

Gear Fault Detection Using Recurrence Quantification Analysis and Support Vector Machine

T. Haj Mohamad¹, Y. Chen², Z. Chaudhry², C. Nataraj¹

¹Villanova Center for Analytics of Dynamic Systems (VCADS), Villanova University, Villanova, PA, USA

²United Technologies Research Center (UTRC), East Hartford, CT, USA

Email: thajmoha@villanova.edu

How to cite this paper: Mohamad, T.H., Chen, Y., Chaudhry, Z. and Nataraj, C. (2018) Gear Fault Detection Using Recurrence Quantification Analysis and Support Vector Machine. *Journal of Software Engineering and Applications*, 11, 181-203.
<https://doi.org/10.4236/jsea.2018.115012>

Received: April 17, 2018

Accepted: May 22, 2018

Published: May 25, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This paper presents the application of recurrence plots (RPs) and recurrence quantification analysis (RQA) in the diagnostics of various faults in a gear-train system. For this study, multiple test gears with different health conditions (such as a healthy gear, and defective gears with a root crack on one tooth, multiple cracks on five teeth and missing tooth) are studied. The vibration data of a gear-train is measured by a triaxial accelerometer installed on the test. Two different support vector machine classifiers are trained and compared. Mutual information is used to rank the extracted features in order to select an optimal subset that provides as much information as possible about the intrinsic dynamics of the system. Results indicate that our approach is quite efficient in diagnosing the status of the health of the gear system and characterizing the dynamic behavior.

Keywords

Gear Diagnostics, Recurrence Analysis, Optimization, Time Series Analysis

1. Introduction

Machine condition monitoring techniques have generated considerable recent research interest due to their important role in preventing consequential damages before they develop into a catastrophic failure. Furthermore, condition monitoring techniques increase lifespans of systems and decrease maintenance costs by shifting maintenance from time-based to event-driven procedures, which offer significant economic benefits. Hence, it is important to develop efficient algorithms and techniques to monitor the status of machines and

identify abnormalities.

A major current focus in fault diagnostics of gears is vibration and acoustic methods due to the valuable information they contain about the condition of rotating machines such as gears [1] [2] [3] [4]. Yet, gear fault diagnostics is still a challenging task because of the highly nonlinear characteristics of faults and its complex nonstationary dynamics [5] [6] [7] [8]. Multi-periodic, quasi-periodic and chaotic responses, as well as bifurcation and limit cycles are some phenomena that can be observed in nonlinear dynamic systems and are reported in industrial machinery [9] [10] [11]. This motivates the use of recurrence plots (RPs) and recurrence quantification analysis (RQA) because of their advantages in dealing with nonlinear dynamical responses and non-stationary data. RPs and RQA have received much attention in recent years in miscellaneous fields of science. In this paper, we introduce the application of RPs and RQA in gear-train transmission systems. The RP is a two-dimensional visualization technique to investigate high-dimensional dynamical systems. It identifies the times when the state space trajectory of the dynamical system visits roughly the same area in the state space. On the other hand, RQA offers a more objective and quantitative method for the investigation of dynamical systems. Thus, RPs and RQA combine and employ higher dimensional phase space trajectories, *i.e.*, all available data (vibrational acceleration in three directions), in which RPs provide a visualization and RQA quantifies these plots and measures their complexity. In addition, RPs and RQA are a modern tool for nonlinear data analysis, which enables us to investigate the various responses of the system (*i.e.*, periodic, quasi periodic and chaotic) and provides valuable information about the dynamics of the system.

Various techniques have been developed to study gear fault detection and diagnostics [12]-[19] and can be time domain, frequency domain, or time-frequency domain methods. However, these methods do not always guarantee stable classification and/or lack generalizability for systems with complex nonlinear responses. For example, time domain techniques such as Time Synchronous Averaging (TSA) are inefficient in different gear faults, especially in the case of multiple simultaneous faults in different gears or in the early phase of faults [20] [21]. In addition, the TSA technique can be time-consuming and is often computationally expensive [22]. Although sideband frequencies analysis (in frequency domain techniques) detects faults in the gearbox, it falls short of distinguishing gear faults, as they may be located in other components of the gearbox [23] [24].

