2017

**Integrated Condition Monitoring and Prognosis Method for Incipient Defect Detection and Remaining Life Prediction of Low Speed Slew Bearings**

Wahyu Caesarendra  
*University of Wollongong, wc026@uowmail.edu.au*

Tegoeh Tjahjowidodo  
*Nanyang Technological University*

Buyung Kosasih  
*University of Wollongong, buyung@uow.edu.au*

Anh Kiet Tieu  
*University of Wollongong, ktieu@uow.edu.au*

Follow this and additional works at: [https://ro.uow.edu.au/eispapers1](https://ro.uow.edu.au/eispapers1)

Part of the [Engineering Commons](https://ro.uow.edu.au/eispapers1), and the [Science and Technology Studies Commons](https://ro.uow.edu.au/eispapers1)

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au
Integrated Condition Monitoring and Prognosis Method for Incipient Defect Detection and Remaining Life Prediction of Low Speed Slew Bearings

Abstract
This paper presents an application of multivariate state estimation technique (MSET), sequential probability ratio test (SPRT) and kernel regression for low speed slew bearing condition monitoring and prognosis. The method is applied in two steps. Step (1) is the detection of the incipient slew bearing defect. In this step, combined MSET and SPRT is used with circular-domain kurtosis, time-domain kurtosis, wavelet decomposition (WD) kurtosis, empirical mode decomposition (EMD) kurtosis and the largest Lyapunov exponent (LLE) feature. Step (2) is the prediction of the selected features' trends and the estimation of the remaining useful life (RUL) of the slew bearing. In this step, kernel regression is used with time-domain kurtosis, WD kurtosis and the LLE feature. The application of the method is demonstrated with laboratory slew bearing acceleration data.

Keywords
condition, incipient, integrated, method, prognosis, monitoring, bearings, slew, speed, low, prediction, life, remaining, detection, defect

Disciplines
Engineering | Science and Technology Studies

Publication Details

This journal article is available at Research Online: https://ro.uow.edu.au/eispapers1/73
Article

Integrated Condition Monitoring and Prognosis Method for Incipient Defect Detection and Remaining Life Prediction of Low Speed Slew Bearings

Wahyu Caesarendra 1,2,*, Tegoeh Tjahjowidodo 3, Buyung Kosasih 1 and Anh Kiet Tieu 1

1 Tribology Laboratory, School of Mechanical, Materials and Mechatronic Engineering, University of Wollongong, Wollongong, New South Wales 2522, Australia; buyung@uow.edu.au (B.K.); ktieu@uow.edu.au (A.K.T.)
2 Mechanical Engineering Department, Diponegoro University, Tembalang, Semarang 50275, Indonesia
3 School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore; ttegoeh@ntu.edu.sg

* Correspondence: w.caesarendra@gmail.com

Academic Editor: Hongrui Cao
Received: 24 February 2017; Accepted: 28 March 2017; Published: 4 April 2017

Abstract: This paper presents an application of multivariate state estimation technique (MSET), sequential probability ratio test (SPRT) and kernel regression for low speed slew bearing condition monitoring and prognosis. The method is applied in two steps. Step (1) is the detection of the incipient slew bearing defect. In this step, combined MSET and SPRT is used with circular-domain kurtosis, time-domain kurtosis, wavelet decomposition (WD) kurtosis, empirical mode decomposition (EMD) kurtosis and the largest Lyapunov exponent (LLE) feature. Step (2) is the prediction of the selected features’ trends and the estimation of the remaining useful life (RUL) of the slew bearing. In this step, kernel regression is used with time-domain kurtosis, WD kurtosis and the LLE feature. The application of the method is demonstrated with laboratory slew bearing acceleration data.

Keywords: condition monitoring; kernel regression; low speed slew bearing; multivariate state estimation technique; prognosis; sequential probability ratio test

1. Introduction

Steelmaking industry has many critical processes which rely on low rotating slew bearings. These bearings are often used in harsh conditions and have high replacement cost with long delivery lead time. Therefore, it is important to monitor their condition at all times. Typical maintenance schedule is replace the bearings which have been in operation of set number of hours. However, due to their low speeds, the bearings run to failure without appropriate condition monitoring method. These can lead to a high level of risk and sub-utilization of the actual bearing useful life. To reduce the risk of sudden bearing failure and optimise the use of the bearing, integrated condition monitoring and prognosis method for low speed slew bearings is needed.

In order to monitor the condition of slew bearings, a method that able to extract the trends related to degradation condition is necessary. Henao et al. [1] presents a review of diagnostic techniques for electrical machines. The summary of recently proposed methods for performing diagnostics in nonstationary conditions were categorised into four: (1) Frequency domain approach using the rotational invariance technique; (2) Time domain analysis-based DWT; (3) Slip-frequency approach based on the description of the instantaneous frequency of an extracted fault component against slip; (4) Diagnosis in the time-frequency domain using STFT, CWT, WVD, Choi-Willians distribution
Machines 2017, 5, 11

(CWD), and Zhao-Atlas-Marks distribution. Currently, artificial intelligence (AI) techniques remain a priority tool to support the condition monitoring and fault diagnosis methods especially in the decision process [1,2]. These techniques have been widely developed in recent years, even though they are not yet applied in the industrial practice [1].

A condition monitoring method for slew bearings using kurtosis extracted from vibration and acoustic emission signals as condition parameters was proposed by Rodgers [3]. A recent slew bearing condition monitoring and fault diagnosis study [4] combined ensemble empirical mode decomposition (EEMD) and multiscale principal component analysis (MSPCA). The method successfully predicted defect frequency of artificially damaged inner race which matched with the ball pass frequency inner race (BPFI). However, the practical challenge is how to detect the incipient defect of slew bearing from low energy vibration of naturally induced defect and predict RUL based on the observation parameters. To date, prognosis methods have been mainly applied in typical rolling element bearings with rotational speed greater than 1000 rpm [5–7]. Lei et al. [8] proposed the application of particle filtering (PF) to estimate model parameters from real-time measured data. The model parameters were updated adaptively, which eventually were used to predict the remaining useful life (RUL) of machineries. The operation conditions of tested bearings are 1500, rpm 1650 rpm and 1800 rpm. Subsequently, Li et al. [9] highlighted the main shortcoming of traditional PF that is attributed to the subjective selection of the first predicting time (FPT), i.e., the triggered time to start the prediction process. The authors, thereafter, improved the methodology by means of adaptive FPT selection via stochastic approach. The bearing was naturally degraded during the experiment with rotational speed of 1800 rpm. The prognosis of low speed bearings with naturally induced defects has not been explored. Prognosis method has not been applied in the study of low rotational speed slew bearing [4,10,11]. This paper present both condition monitoring and prognosis method.

