Original Article

Structural Health Monitoring
2018, Vol. 17(5) 1095–1109

DOI: 10.1177/1475921717736226
journals.sagepub.com/home/shm
SAGE

# Remaining useful life prediction of rolling element bearings using degradation feature based on amplitude decrease at specific frequencies

**Dawn An[1], Joo-Ho Choi[2] and Nam H. Kim[3]**

**Abstract**

This research presents a new method of degradation feature extraction to predict remaining useful life, the remaining time to the maintenance, of rolling element bearings. Since bearing fault is the foremost cause of failure in rotating machinery, there are many studies for evaluating bearings' health status to prevent a catastrophic failure. Most of these studies are based on health monitoring data, such as vibration signals that are indirectly related to bearing fault, from which degradation feature can be extracted. It is, however, challenging to extract a degradation feature that can be applied to all rolling elements. This study focuses on the amplitude decrease at specific frequencies, from which a robust degradation feature is extracted by employing the information entropy. Some important attributes are found from the degradation feature, which is used to predict the remaining useful life of bearings. This method is demonstrated using the real test data provided by FEMTO-ST Institute. The results show that bearings can be used up to 87% of their whole life and 59%–74% of life in average.

**Keywords**

Structural health monitoring, prognostics and health management, degradation feature extraction, rotating machinery, bearing, frequency domain, information entropy, remaining useful life

## Introduction

Bearing spall is the foremost cause of failure in rotating machinery, as 80%–90% of aircraft engine failures are caused by bearing failure. If bearings are not repaired before failure, it can lead to a catastrophic failure. Since bearings cannot be disassembled during operation, it is difficult to measure how much the degradation is progressed, and when the maintenance is required. Instead, indirect system responses, such as vibration signals, oil debris, and thermography, are used to evaluate the level of degradation. Especially, vibration signals have been used for damage evaluation for several decades.[1–3]

Diagnostics techniques for bearings have been developed based on vibration signals, which finds the relationship between failure mechanisms and frequencies. This relationship can be used to estimate the current damage level. There are many effective methods for bearing diagnostics such as discrete/random separation (DRS), minimum entropy deconvolution (MED), and envelope analysis.[4–6] Since damage can grow rapidly, however, sometimes the maintenance is performed too early or too late to prevent the failure especially when only the information from the current damage is considered. For this reason, the prognostics studies focus on predicting the degradation level and remaining useful life (RUL, remaining cycles before the maintenance).

[1]Daegyeong Division/Aircraft System Technology Group, Korea Institute of Industrial Technology, Yeongcheon, Korea
[2]Department of Aerospace & Mechanical Engineering, Korea Aerospace University, Goyang, Korea
[3]Department of Mechanical & Aerospace Engineering, University of Florida, Gainesville, FL, USA

**Corresponding author:**
Dawn An, Daegyeong Division/Aircraft System Technology Group, Korea Institute of Industrial Technology, 57 Yangho-gil, Yeongcheon 38822, Gyeongbuk, Korea.
Email: dawnan@kitech.re.kr

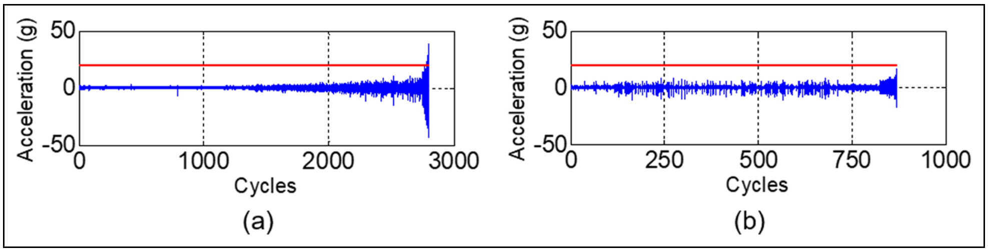


**Figure 1.** Vibration signal from the nominally identical bearing and the same usage conditions: (a) test 1 (Set 1, Condition 1) and (b) test 2 (Set 2, Condition 1).

While studies on bearing diagnostics are well established based on physical phenomena, studies on bearing prognostics depend on the case-by-case application, and no general method can be found even if many efforts have been made to predict bearings' life. There are studies based on a physical model for life prediction of bearings, but they have limitations that uncertainties in service were not considered,[7] or the model is limited to a spall on the outer raceway of a typical roller bearing.[8] Since it is challenging to develop a physical degradation model that includes complicated failure mechanisms, most studies on bearing prognostics rely on extracting degradation features from vibration signals.[9–12] Some of these studies, such as the study of Li et al.,[9] are based on changes in fault-related frequencies, but seeded faults are different from real ones and relatively easy to detect. However, bearing prognostics is difficult when the cause of failure is unknown, especially starting from intact bearings. In such cases, studies focus on finding monotonic degradation features with respect to cycles. Usually a few features distinctly showing monotonic behavior are selected for the purpose of prognostics. For example, Loutas et al.[10] selected the wavelet packet to transform nodal energies and the Wiener entropy based on Spearman's rank correlation coefficient. Kim et al.[11] selected kurtosis, entropy estimation value, and entropy estimation error of time domain signals based on the long-distance evaluation criteria. Sutrisno et al.[12] used the average of the five highest values of absolute acceleration data. Some of the results from previous studies show good RUL prediction results for given vibration signals. However, the vibration signals can significantly be different even from nominally identical bearings under the same usage conditions.

Figure 1 from FEMTO bearing experimental data[13] shows the challenges in predicting bearing failure. The data shown as blue in the two figures are vibration signals measured by accelerometer until failure occurs, and the red horizontal line represents a threshold (20 g). Note that the two quite different signals are obtained under the same usage conditions and from the nominally identical bearings. There is no consistency in the behavior of signals and life span. The life span of test 1 (Figure 1(a)) is around 2800 cycles, but it is 870 cycles for test 2 (Figure 1(b)) even before the signal reaches the threshold.

This article proposes a new method to extract a monotonic degradation feature of bearings from vibration signals having no consistency as shown in Figure 1, which is based on the change in amplitudes at specific frequencies. The nature of the vibration signals changes as the mechanical condition changes, but it is difficult to find any change from raw data that is a result of a combination of system dynamics, damage, and noise signals. Therefore, the signals are decomposed into various frequencies, and those ones that show a change with respect to cycle are selected and analyzed. While most studies have focused on the increase in features, this study pays attention to the decrease in the feature with time. The entropy of each frequency's amplitude is calculated, and specific frequencies showing the decrease in entropy are selected as the degradation feature. Some important attributes are found from the proposed degradation feature, which is used to predict the RUL of bearings. More detailed explanations and procedures are presented in the next section.

FEMTO bearing experimental data[13] are employed to demonstrate the proposed method, whose experimental platform and bearing information are shown in Figure 2 (more detailed explanation of test apparatus can be found in IEEE PHM 2012[14]). Vibration signals are monitored during 0.1 s with 25.6 kHz every 10 s. In this study, the 0.1 s in every 10 s is considered as one cycle, and there are 2560 samples in each cycle. This setting is necessary because otherwise the amount of data would be huge. The information of usage conditions and the number of experimental data are listed in Table 1. Two different usage conditions in terms of the radial force applied to the tested bearing and rotating speed are used in this article. The usage conditions of Conditions 1 and 2 are, respectively, 4 kN and 1800 r/min and 4.2 kN and 1650 r/min. Seven sets of experimental results from each condition are obtained until

**Table 1.** Experimental condition and data usage.

| | Condition 1 | Condition 2 |
|---|---|---|
| Radial force (N) | 4000 | 4200 |
| Rotating speed (r/min) | 1800 | 1650 |
| Number of data set | 7 (Set 1–Set 7) | 7 (Set 1–Set 7) |

failure occurs. Three sets of data from each condition are used as training data to predict RUL of other bearings. Based on the raw vibration signals, it is considered that failure occurs when the magnitude of acceleration reaches 20 g. All raw data of 14 sets with the threshold are shown in Appendix 1.

With the experimental information, this article is organized as follows: in the second section, the method for degradation feature extraction and its attributes are explained; in the third section, prognostics is performed based on the results from the second section; in the fourth section, generalization of the proposed method is discussed; and conclusions and future works with limitations of the proposed method are presented in the final section.