In previous work [25], Recurrence Plots were effectively applied to detect healthy and faulty gear systems. However, the the problem of multiple fault detection was not addressed. The Extended Phase Space Topology (EPST) based on characterizing the phase space topology, developed in our research laboratory, was applied on the same system used in this study [26] [27] [28] [29]. The results indicate that EPST is efficient in diagnosing the status of the health of the gear

system and characterizing the dynamic behavior. However, the EPST features are based on orthogonal functions coefficients that are not well understood and it is a challenging task to track back these coefficients in order to find their physical meaning in the dynamical system.

The present work investigated vibration data of a helicopter gearbox mock-up system (5 m long). For this study, multiple test gears with different health conditions such as healthy gears (H) and defective gears with root crack on one tooth (SCD), multiple cracks on five teeth (MCD) and missing tooth (MTD) are studied. The vibrational signals are recorded using a triaxial accelerometer installed on the test gearbox.

The rest of this paper is organized as follows. In Section 2, the mathematical details of the RPs and RQA are introduced. Section 3 represents the experimental setup of the gear-train and measurement of data. Section 4 discusses the analysis process of the RPs and RQA. In Section 5, different fault classification models are compared and the feature ranking technique is explained. Finally, Section 6 summarizes and concludes the paper.

2. Recurrence Plots and Recurrence Quantification Analysis

The recurrence plot analysis for time series is based on the analysis of a matrix \mathbf{R} whose elements are defined as:

$$\mathbf{R}_{ij} = \begin{cases} 1, & \Phi_i \approx \Phi_j, \\ 0, & \Phi_i \neq \Phi_j, \end{cases} \quad i, j = 1, \dots, N, \quad (1)$$

where $\Phi_i = (\phi_{1i}, \phi_{2i}, \dots, \phi_{mi})$ is a m -dimensional state vector, N is the length of the time series, i and j are the row and column indices of the matrix respectively, and $\Phi_i \approx \Phi_j$ means equality up to an error ϵ . In this paper, the vibrational acceleration of the test gearbox in x , y and z directions is considered as the state vector as follows:

$$\Phi_i = (a_x(i), a_y(i), a_z(i)) \quad (2)$$

The elements of the matrix \mathbf{R} are thus obtained by comparing the state of the system at time i and j with a threshold precision ϵ . Thus, formally, one has:

$$\mathbf{R}_{ij} = \theta(\epsilon - \|\Phi_i - \Phi_j\|), \quad (3)$$

with $\|\cdot\|$ being the Euclidian norm (L_2 -norm) and $\theta(y)$ is the heaviside function defined as:

$$\theta(y) = 1 \text{ for } y > 0 \text{ and } \theta(y) = 0 \text{ for } y < 0$$

The threshold precision ϵ is a crucial parameter in RP analysis. The 5% maximal phase space diameter rule of thumb is used to select the optimal threshold.

Once the \mathbf{R} matrix is constructed, the RP graph is obtained by plotting the \mathbf{R}_{ij} points in the i and j plane with different colors. By definition, RP graphs are always symmetric ($\mathbf{R}_{ij} = \mathbf{R}_{ji}$) and always have a central diagonal.

In order to go beyond the qualitative impression given by RPs, complexity measures have been developed that quantify the structures of RPs and are called recurrence quantification analysis (RQA) [30]. In this paper, the following six RQA parameters were used to quantify the RP of the system under various fault conditions.

- Recurrence rate (*RR*)

The recurrence rate is the simplest RQA parameter and measures the density of recurrence points in a recurrence plot. In other words, *RR* indicates the percentage of the plot occupied by points.

$$RR = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j}(\epsilon) \quad (4)$$

- Determinism (*DET*)

The determinism is the percentage of recurrence points that form diagonal lines in the recurrence plot of minimal length ℓ_{\min} .

$$DET = \frac{\sum_{\ell=\ell_{\min}}^N \ell P(\ell)}{\sum_{\ell=1}^N \ell P(\ell)} \quad (5)$$

where $P(\ell)$ is the frequency distribution of the lengths ℓ of the diagonal lines. In this work, $\ell_{\min} = 2$ is used. This measure is critical in determining the nature of the process (deterministic vs. stochastic). Recurrence plots of a deterministic process usually contain more and longer diagonal lines compared to a stochastic process.

