New statistics for simultaneously machine incipient fault detection and monotonic degradation assessment

Bingchang Hou¹, ², Yikai Chen¹, ², Yanqing Deng¹, ², Yuting Wang¹, ² and Dong Wang¹, ², ³

¹ The State Key Laboratory of Mechanical Systems and Vibration, Shanghai Jiao Tong University, Shanghai 200240, P.R. China
² Department of Industrial Engineering and Management, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, P.R. China
³ E-mail: dongwang4-c@sjtu.edu.cn

Abstract. Machine condition monitoring (MCM) has become an important tool to avoid sudden machine breakdown and gaining more economic profits. Tasks including early fault detection and monotonic degradation assessment are important in MCM. For the incipient fault detection, statistics such as kurtosis, Gini index are widely utilized, but they cannot give an accurately incipient fault detection time, and many fluctuations may exhibit. For the monotonic degradation assessment, root-mean-square are commonly used, however, it is sensitive to energy, and cannot show distinct degradation tendency in an early fault state. Those drawbacks have limited the development of practical MCM algorithms. To address those issues, this paper proposed four parameterized statistics for simultaneously early fault detection and monotonic degradation assessment. The four parameterized statistics can be health indicators and simplify the MCM algorithms, which can be beneficial to the practical MCM applications.

1. Introduction

Machine condition monitoring (MCM) has become an important tool to avoid sudden machine breakdown and gaining more economic profits, which has attracted much attention from academia and industry. Clearly and accurately early fault time detection can be an alarm for fault diagnosis and indicate a start time for remaining useful life (RUL) prediction [2]. Monotonic degradation assessment can be used to establish a model for RUL prediction. Therefore, the early fault detection and monotonic degradation assessment are two major prerequisite tasks to acquire good MCM results.

Faults of bearings and gears usually have features of non-Gaussian and non-stationarity [3], i.e., impulsiveness and cyclo-stationarity. Many statistics are utilized to characterize the two features to distinguish different health states. Antoni et al. [4] proposed to quantify different bandpass-filtered signals with kurtosis to locate at which frequency band the fault signatures exist, and they further constructed fast Kurtogram [5] as an optimal frequency band selection tool. The fast Kurtogram was widely used and gradually became a classic fault diagnosis method [6]. Since the kurtosis is sensitive to non-Gaussian behaviors, Li et al., [2] employed it to determine the first predicting time in a bearing RUL prediction algorithm. But kurtosis cannot exhibit a monotonic degradation assessment tendency. Since root-mean-square (RMS) is an indication of signal’s energy, Li et al., [2] used it to predict bearing RUL. Kong et al. [7] also constructed a modified RMS-based model to predict RUL of bearing. Ahmad et al. [8] also developed a hybrid technique to predict RUL of bearing, the RMS was used in this...
technique. However, the energy increment of incipient fault is not prominent, RMS cannot be used to detect incipient fault time.

Machine learning-based methods have received great concerns for MCM [9, 10]. Feature extraction is one of the most important steps in most machine learning-based methods, and some basic statistical parameters, such as kurtosis, skewness, and entropy, usually play a significant role at this step. Guo et al. [9] established a recurrent neural network-based method to predict the RUL of bearing, in which the entropy, kurtosis, crest factor, and energy ratios of wavelet packages are utilized. Based on statistical features including RMS, crest factor, kurtosis, etc., Liao [10] used genetic programming to generate more effective features to predict the RUL of bearing.

Sparsity measures such as Gini index, kurtosis, and smoothness index, etc., are also essentially statistical parameters. The most known application of kurtosis is the spectral kurtosis developed by Antoni et al [4]. And the smoothness index was constructed by Liang and Bozchalooi [11] to determine wavelet parameters for bearing fault diagnosis. Recently, Wang et al., [12] found that kurtosis can be reformulated as the square of $L_2/L_1$ norm, and the smoothness index is equal to $L_0/L_1$ norm. In this way, a connection between the kurtosis, smoothness index and $L_p/L_q$ norm is established. The Gini index, which is constructed in economic field, was also utilized by Miao et al. [13] to replace kurtosis in the Kurtogram for fault diagnosis. Antoni et al. [14] also studied the negative entropy originally developed in thermodynamics, and then put forward a new tool of spectral negative entropy, it aimed to simultaneously characterize the cyclo-stationarity and impulsiveness of fault signatures.

Though many achievements have been obtained on the statistics for MCM applications nowadays, there are still few works that can be used for simultaneously early fault detection and monotonic degradation assessment. To make some improvements in this aspect, four parameterized statistics are proposed as health indicators to achieve accurately early fault detection and monotonic degradation assessment. The proposed four parameterized statistics are promising to make the MCM algorithms simpler and more reliable.

