Journal of Physics: Conference Series

PAPER • OPEN ACCESS

Vibration Signal for Bearing Fault Detection using Random Forest

To cite this article: Tarek Abedin *et al* 2023 *J. Phys.: Conf. Ser.* **2467** 012017

View the article online for updates and enhancements.

You may also like

- 3D vertebrae labeling in spine CT: an accurate, memory-efficient (Ortho2D) framework
  Y Huang, A Uneri, CK Jones et al.

- Signal quality in cardiorespiratory monitoring
  Gari D Clifford and George B Moody

- Accuracy Degradation Model and Residual Accuracy Life Prediction of CNC Machine Tools Based on Wiener Process
  Shilin Li, Yan Ran, Genbao Zhang et al.

This content was downloaded from IP address 103.53.32.15 on 06/09/2023 at 02:32

Vibration Signal for Bearing Fault Detection using Random Forest

Tarek Abedin[1], S. P. Koh[2*], Chong Tak Yaw[2*], Chen Chai Phing[1], Sieh Kiong Tiong[2], Jian Ding Tan[3], Kharudin Ali[4], K. Kadirgama[5,6,7], and F. Benedict[8]

[1] Department of Electrical and Electronics Engineering, Universiti Tenaga Nasional (The National Energy University), Jalan IKRAM-UNITEN, 43000 Kajang, Selangor, Malaysia.
[2] Institute of Sustainable Energy, Universiti Tenaga Nasional (The National Energy University), Jalan IKRAM-UNITEN, 43000 Kajang, Selangor, Malaysia.
[3] Xiamen University Malaysia, Jalan Sunsuria, Bandar Sunsuria, 43900 Sepang, Selangor Darul Ehsan, Malaysia.
[4] Faculty of Electrical and Automation Engineering Technology, UC TATI, Teluk Kalong, Kemaman 24000, Terengganu, Malaysia.
[5] Advance Nano Coolant-Lubricant (ANCL), College of Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia.
[6] Faculty of Mechanical and Automotive Engineering Technology, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia.
[7] Automotive Engineering Centre, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia.
[8] No.9, Jalan Meranti Jaya 12, Meranti Jaya Industrial Park, 47120, Puchong, Selangor, 47120 Puchong, Selangor, Malaysia.

Corresponding author: Email: chongty@uniten.edu.my; johnnykoh@uniten.edu.my;

Abstract. Based on the chosen properties of an induction motor, a random forest (RF) classifier, a machine learning technique, is examined in this study for bearing failure detection. A time-varying actual dataset with four distinct bearing states was used to evaluate the suggested methodology. The primary objective of this research is to evaluate the bearing defect detection accuracy of the RF classifier. First, run four loops that cycle over each feature of the data frame corresponding to the daytime index to determine the bearing states. There were 465 repetitions of the inner race fault and the roller element fault in test 1, 218 repetitions of the outer race fault in test 2, and 6324 repetitions of the outer race in test 3. Secondly, the task is to find the data for the typical bearing data procedure to differentiate between normal and erroneous data. Out of 3 tests, (22-23) % normal data was obtained since every bearing beginning to degrade usually exhibits some form of a spike in many locations, or the bearing is not operating at its optimum speed. Thirdly, to display and comprehend the data in a 2D and 3D environment, Principal Component Analysis (PCA) is performed. Fourth, the RF algorithm classifier recognized the data frame's actual predictions, which were 99% correct for normal bearings, 97% accurate for outer races, 94% accurate for inner races, and 97% accurate for roller element faults. It is thus concluded that the proposed algorithm is capable to identify the bearing faults.

Keywords: Bearing, Fault Detection, Principal Component Analysis, Random Forest.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

1. Introduction

One of the essential parts of a mechanical spinning system is a bearing. It is a highly standardized, precise mechanical device with a high rate of labour efficiency, minimal friction, and ease of assembly and operation [1]. The inner race, outer race, ball elements, and retainer comprise most rolling bearings. The failure of the inner race, outer race, and ball elements is the most typical sort of rolling bearing fault. Bearing faults often manifest themselves as various component flaws. Statistics show that bearing problems account for more than 40% of motor malfunctions [2]. Figure 1 depicts the rolling bearings' usual life curve. Four phases are involved: Running-in, normal operation, early weak fault onset and healing, and severe fault are the four phases.

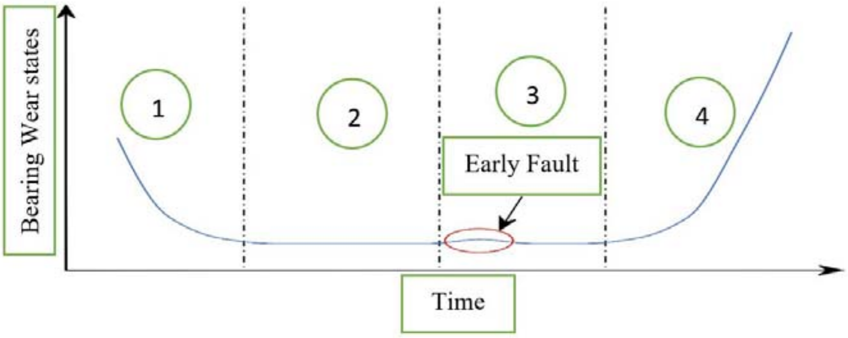


Figure 1. Typical roller-bearing life curve [3].

The traditional signal processing techniques' time, frequency, and time-frequency-domain analyses are all possible [4]. Methods like the fast Fourier transform (FFT) [5], wavelet transformation (WT) [6], empirical mode decomposition (EMD) [7], ensemble empirical mode decomposition (EEMD) [8], empirical wavelet transforms (EWT) [9], wavelet packet transform (WPT) [10], variational mode decomposition (VMD) [11], stochastic resonance, sparse decomposition, etc. are used for the analysis and classification of bearing fault signals. This study measures the root mean square value (RMS), maximum and lowest mean standard deviations (unbiased std), skewness, kurtosis, crest factor, and form factor using the Python programming language. The mode classifier should be developed after feature extraction to achieve automated fault detection. Many studies have reported machine learning (ML) and deep learning (DL)-based detection models. Artificial neural networks (ANN), support vector machines (SVM), extreme learning machines (ELM), K-Nearest Neighbor (KNN), Hidden Markov models (HMMs), auto-encoders, convolutional neural networks (CNNs), deep belief networks, generative adversarial networks, and recurrent neural networks (RNN) are frequently used to create these models [12–20]. To enhance generalisation performance and lessen the overfitting problem of the single decision tree, Breima et al. [21] presented the random forests (RF) classification strategy based on the notion of the bagging technique [22] and integrating random selection [23]. RF has shown strong classification performance when handling high-dimensional and small-sample issues compared to the conventional decision tree approach. Random Forest (RF) has recently shown strong generalization performance for numerous pattern recognition scenarios as a classification approach. A diagnosis method was carried out in [24] employing a classifier based on RF and a genetic algorithm. In [25], the authors utilised feature rankings for several rotating equipment fault categories and used K-nearest neighbours and RF as the classifier. In order to enhance the diversity of classification trees and the performance of individual classification trees, the weighted voting rule was used for random forests [26]. In great detail, studies comparing the weighted RF and traditional RF were also carried out.

According to experimental comparison data, the suggested technique outperforms previous methods in terms of diagnostic defect accuracy. Three features are the key contributions of this work:

- First, 9-time domain characteristics from Tests 1, 2, and 3 were retrieved correspondingly. Using libraries for visualizing, charting, and mathematical feature extraction, such as *matplotlib, scipy, numpy,* and *pandas*, the significant and representative features with sufficient defect information are recovered from the python language.
- Normal data should also be segregated to distinguish between normal and defective data. The probability of having normal data in such a situation is (22%–23%) since any bearing beginning to degrade always exhibits some form of a spike in several locations or is not operating at peak efficiency.
- Finally, the same libraries as before were used to classify errors using the random forest approach. In this investigation, the test side accuracy of the RF approach is fairly excellent, reaching a high classification precision of 95.58%.

1.1. Problem statement

Many methodologies have been put forward in the available literature to identify and categorize bearing problems in IMs. Current methods still have a number of practical shortcomings even if a high classification accuracy has been attained in diagnosing bearing problems. For defect detection in rotating equipment in the time domain, frequency domain (rapid Fourier transform), and time-frequency domain, Chuan Li et al. [27] developed a deep statistical feature learning technique (wavelet packet transform). Fast Fourier Transform (FFT) has spectrum leakage-related issues built-in, nevertheless. In order to identify bearing flaws in IMs, Qingbo He et al. [28] presented novel research to investigate the wavelet packet transform (WPT) flow characteristics of vibration signals. The rolling element-bearing vibration data are pre-processed using empirical model decomposition [29]. (EMD). Even though EMD is a more signal-adaptable algorithm than WPT, it still has mode-mixing and end-effect issues. A unique approach for predicting numerous failure modes in rotating equipment was proposed in [30]. The technique combines non-parametric cumulative incidence functions with a machine learning and pattern recognition approach known as logical analysis of data (LAD). The bearing defect detection approach put forward in this study offers certain benefits over the methods currently being employed for real-time analysis of the suggested method. This technique may reduce the impact of random and asynchronous noise in the vibration signal in the first scenario. Secondly, unlike EMD, this method-based signal analysis is not afflicted by mode mixing and end effect issues. Additionally, refrain from using associated or irrelevant features in machine learning models. In light of its benefits, an RF classifier was chosen because of its ease of deployment, high prediction accuracy, and little need for model adjustment. The RF model has the potential for real-time categorization and is simple to train, evaluate, and apply to a local system. The approach used in this paper to find rolling bearing faults in IMs is based on Python and RF languages.