- Laminarity (*LAM*)

In the same way, the amount of recurrence points forming vertical lines can be quantified by laminarity.

$$LAM = \frac{\sum_{v=v_{\min}}^N v P(v)}{\sum_{v=1}^N v P(v)} \quad (6)$$

where $P(v)$ is the frequency distribution of the lengths v of the vertical lines, which have at least a length of v_{\min} . In this work, $v_{\min} = 2$ is used.

- Longest Diagonal Line (*LMAX*)

LMAX is the length of the longest diagonal line.

$$LMAX = \max(\{l_i; i = 1, \dots, N\}) \quad (7)$$

- Trapping Time (*TT*)

The trapping time measures the average length of the vertical lines.

$$TT = \frac{\sum_{v=v_{\min}}^N v P(v)}{\sum_{v=v_{\min}}^N P(v)} \quad (8)$$

- Entropy (*ENTR*)

The probability that a diagonal line has exactly length ℓ can be estimated with $p(\ell) = \frac{P(\ell)}{\sum_{\ell=\ell_{\min}}^N P(\ell)}$. *ENTR* is the Shannon entropy of this probability,

which reflects the complexity of the RP in respect to the diagonal lines.

$$ENTR = - \sum_{\ell=\ell_{\min}}^N p(\ell) \ln p(\ell) \quad (9)$$

Two RQA parameters were excluded from the study. The divergence parameter which is the inverse of the longest diagonal line $DIV = 1/LMAX$ was not included in the study to avoid any redundancy because of its direct relation with $LMAX$ parameter which is included. Also, trend parameter was not included because of the constant operating speed condition we have in this study where trend has valuable information about the stationarity of the system and that is already considered.

3. Experimental Setup and Data Collection

The current work involved investigating a mock-up of a helicopter gear box system. All of the test data was acquired by collaboration with the United Technologies Research Center (UTRC). The gear-train experimental setup (shown in **Figure 1**) is a large scale machine (5 m long) consisting of a motor, dynamometer and four gearboxes, where each gearbox contains four spur gears. The schematic of the gear-train illustrating the four gear boxes and their components is shown in **Figure 2**. For this study, multiple test gears with different health conditions were studied. The study was implemented by replacing the gear located in gearbox number 3 and shown in green color in **Figure 2** with different test gears while the remaining setup components were kept unchanged. The test gears with 23 teeth that were used in the experiment include one healthy gear (H) and three defective gears with root crack on one tooth (SCD), multiple cracks on five teeth (MCD) and missing tooth (MTD). **Figure 3** illustrates the various gear cracks: a) shows the single root crack of 2 mm, b) shows the locations of the five root cracks on teeth numbers 1, 6, 10, 15, and 19, and c) shows the sizes of the five root cracks ranging from 0.5 mm to 2.5 mm. **Figure 4** shows the test gear with a missing tooth condition that is used in this study.

The vibrational signals were recorded using a triaxial accelerometer installed on gearbox number 3. The vibrational data was measured at the sampling



Figure 1. Gear-train experimental setup.