The remaining part of the present article is structured as follows: Section 2 reports the proposed 4 new statistics, and Section 3 verifies their effectiveness with two public bearing run-to-failure datasets. Conclusions are drawn in Section 4.

2. New statistics for simultaneously early fault detection and monotonic machine condition monitoring

A sampled signal is denoted as $x[n]$, where $n = 1, 2, ..., N$, and $N$ denotes the total signal sample points. By using Hilbert transform $H(\cdot)$, the analytic signal can be obtained as $x_A[n] = x[n] + jH(x[n])$, and square envelope signal is $SE[n] = |x_A[n]|^2$. Statistics are the basis of many MCM algorithms, and four classic sparsity measures used in the field of MCM are given as follows:

$$Kurtosis = \left( \frac{\sum_{n=1}^{N} SE[n]^2/n}{\sum_{n=1}^{N} SE[n]/N} \right)^2,$$

$$Smoothness index = \frac{\sum_{n=1}^{N} SE[n]/N}{\sum_{n=1}^{N} SE[n]^2/n},$$

$$Negative entropy = \left\langle \frac{SE[n]}{\langle SE[n] \rangle} \ln \left(\frac{SE[n]}{\langle SE[n] \rangle}\right)\right\rangle,$$

$$GI = 1 - 2 \sum_{n=1}^{N} \frac{SE_{order}[n]}{\|SE[n]\|_{L_1}} \left(\frac{N-n+0.5}{N}\right)\|SE[n]\|_{L_1},$$

where $\langle \cdot \rangle$ means arithmetic mean operation, and $SE_{order}[n]$ means $SE[n]$ ordered from the smallest to the largest.

Recently, Hou et al., [15] found that six classic sparsity measures including Gini index, $L_p/L_q$ norm, $pq$-mean, negative entropy, smoothness index, and kurtosis can be redefined as the ratio of different quasi-arithmetic means. Based on this new finding [15], ratio of different quasi-arithmetic means (RQAM) is put forward as a new framework to construct new statistics for machine condition
monitoring. Because the RQAM have a form of ratio, they are naturally dimensionless and may be free from varying operation working conditions. Moreover, RQAM has introduced some machine condition-related parameters in the definition of health indicators, which may be beneficial for monotonic degradation assessment. Based on the four sparsity measures and RQAM, Hou et al., [16] further proposed an adaptive weighted signal preprocessing technique (AWSPT) to improve their performance for MCM. Since the sparsity measures are intrinsically affected by impulsive noise, a monotonically decreasing function is designed to weight on the square envelope signal for adaptive damping of impulsive signal points. Further, sparsity measures are used to quantify the weighted square envelope signal. The weighted function is designed as follows:

$$w[n] = \frac{1}{SE[n] \times (SE[n] + m)}$$

(2)

where $m$ is a constant defined by a variance $\nu^2$ of white Gaussian noise under health state and a shifted percentage $SP$ of probability density curve of square envelope:

$$m = -2\nu^2 \ln(1 - SP)$$

(3)

By substituting it into the definition of the four original sparsity measures, four new parameterized statistics can be obtained, and respectively named as PS1, PS2, PS3, and PS4 as follows:

$$PS1 = \left( \sum_{n=1}^{N} \frac{(SE[n] + m)^{-1} - 2[n]/n}{\sum_{n=1}^{N} (SE[n] + m)^{-1} - 1/N} \right)^2$$

$$PS2 = \left( \prod_{n=1}^{N} (SE[n] + m)^{-1} \sum_{n=1}^{N} (SE[n] + m)^{-1} - 1/N \right)^{1/2}$$

$$PS3 = \left( \sum_{n=1}^{N} (SE[n] + m)^{-1} \ln \left( \frac{(SE[n] + m)^{-1}}{(SE[n] + m)^{-1} - 1} \right) \right)$$

$$PS4 = 1 - 2 \sum_{n=1}^{N} (SE[n] + m)^{-1} \frac{\text{order}[n]}{\|SE[n] + m\|_{L_1}} \left( \frac{N-n+0.5}{N} \right)$$

(4)

It should be noted that the constant $m$ is parameterized by the variance of white Gaussian noise under health state, and the variances of vibration signal in different health states are different, the four parameterized statistics are promising for showing monotonic degradation tendency. Therefore, the four statistics will be used as health indicators for MCM in Section 3.