## Degradation feature extraction

As shown in Appendix 1, it is difficult to define common characteristics as a degradation feature from raw data. There is no change in signals even just before the failure in many cases, and the magnitude of the signal at the end of life (EOL) is often less than the threshold value. In this section, a new method is proposed to extract degradation feature from the raw data based on amplitude change at specific frequencies.

### *Amplitude change in frequency domain*

When the vibration signals are decomposed into various frequencies, the amplitudes at some frequencies increase while others decrease as cycle increases. Figure 3(a) and (b) show, respectively, the increasing and decreasing cases of the amplitude at specific frequencies of Set 1, Condition 1 in Figure 1(a). The amplitude in the frequency domain more clearly increases than that in the time domain, which is because the signals in the time domain are the sum of amplitudes of all frequencies including both increasing and decreasing amplitudes. This is more clearly exhibited in the case of Set 2, Condition 1 in Figure 1(b) with Figure 3(c) and (d).

It is expected that the amplitude increases as damage in bearings increases. This is true in the time domain even if it is difficult to observe that at early stages. In the frequency domain, this may be true only for localized defects. In practice, vibration characteristics continue to change as degradation progresses: the natural frequency changes as the stiffness changes due to the fault progression, and the periodic nature of localized defects can diminish since the motion of rolling element becomes irregular and disturbed as damage progresses.[1] Figure 4 shows the change in vibration characteristics in terms of the natural frequency. When a signal changes from blue to red curve with respect to time due to damage, the amplitude at frequency A decreases, while the one at B increases. This explains why amplitude of some frequencies can decrease or increase as cycle increases, as shown in Figure 3.

Based on the vibration characteristics, it seems that both increasing and decreasing amplitudes can be used as a degradation feature. However, there is too large noise to be a degradation feature in the amplitude as shown in Figure 3. In this study, information entropy is employed to reduce the noise and amplify the degradation characteristics (increasing/decreasing trend), which is introduced in the next section.

### *Information entropy as a degradation feature*

Entropy is a measure of disorder and randomness of the system. In physical interpretation, entropy change

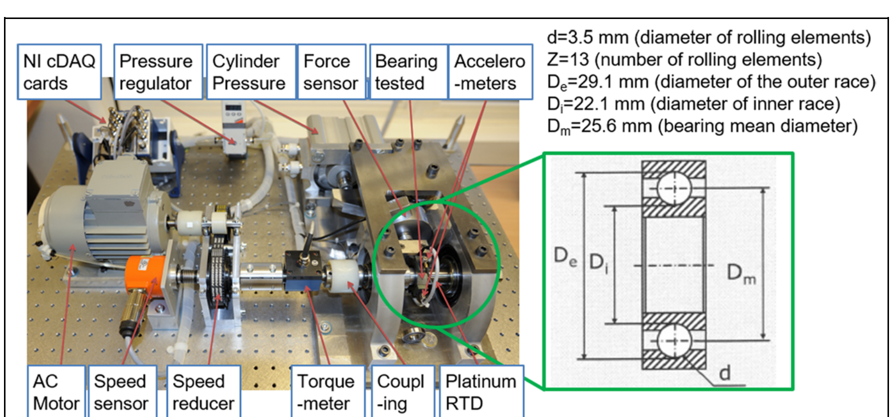


**Figure 2.** Experimentation platform: PRONOSTIA.[13]

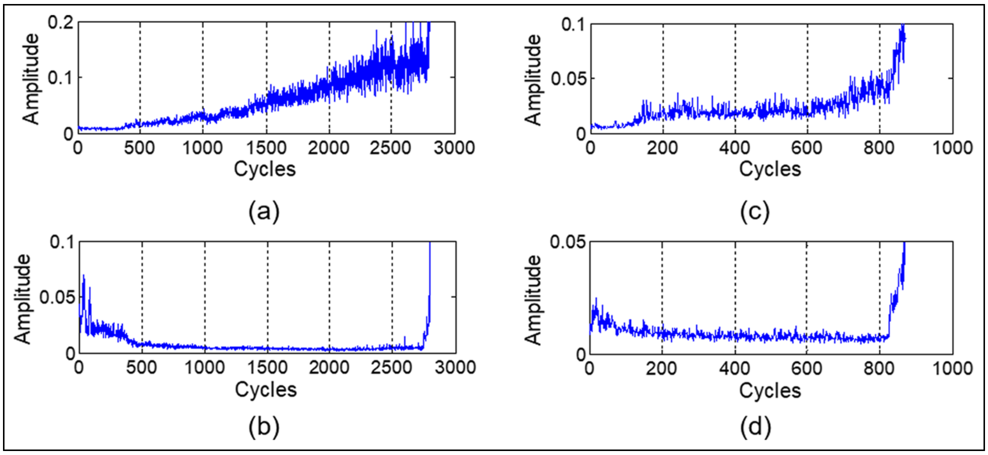


**Figure 3.** Instances of amplitude change in frequency domain: (a) Set 1, Condition 1: frequency #1-1, (b) Set 1, Condition 1: frequency #1-2, (c) Set 2, Condition 1: frequency #2-1, and (d) Set 2, Condition 1: frequency #2-2.

is explained using energy flow. When a system absorbs energy from others, its entropy increases (counterpart decreases). The total entropy of an isolated system, however, never decreases. There is also a mathematical concept of entropy called information entropy or Shannon entropy, which is the average amount of information.[15] In this concept, increase in entropy means an increase in uncertainty by missing the information contained in data. Many researchers argue that physical and information entropy are related to each other,[16,17] but the opposite argument also exists.[18] For the bearing problem, the total energy in bearing increases as the work done by the external force is consumed by damage initiation/propagation,[19] which makes entropy increases.[17] However, the information or data used in this research are the decomposed vibration signals rather than the total thermodynamic energy during the degradation process of bearings because the former fits better. Since it is a debatable topic whether there is a relation between physical and information entropy, information entropy is only considered as a tool to express the changes in vibration signals (Figure 3) rather than to connect physical interpretation with information entropy in this research.

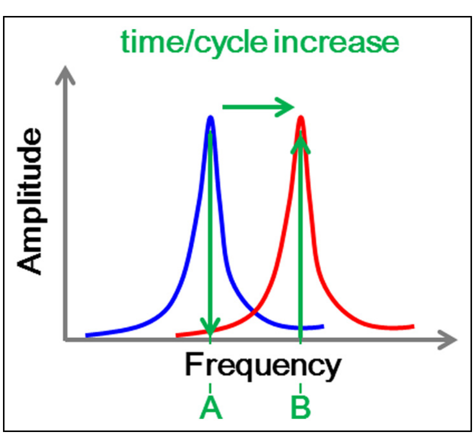


**Figure 4.** Amplitude change at a specific frequency as vibration characteristics changes.

*Information entropy.* In the information theory, entropy is calculated based on the following equation[15]

$$H(X) = -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i) \qquad (1)$$

where $X$ is an information source, $n$ is the number of possible outcomes from $X$, and $p(x_i)$ is the probability of each outcome. In this research, $X$ of bearing problem represents bins that acceleration data belong, which is shown in Figure 5. The range between 0 and 1 is divided into 255 intervals, which are referred to as bins. There are totally 256 bins initially, and the magnitude of data that can be located in each bin is listed at the bottom. Data whose magnitude are less than 1/500 and in between 1/500 and 3/500 are, respectively, included in the first and the second bin, and the same for the last bins (the number of initial bins and the magnitude of data are based on computer usage). After data allocation, the number of bins having non-empty data becomes the number of possible outcomes, $n$; that is, $256 - n$ bins are empty. The probability $p(x_i)$ is the number of data in the $i$th bin divided by the total number of data; this is the same concept as the probability mass function in statistics.

Equation (1) denotes that entropy increases as the number of bins, $n$, increases, and the probability of each bin is even, which corresponds to the case when

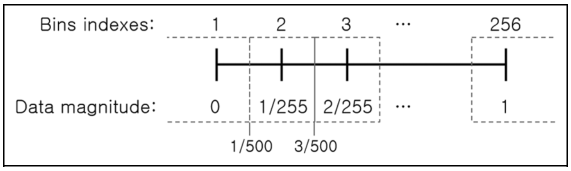


**Figure 5.** Illustration of information source of the bearing problem.

the data within zero to one spread out as illustrated in Figure 6(a). The entropy at $k+1$ cycle is higher than one at $k$ cycle since the amplitudes of data are allocated in more bins with even probability. Figure 6(b) shows the opposite case. The number of bins at $k$ and $k+1$ cycles are the same, but the data at $k+1$ cycle are more located in the middle bins, which makes entropy decrease.

