EXPERIMENTAL STUDIES FOR BEARINGS DEGRADATION MONITORING AT AN EARLY STAGE USING ANALYSIS OF VARIANCE

Djamal ZAROUR 1,2, Salim MEZIANI 1, Marc THOMAS 2

1 Laboratoire de Mécanique, Université des Frères Mentouri – Constantine 1, 25000, Constantine, Algeria, djzarour@gmail.com, salim.meziani@labomecanique-umc.org
2 Department of Mechanical Engineering, École de Technologie Supérieure, 1100, Notre-Dame street West, Montreal, H3C 1K3, Quebec, Canada, marc.thomas@etsmtl.ca

Abstract

This work presents a procedure for bearing degradation monitoring at an early stage. The analysis of variance (ANOVA) coupled with Tukey’s test is used to single out the suitable parameters to follow the fault size evolution ranging from 50 µm to 150µm. The Tukey’s criterion is adopted in this case to study the ability of time and frequency indicators. The rotational speed, centrifugal load and fault size are considered as independent variables while the time and frequency indicators are taken as dependent variables. The experiments are performed on bearings having a fault on outer race. Based on the results of this study, the Kurtosis and Skewness show a good ability to assess the evolution of degradation in the bearings at an early stage. The paper discusses the weakness of the time and frequency indicators.

Keywords: ANOVA, DOE, Time descriptors, Bearings fault.

1. INTRODUCTION

When a fault appears on one of the bearing elements (inner ring, outer ring, roller element or cage), it is necessary to follow its evolution in order to proceed for preventive maintenance and anticipate breakage and avoid production shutdown. Statistics of maintenance service, in the petrochemical industries, show that 52% of rotating machinery failures come from bearing defect. Sometimes, the bearing life given by manufacturers is incorrect and misleading. In fact, it is estimated that only 30% of bearings lifetime of aging caused by the surface fatigue [1]. Consequently, the development of methods for detection bearing fault at an early stage is essential for the prevention of rotating machines. Currently, the most used tools of bearing monitoring are: temperature measurements, acoustic emission, oil analysis (analysis of debris wear) and vibration analysis [1]. The latter is widely used for bearing fault monitoring. I. Bazovsky [2] introduced the use of mathematical optimization method in the philosophy of preventive maintenance. Various signal processing techniques involving time, frequency, and statistical methods have been used to detect and check the progress of the incipient fault [3]. The extraction of meaningful information from these data is always challenging especially due to the presence of noise that masks the interesting information and, therefore, it calls for different approaches to analyze the data. N. Tandon and A. Choudhury studied the application of vibration analysis and acoustic emission in the detection of bearing faults [4]. There are several methods used for bearings monitoring, some of them are easy to apply while others are based on signal processing [5]. Shocks are usually created in the presence of faults and can be analysed either in the time domain [6] (RMS and max-peak amplitude of vibration level, Crest factor and Kurtosis, detection of shock waves method [7], statistical parameters applied to the time signal, Cepstrum [8], etc.); or in the frequency domain (spectral analysis around bearing defect frequencies [9], Spike energy [10], high frequency demodulation [11], Empirical modal decomposition [12], acoustic emission [13], cyclostationnarity [14], time-frequency [15, 16], Fast Fourier Transform [17], Wavelet [18], Kurtogram [19], artificial neural networks [20], etc). Statistical methods are used to assess structural health [21]. The design of experiments (DOE) has been used widely in the field of tribology. The effect of cutting parameters on surface roughness and flank wear was analysed using this method [22]. Other works had exploited the same method for studying the tribological behavior of composites Cu/silica in the presence of the effect of solid braking load, sliding speed and lubricants.

The present work uses the application of method of analysis of variance (ANOVA) alongside Tukey's multiple comparison test to select the good indicator for bearing degradation at an early stage. One-way Analysis of Variance (ANOVA) is used to study the efficacy of time and frequency indicators to follow the evolution of fault size at an early stage ranking from 50 µm to 150 µm. Tim
2. OVERVIEW OF THE ANALYSIS OF VARIANCE (ANOVA) AND TUKEY’S MULTIPLE COMPARISON TESTS

In the first time, we conducted experimental tests using a full plan. Then, all time and frequency indicators were calculated. All data are organized in Table 5. Finally, an analysis of variance coupled with the multiple comparisons using Tukey’s test was run. In the following, the application of this procedure is explained.

\[ H_0: \mu_1 = \mu_2 = \ldots = \mu_k \quad \text{and all } \mu_i \text{ are the same} \]