3. Experimental verification with two bearing run-to-failure datasets

Two public bearing run-to-failure datasets respectively released by the NASA Center for Intelligent Maintenance Systems (NASA IMS) [17] and Xi'an Jiaotong University (XJTU) [18] will be used to demonstrate the effectiveness of the proposed new statistics in this section. According to Reference [16], the parameter $SP$ should be larger than 15%.

![Figure 1. Experimental platform of NASA IMS bearing dataset.](image)
3.1. Verification with NASA IMS bearing run-to-failure dataset

An experimental platform of the NASA IMS dataset is shown in Figure 1. Every 10 mins, a signal was collected at a sampling frequency of 20 kHz with 20480 sample points. During the run-to-failure test, 984 signals were collected and respectively stored in 984 data files. A previous study [19] has checked that the incipient fault happened at file number 532. An outer race failure was finally affirmed.

To monitor the incipient fault time and degradation tendency, the proposed four parameterized statistics including PS1, PS2, PS3, and PS4 were also employed, and the four sparsity measures including negative entropy, smoothness index, kurtosis, and Gini index were respectively used for comparison. The first health signal was used to calculate variance $v^2$ parameter to define constant $m$. When $SP = 15\%$, the variance was 0.0054, and the constant $m$ was 0.0018. To obtain an increasing tendency, the smoothness index and PS2 were multiplied by $-1$ and respectively named as negative smoothness index and NPS2.

The monitoring curves of classic sparsity measures and proposed four parameterized statistics are given in Figure 2(a) to 2(d) and Figure 2(e) to 2(h), respectively. From Figure 2, it can be known that though both sparsity measures and the proposed parameterized statistics can show incipient fault time around file number 532, but only the proposed parameterized statistics (i.e., PS1 to PS4) can show monotonic degradation assessment tendency. And it should also be noticed that the proposed four statistics can exhibit the incipient fault time around file number 532 in a clearer manner.

Figure 2. NASA IMS bearing condition monitoring curves respectively given by the (a) kurtosis, (b) negative smoothness index, (c) Gini index, (d) negative entropy (e) PS1, (f) NPS2, (g) PS3, and (h) PS4.

3.2. Verification with XJTU bearing run-to-failure dataset

The bearing run-to-failure dataset released by XJTU [18] was also utilized to further validate the effectiveness of the new parameterized statistics. Experimental platform of XJTU bearing run-to-failure dataset was given in Figure 3. At a sample frequency of 25.6 kHz, the signal was acquired every 1 min, each sample had a length of 1.28s. After the experimental test, a total of 491 samples were collected and denoted as bearing 2-1 sub-dataset.
Figure 3. An experimental platform of the XJTU bearing run-to-failure dataset.

The four sparsity measures and proposed four parameterized statistics were utilized to monitor the bearing degradation process. Since the variance of the first health signal is 0.1936, and the SP is assigned as 30%, the constant $m$ is 0.0629. And the monitoring curves are given in Figure 4(a) to 4(h), respectively. From Figure 4, it can be observed that though four classic sparsity measures and the proposed parameterized statistics (i.e., PS1 to PS4) can detect an early fault around file number 450, but the sparsity measures can not show monotonic degradation tendency, only the four parameterized statistics (i.e., PS1 to PS4) could simultaneously show clear early fault time and monotonic degradation trend.

Figure 4. XJTU bearing condition monitoring curves respectively given by the (a) kurtosis, (b) negative smoothness index, (c) Gini index, (d) negative entropy (e) PS1, (f) NPS2, (g) PS3, and (h) PS4.
4. Conclusions

To simplify the MCM algorithms and make them more reliable for practical applications, this paper proposed four new parameterized statistics (i.e., PS1 to PS4) as health indicators to simultaneously achieve accurately early fault detection and monotonic degradation assessment. Experiments with two bearing run-to-failure datasets have verified their effectiveness to simultaneously detect incipient fault and monitor degradation trend. As a result, the four parameterized statistics are promising to make MCM algorithms more simple and reliable.

As for the reason why the proposed four new parameterized statistics can simultaneously show early fault time and monotonic degradation tendency: Firstly, because that the parameterized statistics have a form of ratio, they are sensitive to non-Gaussian behaviour and can detect incipient fault time. Secondly, because that the four statistics are parameterized by a variance of vibration signal collected under healthy state, and the variances of signals during the degradation are changed continually, the parameterized four statistics can achieve monotonic degradation assessment. Since two run-to-failure experiments are verified the excellent properties of the four parameterized statistics, future works will be focused on a strictly mathematical proof of the excellent properties.

Acknowledgements

This research work was fully supported by National Natural Science Foundation of China (Project No. 51975355), and the "Zhiyuan Honor Program for Ph.D. student" at SJTU. We would like to thank anonymous reviewers for their valuable and constructive comments on our manuscript.