The method employs a multivariate state estimation technique (MSET) and a sequential probability ratio test (SPRT) to identify subtle changes in bearing condition, and kernel regression to predict the trend of the change and to estimate the RUL of the slew bearing. MSET was first introduced by Singer [12] and developed further by the Argonne National Laboratory [13,14]. Over the years, the technique has been applied successfully in many applications such as in the analysis of nuclear power reactor signals [12–15]. Recently, MSET has been used in prognosis and fault diagnosis studies e.g., the estimation of the remaining useful life of electronic products [16], the monitoring of lithium-ion battery performance [17] and early fault diagnosis of wind turbines [18]. The previous study used MSET to correlate the features of vibration signal acquired from low speed slew bearing [19].

SPRT was initially introduced by Wald [20,21] as a sequential statistical method. This method has been shown to work well either as a stand-alone decision-making method [22–26] or as an integrated method together with MSET [14,27]. When integrated with MSET, SPRT can be used to analyze the MSET result. SPRT can also be combined with auto-associative kernel regression algorithm [28]. Traditionally, kernel regression has been applied in image processing studies such as in image de-noising and enhancement [29] and image reconstruction [30]. It is an appropriate prediction method [31] and adopted in this paper.

This paper is organized as follows. Section 2 describes how features such as circular-domain kurtosis, time-domain kurtosis, WD kurtosis, EMD kurtosis and LLE feature are computed and monitored using MSET and SPRT to detect the incipient slew bearing defect and used in the kernel regression computation to predict the trends of these features. Section 3 describes the experimental set-up. Section 4 presents the results and discusses the proposed integrated condition monitoring and prognosis method including the estimation of the RUL. Finally, Section 5 presents the conclusions of the current work.
2. Integrated Condition Monitoring and Prognosis Method for Low Speed Slew Bearings

The proposed integrated condition-monitoring and prognosis method is implemented as follows: Step (1) is the detection of the incipient slew bearing defect where combined MSET and SPRT is used to process circular-domain kurtosis [32], time-domain kurtosis [32], WD kurtosis [32], EMD kurtosis [32] and LLE feature [33]. Step (2) predicts the trends of the selected features and estimates the RUL of the slew bearing. In this step, kernel regression is used to predict the trend of the time-domain kurtosis, WD kurtosis and the LLE features. Step (2) is initiated when the incipient bearing defect from Step (1) is registered. Figure 1 illustrates the computational implementation of the method.

![Diagram of the condition monitoring and prognosis method](image)

**Figure 1.** The integrated condition monitoring and prognosis method for low speed slew bearing.

Examples of features extracted from slew bearing acceleration data are shown in Figure 2. How these features were extracted has been described in Refs. [32, 33]. It is apparent that the features do not display any clear trend. In this paper, circular kurtosis, time-domain kurtosis, WD kurtosis and EMD kurtosis are employed to obtain the bearing degradation parameter for prognosis method.

![Graph of circular kurtosis over time](image)

**Figure 2.** Cont.
Figure 2. (a) circular kurtosis [32]; (b) time-domain kurtosis [32]; (c) wavelet kurtosis [32]; (d) EMD kurtosis [32]; (e) the LLE feature [33]. Note: $t = \text{day}$. 
2.1. Step (1) Detection of the Incipient Slew Bearing Defect Using MSET and SPRT

2.1.1. The Multivariate State Estimation Technique (MSET)

MSET is a technique that typically used to analyse anomalous process based on sensor reading and input parameters during the online condition monitoring. This technique has been applied in process monitoring, signal validation and surveillance applications [12–19,27]. The previous study used MSET to correlate the features of vibration signal acquired from low speed slew bearing such as largest Lyapunov exponent, approximate entropy, margin factor and impulse factor. In the present study, MSET is employed for different features i.e., circular-domain kurtosis, time-domain kurtosis, wavelet decomposition-kurtosis and empirical mode decomposition-kurtosis.

The MSET algorithm is applied systematically in 5 steps [16] as shown in Figure 3 in the first and second blocks. First, data matrix \( P \) is created where the row elements of \( P \) are the extracted features and the columns are the measurement time i.e., days. Data matrix \( P \) of size \( m \) features \( \times n \) days is defined in Equation (1).

\[
P = \begin{bmatrix}
    f_1(1) & f_1(2) & f_1(3) & \ldots & f_1(n) \\
    f_2(1) & f_2(2) & f_2(3) & \ldots & f_2(n) \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    f_m(1) & f_m(2) & f_m(3) & \ldots & f_m(n)
\end{bmatrix}
\]

Second, data matrix \( P \) is subdivided into two matrices: training matrix \( T \) and observation matrix \( P_{\text{obs}} \). \( T \) is further divided into memory matrix \( D \) and remaining training matrix \( L \). It should be noted that \( T \) holds data from the normal state while \( P_{\text{obs}} \) holds the monitored state, as shown in Equation (2).

\[
P = \begin{bmatrix}
    f_1(1) & f_1(2) & \ldots & f_1(n_D) & f_1(n_D + 1) & \ldots & f_1(n_L) & f_1(n_L + 1) & f_1(n_L + 2) & \ldots & f_1(n) \\
    f_2(1) & f_2(2) & \ldots & f_2(n_D) & f_2(n_D + 1) & \ldots & f_2(n_L) & f_2(n_L + 1) & f_2(n_L + 2) & \ldots & f_2(n) \\
    \vdots & \vdots & \ldots & f_3(n_D) & f_3(n_D + 1) & \ldots & f_3(n_L) & f_3(n_L + 1) & f_3(n_L + 2) & \ldots & f_3(n) \\
    \vdots & \vdots & \ldots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
    f_m(1) & f_m(2) & \ldots & f_m(n_D) & f_m(n_D + 1) & \ldots & f_m(n_L) & f_m(n_L + 1) & f_m(n_L + 2) & \ldots & f_m(n)
\end{bmatrix}
\]

where \( n_D \) is the number of days in \( D \) matrix and \( n_L \) is the number of days in \( L \) matrix.