Where: \( H_0 \) and \( H_1 \) are the null and alternative hypothesis, respectively. The significance level of the test in this paper is set to \( \alpha = 0.05 \). For a completely randomized multi-factor experiment having \( k \) groups and \( l \) trials in each group. The total number of measurements is:

\[ M_g = \mu + \varepsilon_g \quad i = 1,2,\ldots,k \text{ and } j = 1,2,\ldots,l \]

Here, \( M \) is termed the root mean square deviation (RMSD) and \( \varepsilon \) is the deviation from the mean of each group. In the analysis of variance ANOVA, the total variability can result either from the variability between different groups (between-groups) or from the variability within each group (within-the-group). The mean square error within groups \( S_{wg} \) can be given by:

\[ S_{wg} = \frac{\sum_{i=1}^{K} \sum_{j=1}^{l} (M_g - \mu)^2}{N - k} \]

While the mean square error between groups \( S_{bg} \) is given by:

\[ S_{bg} = \frac{\sum_{i=1}^{l} l_i (\mu_i - \mu)^2}{k - 1} \]

Where:

\[ \mu = \left( \frac{1}{N} \right) \sum_{i=1}^{K} \sum_{j=1}^{l} M_g , \text{ is the mean of all the measured results. Once the total variability is decomposed (within + between), we compare these two components. The scalar value used to assess this is the F-ratio given by:} \]

\[ F-ratio = \frac{S_{bg}}{S_{wg}} \]

After obtaining the F-ratio, the probability value \( p-value \) is obtained through a comparison with the F-distribution \( F(\alpha;k-1,N-k) \). If some of the mean values are different, the numerator in Eq. (5) tends to be larger. If \( F > F(\alpha;k-1,N-k) \), then \( H_0 \) can be accepted. However, this is not sufficient to reject the \( H_0 \) hypothesis. In case of \( p-value \) smaller than alpha the null hypothesis should be rejected, otherwise, there is no statistical significance to reject the null hypothesis.

Furthermore, we need to determine which pairs are significantly different. In this regard, the Tukey’s test provides the ability to perform multiple comparisons between several pairs of groups to determine significant differences between the means [24]. This latter is coupled with the ANOVA method. It is worthy to be noted that there is a test similar to that of Tukey called t-test; however this latter test can only compare the means between two groups. In our case, the Tukey test is preferred because it compares the means of a set of pairs of groups. Generally, the confidence interval is calculated by:

\[ S_{wg} \alpha \quad \sqrt{\frac{(\alpha;k-1,N-k) \cdot S_{wg}}{l}} \]

This interval is used to quantify the effectiveness of an estimator. The difference between two means is computed using the following equation:

\[ (\mu_1 - \mu_2) \pm q(\alpha;k-1,N-k) \cdot \sqrt{\frac{S_{wg}}{l}} \]

Where \( q(\alpha;k,N-k) \) is the \( 100 \times (1-\alpha) \) percentage of the considered ranges distribution for comparing \( k \) mean values. The major disadvantage of this method is that it is valid only for samples in equal size. Using only the ANOVA method cannot provide any information on significant differences. To circumvent this obstacle, we need to couple the ANOVA with the Tukey test. The introduction of the Tukey test using a specific confidence interval allows the determination of the means that are statistically difference.

3. MATERIALS AND METHODS

The test bench is composed of a shaft driven by an electric motor (Fig.1.A) with speed controlled by a drive controller. The shaft is guided in rotation by two bearings and connected to the motor by a flanged coupling bolted rubber. A disc weight of 4.7 kg can be maintained and can be loaded with unbalance mass (which provides a rotating load and thus a centrifugal force).

The bearings used are "ball bearings cylindrical and tapered bore" SKF brand and 1210 EKTN9 model. The bearings have induced fault (groove) of different sizes on the outer race. Since the objective of this study is the early detection of fault grooves with very small width were artificially created by electro-erosion and measured with a microscope. Three defective bearings with groove widths of 50, 100 and 150 micrometers were used. The frequencies [1] symptomatic of the fault of the bearings are shown in (Table 1).
Zarour D, Meziani S, Thomas M: Experimental studies for bearings degradation monitoring...

The acquisition chain is shown in Fig 2. It is composed of piezoelectric sensors (352C34) with a sensitivity of 100 mV/g for the measurement of vibrations. Sensor is connected to an analogue digital converter THOR PRO Analyzer: DT9837-13310 with a sampling frequency of 48 kHz. Each recording lasted 5 seconds, which means that each time data file contains 240,000 samples. Using a tachometer was also necessary to verify that the actual speed of the shaft corresponds to that displayed on the inverter.