1.2. Research Objectives

- To predict the faults in a bearing applying Python language and RF classifier that are used to train machine learning algorithm.
- To evaluate the meaningful feature extraction from obtained data using matplotlib, scipy, numpy, pandas' libraries which are usually used for, visualizing, plotting and mathematical.
- To convert this problem as a classification task data is visualized and Fault-labels are created.
- To visualize and better understand the data in a 2D, and 3D space dimensionality reduction technique, more specifically principal compact analysis (PCA) is used.
- To evaluate the accuracy of bearing faults and performance on the test run by trained RF algorithm.

2. Theoretical Background

2.1 Bearing Faults Signature

The most frequent issues with induction motors (IMs) account for around 40% of all faults, according to failure surveys performed by the Electric Power Research Institute (EPRI) [31–33]. The inner and outer races of a rolling bearing are separated from one another by rolling components like balls or cylindrical rollers [34]. These components may experience material fatigue or wear, resulting in flaking and pitting [35]. Single-point flaws and generalized roughness are the two categories into which bearing faults may be divided. Single-point flaws are localized and divided into rolling element flaws, inner raceway flaws, and outer raceway flaws. The bearing components and three of the most typical localized bearing faults—outer race, inner race, and rolling element faults—are shown in Figure 2.

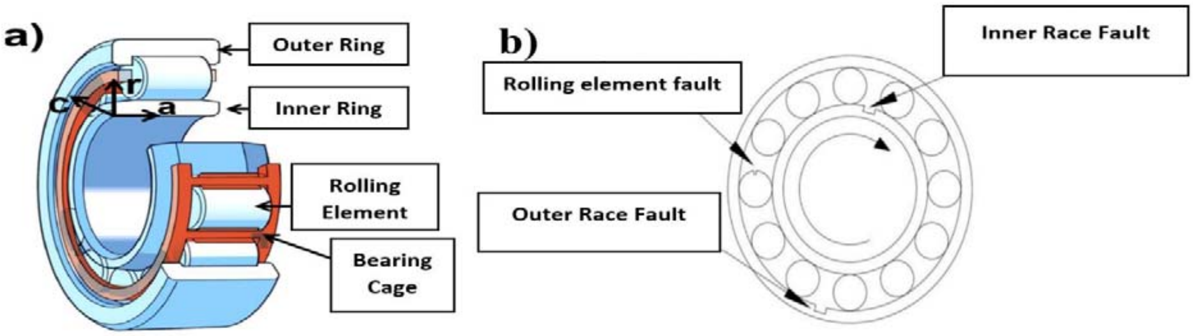


Figure 2. a) Bearing components [36]. b) Schematic illustrates localized flaws on a rolling element and in the inner and outer races [37].

When a ball travels through an area damaged in a bearing, shock pulses with a certain frequency occur. The rolling element's shape and rotational frequency f_m may be used to derive the characteristic frequency. The defect introduces abnormalities in the air-gap flux density via vibrations, which results in an aberration in the current signal's harmonics. The current signal's variation reveals the existence of bearing problems. Below is a list of the various characteristic frequencies for each defect specified in [35].

- Inner race fault: $f_{inner} = f_m \frac{N_{ball}}{2}(1 + \frac{D_{ball}}{D_{cage}} cos\beta)$ (1)
- Outer race fault: $f_{outer} = f_m \frac{N_{ball}}{2}(1 - \frac{D_{ball}}{D_{cage}} cos\beta)$ (2)
- Ball or rolling element fault: $f_{ball} = f_m \frac{D_{cage}}{D_{ball}} (1 - \frac{D_{ball}^2}{D_{cage}^2} cos^2\beta)$ (3)
- Cage fault: $f_{cage} = f_m \frac{1}{2}(1 - \frac{D_{ball}}{D_{cage}} cos\beta)$ (4)

where β is the contact angle of the balls, N_{ball} is the number of balls, D_{ball} is the diameter of the ball, and D_{cage} is the cage diameter, commonly referred to as the ball or roller pitch diameter.

Due to bearing degradation, the stator and rotor move radially, causing oscillations that introduce recognisable fault frequencies into the current signals. The cause of this is the rotor's radial displacement concerning the stator as a result of the bearing problems. It causes the spinning eccentricity and load torque to fluctuate. This causes the machine inductances to fluctuate, which modulates the motor-current signals' amplitude, frequency, and phase. According to [2][38], the motor current equation for a bad bearing is as follows:

- $i(t) = \sum_{k=1}^{\infty} i_k \cos(\omega_{c_k} t + \varphi)$ (5)

Together with phase angle φ and angular velocity ω_{c_k}, an expression for the angular velocity is:

- $\omega_{c_k} = \frac{2\pi f_{bearing}}{p}$ (6)

Here, p is the machine's specific pole pair number, and $f_{bearing}$ is the fault current harmonic frequency. Notably, f bearing may be expressed as follows:

- $f_{bearing} = |f_s \pm m f_v|$ (7)

Where f_v may either be the inner race defect frequency (f_{inner}) or the outer race defect frequency (f_{outer}), and f_s is the supply or fundamental frequency, $m = 1,2,3, \ldots\ldots.$ and indicates the harmonic indexes. The frequency auto-search technique described in [39] may be used to determine the estimated fault signature frequencies. The harmonics generated by bearing failures may sometimes overlap or be near noise frequencies, making it challenging to tell them apart and bearing fault identification [40]. Therefore, it is challenging to identify the bearing faults in an IM when the bearing specifications are unknown, and the inverter frequency fluctuates.

2.2 Feature Extraction

By dividing the original raw data into smaller, easier-to-process groupings, feature extraction minimises the data size. The raw dataset often has a lot of variables. Therefore, processing it takes a lot of CPU power. Because of its capacity to identify developing flaws, online diagnostics for condition monitoring has recently attracted interest. Directly measured signals are insufficient for online usage since a tiny quantity of data is insufficient for diagnosis. A large data sample is necessary for efficient defect identification. Therefore, feature extraction becomes an essential phase that saves important information for reaching the final selection to simplify the computation [41].

The vibration signal is first segmented in a time-domain technique before the specific statistical properties of each segment are assessed. These characteristics could provide a statistical analysis of the signal, which might assist in elucidating certain hidden data from the unprocessed vibration signals of a sound or damaged bearing [42] . This study used time domain data to generate nine traditional statistical feature parameters for condition evaluation. Maximum, minimum, mean, median, standard deviation, skewness, kurtosis, root mean square, and crest factor are the temporal domain characteristics retrieved from the signal. These characteristics look at the signal's probability density function (PDF). The median is the number in the center when the data points are sorted in ascending order, while the mean is the average of the available data points. The standard deviation may be calculated by calculating the variance's square root. It is a statistic that measures the spread of observations over data collection. It is healthy knowledge that as a bearing's state changes, the PDF likewise does, making variations in skewness and kurtosis detectable. The signal's skewness may also be utilised to determine whether it is favourably or negatively skewed. The PDF's peak value is measured by kurtosis, which shows that the signal is impulsive. According to the definition of the moment, the skewness of a signal with a normal distribution, or a normal bearing signal, is equal to zero. Positive values suggest non-symmetry towards higher values, whereas negative values are caused by skewness towards lower values [43]. Kurtosis is a further measurement that may be made since skewness is determined by the mean value of the vibration signal's PDF, while the PDF's peak is used to determine the fourth-order moment. It is well known that a standard bearing's vibration signal has a skewness value of around zero and a kurtosis value of three [80]. The kurtosis value will then rise to over three when the vibration signal changes due to defects, and the skewness value will change to either negative or positive. The RMS value will steadily rise as the fault develops. RMS, however, has the drawback of being unable to offer information on the developing fault stage while it grows with the development of the issue. Last, the crest factor shows how a dataset's peak value compares to its practical value. The crest factor quantifies the impact. For "spiky signals," crest factor is adequate [43]. Table 1 contains the mathematical formulas for the statistical aspects that have been mentioned. Here, x_i, $i = 1,2,3, \ldots, N$ represents the motor current signal. N, μ, and σ represent the number of data points, mean and standard deviation respectively. The time-domain characteristics are of high quality because of their sensitivities and because they provide statistical information about the current signal [44].