*Entropy calculation results.* The information entropy of amplitude at specific frequencies in Figure 3 is calculated, which is shown in Figure 7. It is shown that entropy calculation has much less noise and more apparent degradation feature than amplitudes in Figure 3. Based on these results, it seems that entropy change of both increasing and decreasing amplitudes can be used as a degradation feature. After examining 14 FEMTO data sets, it is found that the trends of decrease in entropy are consistent for all data sets, and some important attributes can be found. However, no common characteristics are found from the increase in entropy. Even though the entropy clearly increases, its behavior is unpredictable. This may be because increasing amplitudes can appear in the middle of degradation process when the level of damage can be detectable at a specific frequency, whereas the decreasing ones have existed from the beginning and are gradually affected by other frequencies. Thus, decreasing amplitudes are more stable and consistent than increasing ones. In addition, it is found that the frequencies showing a decrease in entropy are around 4 kHz for all data sets, which may depend on the structure of the systems, such as the bearing types and the number of rolling elements. Therefore, the entropy decrease in the frequency domain is used as degradation feature. The detailed procedure to extract the degradation feature based on entropy change in the frequency domain is explained in the following section.

## Procedure of degradation feature extraction

The procedure of the proposed method of extracting the degradation feature is illustrated in Figure 8, whose detailed explanation is as follows:

*Step 1*. Convert signals in the time domain into the frequency domain using fast Fourier transform (FFT)—as mentioned in section "Introduction"—one cycle includes 2560 samples of vibration signals (data acquisition during 0.1 s with a sampling rate of 25.6 kHz), which is converted into the frequency domain using FFT.[20] There are 1401 different FFT results from different cycles (0–1400 cycles) for the example in Figure 8.
*Step 2*. Reshape FFT results frequency-wise (frequency-wise plot)—the amplitude at a fixed frequency (e.g. Frq: 1) changes at different cycles. Therefore, amplitudes at different cycles are collected at a fixed frequency, which is called a frequency-wise plot here. Since FFT results are symmetric with 2560 different frequencies between 0 Hz and 25.6 kHz, there are 1280 frequencies to be considered as candidates for degradation feature, that is, there are 1280 amplitude cycle plots.
*Step 3*. Calculate entropy and select specific frequencies showing entropy decrease—the graphs of frequency-wise amplitude versus cycle from Step 2 are used to calculate entropy using equation (1), which ends up 1280 graphs for entropy versus cycle, and different traces of entropy change in terms of cycles are illustrated in Step 3 in Figure 8. Among them, specific frequencies showing the entropy decrease (like Frq: 1) are selected. If multiple frequencies are decreased among 1280 frequencies, then a median of these entropies is taken as a damage feature.

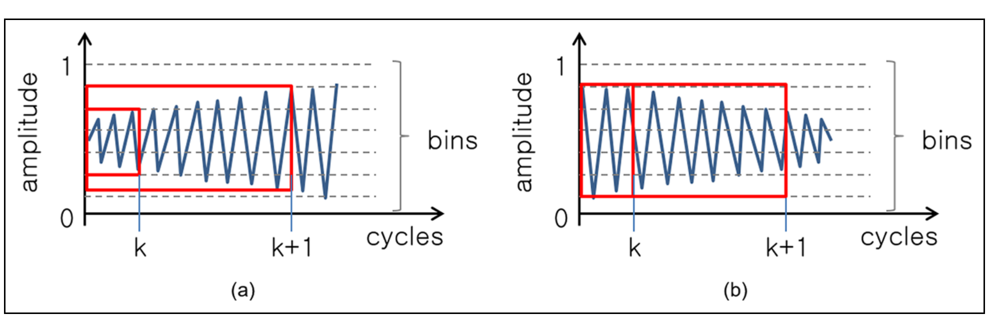


**Figure 6.** Illustration of two cases of entropy change: (a) entropy increase and (b) entropy decrease.

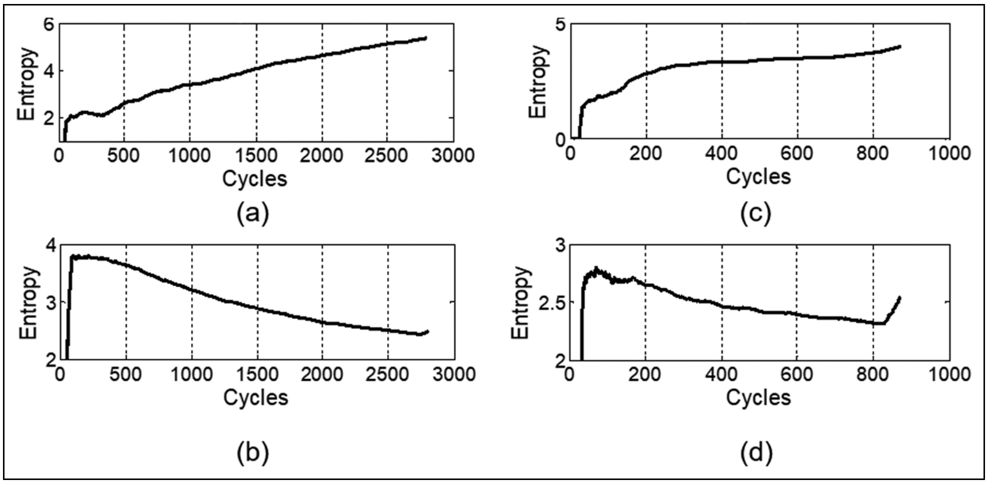


**Figure 7.** Entropy calculation with amplitude in Figure 3: (a) Set 1, Condition 1: frequency #1-1, (b) Set 1, Condition 1: frequency #1-2, (c) Set 2, Condition 1: frequency #2-1, and (d) Set 2, Condition 1: frequency #2-2.