In this work, the considered factors are the fault size (3 sizes), the centrifugal load (3 loads) and rotational speed (3 speeds). Table 2 presents the factors and their levels. A full factorial design was selected for accuracy and the experimental plan included all possible combinations. This means that we have $3 \times 3 \times 3 = 27$ trials.

In order to follow the evolution of the fault size, we maintain for each sample the same conditions (rotational speed and centrifugal load) and we change only the fault size ranging from 50 µm to 150 µm. For this reason, we establish all the possible combinations between shaft speed and centrifugal load and we vary for each combination the fault size.

Table 2. Summary of experiments

<table>
<thead>
<tr>
<th>Group</th>
<th>Bearing default size (µm)</th>
<th>Rotational speed (rpm)</th>
<th>Centrifugal load (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>50</td>
<td>300 600 900</td>
<td>50 130 210</td>
</tr>
<tr>
<td>Group 2</td>
<td>100</td>
<td>300 600 900</td>
<td>50 130 210</td>
</tr>
<tr>
<td>Group 3</td>
<td>150</td>
<td>300 600 900</td>
<td>50 130 210</td>
</tr>
</tbody>
</table>

Equation 7 was used to calculate the effect of the centrifugal load. Note that the speed and rotating load are dependent on each other. To study the effects of these two parameters separately, a method based on the use of multiple masses for adjusting the rotational force effect was used.

$$f = m \times R \times \omega^2$$

The objective is to maintain 3 force levels for each speed. Table 3 summarizes the 9 different masses to be applied to offset the effects of speed and keep the three load levels. The intention in this paper is not to study this phenomenon. It is rather to satisfy the application conditions of ANOVA. The results for each group and for each indicator are given in table 5.

Table 3. Summary of compensation masses. (V - rotational speed, m - applied masses and L - centrifugal load)

<table>
<thead>
<tr>
<th>V</th>
<th>m</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>441</td>
<td>50</td>
</tr>
<tr>
<td>600</td>
<td>1145</td>
<td>130</td>
</tr>
<tr>
<td>900</td>
<td>1850</td>
<td>210</td>
</tr>
</tbody>
</table>

4. RESULTS AND DISCUSSIONS

4.1 Time descriptor

In the time domain, the statistical indicators considered as dependent variables are: RMS, Kurtosis, Skewness, Peak and crest factor. They are detailed in Table 4. Using vibrations measurement, the effectiveness of these indicators is analyzed in order to monitor bearings degradation at an early stage. Table 5 shows the experimental results for each indicator.

Table 4. Time indicators.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS</td>
<td>$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2}$</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>$Kurtosis = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2$</td>
</tr>
<tr>
<td>Peak</td>
<td>Crest factor = Peak/RMS</td>
</tr>
<tr>
<td>Skewness</td>
<td>$Skewness = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{X_i - \bar{X}}{RMS} \right)^3$</td>
</tr>
</tbody>
</table>

4.2 Frequency indicator BPFO

In the frequency domain, we compute the first peak amplitude of BPFO (Ball Pass Outer Race) to evaluate the bearings degradation. To do so, build a numerical program which is used the theoretical fault frequency given by the constructor to determine the interval in which the real fault frequency is located. This latter is found by searching the maximum of BPFO amplitude within this interval (BPFO ±f). An example of raw vibratory signal measured during one second are shown on (Fig. 8).
The first step, we used the Lilliefors’s test to confirm that the population distribution was normal [25]. Therefore, the statistical toolbox of Matlab is used to calculate the Lilliefors’s test. The result is 1 if the test rejects the null hypothesis at the 5% significance level. Otherwise, it takes the value 0.

In all cases of table 6, the test statistic k is less than the critical value c, so the value of h=0 indicate that lillietest does not reject the hypothesis at 5% significance level [26]. To find more accurate of p-value, the Monte Carlo approximation is used [26].

$$SE = \sqrt{\frac{(p) - (1-p)}{mcreps}}$$

Where: $p$ is the estimated p-value of the hypothesis test, and $mcreps$ is the number of Monte Carlo replications performed. The number of Monte Carlo replications, $mcreps$ is determined such that the Monte Carlo standard error for less than the value specified for Monte Carlo approximation.
Additionally, a Levene test (table 7), showed that samples with equal variance could not be rejected [27]. Thus, the ANOVA could be executed on the experimental data. Levene’s test, displays P-value less than 0.05 leads to rejection of the hypothesis of equal sigmas at the 5% significance level. In this case, the standard deviations are not significantly different from one to another. Since the P-value is well above 0.05.