Third, the weight vector for the normal state (\( w_1 \)) is calculated as follows [16]

\[
w_1 = (D^T \otimes D)^{-1} \bullet (D^T \otimes L)
\]

and the weight vector for the observation state (\( w_2 \)) is calculated as follows [16]

\[
w_2 = (D^T \otimes D)^{-1} \bullet (D^T \otimes P_{\text{obs}})
\]

Operator \( \otimes \) is a non-linear operator. Several non-linear operators can be selected for MSET [34]. The most popular operator is the Euclidean distance, which is the one adopted in this paper.

Fourth, the normal estimate matrix (\( L_{\text{est}} \)) is computed by Equation (5)

\[
L_{\text{est}} = D \bullet w_1
\]

and the estimate matrix of monitored state (\( P_{\text{est}} \)) is calculated by Equation (6)

\[
P_{\text{est}} = D \bullet w_2
\]
Fifth, the normal residual matrix ($R_N$) is the difference between $L_{est}$ and $L$ and calculated by Equation (7).

$$R_N = L_{est} - L$$  

(7)

And the residual matrix of monitored state ($R_M$) is the difference between $P_{est}$ and $P_{obs}$ and calculated by Equation (8).

$$R_M = P_{est} - P_{obs}$$  

(8)

2.1.2. The Sequential Probability Ratio Test (SPRT)

SPRT is a sequential statistical binary hypothesis technique [20,21] that has been used as a simple and effective stand-alone decision-making technique for engineering systems such as the surveillance of nuclear power plant components [23], statistical damage classification of an automotive system [24], long-term radiation monitoring [25], and multiple fault recognition in gearboxes [26]. SPRT has been integrated with MSET to analyse the output of the MSET and digitally assess the condition of the system being monitored [14,27]. With this capability SPRT is a potential early warning method embeddable in online monitoring systems.

To detect the incipient slew bearing defect, SPRT utilizes the outputs of MSET, the normal residual matrix ($R_N$) and the residual matrix of monitored state ($R_M$) as shown in Figure 3. The MSET outputs are then subjected to binary hypothesis test, where the objective is to assess whether the bearing is still in normal condition or already in defective condition. The decision making process is based on a set of hypotheses: normal condition denoted by a null hypothesis ($H_0$) and defect condition denoted by ($H$). In practice, there could be more than one hypothesis ($H$), i.e., $H_1$, $H_2$, . . . , $H_z$ (see Table 1). It should be noted that with more hypotheses used, the greater the accuracy of the decision making will be.

Table 1. Hypotheses for SPRT.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Statistical Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Normal condition ($H_0$)</td>
<td>$M_{normal}$ (Table 5)</td>
</tr>
<tr>
<td>2. Abnormal condition 1 ($H_1$)</td>
<td>$M_{monitored} &gt; M_{max}$</td>
</tr>
<tr>
<td>3. Abnormal condition 2 ($H_2$)</td>
<td>$M_{monitored} &lt; M_{min}$</td>
</tr>
<tr>
<td>4. Abnormal condition 3 ($H_3$)</td>
<td>$M_{min} &lt; M_{monitored} &lt; M_{max}$</td>
</tr>
<tr>
<td>5. Abnormal condition 4 ($H_4$)</td>
<td>$M_{min} &lt; M_{monitored} &lt; M_{max}$</td>
</tr>
</tbody>
</table>

Note: $M_{min} < M_{normal} < M_{max}$

In SPRT analysis (see Figure 3), $L_i$ the ratio of the probability of $H_0$ and the probability of $H$ is calculated. The decision is then made based on the SPRT index calculated as follows [35]:

$$SPRT_{index} = \ln(L_i)$$  

(9)

where $L_i$ is given by

$$L_i = \frac{\text{probability of statistical properties } \{S_i\} \text{ given } H \text{ is true}}{\text{probability of statistical properties } \{S_i\} \text{ given } H_0 \text{ is true}}$$  

(10)

where $S_i$ are the statistical properties e.g., mean, $M$ and standard deviation, $\sigma$ of observation features at the monitored measurement day.
2.2. Step (2) Prediction of the Future State of Monitored Parameters Using Modified Kernel Regression

Kernel regression [36–38] is a non-parametric regression technique used to estimate regression function \( f(x, y) \) that best fit non-linear data set \( (X_j, Y_j) \). \( X_j \) in this paper refers to the measurement day and \( Y_j \) is the value of the extracted feature. Unlike linear regression or polynomial regression, kernel regression does not assume any underlying distribution to estimate the regression function [39]. The kernel regression uses a set of identical weighted function called local kernels to each observation.
data point. The kernel basis function depends only on the kernel width from local data point \( X \) to a set of neighboring locations \( x \). The procedures of the proposed kernel regression-based prognostic are implemented in three systematic steps.

Firstly, three of the five features presented in Figure 2 are selected based on the evaluation criteria (described in Section 4.2). Once Step (1) has triggered the incipient defect, initial threshold is set up. The initial threshold is set at four times the kurtosis value of the normal bearing condition. In Step (2), the data point that exceeds the predetermined initial threshold of each selected feature is saved and used for prediction. From the laboratory slew bearing data, it was found that the minimum number of data points used for prediction is 3. Time-domain kurtosis is used as an illustration in Figure 4. The measurement day and the kurtosis level of the first three non-linear data points which exceed the threshold are presented in Table 2. We refer these points as 'non-linear' data because the interval between one point and the next is not equal and the level of kurtosis does not always increase linearly.