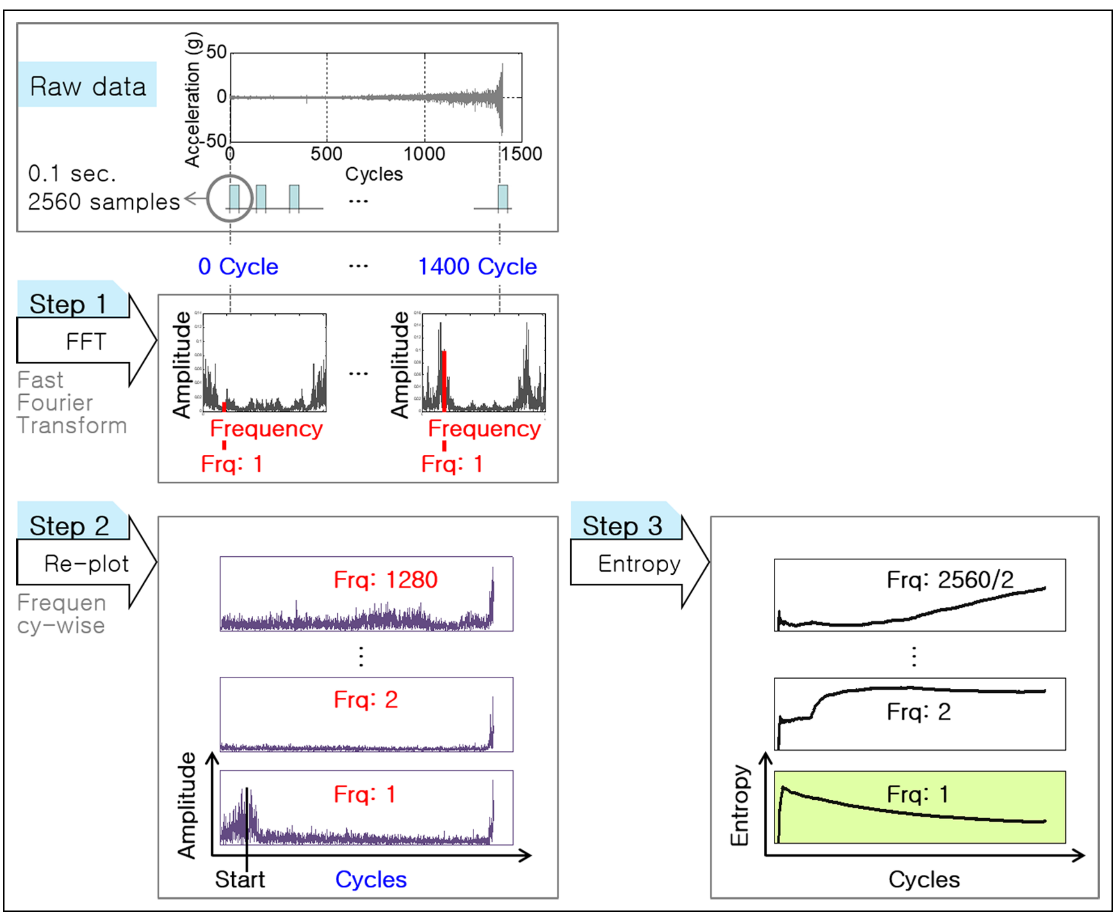


**Figure 8.** Illustration of degradation feature extraction.

There are two points to be considered in the calculation of entropy. First, a starting cycle for entropy calculation needs to be determined based on the amplitude at frequency domain, in order to avoid the initial effect such as the possibly misaligned systems making large vibration, as shown in Frq: 1 at Step 2 of Figure 8. To

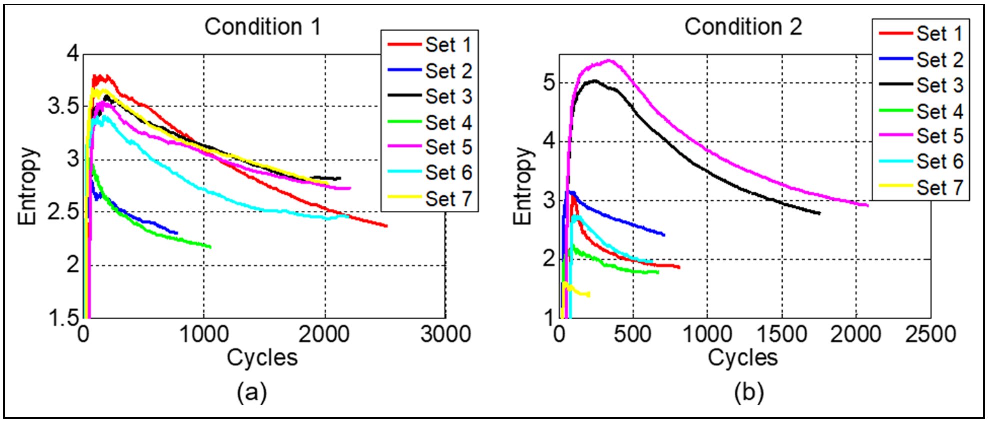


**Figure 9.** Results of the degradation feature: (a) Condition 1 and (b) Condition 2.

this end, a large amplitude or trend changing point at an early cycle is selected as the starting cycle. The starting cycle can differ for different sets, but it is usually located between 12 and 100 cycles. Second, the method of selecting frequencies is based on a magnitude of the slope of linear regression using entropy data. If frequencies whose magnitude of the slope is large are selected, then the degradation feature becomes clear, but at the same time, instabilities increase. In this study, frequencies whose magnitude of entropy slope is at top 25 are selected. The starting cycle and selected frequencies have an effect on calculation results of a particular data set, but overall attributes (will be discussed in section "Results of feature extraction and its attributes") of bearing problem from the proposed methods do not change.

## *Results of feature extraction and its attributes*

Degradation features extracted based on the procedure in Figure 8 are shown in Figure 9, in which entropy of all 14 sets decreases as cycle increases. Each curve is obtained by taking a median of 25 entropy values at each cycle, which are calculated from selected frequencies, showing a consistent decrease in entropy. From the entropy curves, the maximum and minimum entropy and EOL are defined as shown in Figure 10.

There are two main findings when the proposed method is used. First, EOL is proportional to the maximum entropy, which is shown in Figure 11(a). It is possible that the higher energy at the initial stage relates the longer life. The RUL can be predicted by utilizing the linear relation between the maximum entropy and EOL based on training data. Second, when the degradation rate is defined using the maximum and minimum entropy as

$$dr = 1 - \frac{\min.\text{Entropy}}{\max.\text{Entropy}} \tag{2}$$

the degradation rate is classified into two groups, as shown in Figure 11(b). It is possible that the two groups

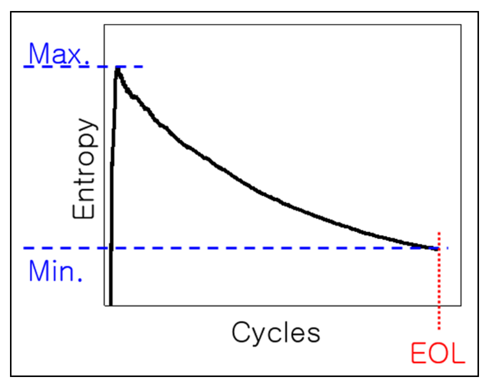


**Figure 10.** Definition of maximum and minimum entropy and EOL.

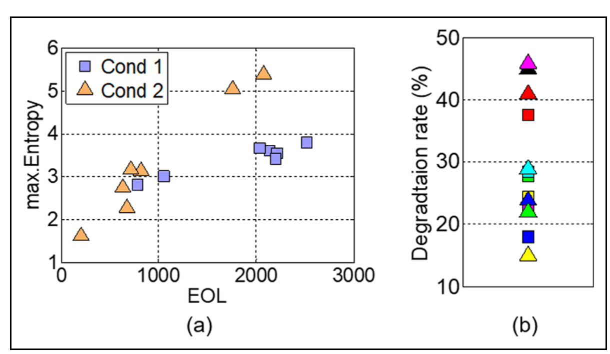


**Figure 11.** Two important attributes from the extracted feature: (a) linear relation between maximum entropy and EOL and (b) two groups of degradation rate.

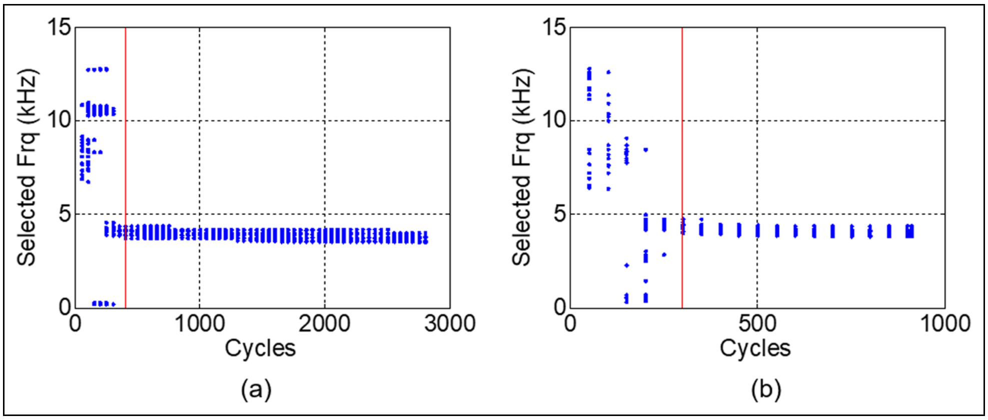


**Figure 12.** Selected frequencies plot as cycle increases: (a) Set 1, Condition 1 and (b) Set 1, Condition 2.

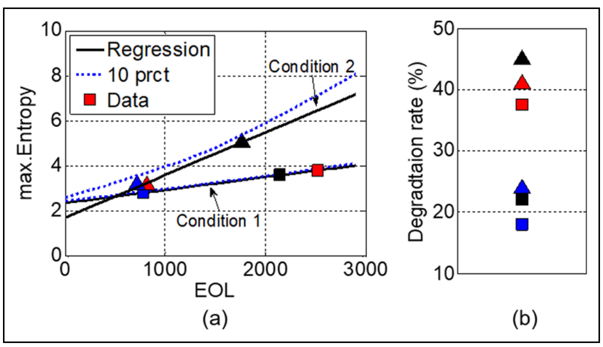


**Figure 13.** Information from three sets of training data: (a) linear relation between maximum entropy and EOL and (b) degradation rate (threshold).

are related to different failure mechanisms, but it cannot be verified with the given data because the cause of failure is unknown. The two groups of the degradation rate are distributed around 20% and 40%, which can be used for a threshold. In the following section, RUL is predicted in the two ways based on the two findings, one using the linear relation and the other the degradation rate with the entropy trend.