Table 1. Features for the feature matrix were extracted from time domains (x is the current signal)

Feature name	Formula
Maximum	$max\|x_i\|$
Minimum	$min\|x_i\|$
Mean	$\mu = \frac{1}{N}\sum_{i-1}^{N} x_i$
Median	$median = \left(\frac{(N+1)^{th}}{2}\right)$
Standard deviation	$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N-1}}$
Skewness	$skewness = \frac{1}{N}\frac{\sum_{i=1}^{N}(x_i-\mu)^3}{\sigma^3}$
Kurtosis	$Kurtosis = \frac{1}{N}\frac{\sum_{i=1}^{N}(x_i-\mu)^4}{\sigma^4}$
Root mean square	$X_{rms} = \frac{\sqrt{\sum_{i=1}^{N} x_i^2}}{N}$
Crest factor	$C_f = \frac{X_{max}}{X_{min}}$
Form factor	$F_f = \frac{X_{rms}}{mean}$

2.3 Dimensionality Reduction Technique

Visualizing the training set and working on it becomes more challenging as the number of characteristics increases. These traits may sometimes be redundant or connected. Dimensionality reduction is thus an essential technique for ML applications (DR). There are two primary methods for performing DR: (I) keeping only the most pertinent features from the original dataset (generally referred to as feature selection); and (II) shrinking the original dataset into a new one through analysis or combination of the input variables, where the new dataset essentially contains the same information as the original (generally referred to as dimensionality reduction) [45]. Projective approaches and manifold modelling are two different categories of DR techniques. Principal component analysis (PCA) [46], probabilistic PCA [47], and Gaussian generalised discriminant analysis (GDA) [48] are projective approaches. The following methods are taken into consideration for the various techniques: local linear embedding (LLE) [47], stochastic neighbour embedding (SNE) [49], t-distributed stochastic neighbour embedding (t-SNE) [49], neighbourhood preserving embedding (NPE) [50], locality preserving projection (LPP) [51], stochastic proximity embedding (SPE), and isometric feature mapping (Isomap) [52]. PCA was employed in this study to display the data in both 2D and 3D formats. Primary component analysis (PCA), a standard statistical analytic technique, may be used to identify the principal components. This technique is often utilised in data dimension reduction fault diagnosis, feature extraction and fusion fault diagnosis, high-dimensionality visualisation fault diagnosis, data regression fault diagnosis, etc. It is possible to think about PCA as a transformation that projects the original data into a smaller-dimensional new space.

2.4 Algorithm Development for Random Forests (RF)

A classifier called random forest is made up of many decision-tree classifiers. The creation of the algorithm is broken down into the following 3 phases:

1. From the initial data collection, T, training samples are taken and returned using the Bootstrap sampling technique. Similar to the first data set, there are the same number of samples. In the event when X is a collection of data with n samples, $\{x_1, x_2, \ldots. x_n\}$, a sample $x_i(i = 1{,}2\ldots.n)$ is taken from

the primary data collection X. To merge it into a new set X^*, n times are required. Without a sample, x_j, the probability of X^* is therefore:

- $p = \left(1 - \frac{1}{n}\right)^n$ (8)
- $\lim_{n\to\infty} p = \lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^n = e^{-1} \approx 0.368$ (9)

About 36.8% of the samples in the original data set won't be retrieved when n is big enough. When this is the case, the random forest's decision tree cannot find a local optimum solution. As a result, it can acquire a stronger classifier and successfully prevent that anomalous data from showing up in the sample set. Meanwhile, the generalization error, correlation coefficient, and decision-tree intensity are estimated using the undiscovered Out-Of-Bag (OOB). As a result, the classification accuracy of the algorithm may be measured.

2. The T decision-tree models $h_i(X^*, \Theta_k)$ are built for the T training samples $(X_1^*, X_2^*, \ldots.. X_T^*)$ in which $i = 1,2 \ldots . T$, $K = 1,2 \ldots .$

The decision tree model stated in [41] is shown in Equation (10).

- $c(x_1, x_2, \ldots . x_n, h_t) = \begin{cases} label(h_t) & h_t\ is\ the\ leaf\ node \\ c(x_1, x_2, \ldots . x_n, h_t) & h_t\ is\ the\ inner\ node \end{cases}$ (10)
- $h_i(X^*, \Theta_k) = c(x_1, x_2, \ldots . x_n, root(h_i))$ (11)

Where $root(h_i)$ is the root node of the decision tree $h_i(X^*, \Theta_k)$. $c(x_1, x_2, \ldots . x_n, h_t)$ is the segmentation criterion of the decision tree $h_i(X^*, \Theta_k)$.

The random forest generated by T decision trees is used to categorise the test sample. Each tree has a vote privilege that determines how the categorisation will turn out. The decision-tree output categories that have been most fully classified make up the final classification result.

- $H(x) = arg \max_Y \sum_{i=1}^{T} I(h_i(X^*, \Theta_k) = Y)$ (12)

where Y is the output tag variable; I (*) is the indicator function; and $h_i(X^*, \Theta_k)$ is the single decision tree. The design and testing of the random forest are shown in Figure 3.

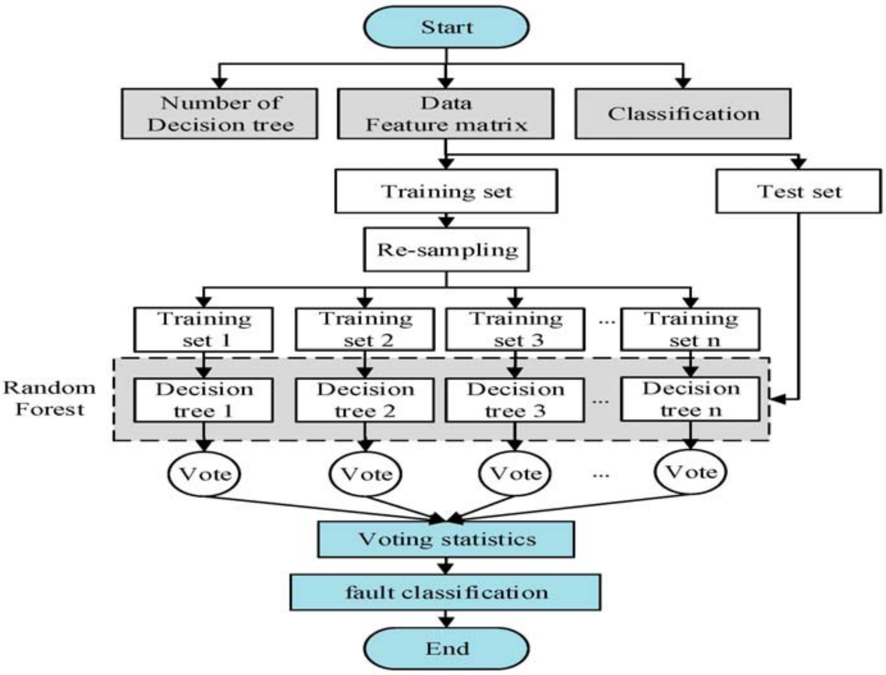


Figure 3. Flow chart of random forests [53].

2.5 Accuracy Prediction

Generally speaking, the prediction accuracy improves with the number of estimators. Beyond a threshold number of trees, adding additional trees results in no appreciable speed increase and raises computational demand. The number of trees mentioned in the literature is 128 [54], 200 [55], or 250 [56]. A number of classification criteria are utilised to evaluate the trained RF algorithm's prediction accuracy [57][58]. The accuracy (q_a), which is the ratio of the number of accurate predictions (N_T) to the total number of samples (N), is the easiest statistic to understand, i.e.,

- $q_a = \frac{N_T}{N}$ (13)

A classification is considered True Positive (N_{TP}) if a sample labelled as positive is also anticipated to be positive. False Negative (N_{FN}) is the categorization if it is expected to be negative. False Positives (N_{FP}) and True Negatives (N_{TN}) are defined similarly. A 2×2 confusion matrix C may be used to show these four integers. The implementation used in this work is based on scikit-learn [59]; other sources, such as [60], may use a transposed version.

- $C = \begin{matrix} N_{TP} & N_{FN} \\ N_{FP} & N_{TN} \end{matrix}$ (14)

The equation gives the number of accurate predictions using the four definitions (15)

- $N_T = N_{TP} + N_{TN}$ (15)

The precision (q_p) or confidence is defined as the fraction of all positively predicted samples (N_{PP}), which are labelled as positive (N_{TP}), i.e.,

The proportion of all positively predicted samples (N_{PP}) that are labelled as positive (N_{TP}), or the precision (q_p) or confidence, is defined as the precision (q_p), i.e.,

- $q_p = \frac{N_{TP}}{N_{PP}} = \frac{N_{TP}}{N_{TP}+N_{FP}}$ (16)

Conversely, the recall (q_r) or sensitivity gives the fraction of all positively labelled samples (N_{PL}), which are correctly identified as positive, i.e.,

The recall (q_r) or sensitivity, on the other hand, provides the percentage of all positively labelled samples (N_{PL}), that are accurately detected as positive, i.e.