![Figure 4. Kurtosis extracted from the original vibration signal.](image)

**Table 2. Non-linear data points.**

<table>
<thead>
<tr>
<th>Data Point</th>
<th>Measurement Day (( X_j ))</th>
<th>Feature Level (( Y_j ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>92</td>
<td>19</td>
</tr>
<tr>
<td>B</td>
<td>103</td>
<td>24.12</td>
</tr>
<tr>
<td>C</td>
<td>114</td>
<td>36.83</td>
</tr>
</tbody>
</table>

where \( j \) represents points A, B or C

Secondly, the non-linear regression model is built in 4 steps. (1) Calculate data point \( x_j \) with small step \( dx \). \( dx \) equal to 0.2 is used throughout this paper. The result is the vector \( X \). (2) Set the kernel width \( \alpha \). \( \alpha \) for time-domain kurtosis prediction is 6. \( dx \) and \( \alpha \) are selected by trial and error to get the optimum regression model; (3) Apply the Gaussian kernel to each data point \( X_j \) using the following equation

\[
K_j(x, X_j) = \exp\left(-\frac{(x - X_j)^2}{2\alpha^2}\right)
\]

where subscript \( j \) denotes non-linear data point. According to the example given in Table 2, \( X_j \) are \( X_A, X_B \) and \( X_C \). (4) Compute the weight vector \( w = (w_1, w_2, \ldots, w_p) \) using the least square method by minimizing the sum square error between predicted \( \hat{Y}_j \) and monitored \( Y_j \) as follows:
where $p$ is the number of non-linear data points to be regressed and predicted. In this example, $p = 3$. lsqcurvefit MATLAB function is used to obtain the weight vector $w$. The result of the non-linear regression model is presented in Table 3 and shown in Figure 5.

\[ \epsilon = \sum_{j=1}^{p} (\hat{Y}_j - Y_j) \]  

(12)

Figure 5. Regression model of three data points and one-step-ahead prediction.

Table 3. Kernel regression model for 1 step-ahead prediction of time-domain kurtosis.

<table>
<thead>
<tr>
<th>Data Point (x-axis)</th>
<th>x</th>
<th>$K_1(x_iX_A)$</th>
<th>$K_2(x_iX_B)$</th>
<th>$K_3(x_iX_C)$</th>
<th>$w_1$</th>
<th>$K_1$</th>
<th>$w_2$</th>
<th>$K_2$</th>
<th>$w_3$</th>
<th>$K_3$</th>
<th>$\sum_{j=1}^{p} w_j K_j$</th>
<th>$\sum_{j=1}^{p} K_j$</th>
<th>Data Point (y-axis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>XA = 92</td>
<td>92</td>
<td>0.186</td>
<td>0.0012</td>
<td>18.354</td>
<td>4.158</td>
<td>0.047</td>
<td>22.560</td>
<td>1.187</td>
<td>YA = 18.998</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>92.2</td>
<td>0.999</td>
<td>0.197</td>
<td>0.0014</td>
<td>18.344</td>
<td>4.417</td>
<td>0.053</td>
<td>22.815</td>
<td>1.198</td>
<td>19.033</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>92.4</td>
<td>0.997</td>
<td>0.210</td>
<td>0.0015</td>
<td>18.313</td>
<td>4.688</td>
<td>0.060</td>
<td>23.062</td>
<td>1.209</td>
<td>19.070</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>102.6</td>
<td>0.210</td>
<td>0.997</td>
<td>0.164</td>
<td>3.854</td>
<td>22.273</td>
<td>6.050</td>
<td>32.633</td>
<td>1.372</td>
<td>23.780</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>102.8</td>
<td>0.197</td>
<td>0.999</td>
<td>0.175</td>
<td>3.632</td>
<td>22.310</td>
<td>6.927</td>
<td>32.869</td>
<td>1.372</td>
<td>23.949</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XB = 103</td>
<td>103</td>
<td>0.186</td>
<td>1</td>
<td>0.186</td>
<td>3.418</td>
<td>22.322</td>
<td>7.367</td>
<td>33.109</td>
<td>1.372</td>
<td>24.122</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>103.2</td>
<td>0.175</td>
<td>0.999</td>
<td>0.197</td>
<td>3.214</td>
<td>22.310</td>
<td>7.827</td>
<td>33.352</td>
<td>1.372</td>
<td>24.301</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>103.4</td>
<td>0.164</td>
<td>0.997</td>
<td>0.210</td>
<td>3.018</td>
<td>22.273</td>
<td>8.307</td>
<td>33.599</td>
<td>1.372</td>
<td>24.484</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>113.6</td>
<td>0.0015</td>
<td>0.210</td>
<td>0.997</td>
<td>0.0282</td>
<td>4.688</td>
<td>39.466</td>
<td>44.183</td>
<td>1.209</td>
<td>36.535</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>113.8</td>
<td>0.0014</td>
<td>0.197</td>
<td>0.999</td>
<td>0.0250</td>
<td>4.417</td>
<td>39.532</td>
<td>43.975</td>
<td>1.198</td>
<td>36.685</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XC = 114</td>
<td>114</td>
<td>0.0012</td>
<td>0.186</td>
<td>1</td>
<td>0.0221</td>
<td>4.158</td>
<td>39.354</td>
<td>43.734</td>
<td>1.187</td>
<td>36.830</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>114.2</td>
<td>0.0011</td>
<td>0.175</td>
<td>0.999</td>
<td>0.0195</td>
<td>3.909</td>
<td>39.532</td>
<td>43.461</td>
<td>1.175</td>
<td>36.968</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>114.4</td>
<td>0.0009</td>
<td>0.164</td>
<td>0.997</td>
<td>0.0173</td>
<td>3.671</td>
<td>39.466</td>
<td>43.135</td>
<td>1.163</td>
<td>37.100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>119.6</td>
<td>0.000025</td>
<td>0.021</td>
<td>0.646</td>
<td>0.00046</td>
<td>0.486</td>
<td>25.588</td>
<td>26.074</td>
<td>0.668</td>
<td>38.992</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>119.8</td>
<td>0.000021</td>
<td>0.019</td>
<td>0.626</td>
<td>0.0004</td>
<td>0.442</td>
<td>24.790</td>
<td>25.233</td>
<td>0.646</td>
<td>39.025</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XD = 120</td>
<td>120</td>
<td>0.000018</td>
<td>0.018</td>
<td>0.606</td>
<td>0.00034</td>
<td>0.403</td>
<td>23.991</td>
<td>24.394</td>
<td>0.624</td>
<td>39.055</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: $w = [w_1; w_2; w_3] = [18.35; 22.32; 39.55]$. 
Thirdly, the future value of the feature is predicted. The predicted value of $Y_j$ at future $X_j$ is given by the kernel regression formula (also called Nadaraya-Watson kernel weighted average) \cite{39} as follows:

$$
\hat{Y}_j = \frac{\sum_{i=1}^{p} w_i K_{ij}}{\sum_{i=1}^{p} K_{ij}} \quad (13)
$$

The one-step-ahead prediction of time-domain kurtosis is presented in Figure 6.