## Prognosis

Even though the vibration signals until the failure occurs (true EOL) are given in Appendix 1, it is assumed that EOL is 90% of the true EOL for the purpose of maintenance scheduling. The procedure illustrated in Figure 8 is repeated every 50 cycles to select featured frequencies. Even though the selected frequencies can be different at each cycle, they gradually become consistent and converge to around 4 kHz as cycles increase, as shown in Figure 12. Therefore, RUL prediction is performed after the cycle that the selected frequencies are converged, as shown in the red vertical line in the figure. RUL is predicted in two ways based on (1) the linear relation between maximum entropy and EOL and (2) entropy trend with a threshold.

### *Max.E-EOL method: the relation between maximum entropy and EOL*

As mentioned before, three sets of data (Sets 1, 2, and 3) from each condition are used as the training data. These sets are utilized to construct the linear relation between maximum entropy and EOL, which is shown in Figure 13(a). In the figure, three different colors represent different sets (red: Set 1, blue: Set 2, and black: Set 3), and square and triangle markers, respectively, represent Conditions 1 and 2. The black solid and blue dashed lines are, respectively, mean of regression results and 10-percentile lower confidence bound using three data from each condition, and the latter is used for a conservative prediction. For example, when the maximum entropy is obtained as 4 at the current 700 cycles, the EOL and the RUL are predicted as 1000 and 300 cycles, respectively.

### *E.trend method: entropy trend with threshold*

The future behavior of entropy is predicted based on the following model

$$\text{Entropy} = \beta_1 \exp\left(\beta_2 \text{Cycle}^{\beta_3}\right) \quad (3)$$

whose expression can follow the entropy trend in Figure 9. Three unknown parameters, $\beta_1, \beta_2, \beta_3$ in the equation, are identified based on nonlinear regression from the data between the maximum entropy and cycles. The RUL is predicted by extrapolating the model with identified parameters until it reaches the threshold that is determined from the degradation rate from three sets of data (Sets 1, 2, and 3) from each

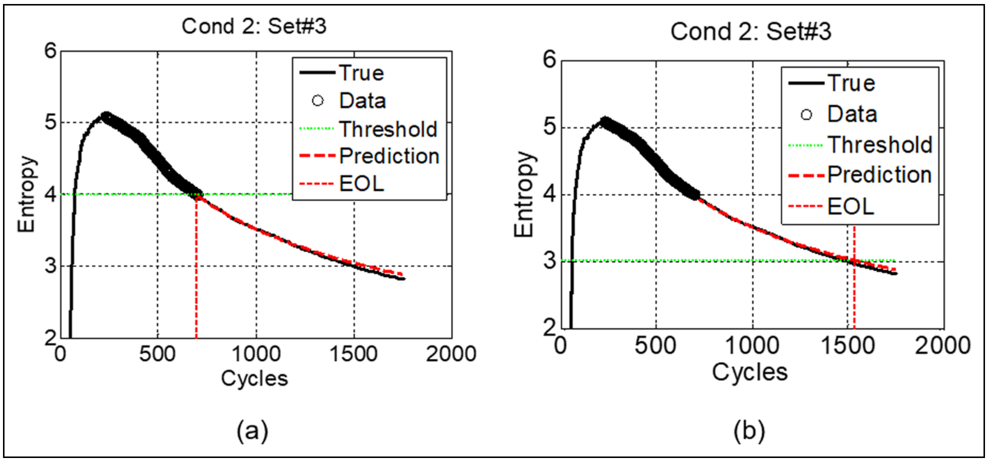


**Figure 14.** Results of degradation prediction: (a) threshold 21% and (b) threshold 41%.

condition, as shown in Figure 13(b). Degradation rates of total six training data are classified into two groups, and the mean of each group is considered as a threshold: that is, 21% and 41%.

It can be determined whether the current data belongs to 21% or 41% based on the linear regression model, as shown in Figure 13(a). As an illustration, the prediction result of future entropy trend of Set 3, Condition 2 is shown in Figure 14. In Figure 14(a), black solid curve, black circles (looks like thick black curve), and green dotted horizontal line are, respectively, true trend up to EOL, used data to identify $\beta_1, \beta_2, \beta_3$ in equation (3) and the threshold calculated from 21% degradation rate. The red dashed curve from the current cycle (700 cycles) is a predicted entropy trend based on equation (3) with identified parameters ($\beta_1 = 6.15\times10^6$, $\beta_2 = -12.14$, $\beta_3 = 0.0244$). The intersection of the predicted entropy trend and the threshold is a predicted EOL, which is shown as the red dashed vertical line. This result shows RUL is $-9$ cycles using 21% threshold (i.e. a total of 691 lifecycles), which does not make sense and the result can be changed when the threshold is estimated using the linear regression in Figure 13(a). The maximum entropy is 5.08 from Figure 14, which corresponds to 1780 cycles EOL from the mean of the regression model (max.Entropy = $1.70 + 0.0019 \cdot$ EOL; since this is just for threshold classification, the solid line is used) in Figure 13(a). Compared to this result, 691 cycles is too short as the EOL. The new threshold is estimated as 43% by considering minimum entropy at 1780 cycles, which is obtained by extrapolating entropy trend (red dashed curve in Figure 14(a)) to 1780 cycles. Since the newly estimated threshold is close to the threshold in the other group (41%) from the training data, the RUL is re-predicted using 41% threshold, which is shown in Figure 14(b). With 41% threshold, the RUL is predicted as 833 cycles whose error with the true RUL (1059 cycles) is 0.2137, which is calculated by dividing the true RUL minus the predicted one by the true one.

## *RUL prediction results*

The results of RUL prediction for all prediction sets are shown in Figure 15, in which black solid lines are true RUL, the green vertical line indicates the cycles when the frequencies converge, and blue dotted lines and red dashed lines are prediction results based on Max.E-EOL and E.trend methods, respectively. The results from the Max.E-EOL method of Condition 1 are closer to true RUL than ones from the E.trend method, while the Max.E-EOL method does not perform well for Condition 2 since EOL of Condition 2 is short and maximum entropy is small, but 10th percentile at small entropy gives very conservative prediction (see in Figure 13(a)).

It is considered that the RUL results can be reliable after selected frequencies are converged (the green vertical lines), however, the RUL prediction of some cases change from negative value to positive; for example, E.trend result of Set 5 in Figure 15(a) is negative at 1500 cycles, but it becomes positive at 1800 cycles. Since RUL at future cycles is unknown, it is assumed that when RUL is less than 50 cycles, maintenance is ordered. In this case, the used life is calculated at the cycle when maintenance is ordered. The ratio of used life to the EOL is listed in Table 2; the higher the value is, the better the prediction is. However, it is failed to predict the RUL of Set 7, Condition 2, in which failure occurs before the selected frequencies are converged. By considering the EOL of this set is very short, it seems that there was a significant initial defect in this bearing. Except for this bearing, the mean of the ratio from seven results is calculated. The mean of the ratio