- $q_r = \frac{N_{TP}}{N_{PL}} = \frac{N_{TP}}{N_{TP}+N_{FN}}$ (17)

In the case of multi-label classification, precision and recall values are calculated separately for each class, with 'positives' meaning samples belonging to the respective class. Each row in the confusion matrix represents a 'true' class, with the 'predicted' class labels as columns. In this case, the confusion matrix contains the number of correct predictions of each class in the diagonal, and false predictions are contained in the respective off-diagonal elements. Given a classification with N labels, the precision and recall can be calculated separately for each class (denoted by index $i, i = 1 \ldots N$ from the coefficients of the $N \times N$ confusion matrix as follows:

- $q_p^{(i)} = \frac{C_{ii}}{\sum_{j=1}^{N} C_{ji}}$ (18)
- $q_r^{(i)} = \frac{C_{ii}}{\sum_{j=1}^{N} C_{ij}}$ (19)

3. Experimental Setup and Methodology

3.1 Test Rig Setup

A shaft has four bearings placed on it. An AC motor connected to the shaft by rub belts controlled the rotation speed to remain constant at 2000 RPM. A spring mechanism provides a radial force of 6000 lbs on the shaft and bearing. All bearings are greased by force. As shown in Figure 4, Rexnord ZA-2115 double-row bearings were mounted on the shaft. PCB 353B33 High-sensitivity Quartz ICP accelerometers were placed on the bearing housing (two accelerometers for each bearing [x- and y-axes] for data set 1, one accelerometer for each bearing for data sets 2 and 3). Figure 4 shows displays the positioning of the sensors. All failures happened after the bearing had completed more than its intended lifespan of more than 100 million rotations.

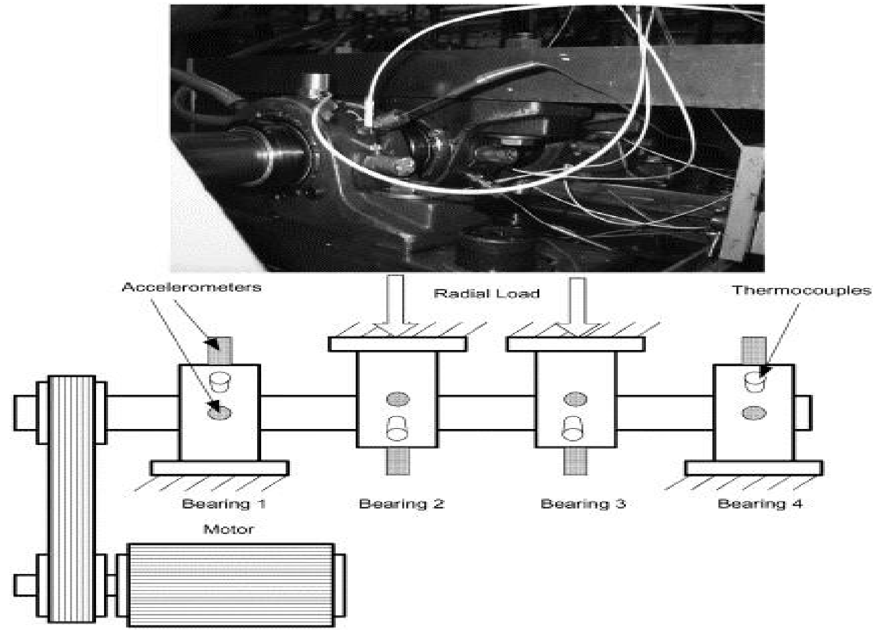


Figure 4. Illustration of a sensor installation and a bearing test setup.

Table 2. Test characteristics

Test no.	Quantity of accelerometers	Test Period	Faults
Test 1	8	49680 min 34 days 12h	Inner race [Bearing 3] Rolling element [Bearing 4]
Test 2	4	9840 min 6 days 20h	Outer race [Bearing 1]
Test 3	4	44480 min 31 days 10h	Outer race [Bearing 3]

3.2 Data Structure

The data packet contains three (3) data sets (IMS-Rexnord Bearing Data.zip). A test-to-failure experiment is described in each piece of data. Each data set comprises individual files, including snapshots of the vibration signal taken every second at predetermined intervals. Each file has a sampling rate of 20 kHz and a total point count of 20,480. The file name indicates the data collection date. A data point is a record (row) in a data file. NI DAQ Card 6062E enabled data collecting. Greater time stamp intervals indicate that the experiment will resume the next working day.

Table 3. Bearing characteristics

Rexnord ZA-2115 variables	Values
Pitch diameter (mm)	71.5
Rolling element diameter (mm)	8.4
Number of rolling elements per row	16
Contact angle (◦)	15.17

3.3 Methodology

This test set illustrates how a machine-learning method may be used to foresee bearing failures. In the first section, look at the data pre-treatment phase and learn to take data from a text file and extract features from the resulting data to create useful features. In this instance, time domain characteristics were retrieved from the provided acceleration data and stored in a CSV file. The pre-processed data will be used later in this series to train machine learning systems that can anticipate bearing failures. Random forests are supervised learning techniques. With it, one can do both classification and regression. The algorithm's flexibility and usability make it the best choice. Random Forest creates decision trees on randomly selected data samples, gets a prediction from each tree, and selects the best alternative. It also provides a reasonably precise measurement of the utility of the feature. Figure 5 depicts the complete method of bearing faults detection.

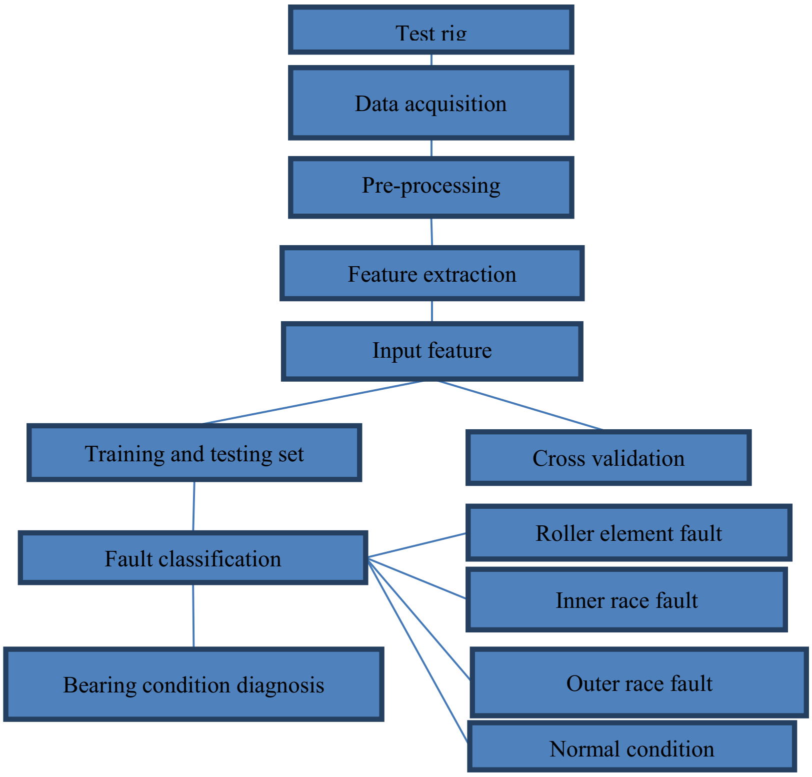


Figure 5. Proposed bearing fault diagnosis strategy.