The result shown in Figure 6 is one-step-ahead prediction of time-domain kurtosis based on modified kernel regression method presented in Section 2.2. The modified kernel regression is proposed to calculate a model based on 3 points and predict one-point-ahead. This method can predict one-point-ahead of time-domain kurtosis feature based on 3 points with similar spacing between the points as illustrated in Figure 4. The proposed method still works although the spacing between 3 points is different.

![Figure 6. One-step-ahead prediction of time-domain kurtosis feature.](image)

3. Experimental Set-Up

The vibration data of the accelerated wear test used in this paper was acquired from a laboratory slew bearing test rig shown in Figure 7. The schematic of the slew bearing test rig showing the main drive reducer, the hydraulic load and how the bearing is attached is shown in Figure 7a. A 30 tonnes axial load was hydraulically applied. The test rig was operated in continuous rotation at 1 rpm. The slew bearing used was INA YRT260 type axial/radial bearing with inner and outer diameter of 260 mm and 385 mm, respectively. The vibration data was acquired from four accelerometers installed on the inner radial surface at 90 degrees to each other (see Figure 7b) with 4880 Hz sampling rates. The accelerometers were IMI608 A11 ICP type sensors. The accelerometers were connected to a high speed Pico scope DAQ (PS3424). 139 days data were collected. In order to accelerate the damage and reduce the bearing service life, coal dust was injected into the bearing. In practice, especially in steel making companies, slew bearing is located in open air and exposed to dusty environment. To replicate this condition, coal dust was inserted in day 58.
4. Results and Discussion

4.1. Step (1) Detection of the Incipient Slew Bearing Defect Using MSET and SPRT

Twelve features were extracted from four different methods [31]. Not all features were sensitive to the change of the bearing condition. Consequently, the method only uses five of the most sensitive features as the input to the MSET. The five features were (1) kurtosis extracted from accelerometer signal, \( f_1(t) \); (2) kurtosis extracted from the piecewise aggregate approximation (PAA) combined with circular analysis, \( f_2(t) \); (3) kurtosis extracted from the wavelet decomposition, \( f_3(t) \); (4) kurtosis extracted from the EMD, \( f_4(t) \); and (5) LLE feature obtained from the largest Lyapunov exponent algorithm, \( f_5(t) \) [32]. The five features were extracted over the 139 days test period. The plot of the aforementioned features has been shown in Figure 2. These features are used to establish data matrix \( P \). Thus, the dimension of data matrix \( P \) in Equation (1) is 5 by 139, where 5 is the number of features and 139 is the number of days.

To detect the incipient slew bearing defect (Step 1), the MSET algorithm calculates the normal residual matrix (\( R_N \)) and the residual matrix of monitored state (\( R_M \)) (see Figure 3). In Step (1), the following inputs are required: training matrix \( T \), memory matrix \( D \), remaining training matrix \( L \) and observation matrix \( P_{\text{obs}} \). In this paper, \( T \) of 50 days was used. In the computation, the sizes of matrices, \( D \) and \( L \) have to be half of \( T \) i.e., 25 days each. The reason why 50 days was used is because coal dust was not injected until day 58. In other words, during the first 50 days, the vibration signal had been acquired when the condition of the bearing was still normal. In practice, matrix \( P_{\text{obs}} \) is obtained from daily observation or measurement. The number of measurement days from normal to failure has been set as 139 days. \( P_{\text{obs}} \) is the matrix from day 51 to day 139.

After \( D \) and \( L \) have been determined, the weight of normal state vector (\( w_1 \)) is then calculated using Equation (3) by employing the Euclidean distance as the non-linear operator. Vector \( w_1 \) represents the weight of normal data points in the training data. \( w_1 \) is used to calculate the normal estimate matrix (\( L_{\text{est}} \)) using Equation (5). And the normal residual matrix (\( R_N \)) is estimated by taking the difference between \( L_{\text{est}} \) and \( L \), as shown in Equation (7). It should be noted that prior to the calculation of the statistical properties, matrix \( R_N \) is normalized with a zero mean. This is necessary for the next process in the SPRT method because under normal bearing conditions the mean value of the normal residual matrix (\( R_N \)) is expected to be \( \approx 0 \). The result of normalized matrix (\( R_N \)) is presented in Table 4. A similar procedure is also employed to obtain the residual matrix of monitored state (\( R_M \)). Equation (4) is used to calculate the weight of monitored state vector (\( w_2 \)), Equation (6) is then used to calculate the estimate matrix of monitored state (\( L_{\text{est}} \)) and Equation (8) to calculate the residual matrix of monitored state (\( R_M \)). Matrix \( R_M \) is then normalized within a range between maximum value and
minimum value of the normalized normal residual matrix \((R_N)\). The calculation of \(R_N\) and \(R_M\) is the final step of MSET. Both results are then fed into the SPRT.

**Table 4. Normalized \(R_N\) matrix.**

<table>
<thead>
<tr>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Day 5</th>
<th>Day 6</th>
<th>Day 7</th>
<th>Day 8</th>
<th>Day 9</th>
<th>Day 10</th>
<th>Day 11</th>
<th>Day 12</th>
<th>Day 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.044</td>
<td>0.042</td>
<td>0.042</td>
<td>0.042</td>
<td>0.042</td>
<td>0.043</td>
<td>0.042</td>
<td>0.043</td>
<td>0.044</td>
<td>0.042</td>
<td>0.042</td>
<td>0.042</td>
<td>0.043</td>
</tr>
<tr>
<td>0.025</td>
<td>0.021</td>
<td>0.021</td>
<td>0.022</td>
<td>0.022</td>
<td>0.023</td>
<td>0.021</td>
<td>0.024</td>
<td>0.025</td>
<td>0.022</td>
<td>0.022</td>
<td>0.021</td>
<td>0.022</td>
</tr>
<tr>
<td>0.016</td>
<td>0.015</td>
<td>0.015</td>
<td>0.015</td>
<td>0.015</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.016</td>
<td>0.015</td>
<td>0.015</td>
<td>0.015</td>
<td>0.015</td>
</tr>
<tr>
<td>0.018</td>
<td>0.017</td>
<td>0.017</td>
<td>0.017</td>
<td>0.017</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.017</td>
<td>0.017</td>
<td>0.017</td>
<td>0.018</td>
</tr>
<tr>
<td>(-0.955)</td>
<td>(-0.957)</td>
<td>(-0.957)</td>
<td>(-0.957)</td>
<td>(-0.957)</td>
<td>(-0.956)</td>
<td>(-0.957)</td>
<td>(-0.956)</td>
<td>(-0.955)</td>
<td>(-0.957)</td>
<td>(-0.957)</td>
<td>(-0.957)</td>
<td>(-0.957)</td>
</tr>
</tbody>
</table>