References

[1] Lei Y, Li N, Guo L, Li N, Yan T and Lin J 2017 Machinery health prognostics: A systematic review from data acquisition to RUL prediction Mech. Syst. Signal Process 104(2018) 799-834 https://doi.org/10.1016/j.ymssp.11.016
[2] Li N, Lei Y, Lin J and Ding S X 2015 An improved exponential model for predicting remaining useful life of rolling element bearings IEEE Trans. Ind. Electron 62(2015) 7762-7773 https://doi.org/10.1109/TIE.2455055
[3] Antoni J and Borghesani P 2018 A statistical methodology for the design of condition indicators Mech. Syst. Signal Process 114(2019) 290-327 https://doi.org/10.1016/j.ymssp.05.012
[4] Antoni J 2004 The spectral kurtosis: A useful tool for characterising non-stationary signals Mech. Syst. Signal Process 20(2006) 282-307 https://doi.org/10.1016/j.ymssp.09.001
[5] Antoni J 2005 Fast computation of the kurtogram for the detection of transient faults Mech. Syst. Signal Process 21(2007) 108-124 https://doi.org/10.1016/j.ymssp.12.002
[6] Wang Y, Xiang J, Markert Ra and Liang M 2015 Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications Mech. Syst. Signal Process 66-67(2016) 679-698 https://doi.org/10.1016/j.ymssp.04.039
[7] Kong X and Yang J 2019 Remaining Useful Life Prediction of Rolling Bearings Based on RMS-MAVE and Dynamic Exponential Regression Model IEEE Access 7(2019) 169705-169714 https://doi.org/10.1109/ACCESS.2954915
[8] Ahmad W, Ali Khan S and Kim J M 2017 A hybrid prognostics technique for rolling element bearings using adaptive predictive models IEEE Trans. Ind. Electron. 65(2017) 1577-1584 https://doi.org/10.1109/TIE.2733487
[9] Guo L, Li N, Jia F, Lei Y and Lin J 2017 A recurrent neural network based health indicator for remaining useful life prediction of bearings Neurocomputing 240(2017) 98-109 https://doi.org/10.1016/j.neucom.02.045
[10] Liao L 2013 Discovering prognostic features using genetic programming in remaining useful life prediction IEEE Trans. Ind. Electron. 61(2014) 2464-2472 https://doi.org/10.1109/TIE.2270212
[11] Bozchalooi I S and Liang M 2007 A smoothness index-guided approach to wavelet parameter selection in signal de-noising and fault detection J. Sound Vib. 308(2007) 246-267
[12] Wang D 2018 Some further thoughts about spectral kurtosis, spectral L2/L1 norm, spectral smoothness index and spectral Gini index for characterizing repetitive transients Mech. Syst. Signal Process 108(2018) 58-72 https://doi.org/10.1016/j.ymssp.02.034

[13] Miao Y, Zhao M and Lin J 2017 Improvement of kurtosis-guided-grams via Gini index for bearing fault feature identification Meas. Sci. Technol. 28(2017) https://doi.org/10.1088/1361-6501/aa8a57

[14] Antoni J 2015 The infogram: Entropic evidence of the signature of repetitive transients Mech. Syst. Signal Process 74(2016) 73-94 https://doi.org/10.1016/j.ymssp.04.034

[15] Hou B, Wang D, Xia T, Wang Y, Zhao Y and Tsui K 2020 Investigations on quasi-arithmetic means for machine condition monitoring Mech. Syst. Signal Process 151(2021) 107451 https://doi.org/10.1016/j.ymssp.107451

[16] Hou B, Wang D, Wang Y, Yan T, Peng Z and Tsui K L 2020 Adaptive Weighted Signal Preprocessing Technique for Machine Health Monitoring IEEE Trans. Instrum. Meas. 70(2021) https://doi.org/10.1109/TIM.3033471

[17] Qiu H, Lee J, Lin J and Yu G 2005 Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics J. Sound Vib. 289(2006) 1066-1090 https://doi.org/10.1016/j.jsv.03.007

[18] Lei Y, Han T, Wang B, Li N, Yan T and Yang J 2019 XJTU-SY Rolling Element Bearing Accelerated Life Test Datasets: A Tutorial J. Mech. Eng. 55(2019) 1-6 https://doi.org/10.3901/JME.16.001

[19] Wang D and Tsui K L 2017 Theoretical investigation of the upper and lower bounds of a generalized dimensionless bearing health indicator Mech. Syst. Signal Process 98(2018) 890-901 https://doi.org/10.1016/j.ymssp.2017.05.040