4. Results and Discussion

4.1 Feature Extraction

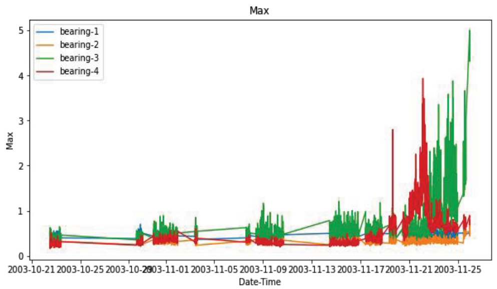
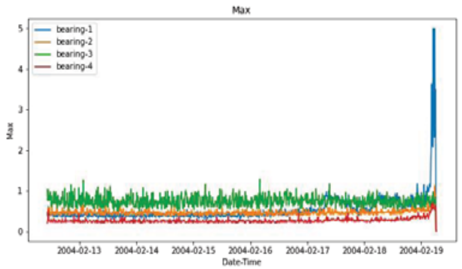
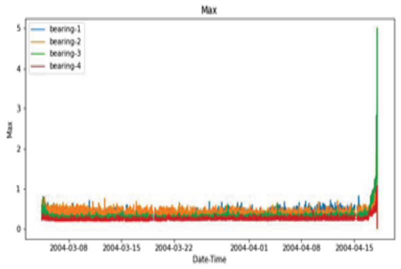
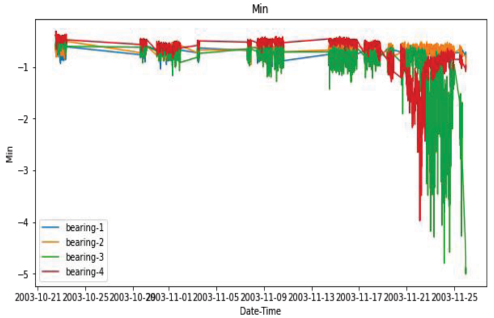
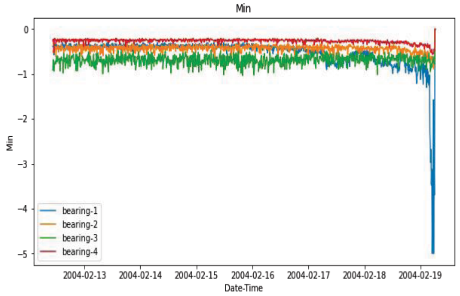
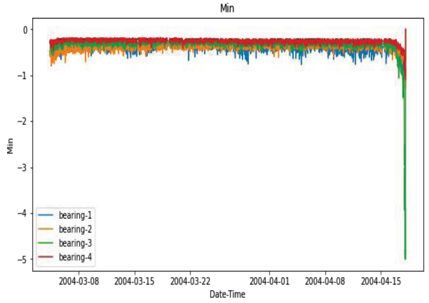
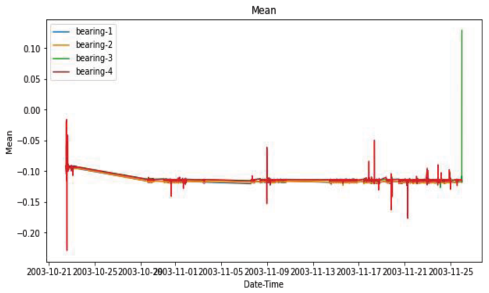
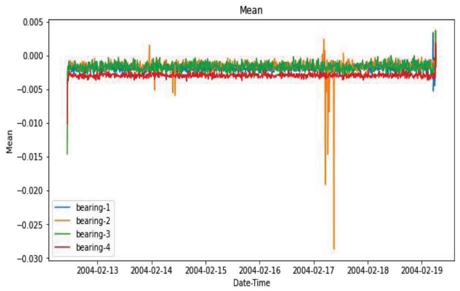
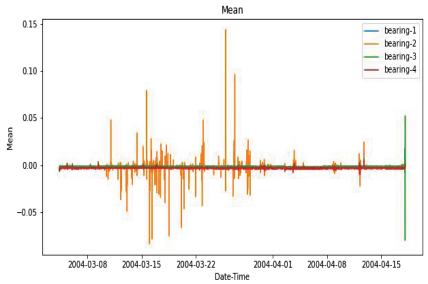
Using Table 3, Two accelerometers are used in Test 1 for each bearing, one along the x-axis and the other along the y-axis, both in radial directions but orthogonal to one another. Because there is little change in the data along the x and y axes, just one axis of the accelerometers for each bearing is used in Dataset 2. Once all the data has been extracted, it is stored in data frames with daytime indexes so that it may be plotted. The first training set disclosed that the third bearing experienced an inner race fault, and the fourth bearing experienced a roller element defect. Therefore, till the completion of the test, bearings 1 and 2 do not exhibit any form of flaw. Therefore, it is important to consider whether or not the extraction characteristics for bearings 3 and 4 may reveal any minimal variances. By looking at the first characteristic, the maximum value, it is simple to see that bearing 3 has a larger deviation and that bearing

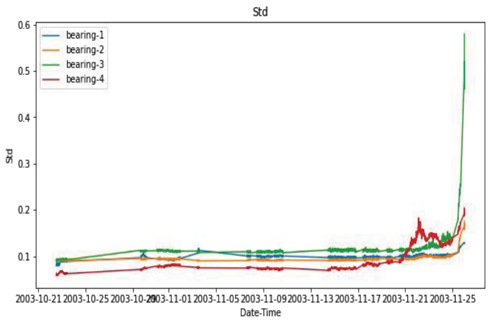
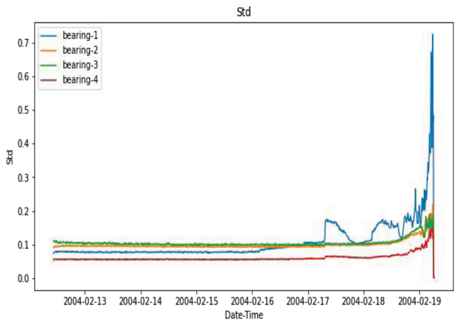
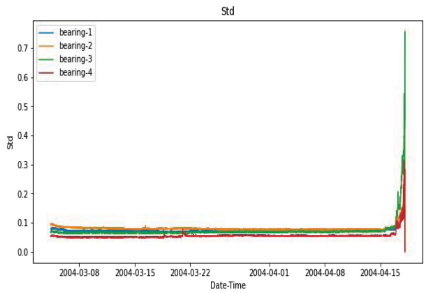
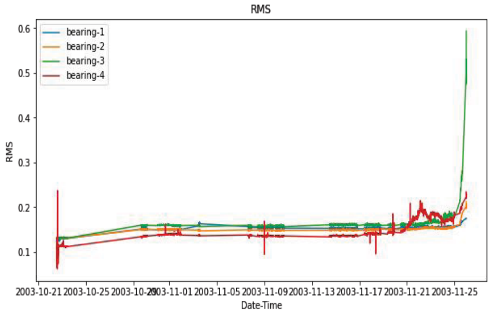
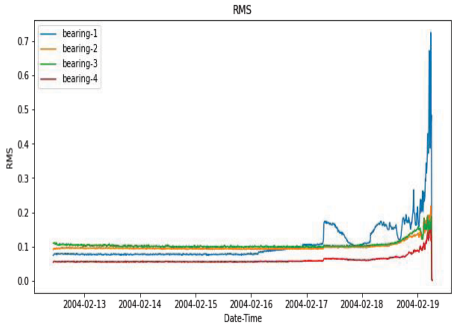
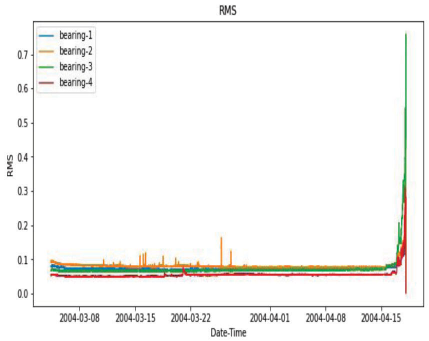
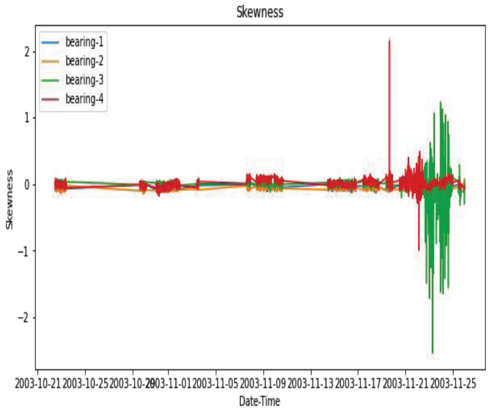
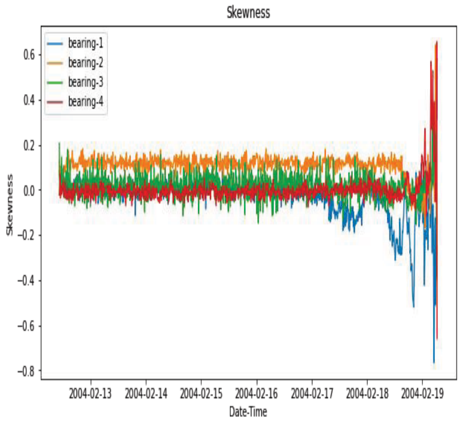
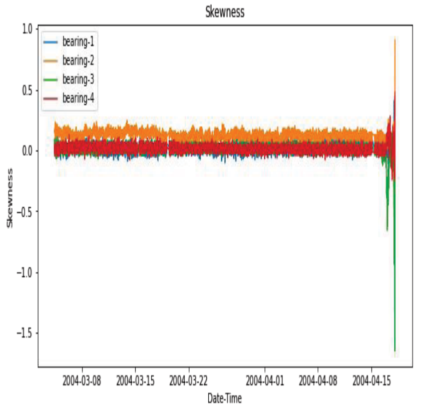
4 has faults that are more noticeable as deterioration progresses when compared to bearings 1 and 2, which are colored blue and orange, respectively. The minimum feature value might again see bearings 3 and 4 deviate much from the standard bearings 1 and 2. Only Bearing 3 exhibits a considerable variation from the mean value of the bearings; the other bearings do not. As can be observed from the standard deviation, bearing 3 similarly exhibits significant variation near the graph's edge, but bearing 4 exhibits very little departure from standard bearings. When the data visualization portion is complete, it is possible to observe where the flaws first appeared. The defects do not start at the beginning of the degradation phase; they only appear towards the finish. You should determine the maximum value and the point at which the deterioration value first became significant before classifying those data sets as flawed. The primary objective of this extraction is to create a massive classification model that can categorize various fault kinds from various bearings. While making the faulty data set for inner race fault will take the data for the 3rd bearing from this '2003-11-21 00:32:00' to '2003-11-24 18:22:00' daytime. That is because from 003-11-21 00:32:00' time the bearings 3 and 4 started to show deviation from the normal bearings. During the mentioned interval for test 1, (465 rows × 10 columns) times inner race and (465 rows × 10 columns) roller element fault will be repeated.