**Table 5. Mean (\(M\)) values of normal state.**

<table>
<thead>
<tr>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Day 5</th>
<th>Day 6</th>
<th>Day 7</th>
<th>Day 8</th>
<th>Day 9</th>
<th>Day 10</th>
<th>Day 11</th>
<th>Day 12</th>
<th>Day 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-0.169)</td>
<td>(-0.172)</td>
<td>(-0.172)</td>
<td>(-0.171)</td>
<td>(-0.171)</td>
<td>(-0.172)</td>
<td>(-0.170)</td>
<td>(-0.170)</td>
<td>(-0.171)</td>
<td>(-0.172)</td>
<td>(-0.172)</td>
<td>(-0.171)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(-0.171)</td>
<td>(-0.172)</td>
<td>(-0.170)</td>
<td>(-0.171)</td>
<td>(-0.171)</td>
<td>(-0.170)</td>
<td>(-0.170)</td>
<td>(-0.171)</td>
<td>(-0.170)</td>
<td>(-0.172)</td>
<td>(-0.172)</td>
<td>(-0.171)</td>
</tr>
<tr>
<td>(-0.957)</td>
<td>(-0.957)</td>
<td>(-0.956)</td>
<td>(-0.956)</td>
<td>(-0.956)</td>
<td>(-0.957)</td>
<td>(-0.957)</td>
<td>(-0.957)</td>
<td>(-0.957)</td>
<td>(-0.957)</td>
<td>(-0.957)</td>
<td>(-0.957)</td>
</tr>
</tbody>
</table>

Note: \(M_{\text{min}} = -0.172, M_{\text{max}} = -0.169\).

**Table 6. Standard deviation (\(\sigma_{\text{normal}}\)) values of normal state.**

<table>
<thead>
<tr>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
<th>Day 4</th>
<th>Day 5</th>
<th>Day 6</th>
<th>Day 7</th>
<th>Day 8</th>
<th>Day 9</th>
<th>Day 10</th>
<th>Day 11</th>
<th>Day 12</th>
<th>Day 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
</tr>
<tr>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
<td>0.439</td>
</tr>
</tbody>
</table>

This comparison is repeated to the 51 to 139 columns of the \(P_{\text{obs}}\) matrix. If the mean value or the standard deviation in each day lies outside the range of the reference mean values (Table 5) or standard deviation values (Table 6), incipient defect is considered to have occurred. This step is part of the SPRT procedure. To make a comparison, the list of hypotheses must be determined first including the normal condition \((H_0)\) represented by mean value = \(M_{\text{normal}}\) and the standard deviation = \(\sigma_{\text{normal}}\). Alternatives hypotheses \(H_1\) to \(H_4\) indicate the abnormal conditions. The set of hypotheses is given in Table 1.

To measure the comparison mathematically, the ratio of the probability of alternative hypotheses \((H_1\) to \(H_4\)) and the probability of the null hypothesis \((H_0)\) is calculated, Equation (10). Once probability ratio \(L_i\) is obtained, the SPRT index can be calculated by taking the natural logarithm of probability ratio \(L_i\). The result of Step (1) is presented in Figure 8. It can be seen from the monitored data of day 51...
to day 89 that the bearing is still in normal condition. The impending deterioration of the bearing is identified on day 90. Note that the impending deterioration is needed in Step (2) to predict the future state and to estimate the RUL estimation. A sample calculation of $L_i$ for a day (day 51 and day 90) is presented in Table 7. When the condition of the bearing is still normal, the probability of statistical properties $\{S_i\}$ that fall within the hypotheses $H$ ($p(\{S_i\}/H)$) is much lower than the probability of statistical properties $\{S_i\}$ of the null hypothesis $H_0$ ($p(\{S_i\}/H_0)$). To illustrate the normal condition (i.e., day 50), the natural logarithmic of $L_i$ is given in Table 7 and shown in Figure 8. On the contrary, on day 90 the condition of the bearing started to deteriorate, the probability of statistical properties $\{S_i\}$ that fall within the hypotheses $H$ ($p(\{S_i\}/H)$) will increase and the probability of statistical properties $\{S_i\}$ that fall in the null hypothesis $H_0$ ($p(\{S_i\}/H_0)$) will decrease, the natural logarithmic is then given in Table 7 and shown in Figure 8.

![Figure 8. The SPRT index result for detecting impending bearing deterioration.](image)

Table 7. Example calculation of $L_i$ in one day observation.

<table>
<thead>
<tr>
<th>Dayth Observation</th>
<th>$p({S_i}/H)$</th>
<th>$p({S_i}/H_0)$</th>
<th>$L_i$ (Equation 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>day 51</td>
<td>0.04</td>
<td>0.99</td>
<td>-3.209</td>
</tr>
<tr>
<td>day 90</td>
<td>0.08</td>
<td>0.98</td>
<td>-2.506</td>
</tr>
</tbody>
</table>

Note: $p(\{S_i\}/H)$ is probability of statistical properties $\{S_i\}$ given $H$ is true and $p(\{S_i\}/H)$ is probability of statistical properties $\{S_i\}$ given $H_0$ is true.

4.2. Step (2) Prediction of the Future State of Selected Features Using Kernel Regression

Features extracted from slew bearing vibration signals have different characteristic to those extracted from accelerometers signals of typical rolling element bearings. In typical rolling element bearings, when the bearing condition deteriorates, the features values increase gradually. In slew bearings, the changes of the slew bearing condition can be detected from a sharp increase of feature value. However, this value does not increase steadily as it does in typical high speeds roller bearings. The extracted features of the slew bearing signal usually revert to the normal level and rise again as the condition has deteriorated significantly. This condition is referred as ‘self-healing’ characteristic [41].