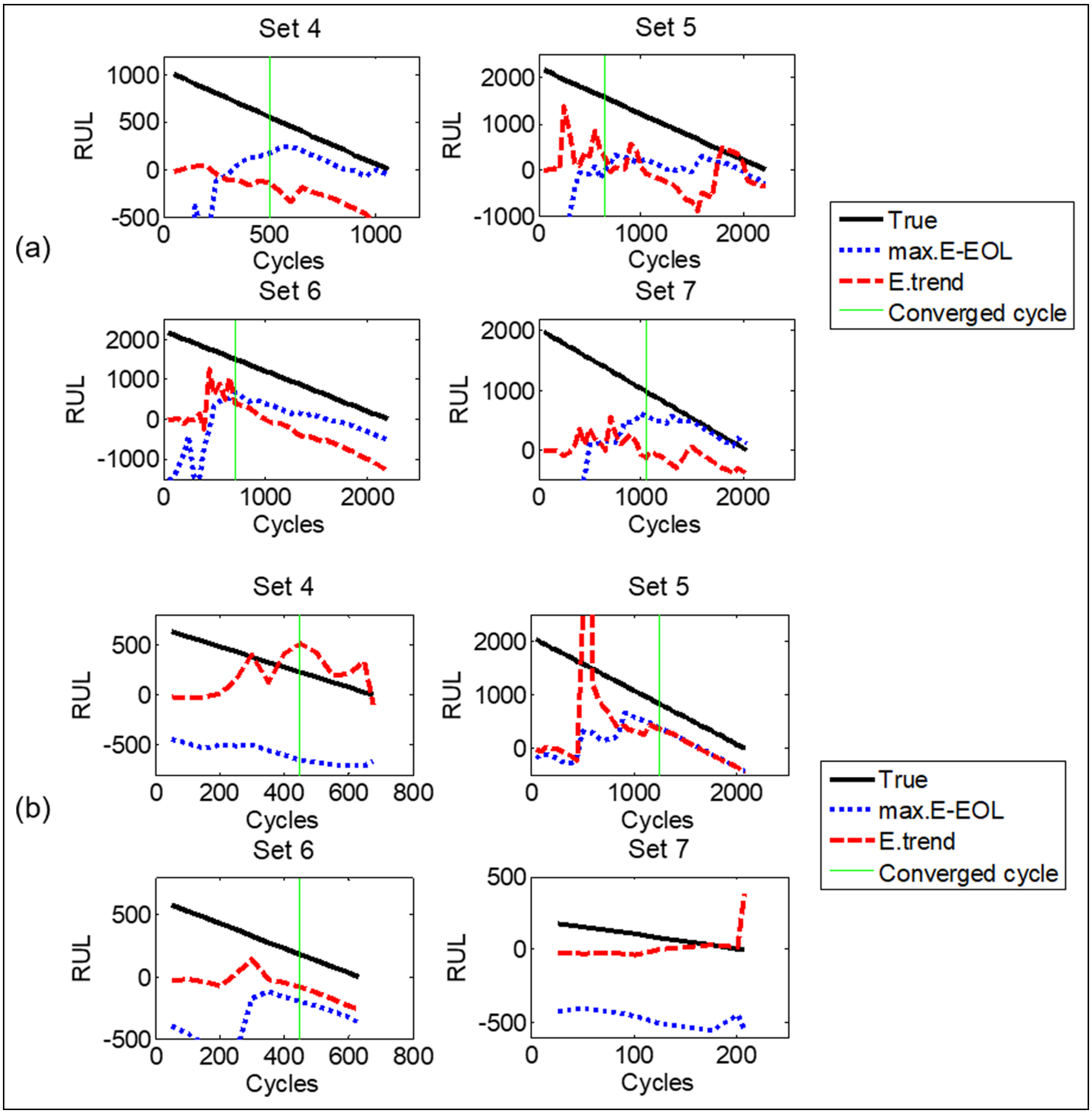


**Figure 15.** RUL prediction using three training data: (a) Condition 1 and (b) Condition 2.

of used life is 0.56 by considering conservative results between the Max.E-EOL and the E.trend methods (0.47, 0.29, 0.45, 0.52 from Condition 1 and 0.67, 0.78, 0.71 from Condition 2). However, when the optimistic results are selected by ignoring conservative results and going to next prediction steps, the mean is 0.78 (0.81, 0.32, 0.72, 1.08 from Condition 1 and 1.04, 0.79, 0.71 from Condition 2). That is, bearings can be used 56% or 78% of their whole life in average. The choice between the Max.E-EOL and the E.trend methods can be made depending on the trade-offs between maintenance cost and risk.

Finally, the prediction results from different combinations of three training data sets are provided to validate the proposed method, which is listed in Table 3. There are totally 10 cases by adding 9 additional cases among total 35 possible combinations (selecting 3 out of 7). In each case, three training sets are randomly selected, but the combinations that three data do not show a proportional relation between maximum entropy and EOL are excluded. From the three sets from each condition, threshold levels and the linear relation between maximum entropy and EOL (see Figure 13) are determined and utilized to predict RUL of remaining four bearings in each condition, which is repeated for all cases.

The mean of conservative and optimistic results from each case are listed in Table 3, which is calculated in the same manner as in Table 2. The statistical results (three percentiles, minimum and maximum values) from the 10 cases are also calculated. Based on these results, 56%–64% and 71%–80% of bearings' life from

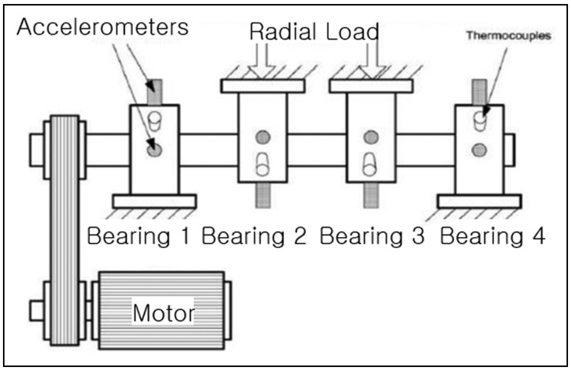


**Figure 16.** Illustration of IMS bearing.[21]

**Table 2.** Ratio of used life to EOL using three training data.

| | | Set 4 | Set 5 | Set 6 | Set 7 |
|---|---|---|---|---|---|
| Condition 1 | Max.E-EOL | 0.81 | 0.29 | 0.72 | 1.08 |
| | E.trend | 0.47 | 0.32 | 0.45 | 0.52 |
| Condition 2 | Max.E-EOL | 0.67 | 0.79 | 0.71 | fail |
| | E.trend | 1.04 | 0.78 | 0.71 | fail |
| Conservative mean | | 0.56 | | | |
| Optimistic mean | | 0.78 | | | |

EOL: end of life.

the conservative way and optimistic way, respectively, can be used in average. However, note that there are several sets that are failed to be predicted by relying on the optimistic results, and the mean results in Table 3 are obtained by excluding those ones (the numbers in the parenthesis denote the total number of bearings to calculate the mean values). It is required to be very careful when relying on the optimistic results with high safety systems. Also, the proposed method cannot predict the RUL for one set in Case 10.

## Discussions on generality of the proposed method

The main findings from the proposed method are that the entropy gradually decreases to a certain level of thresholds, and the EOL is proportional to the maximum entropy. If these attributes are also found in the other application, the proposed method might be widely applicable. In addition, if the effect of usage conditions on these attributes is found, it is expected that more accurate prediction is possible with less number of training data. In this section, these three issues are discussed.

### *Another bearing application*

Additional experimental results from another bearing application are employed to validate the attributes of the proposed method, which are provided by the NSF I/UCR Center for Intelligent Maintenance Systems (IMS).[21] Four double row bearings (16 rollers) are installed on a shaft, and the rotating speed and radial load are, respectively, 2000 r/min and 6000 lbs, which is illustrated in Figure 16. The test stops when the accumulated debris exceeds a certain level, and there are three sets of experimental data available. A run-and-stop operation is repeated for experimental Set 1 and Set 3 (Set 2 is continuously monitored to EOL). Since the number of data set is too small, the run-and-stop of the operation has an effect on vibration signal and entropy calculation, and four bearings on a shaft can interact with each other in the fault progression; it is difficult to make any definite conclusions. Therefore, it is observed how consistent results are obtained compared to the previous case with the same method.