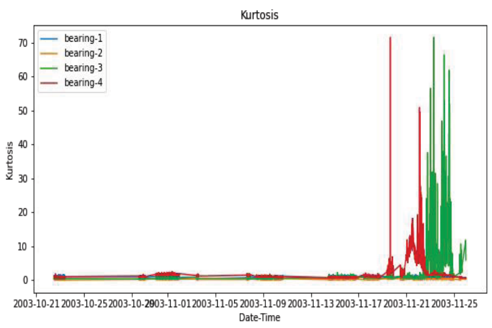
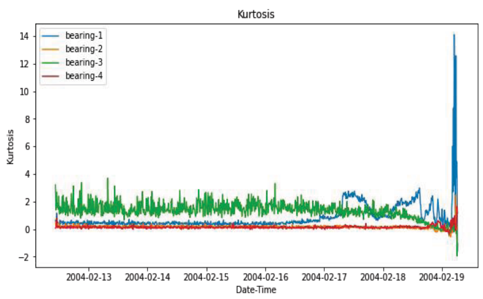
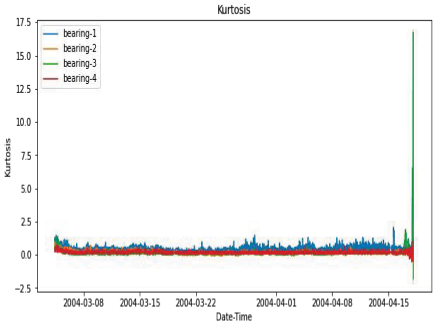
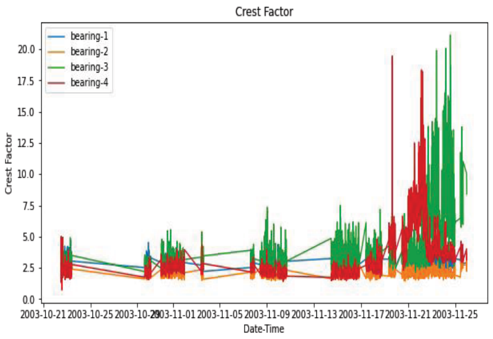
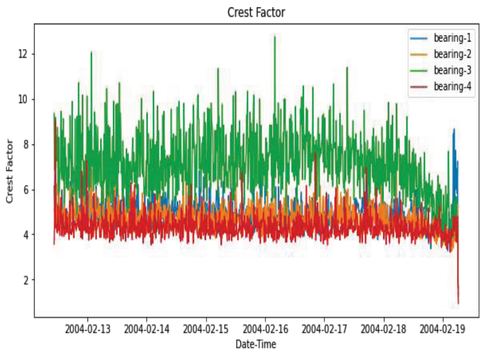
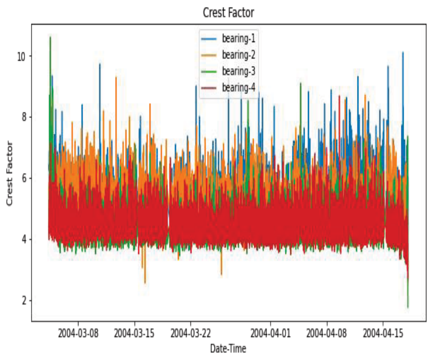
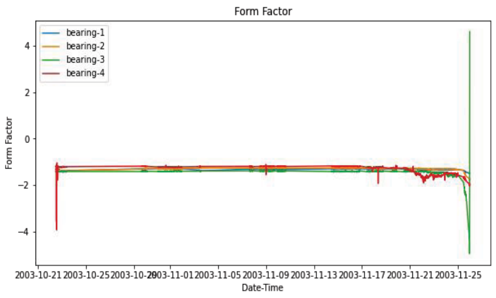
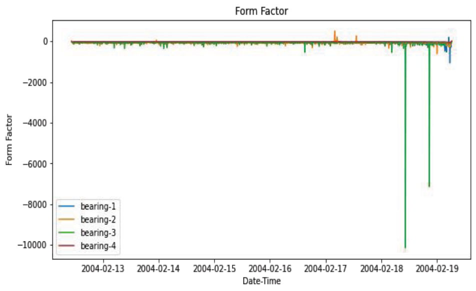
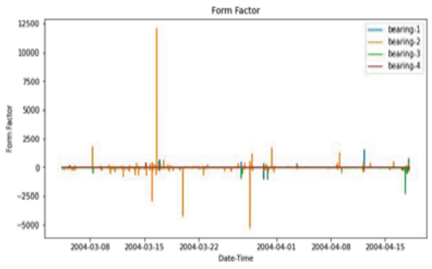
The same data is presented in test 2 in the same manner, and fault labels are generated to transform this issue into a classification challenge. After the test-to-fail experiment, bearing 1 experienced outer race failure. From the maximum value, it is obvious that bearing 1 exhibit a significant departure from the acceleration data of the other bearings after the test. It is also indicated in the description that bearing 1 exhibits outer race failure. From the RMS value, it is evident that the date (2004-02-17) bearing 1 began to exhibit some form of anomalous behaviour, which continued to the conclusion of the graph. Therefore, a period must be chosen from the date (2004-02-17) to determine the outer race failure (2004-02-19). The outer race fault will occur 218 rows by 10 columns times within the time frame specified for test 2.

As in Tests 1 and 2, it is possible to transform any data frame index to a daytime index in Test 3 so that it will aid in plotting. To display the features for all 4 bearings and the time on the x-axis, the Run 4 loop will cycle over each column or feature of the data frame. It is evident from the table that the max or mean graph displays a particular equilibrium before (2004-03-04 09:27:00). However, given that the identical error happened in test 2, it then exhibits a significant variance until the test is complete. In contrast, test 2 had a defect in bearing 1, test 3 faulted bearing 3, while the other two were balanced as normal. You must use the time from (2004-03-04 09:27:00) to (2004-04-18 02:42:00) to get the outer race fault timings. During this interval it will give (6324 rows × 9 columns) times of outer race fault.

Table 4. Feature extraction of bearing test (1,2,3)

Feature Extraction			
Feature	**Test 1**	**Test 2**	**Test 3**
Maximum			
Minimum			
Mean			

Standard deviation (unbiased std)			
Root mean square value (RMS)			
Skewness			

Kurtosis			
Crest factor = $\frac{Max}{RMS}$			
Form factor $=\frac{RMS}{Mean}$			

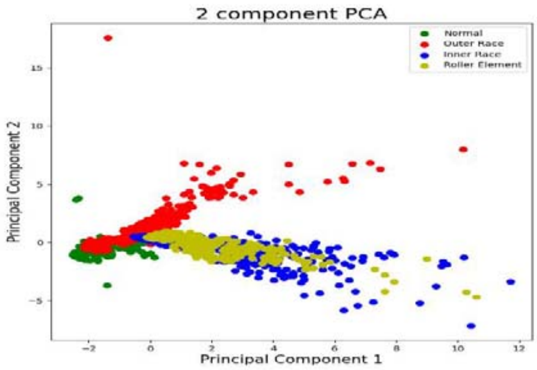
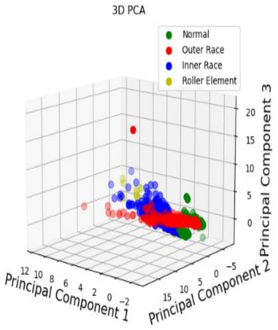
4.2 Normal Fault Data Collection

The deviation data or the time of the interval data was used to identify the flaws in the earlier tests. The task is to identify the data for the normal bearing data operation so that normal and incorrect data may be distinguished. Each test's data frame must be read to have the standard data frame. The probability of having normal data in such a situation is (22–23%) since any bearing beginning to degrade always exhibits some form of a spike in several locations or is not operating at peak efficiency. Therefore, in this instance, the data must be taken to a location where it is in excellent shape and is already coping well. After extraction, 22-23% of 3 tests' (768 rows 10 columns) periods of normal data would be obtained.

4.3 Dimensionality Reduction and Visualization

In order to better comprehend the data in a 2D and 3D environment, the dimensionality reduction approach, more especially Principal Component Analysis (PCA), is shown in Table 5. PCA is justified since data visualization is beneficial for many machine learning applications. When two or three characteristics are available, it is simple to plot the data and see it. However, when more than 3 features are available (from 3 tests, 9 features were collected), it is quite challenging to plot the data. For this reason, the dimensional reduction approach was required to reduce the 9 features to 2 or 3, which should help you comprehend the data better.

Table 5. Graphical visualization using PCA

Principal compact analysis	
2D	**3D**
	

Here, different types of flaws are grouped in various regions, as can be seen in 2D. Green stands in for the normal fault, red for the outer race fault, yellow for the roller element, and blue for the inner race fault, respectively. The first two main components account for 0.65, or 65% of the variance (the first principal component accounts for 0.47, or 47%, and the second principal component accounts for 0.18, or 18%). One component must be added to the other two major components to form a data frame with three principal components. One thing that can be analysed clearly in 3D is that normal bearing is extremely distinct from other flaws or seems to blend in very well. However, the yellow roller element defect and the blue inner race fault are well aligned. Therefore, it is challenging to discern between roller element defects and inner race faults. On the other hand, the variance of the outer race fault and normal bearing can be distinguished by 0.80, or 80%, of the variance (47% by principal component 1, 18% by principal component 2, and the final one is 14.5%), which is excellent.

