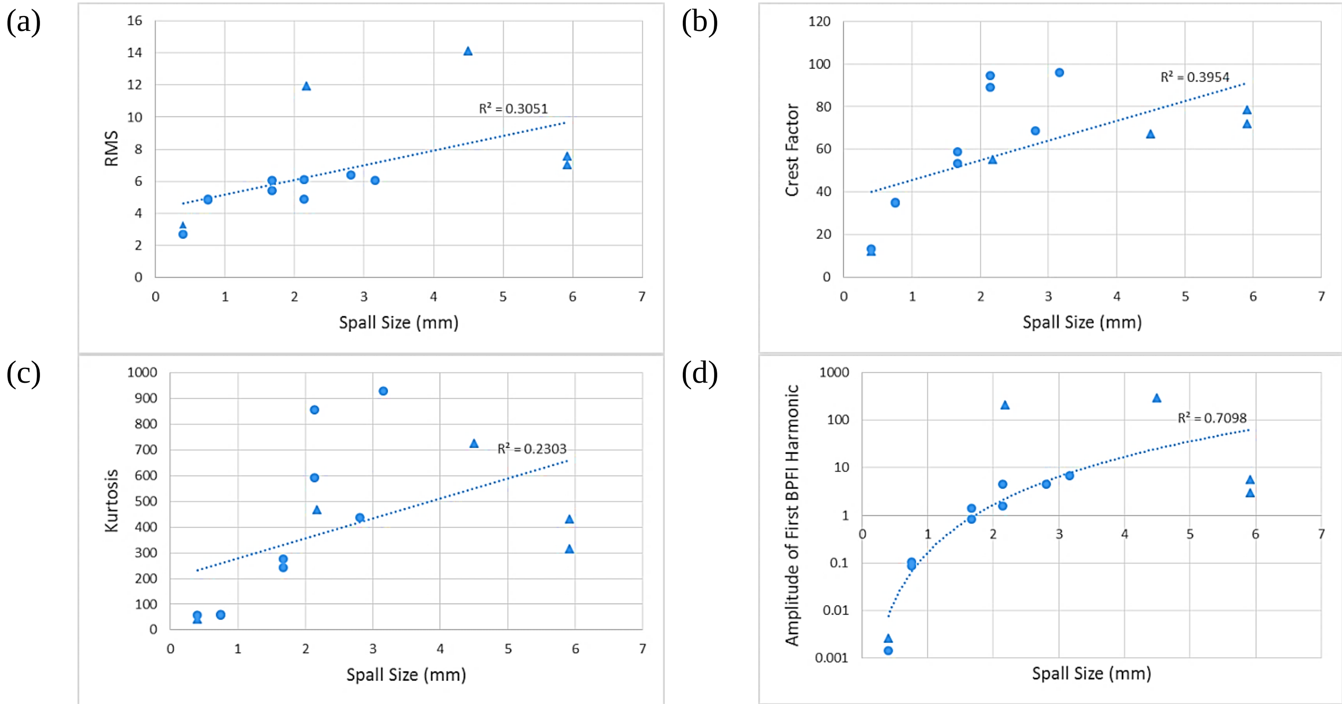

Figure 9. Plots of trending parameter against mean spall size. (a) RMS; (b) Crest factor; (c) Kurtosis; (d) Amplitude of first BPFI harmonic. Triangles indicate Test 1 and dots indicate Test 2.

The results in Figure 9 indicate that the amplitude of the BPFI harmonic has a good correlation with spall size. In contrast, RMS and kurtosis show a poor correlation with spall propagation.

It should be noted that this correlation attempt only serves as a crude starting point and is thus subject to a number of limitations. For example, the underlying relationship between the parameter and the spall size might not be linear. In addition, the number of data points might be suboptimal for a strong conclusion to be made. In this regard, it was noted that if the two peculiar outliers were removed in the RMS plot, the R-squared value will increase from 0.3051 to 0.7015, which is a very significant change.

5. Conclusions

The results from this study identified several potential parameters useful in the tracking of fault severity. Among them are RMS, kurtosis, and amplitude of the defect frequency component in the envelope spectrum. RMS appears to be quite promising in this respect, however it only undergoes significant change when the fault is already well-advanced. On the other hand, kurtosis is able to detect fault propagation earlier, but it is not as stable as RMS and could experience fluctuations. The amplitude of the defect frequency component (BPFI in this case) has demonstrated good trending behaviour in this test and will be studied more in the future.