When we compare multiple samples, it is usually to perform a one-way analysis of ANOVA. The ANOVA is used to test the hypothesis of equal population means by choosing between null hypothesis and alternative hypothesis. The results of the analysis of variance “ANOVA” for the different hypothesis and alternative hypothesis. The results of the analysis of variance “ANOVA” for the different groups, whereas the RMS, Crest factor, BPFO and Peak does not show any significant difference between the first cases 50 µm and 100 µm, but it is not able to check the difference between 100 µm and 150 µm. The Skewness is able to distinguish the difference between 50 µm and 100 µm, it is not able to check the difference between 100 µm and 150 µm but it is not able to check the difference between 50 µm and 100 µm.

It can be seen from Fig. 5 that only the kurtosis and Skewness have a significant difference for two groups, whereas the RMS, Crest factor, BPFO and Peak does not show any significant difference between the groups. In this regard, all descriptors cannot estimate the difference between all groups. For example, The Kurtosis is able to distinguish the difference between the first cases 50 µm and 100 µm, but it is not able to check the difference between 100 µm and 150 µm. The Skewness is able to distinguish between 100 µm and 150 µm but it is not able to check the difference between 50 µm and 100 µm.

Table 6. Lilliefors’s test (h: hypothesis test result, p: p-value, k: Test statistic, c: Critical value)

<table>
<thead>
<tr>
<th>Time Indicators</th>
<th>50 µm</th>
<th>100 µm</th>
<th>150 µm</th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>k</td>
<td>0.14</td>
<td>0.18</td>
<td>0.30</td>
</tr>
<tr>
<td>c</td>
<td>0.24</td>
<td>0.23</td>
<td>0.21</td>
</tr>
</tbody>
</table>

Table 7. Levene’s test results

<table>
<thead>
<tr>
<th>Time indicators</th>
<th>50, µm</th>
<th>100, µm</th>
<th>150, µm</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS</td>
<td>Mean</td>
<td>Std Dev</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>0,150</td>
<td>0,120</td>
<td>0,174</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>15,747</td>
<td>4,870</td>
<td>25,736</td>
</tr>
<tr>
<td>Skewness</td>
<td>2,616</td>
<td>0,364</td>
<td>2,832</td>
</tr>
<tr>
<td>C-factor</td>
<td>15,424</td>
<td>2,615</td>
<td>16,645</td>
</tr>
<tr>
<td>BPFO</td>
<td>0,0145</td>
<td>0,0133</td>
<td>0,0190</td>
</tr>
</tbody>
</table>

Table 7. ANOVA results

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Swg</th>
<th>Sbg</th>
<th>F-ratio</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMS</td>
<td>0,006</td>
<td>0,018</td>
<td>0,335</td>
<td>0,719</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>321,954</td>
<td>24,152</td>
<td>13,330</td>
<td>0,000*</td>
</tr>
<tr>
<td>Skewness</td>
<td>3,012</td>
<td>0,152</td>
<td>19,829</td>
<td>0,000*</td>
</tr>
<tr>
<td>Crest factor</td>
<td>13,908</td>
<td>10,698</td>
<td>1,300</td>
<td>0,291</td>
</tr>
<tr>
<td>Peak</td>
<td>4,246</td>
<td>2,953</td>
<td>1,438</td>
<td>0,257</td>
</tr>
<tr>
<td>BPFO</td>
<td>0,0001</td>
<td>0,0002</td>
<td>0,560</td>
<td>0,579</td>
</tr>
</tbody>
</table>
5. CONCLUSION

Monitoring the evolution of bearings fault size at an early stage is crucial. In this paper, time and frequency indicators are used to estimate the variation of default size at a macroscopic scale.

The method of the one-way analysis of variance "ANOVA" alongside Tukey’s multiple comparison tests proved to be an efficient method in selecting the good indicators to monitor the bearing degradation ranging from 50 µm to 150 µm. In this regard, the Skewness and Kurtosis indicators show a good ability to assess defectiveness in the bearings at an early stage; the disadvantage of other indicators such as the RMS, Peak, Crest factor and BPFO cannot detect the changes in the fault size of bearings varying from 50 µm to 150 µm. However, the "ANOVA" seems to be unsuitable to compare between time indicators since they do not fulfill the conditions of this method.

Finally, the procedure proposed in this paper can also be useful to study the efficiency of the parameters of other problems.

6. REFERENCE


23. Ram Prabhu T. Effects of solid lubricants, load, and sliding speed on the tribological behavior of silica reinforced composites using design of experiments. Materials and Design. 2015; 77:149–160. https://doi.org/10.1016/j.matdes.2015.03.059


Received 2018-04-24
Accepted 2018-09-08
Available online 2018-10-30