4.4 Fault Classification

Try to read the faults (Normal, outer race fault test2, outer race fault test3, inner race fault and roller element fault test1) data set to classify the various sorts of faults that have been issued to them in the past. A 20% test size and the random forest technique were utilized to produce the test split. The fault categories are shown in Figure 6.

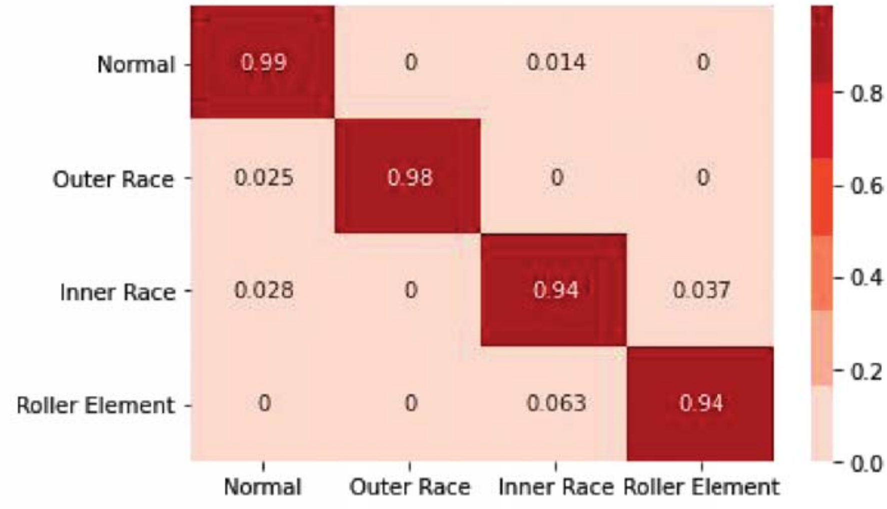


Figure 6. True prediction of data fame.

Figure 6 shows that the genuine predictions were accurate and could anticipate 97% of the outer race, 94% of the inner race, and 99% of the roller element fault. Having said that, the test side accuracy is pretty excellent, with an average prediction accuracy of 96.26%.

5. Conclusion

In this research, a thorough investigation of the use of the random forest method and the Python programming language to identify bearing problems is conducted using appropriate libraries. It is efficient in procedures dealing with enormous volumes of data. Effective feature extraction, standard data separation, dimensionally reduction, and visualization have been carried out for the categorization of bearing faults. Out of 20,480 data points, 9 domain features were extracted at a sampling rate of 20 kHz. After extraction, periods of typical data (768 rows x 10 columns) would be retrieved. The variance obtained using PCA for 2D and 3D is 65% and 80% respectively. The accuracy of the RF model with feature selection was > 95%, according to experimental findings, while the model's stability and performance may be enhanced. The suggested approach for bearing failure identification in IMs may be realistically used. Automatic input parameter selection for dimensionality reduction methods can be the subject of future study.

6. Acknowledgments

The authors also acknowledge the contribution of the facilities of Universiti Tenaga Nasional Sdn. Bhd, Malaysia (UNITEN), AAIBE Chair of Renewable Energy (ChRE) and BOLDREFRESH2025 (J510050002 (IC-6C))) for providing all out-laboratory support through the research grant number 202101KETTHA.

Reference

[1] Y. Chen, T. Zhang, Z. Luo, and K. Sun, “A Novel Rolling Bearing Fault Diagnosis and Severity Analysis Method,” *Applied Sciences 2019, Vol. 9, Page 2356*, vol. 9, no. 11, p. 2356, Jun. 2019.

[2] M. Cerrada *et al.*, “A review on data-driven fault severity assessment in rolling bearings,” *Mech Syst Signal Process*, vol. 99, pp. 169–196, Jan. 2018.

[3] X. Zhang, Y. Cong, Z. Yuan, T. Zhang, and X. Bai, “Early Fault Detection Method of Rolling Bearing Based on MCNN and GRU Network with an Attention Mechanism,” *Shock and Vibration*, vol. 2021.

[4] D. Neupane and J. Seok, “Bearing fault detection and diagnosis using case western reserve university dataset with deep learning approaches: A review,” *IEEE Access*, vol. 8, pp. 93155–93178, 2020, doi: 10.1109/ACCESS.2020.2990528.

[5] M. Jalayer, C. Orsenigo, and C. Vercellis, “Fault detection and diagnosis for rotating machinery: A model based on convolutional LSTM, Fast Fourier and continuous wavelet transforms,” *Comput Ind*, vol. 125, p. 103378, Feb. 2021.

[6] X. Tao *et al.*, “Bearings fault detection using wavelet transform and generalized Gaussian density modeling,” *Measurement*, vol. 155, p. 107557, Apr. 2020, doi: 10.1016/J.MEASUREMENT.2020.107557.

[7] J. Gu and Y. Peng, “An improved complementary ensemble empirical mode decomposition method and its application in rolling bearing fault diagnosis,” *Digit Signal Process*, vol. 113, p. 103050, Jun. 2021.

[8] R. N. Toma, C. H. Kim, and J. M. Kim, “Bearing Fault Classification Using Ensemble Empirical Mode Decomposition and Convolutional Neural Network,” *Electronics 2021, Vol. 10, Page 1248*, vol. 10, no. 11, p. 1248, May 2021.

[9] Q. Liu, J. Yang, and K. Zhang, “An improved empirical wavelet transform and sensitive components selecting method for bearing fault,” *Measurement*, vol. 187, p. 110348, Jan. 2022.

[10] F. He and Q. Ye, “A Bearing Fault Diagnosis Method Based on Wavelet Packet Transform and Convolutional Neural Network Optimized by Simulated Annealing Algorithm,” *Sensors 2022, Vol. 22, Page 1410*, vol. 22, no. 4, p. 1410, Feb. 2022.

[11] J. Gai, J. Shen, Y. Hu, and H. Wang, “An integrated method based on hybrid grey wolf optimizer improved variational mode decomposition and deep neural network for fault diagnosis of rolling bearing,” *Measurement*, vol. 162, p. 107901, Oct. 2020.

[12] T. Ali, D. #1, A. Abdulhady, and J. #2, “Bearing Fault Diagnosis Using Motor Current Signature Analysis and the Artificial Neural Network,” vol. 10, no. 1, 2020.

[13] L. Yuan, D. Lian, X. Kang, Y. Chen, and K. Zhai, “Rolling Bearing Fault Diagnosis Based on Convolutional Neural Network and Support Vector Machine,” *IEEE Access*, vol. 8, pp. 137395–137406, 2020.

[14] M. J. Hasan and J. M. Kim, “Fault Detection of a Spherical Tank Using a Genetic Algorithm-Based Hybrid Feature Pool and k-Nearest Neighbor Algorithm,” *Energies 2019, Vol. 12, Page 991*, vol. 12, no. 6, p. 991, Mar. 2019.

[15] S. Wang, J. Xiang, Y. Zhong, and Y. Zhou, “Convolutional neural network-based hidden Markov models for rolling element bearing fault identification,” *Knowl Based Syst*, vol. 144, pp. 65–76, Mar. 2018.

[16] H. Shao, H. Jiang, H. Zhao, and F. Wang, “A novel deep autoencoder feature learning method for rotating machinery fault diagnosis,” *Mech Syst Signal Process*, vol. 95, pp. 187–204, Oct. 2017.

[17] J. Zhao, S. Yang, Q. Li, Y. Liu, X. Gu, and W. Liu, “A new bearing fault diagnosis method based on signal-to-image mapping and convolutional neural network,” *Measurement*, vol. 176, p. 109088, May 2021.

[18] X. Yan, Y. Liu, and M. Jia, “Multiscale cascading deep belief network for fault identification of rotating machinery under various working conditions,” *Knowl Based Syst*, vol. 193, p. 105484, Apr. 2020.

[19] P. Liang, C. Deng, J. Wu, and Z. Yang, "Intelligent fault diagnosis of rotating machinery via wavelet transform, generative adversarial nets and convolutional neural network," *Measurement*, vol. 159, p. 107768, Jul. 2020, doi: 10.1016/J.MEASUREMENT.2020.107768.

[20] S. Liu, C. Shen, Z. Chen, W. Huang, and Z. Zhu, "A sudden fault detection network based on Time-sensitive gated recurrent units for bearings," *Measurement*, vol. 186, p. 110214, Dec. 2021.

[21] L. Breiman, "Random Forests," *Machine Learning 2001 45:1*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[22] L. Breiman, "Using Iterated Bagging to Debias Regressions," *Machine Learning 2001 45:3*, vol. 45, no. 3, pp. 261–277, Dec. 2001.

[23] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Trans Pattern Anal Mach Intell*, vol. 20, no. 8, pp. 832–844, 199.