Further work is already underway in order to further investigate and develop a trending metric for reliable prognostics of rolling element bearings. It may certainly be true that not a single parameter could serve as an accurate health indicator across all stages of wear development and an effective combination of different parameters is the key to success.

Acknowledgements

This research was supported by the Australian Research Council and SpectraQuest, through Linkage Project LP110200738. The authors would like to acknowledge the UNSW mechanical workshop for their help in the project, in particular Ian Cassapi and Andy Higley.

References

[1] Randall, R.B. and Antoni, J. "Rolling element bearing diagnostics—A tutorial", *Mechanical Systems and Signal Processing*, **25**, 485-520 (2011).

[2] Salam, I., Tauqir, A., Ul Haq, A., and Khan, A.Q. "An air crash due to fatigue failure of a ball bearing", *Engineering Failure Analysis*, **5**, 261-269 (1998).

[3] Randall, R.B. "Computer aided vibration spectrum trend analysis for condition monitoring", *Maintenance Management International*, **5**, 161-167 (1985).

[4] Shakya, P., Darpe, A.K., and Kulkarni, M.S. "Vibration-based fault diagnosis in rolling element bearings: ranking of various time, frequency and time-frequency domain data-based damage identification parameters", *The International Journal of Condition Monitoring*, **3**, 1-10 (2013).

[5] Zhang, B., Zhang, L., and Xu, J. "Degradation feature selection for remaining useful life prediction of rolling element bearings", *Quality and Reliability Engineering International*, (2015).

[6] Mathew, J. and Alfredson, R.J. "The condition monitoring of rolling element bearings using vibration analysis", *Journal of Vibration, Acoustics, Stress, and Reliability in Design*, **106**, 447-453 (1984).

[7] Lybeck, N., Marble, S., and Morton, B. "Validating prognostic algorithms: A case study using comprehensive bearing fault data" in *IEEE Aerospace Conference*, Big Sky, MT, 3-10 March 2007.

[8] Kogan, G., Itzhak, I., Shaharabany, S., Bortman, J., and Klein, R. "Towards model based prognostics - Characterization of fault size in bearings" in *The Tenth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies*, Krakow, Poland, 17-20 June 2013.

[9] Smith, J.D., *Vibration measurement and analysis*, Butterworths, 1989.

[10] Norton, M.P. and Karczub, D.G., *Fundamentals of noise and vibration analysis for engineers*, 2nd edition, Cambridge University Press, 2003.

[11] Dyer, D. and Stewart, R.M. "Detection of rolling element bearing damage by statistical vibration analysis", *Journal of Mechanical Design*, **100**, 229-235 (1978).

[12] Bolaers, F., Cousinard, O., Marconnet, P., and Rasolofondraibe, L. "Advanced detection of rolling bearing spalling from de-noising vibratory signals", *Control Engineering Practice*, **12**, 181-190 (2004).

[13] Rogers, L.M. "The application of vibration signature analysis and acoustic emission source location to on-line condition monitoring of anti-friction bearings", *Tribology International*, **12**, 51-58 (1979).

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

[15] Dron, J.P., Bolaers, F., and Rasolofondraibe, l. "Improvement of the sensitivity of the scalar indicators (crest factor, kurtosis) using a de-noising method by spectral subtraction: application to the detection of defects in ball bearings", *Journal of Sound and Vibration*, **270**, 61-73 (2004).

[16] Sawalhi, N. and Randall, R.B. "Signal prewhitening using cepstrum editing (liftering) to enhance fault detection in rolling element bearings" in *COMADEM*, Stavanger, Norway, 30 May-1 June 2011.

[17] Ovacikh, A.K., Paajarvi, P., and LeBlanc, J.P. "Skewness as an objective function for vibration analysis of rolling element bearings" in *8th International Symposium on Image and Signal Processing and Analysis*, Trieste, Italy, 4-6 September 2013.

[18] de Almeida, R.G.T., da Silva Vicente, S.A., and Padovese, L.R. "New technique for evaluation of global vibration levels in rolling bearings", *Shock and Vibration*, **9**, 225-234 (2002).

[19] Wang, H. and Chen, P. "A feature extraction method based on information theory for fault diagnosis of reciprocating machinery", *Sensors (Basel)*, **9**, 2415-36 (2009).

[20] Hemmati, F., Orfali, W., and Gadala, M.S. "Rolling element bearing condition monitoring using acoustic emission technique" in *25th International Conference on Noise and Vibration Engineering*, Leuven, Belgium, 17-19 September 2012.