Threshold, EOL, and maximum entropy of the failed bearings in each set are listed in Table 4. The maximum

**Table 3.** Ratio of used life to true EOL using different combinations of three training data.

| | Case 1 | Case 2 | Case 3 | Case 4 | Case 5 |
|---|---|---|---|---|---|
| Training set # | [1 2 3] | [3 4 5] | [1 3 4] | [2 5 7] | [1 3 7] |
| Conservative mean | 0.56 (7) | 0.67 (7) | 0.60 (7) | 0.58 (8) | 0.54 (8) |
| Optimistic mean | 0.78 (7) | 0.71 (5) | 0.72 (7) | 0.71 (7) | 0.68 (8) |
| Fail to be predicted with the optimistic result | | two sets | | one set | |
| | Case 6 | Case 7 | Case 8 | Case 9 | Case 10[a] |
| Training set # | [1 2 5] | [2 3 5] | [1 4 5] | [2 6 7] | [4 5 6] |
| Conservative mean | 0.63 (7) | 0.56 (7) | 0.70 (7) | 0.45 (8) | 0.64 (6) |
| Optimistic mean | 0.87 (7) | 0.84 (6) | 0.75 (6) | 0.55 (6) | 0.80 (3) |
| Fail to be predicted with the optimistic result | | one set | one set | two sets | three sets |
| | Min. | 25th percentile | median | 75th percentile | Max. |
| Conservative mean | 0.45 | 0.56 | 0.59 | 0.64 | 0.70 |
| Optimistic mean | 0.55 | 0.71 | 0.74 | 0.80 | 0.87 |

EOL: end of life.
[a]One set fails to be predicted with the proposed method.

**Table 4.** Failure summary of three data sets.

| Data set | Minimum Max.E | Failed bearing | Failed element | Threshold | EOL (cycle) | Max.E |
|---|---|---|---|---|---|---|
| Set 1 | Bearing 4 | Bearing 3 | Inner race | 48% | 1940 | 1.37 |
| | | Bearing 4 | Roller element | 65% | 1940 | 1.00 |
| Set 2 | Bearing 1, Bearing 4 | Bearing 1 | Outer race | 73% | 886 | 0.81 |
| Set 3 | Bearing 3 | Bearing 3 | Outer race | 71% | 4003 | 0.68 |

EOL: end of life.

entropy of Set 3 is too small compared to its EOL. However, it is still early to conclude that the EOL is not proportional to the maximum entropy since three sets of data are very small, and there were several run-and-stop during the operation. At least, failure occurs at bearings having the lowest maximum entropy among four bearings in each set (see the second and third columns in Table 4). Also, the four thresholds seem to be grouped into two: 48% is one of them, and 65%, 73%, and 71% are another one. The difference in magnitude between the first group and the mean of the second group is similar to that in FEMTO bearing (20% and 40%). Finally, two of the three failed elements occur at the outer race in the second threshold group (see the fourth and fifth columns in Table 4), which is unknown for the FEMTO bearing. This slightly increases the possibility that the different groups might be related to different failure mechanisms.

### *The relation between threshold/Max.E-EOL and usage conditions*

In the previous section, the threshold of the IMS bearing is around 50%–70%, which differs from one of the FEMTO bearings, 20%–40%. The cycle of the FEMTO bearing is based on seconds, and the EOL is around 2–7 hours, while the EOL of IMS bearing is around 7–30 days. That means, FEMTO bearings are under the accelerated test conditions, while IMS bearings are under the nominal operating conditions. Therefore, it seems that there is a relationship between the degradation rate (threshold) and the applied load, which cannot be found in the research due to the lack of test data. However, if this relationship can be established, then it can help set a threshold with a small number of training data.

As another effect of usage conditions, if it is related to the slope of the relationship between maximum entropy and EOL, more information can be used. It seems like the magnitude of the slope is proportional to the magnitude of the load based on Condition 1 and Condition 2 in Figure 13(a). If a relation between the magnitude of slope and usage conditions are found, it can help decision making to utilize the relation between maximum entropy and EOL for prediction. However, two conditions are not enough to validate the relation. This will be considered in near future with more data sets under various usage conditions.

## Conclusion and future works

A new method is proposed based on entropy changes at specific frequencies to extract the degradation feature from vibration signals and to predict the RUL of bearing applications. The main contributions and attributes of the proposed method are as follows:

- Degradation feature having a generality is found from vibration signals, which gradually decreases as cycle increases for all experimental sets.
- Degradation rates of different experimental sets from the same application are similar each other, which is used as a threshold.
- EOL is proportional to the maximum entropy, which can be used as another prediction method without a threshold.

The proposed method is demonstrated using 14 sets of bearing experimental data under two different conditions. By considering used life, 59%–74% of bearings' life can be used based on the proposed method.

The results from the proposed method are noticeable by considering the current level of study, but there are still several limitations to be solved. First, the entropy trend exponentially decreases, which can make a large difference in life prediction with a small perturbation of threshold. Second, the proposed method is based on the accumulated vibration data. Since the bearing systems under real operating conditions may last very long time, a tremendous amount of data should be stored. Finally, the proposed method has been developed based on the vibration signal. Even if the basis of the found attributes is tried to be explained based on physical interpretation, they are not proved yet, such as why the increasing amplitude is not proper to be a degradation feature, why the selected frequencies are at around 4 kHz, why EOL is proportional to the maximum entropy, and why there are different groups of the degradation rate.

As the future works, the proposed method will be improved by resolving the aforementioned limitations. Also, a generality of the method will be demonstrated by studying the effect of usage conditions on the level of threshold and the slope of Max.E-EOL. Even though the proposed method has not been demonstrated using the other bearing applications yet, the general usage of this method is promising by judging the results from 14 sets of bearing data, and the results in another bearing application show the possibility.

## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Korea Agency for Infrastructure technology Advancement (KAIA) in 2017 (17RTRP-B104370-04-000000).

## References

1. Haward I. *A review of rolling element bearing vibration "detection, diagnosis and prognosis."* Report no. DSTO-RR-0013, October 1994. Canberra, ACT, Australia: Defense Science and Technology Organization.
2. Carden EP and Fanning P. Vibration based condition monitoring: a review. *Struct Health Monit* 2004; 3(4): 355–377.
3. Cui H, Qiao Y, Yin Y, et al. An investigation on early bearing fault diagnosis based on wavelet transform and sparse component analysis. *Struct Health Monit* 2017; 16(1): 39–49.
4. Randall RB and Antoni J. Rolling element bearing diagnostics—a tutorial. *Mech Syst Signal Pr* 2011; 25: 485–520.
5. El-Thalji I and Jantunen E. A summary of fault modelling and predictive health monitoring of rolling element bearings. *Mech Syst Signal Pr* 2015; 60–61: 252–272.
6. Reuben LCK and Mba D. Bearing time-to-failure estimation using spectral analysis features. *Struct Health Monit* 2014; 13(2): 219–230.
7. Yu WK and Harris TA. A new stress-based fatigue life model for ball bearings. *Tribol T* 2001; 44(1): 11–18.
8. Bolander N, Qiu H, Eklund N, et al. Physics-based remaining useful life prediction for aircraft engine bearing prognosis. In: *Proceedings of the annual conference of the prognostics and health management society*, San Diego, CA, 27 September−1 October 2009.
9. Li R, Sopon P and He D. Fault features extraction for bearing prognostics. *J Intell Manuf* 2012; 23: 313–321.
10. Loutas TH, Roulias D and Georgoulas G. Remaining useful life estimation in rolling bearings utilizing data-driven probabilistic E-support vectors regression. *IEEE T Reliab* 2013; 62(4): 821–832.
11. Kim HE, Tan A, Mathew J, et al. Bearing fault prognosis based on health state probability estimation. *Expert Syst Appl* 2012; 39: 5200–5213.
12. Sutrisno E, Oh H, Vasan ASS, et al. Estimation of remaining useful life of ball bearings using data driven methodologies. In: *Proceedings of the IEEE international conference on prognostics and health management*, Denver, CO, 18−21 June 2012.
13. Nectoux P, Gouriveau R, Medjaher K, et al. PRONOSTIA: an experimental platform for bearings accelerated degradation test. In: *Proceedings of the IEEE international conference on prognostics and health management*, Denver, CO, 18−21 June 2012.
14. IEEE PHM 2012 prognostic challenge: outline, experiments, scoring of results, winners, http://www.femto-st.fr/f/d/IEEEPHM2012-Challenge-Details.pdf
15. Shannon CE. A mathematical theory of communication. *Bell Syst Tech J* 1948; 27: 379–423.
16. Brillouin L. *Science and information theory*. Mineola, NY: Dover Publications, 1956.
17. Bao T, Peng Y, Cong P, et al. Analysis of crack propagation in concrete structures with structural information entropy. *Sci China Technol Sc* 2010; 53(7): 1943–1948.
18. Frigg R and Werndl C. Entropy—a guide for the perplexed. In: Beisbart C and Hartmann S (eds) *Probabilities in physics*. Oxford: Oxford University Press, 2010, pp. 115–142.
19. Wang Z and Nakamura T. Simulations of crack propagation in elastic–plastic graded materials. *Mech Mater* 2004; 36: 601–622.
20. Walker JS. *Fast Fourier transforms*. 2nd ed. Boca Raton, FL: CRC Press, 1996.
21. Lee J, Qiu H, Yu G, et al.; Rexnord Technical Services. IMS, University of Cincinnati. Bearing data set. NASA Ames Prognostics Data Repository. Moffett Field, CA: NASA Ames Research Center, 2007, http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/