[24] M. Cerrada, G. Zurita, D. Cabrera, R. V. Sánchez, M. Artés, and C. Li, "Fault diagnosis in spur gears based on genetic algorithm and random forest," *Mech Syst Signal Process*, vol. 70–71, pp. 87–103, Mar. 2016.

[25] R. V. Sánchez, P. Lucero, R. E. Vásquez, M. Cerrada, J. C. Macancela, and D. Cabrera, "Feature ranking for multi-fault diagnosis of rotating machinery by using random forest and KNN," *Journal of Intelligent & Fuzzy Systems*, vol. 34, no. 6, pp. 3463–3473, Jan. 2018.

[26] Y. Liu and Z. Ge, "Weighted random forests for fault classification in industrial processes with hierarchical clustering model selection," *J Process Control*, vol. 64, pp. 62–70, Apr. 2018.

[27] C. Li, R. V. Sánchez, G. Zurita, M. Cerrada, and D. Cabrera, "Fault Diagnosis for Rotating Machinery Using Vibration Measurement Deep Statistical Feature Learning," *Sensors 2016, Vol. 16, Page 895*, vol. 16, no. 6, p. 895, Jun. 2016.

[28] Q. He, "Vibration signal classification by wavelet packet energy flow manifold learning," *J Sound Vib*, vol. 332, no. 7, pp. 1881–1894, Apr. 2013.

[29] Z. Liu, H. Cao, X. Chen, Z. He, and Z. Shen, "Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings," *Neurocomputing*, vol. 99, pp. 399–410, Jan. 2013, doi: 10.1016/J.NEUCOM.2012.07.019.

[30] A. Ragab, S. Yacout, M. S. Ouali, and H. Osman, "Prognostics of multiple failure modes in rotating machinery using a pattern-based classifier and cumulative incidence functions," *Journal of Intelligent Manufacturing 2016 30:1*, vol. 30, no. 1, pp. 255–274, Jul. 2016.

[31] A. Heng, S. Zhang, A. C. C. Tan, and J. Mathew, "Rotating machinery prognostics: State of the art, challenges and opportunities," *Mech Syst Signal Process*, vol. 23, no. 3, pp. 724–739, Apr. 2009.

[32] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors - A review," *IEEE Transactions on Energy Conversion*, vol. 20, no. 4, pp. 719–729, Dec. 2005.

[33] J. Lee, M. Ghaffari, and S. Elmeligy, "Self-maintenance and engineering immune systems: Towards smarter machines and manufacturing systems," *Annu Rev Control*, vol. 35, no. 1, pp. 111–122, Apr. 2011.

[34] C. P. Mboo and K. Hameyer, "Fault diagnosis of bearing damage by means of the linear discriminant analysis of stator current features from the frequency selection," *IEEE Trans Ind Appl*, vol. 52, no. 5, pp. 3861–3868, Sep. 2016.

[35] R. N. Toma, A. E. Prosvirin, and J. M. Kim, "Bearing Fault Diagnosis of Induction Motors Using a Genetic Algorithm and Machine Learning Classifiers," *Sensors 2020, Vol. 20, Page 1884*, vol. 20, no. 7, p. 1884, Mar. 2020.

[36] M. E. Curd, T. L. Burnett, J. Fellowes, P. Yan, and P. J. Withers, "Redistribution of carbon caused by butterfly defects in bearing steels," *Acta Mater*, vol. 183, pp. 390–397, Jan. 2020.

[37] J. B. Wood and I. Howard, "An Investigation Into Bearing Fault Diagnostics for Condition Based Maintenance Using Band-Pass Filtering and Wavelet Decomposition Analysis of Vibration Signals".

[38] M. Blodt, P. Granjon, B. Raison, and G. Rostaing, "Models for bearing damage detection in induction motors using stator current monitoring," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 4, pp. 1813–1822, Apr. 2008.

[39] J. H. Jung, J. J. Lee, and B. H. Kwon, "Online diagnosis of induction motors using MCSA," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 6, pp. 1842–1852, 2006.

[40] T. Yang, H. Pen, Z. Wang, and C. S. Chang, "Feature Knowledge Based Fault Detection of Induction Motors Through the Analysis of Stator Current Data," *IEEE Trans Instrum Meas*, vol. 65, no. 3, pp. 549–558, Mar. 2016.

[41] R. N. Toma, A. E. Prosvirin, and J. M. Kim, "Bearing Fault Diagnosis of Induction Motors Using a Genetic Algorithm and Machine Learning Classifiers," *Sensors 2020, Vol. 20, Page 1884*, vol. 20, no. 7, p. 1884, Mar. 2020.

[42] M. Farajzadeh-Zanjani, R. Razavi-Far, and M. Saif, "A critical study on the importance of feature extraction and selection for diagnosing bearing defects," *Midwest Symposium on Circuits and Systems*, vol. 2018-August, pp. 803–808, Jan. 2019.

[43] W. Caesarendra, "Vibration and acoustic emission-based condition monitoring and prognostic methods for very low speed slew bearing," *University of Wollongong Thesis Collection 1954-2016*, Jan. 2015.

[44] M. Kang, M. R. Islam, J. Kim, J. M. Kim, and M. Pecht, "A Hybrid Feature Selection Scheme for Reducing Diagnostic Performance Deterioration Caused by Outliers in Data-Driven Diagnostics," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3299–3310, May 2016.

[45] L. C. Brito, G. A. Susto, J. N. Brito, and M. A. V. Duarte, "Fault Detection of Bearing: An Unsupervised Machine Learning Approach Exploiting Feature Extraction and Dimensionality Reduction," *Informatics 2021, Vol. 8, Page 85*, vol. 8, no. 4, p. 85, Nov. 2021, doi: 10.3390/INFORMATICS8040085.

[46] G. Vashishtha and R. Kumar, "Pelton Wheel Bucket Fault Diagnosis Using Improved Shannon Entropy and Expectation Maximization Principal Component Analysis," *Journal of Vibration Engineering & Technologies 2021 10:1*, vol. 10, no. 1, pp. 335–349, Sep. 2021, doi: 10.1007/S42417-021-00379-7.

[47] C. J. C. Burges, "Dimension Reduction: A Guided Tour," *Foundations and Trends® in Machine Learning*, vol. 2, no. 4, pp. 275–365, 2010, doi: 10.1561/2200000002.

[48] G. Baudat and F. Anouar, "Generalized Discriminant Analysis Using a Kernel Approach," *Neural Comput*, vol. 12, no. 10, pp. 2385–2404, Oct. 2000..

[49] Y. Tian, Z. Wang, L. Zhang, C. Lu, and J. Ma, "A subspace learning-based feature fusion and open-set fault diagnosis approach for machinery components," *Advanced Engineering Informatics*, vol. 36, pp. 194–206, Apr. 2018.

[50] S. Ghosh, A. Rana, and V. Kansal, "A Hybrid Nonlinear Manifold Detection Approach for Software Defect Prediction," *2018 7th International Conference on Reliability, Infocom Technologies and Optimization: Trends and Future Directions, ICRITO 2018*, pp. 453–459, Aug. 2018..

[51] R. Wang, F. Nie, R. Hong, X. Chang, X. Yang, and W. Yu, "Fast and Orthogonal Locality Preserving Projections for Dimensionality Reduction," *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 5019–5030, Oct. 2017.

[52] A. Najafi, A. Joudaki, and E. Fatemizadeh, "Nonlinear Dimensionality Reduction via Path-Based Isometric Mapping," *IEEE Trans Pattern Anal Mach Intell*, vol. 38, no. 7, pp. 1452–1464, Jul. 2016.

[53] J. Tian, L. Liu, F. Zhang, Y. Ai, R. Wang, and C. Fei, "Multi-Domain Entropy-Random Forest Method for the Fusion Diagnosis of Inter-Shaft Bearing Faults with Acoustic Emission Signals," *Entropy 2020, Vol. 22, Page 57*, vol. 22, no. 1, p. 57, Dec. 2019.

[54] T. M. Oshiro, P. S. Perez, and J. A. Baranauskas, "How many trees in a random forest?," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7376 LNAI, pp. 154–168, 2012.

[55] P. Probst and A.-L. Boulesteix, “To Tune or Not to Tune the Number of Trees in Random Forest,” *Journal of Machine Learning Research*, vol. 18, pp. 1–18, 2018.
[56] T. Fawcett, “An introduction to ROC analysis,” *Pattern Recognit Lett*, vol. 27, no. 8, pp. 861–874, Jun. 2006.
[57] F. Pedregosa FABIANPEDREGOSA *et al.*, “Scikit-learn: Machine Learning in Python Gaël Varoquaux Bertrand Thirion Vincent Dubourg Alexandre Passos PEDREGOSA, VAROQUAUX, GRAMFORT ET AL. Matthieu Perrot,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
[58] J. Lee, H. Qiu, G. Yu, J. Lin, and Rexnord Technical Services, “IMS bearing data,” pp. 2–3, 2007.
[59] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, and G. Tripot, “Hidden Markov Models for failure diagnostic and prognostic,” *2011 Prognostics and System Health Management Conference, PHM-Shenzhen 2011*.