This is one of the difficulties in defect prediction especially for data-driven prognosis methods [42–44]. In this paper, the problem has been overcome using kernel regression. Prior to the use of kernel regression, all features have to be evaluated. Caesarendra et al. [32] presented four evaluation criteria.
to track the progressive bearing defect. In the present study, another evaluation criterion based on exponential function is proposed. It is well known that as bearing deteriorates, certain features will increase exponentially. Hence, an evaluation criterion \( E \) obtained from the coefficient of the exponential curve fitting of the feature being evaluated is used as illustrated below. Suppose \( f(t) \) is the monitored evaluation function.

\[
f(t) = a e^{b(t)}
\]

where \( a \) and \( E \) are the curve fitting exponential coefficients, and \( t \) is the measurement days. The calculation of \( E \) for the time-domain kurtosis feature and the LLE feature is presented in Figure 9. The two figures show clearly that the LLE feature shows a greater exponential increasing trend than the time-domain kurtosis feature. The dotted line in Figure 9a,b is the exponential curve fitted from day 80 to day 139.

![Figure 9](image)

**Figure 9.** Evaluation criterion \( (E) \) calculation: (a) time domain kurtosis feature; (b) LLE feature.

In Step (2), kernel regression is employed to predict the future bearing condition and to estimate the RUL of the slew bearing. This prediction starts when there is warning of impending deterioration of the bearing condition. The basic kernel regression algorithm is improved by adding a \( k \)-step-ahead subroutine. At each \( k \)-step, non-linear regression model and posterior (predicted) point is estimated and updated. The purpose of this process is to estimate a one-step-ahead prediction of the kernel regression model based on past non-linear data points.

Kernel regression is an effective method to build a model from non-linear data. The process is done by taking the highest data point of the current kernel regression model and then using it as the next data point (e.g., predict the 4th data point if the initial model is built from three data points). The new data point is then used to build a new kernel regression model and estimate the next data point. The process is repeated until the specified \( k \)-step-ahead prediction is reached. To illustrate the process, a time domain kurtosis feature is used as the monitored parameter being predicted.

Kurtosis is selected as an input of kernel regression because it has a certain value for normal bearing condition and the kurtosis value for defect condition. This is necessary for the detection of incipient defect in Step (1), which is set with the predetermined warning threshold and for prediction in Step (2), which is set with the damage bearing threshold. The kurtosis value for normal bearing signals is approximately 3 [40]. The increase of this value indicates that the bearing condition has changed. Some research mentioned that the kurtosis value will reach about 50 when high impact occur [45]. Aye [46] studied the kurtosis of normal and faulty tapered bearings running at speed between 409 RPM and 1200 RPM. It has been found that the kurtosis value for normal bearings is 2.89. Surprisingly, the kurtosis value for faulty bearings with 409 RPM increases significantly to about 30.26 and for 1200 RPM drops to 1.23. The same author [46] demonstrated that the kurtosis value of bearing damage with low rotational speed (low RPM) can be much higher than the normal value. Guo et al. [47] presented a method for recovering the bearing signal from large noisy signal using
a hybrid method based on spectral kurtosis and ensemble EMD. It was found that the kurtosis value of each bearing fault condition was recovered and identified more easily. For instance, the kurtosis value of a bearing with an inner race defect and a ball defect is 18.35 and 41.20, respectively.

Supposed 3 points are available denoted as point 1, 2, and 3 as shown in Figure 10. Firstly, the method predict one-step-ahead of degradation parameter denoted as point 4 (red triangle) based on 3 points i.e., point 1, 2 and 3. Secondly, to predict next degradation parameter denoted as point 5, the method used point 2, 3 and 4. Thirdly, to predict point 6 the method used point 3, 4 and 5. According to the Figure 10, the plotted curve shows an increasing trend, it is due to the first 3 points used has an increasing trend. Thus for next prediction the method employed both actual value and predicted value to predict the next degradation parameter. In practice, when the kurtosis value is acquired from online condition monitoring, the points can be replaced by actual kurtosis value. The method will use 3 updated kurtosis values to do one-step-ahead prediction.

![Figure 10. Kernel regression model and three-step-ahead prediction.](image)

Setting a threshold level are depends on the different situations and types of bearings. The information can be obtained from the maintenance engineer or past condition monitoring (CM) data with has similar type of bearing and load condition. As has been done in past works [45–47], we have set the threshold kurtosis value of 45 for the time-domain kurtosis and the WD feature, while the threshold for the LLE feature is set at 35. The proposed kernel regression method was applied up to 120 days, as by day 120 the kurtosis value has exceeded 4 times the normal value and three data points greater than the alarm threshold level have been detected. It should be noted that three data points is the minimum requirement in the initial kernel regression model. Note that the levels of the three selected features up to day 120 are still below the bearing damage threshold of 45. The predictions of the kernel regression of the three features mentioned above are presented in Figure 11a–c. The four step-ahead of kernel regression are used (k = 4) and thus the 4 step-ahead predicted data points are shown with ‘Δ’ symbol, while ‘o’ represents the actual value. It can be seen that the predicted future values increases. When the data point exceeds the damage threshold level, the last predicted day is noted and used to estimate the histogram of final failure. Once the histogram of final failure is obtained, the RUL can be easily estimated by taking the difference between the predicted final failure time and the last measurement day (the last measurement day is 120).

Based on this analysis, the predicted failure day of the bearing is shown to be dependent on the features and summarized in Table 8. Detail information of the three non-linear data used to build the initial model and the predicted data point of the modified kernel regression method for each feature is shown in Tables 9–11 (point A–D). It is shown in Figure 11b,c is the WD kurtosis feature and LLE feature, respectively. For WD kurtosis feature the spacing is different (day 104th – day 91th = 13 days and day 107th – day 104th = 3 days). The LLE Feature also has different spacing (day 104th – day 92th = 12 days and day 114th – day 104th = 10 days). Points D–G are the predicted features values. These points are estimated based on the 4-step-ahead prediction of the kernel regression. According to
the results in Tables 9–11, the y-axis values of the fourth predicted points exceed the bearing damage threshold. The day of each fourth predicted point (point G) is noted as the final failure of the slew bearing. Three predicted days (point G) are then inputted to the histogram. The result of this is presented in Figure 11d. It should be noted that with more features used, the more accurate the distribution fit of the histogram will be. It can be seen from Figure 11d that the predicted final failure is day 135. Further, the RUL of the bearing can be estimated by taking the difference between the predicted failure day and the last measurement day i.e., $135 - 120 = 15$ day. The final failure histogram was calculated using a common MATLAB function ‘histfit’. The probability curve is plotted based on the predicted final failure from three methods: (a) time-domain kurtosis at 137.6 day; (b) WD kurtosis at 130.8 day; and (c) LLE feature at 137 day.