## Appendix 1

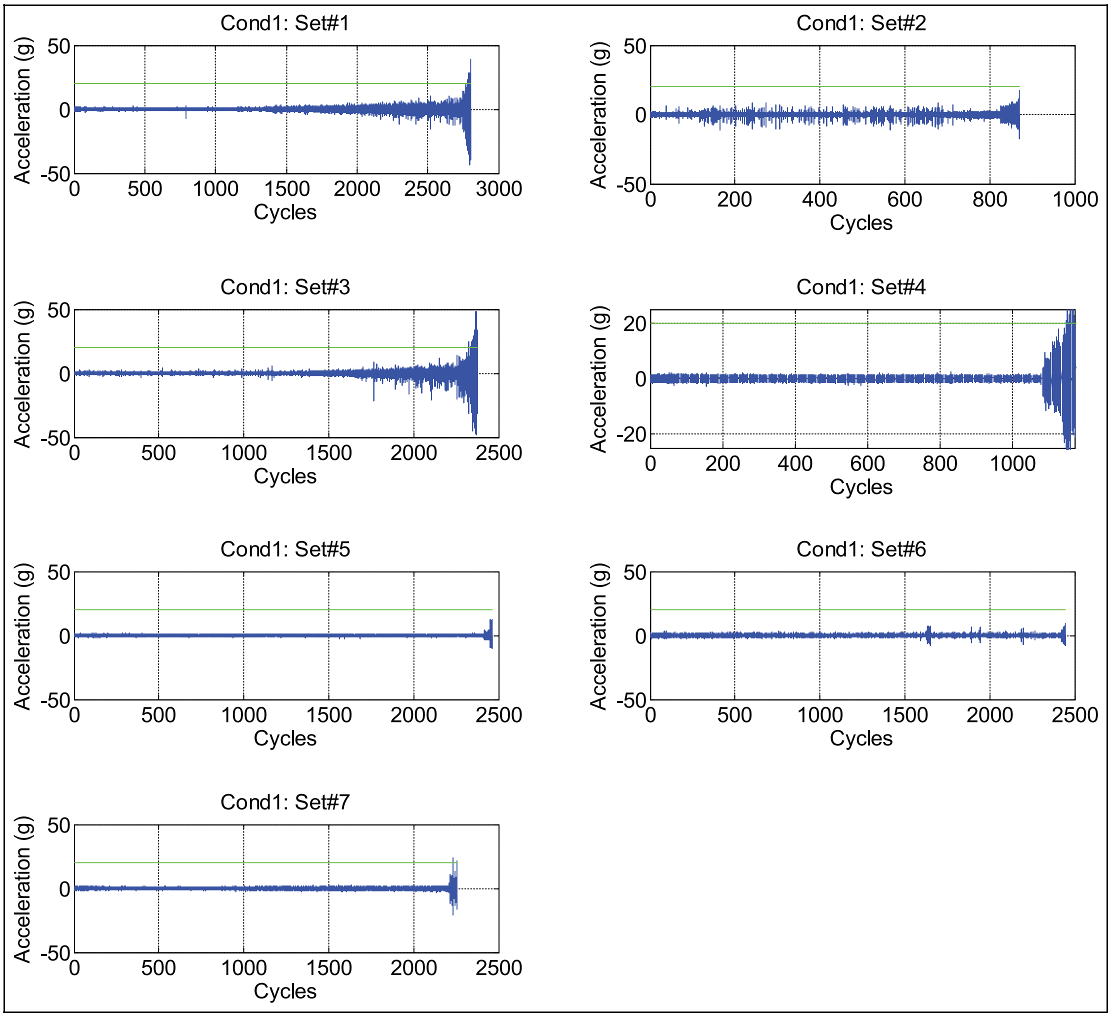


**Figure 17.** Raw data from horizontal axis: Condition 1.

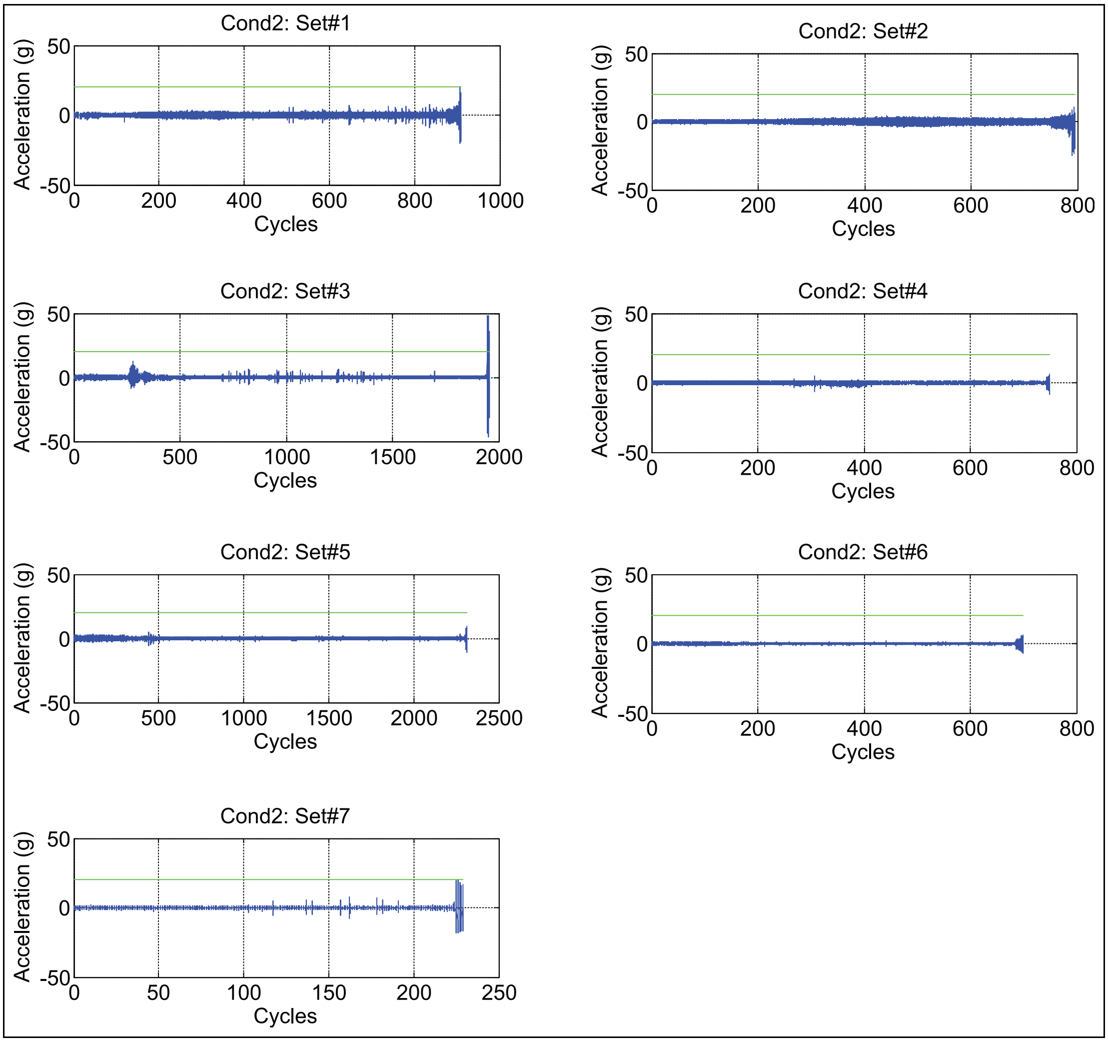


**Figure 18.** Raw data from horizontal axis: Condition 2.