During the slew bearing lab experiment, complete failure is unpredictable. When a burst vibration signal on day 141 was detected, the test-rig had to be shut down. To inspect the damage and verify the result of the proposed prognosis method, the slew bearing was dismantled and inspected. Some of the defective rollers and outer race can be clearly seen in Figure 12. The actual final failure day of lab slew bearing is considered one day before the severe burst vibration signal i.e., day 140. The accuracy of prediction is estimated as follows:

$$A_p = \left(1 - \frac{|t_a - t_p|}{t_a}\right) \times 100\% = \left(1 - \frac{|140 - 135|}{140}\right) \times 100\% = 96.43\%$$ (15)

where $t_a$ is actual final failure day and $t_p$ is predicted failure day.

![Figure 11](image1.png)

Figure 11. Prediction of modified Kernel regression: (a) Time-domain kurtosis; (b) Wavelet kurtosis; (c) LLE feature; (d) PDF of the predicted failure day of the slew bearing. 135.2 is the peak value of the normal distribution.
Table 8. The prediction of failure day.

<table>
<thead>
<tr>
<th>Feature</th>
<th>nth Day Failure (x-axis)</th>
<th>Feature Level (y-axis)</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time-domain kurtosis</td>
<td>137.6</td>
<td>55</td>
<td>45</td>
</tr>
<tr>
<td>WD kurtosis</td>
<td>130.8</td>
<td>46</td>
<td>45</td>
</tr>
<tr>
<td>LLE feature</td>
<td>137</td>
<td>39</td>
<td>35</td>
</tr>
</tbody>
</table>

Table 9. Original non-linear data points and predicted data points of time domain kurtosis.

<table>
<thead>
<tr>
<th>Data Point</th>
<th>Measurement Day</th>
<th>Feature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>92</td>
<td>19</td>
</tr>
<tr>
<td>B</td>
<td>103</td>
<td>24.12</td>
</tr>
<tr>
<td>C</td>
<td>114</td>
<td>36.83</td>
</tr>
<tr>
<td>D</td>
<td>120</td>
<td>39</td>
</tr>
<tr>
<td>E</td>
<td>128.6</td>
<td>43</td>
</tr>
<tr>
<td>F</td>
<td>129.2</td>
<td>47</td>
</tr>
<tr>
<td>G</td>
<td>137.6</td>
<td>55</td>
</tr>
</tbody>
</table>

Table 10. Original non-linear data points and predicted data points of WD kurtosis.

<table>
<thead>
<tr>
<th>Data Point</th>
<th>Measurement Day</th>
<th>Feature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>91</td>
<td>7.24</td>
</tr>
<tr>
<td>B</td>
<td>104</td>
<td>8.84</td>
</tr>
<tr>
<td>C</td>
<td>107</td>
<td>14</td>
</tr>
<tr>
<td>D</td>
<td>116.4</td>
<td>13</td>
</tr>
<tr>
<td>E</td>
<td>122.2</td>
<td>32</td>
</tr>
<tr>
<td>F</td>
<td>125</td>
<td>46</td>
</tr>
<tr>
<td>G</td>
<td>130.8</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 11. Original non-linear data points and predicted data points of LLE feature.

<table>
<thead>
<tr>
<th>Data Point</th>
<th>Measurement Day</th>
<th>Feature Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>92</td>
<td>12.36</td>
</tr>
<tr>
<td>B</td>
<td>104</td>
<td>12.22</td>
</tr>
<tr>
<td>C</td>
<td>114</td>
<td>25.69</td>
</tr>
<tr>
<td>D</td>
<td>117.8</td>
<td>26</td>
</tr>
<tr>
<td>E</td>
<td>118.8</td>
<td>28</td>
</tr>
<tr>
<td>F</td>
<td>128</td>
<td>35</td>
</tr>
<tr>
<td>G</td>
<td>137</td>
<td>39</td>
</tr>
</tbody>
</table>

Figure 12. Slew bearing damage: (a) A view of damaged rollers in the axial plane; (b) Outer raceway damage.

5. Conclusions

An integrated condition monitoring and prognosis method of low speed slew bearing has been presented. The method employs combined MSET and SPRT to detect the incipient bearing defect,
and kernel regression to predict the future state and to estimate the RUL. The method has been implemented in the monitoring and prediction of laboratory slew bearing condition, where the test was run with new bearing condition until failure. Combined MSET and SPRT method was used to analyse the vibration features calculated from the accelerometer data of 139 days measurement. Based on the set of hypothesis in SPRT method, it was found that the incipient defect of slew bearing occurred in day 90. After the incipient defect was detected, the kernel regression calculation started to predict the future state. To validate the method, 120 measurement days were analysed. The predicted failure days were constructed and the final failure day (day 135) was predicted based on the peak of the normal distribution fit of the histogram. The RUL was estimated by to be 15 days (135 – 120 days).

Acknowledgments: The first author thanks the University of Wollongong, Australia for the financial support through University Postgraduate Award (UPA) and International Postgraduate Tuition Award (IPTA) during this study. The authors gratefully acknowledge the help of Madeleine Strong Cincotta in the final English editing of this paper.

Author Contributions: Wahyu Caesarendra conducted Data curation, investigation, data processing, formal analysis and original draft preparation. Tegoeh Tjahjowidodo provided Funding acquisition, formal analysis, review and editing. Buyung Kosasih conducted Conceptualization, resources, review and editing. Anh Kiet Tieu provided Funding acquisition, conceptualization, resources, review and editing.

Conflicts of Interest: The authors declare no conflict of interest. The authors of this manuscript certify that they have NO affiliations with or involvement in any organization or entity with any financial interest.

References


47. Guo, W.; Tse, P.W.; Djordjevich, A. Faulty bearing signal recovery from large noise using a hybrid method based on spectral kurtosis and ensemble empirical mode decomposition. Measurement 2012, 45, 1308–1322. [CrossRef]

© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).