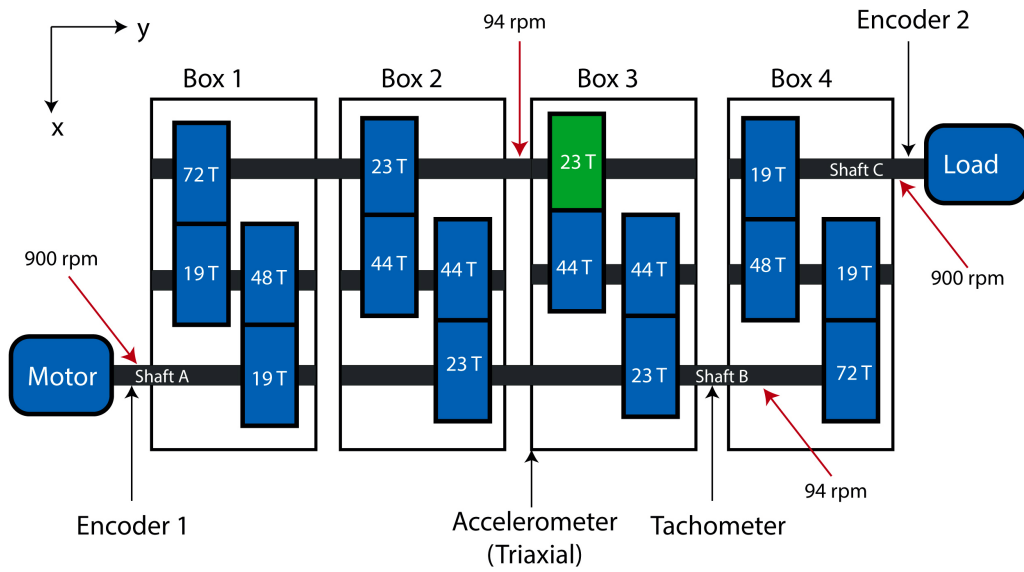


Figure 2. Gear-train schematic.

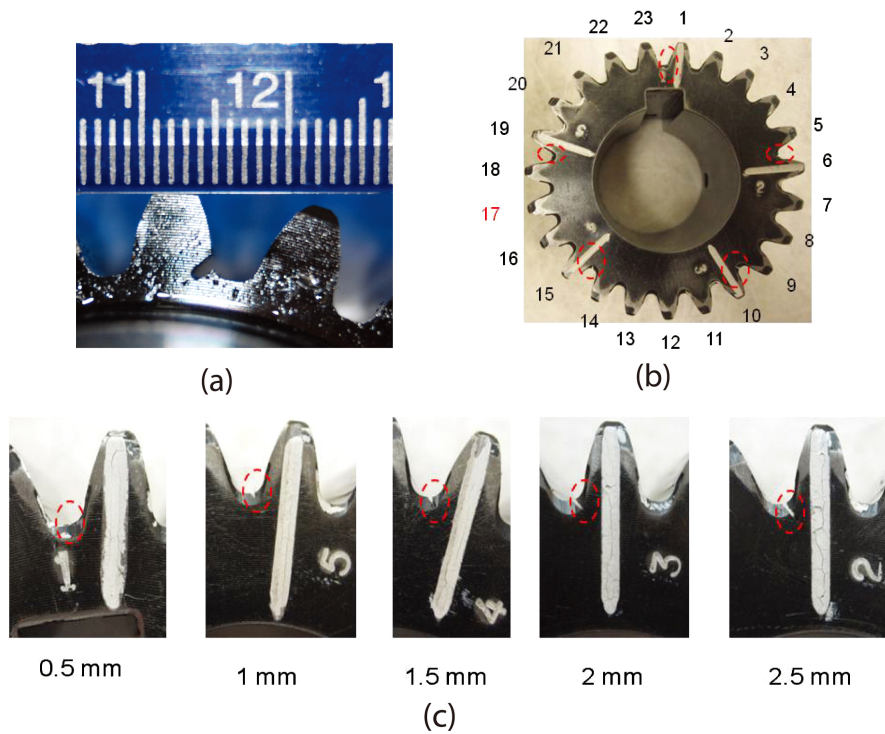


Figure 3. Gear defects: (a) crack on one tooth; (b) crack on 5 teeth; (c) crack sizes on 5 teeth.



Figure 4. Test gear with a missing tooth.

frequency of 102,400 Hz. The rotational speeds of shafts A, C, and B were measured using two encoders and a tachometer. Two encoders were installed at shaft “A” (input shaft) and shaft “C” (output shaft) to measure their rotational speeds with a 360 pulse/rev resolution. The tachometer on shaft “B” was used to measure the shaft rotational speed at a rate of 1 pulse/rev. Due to the gear teeth ratio, the test gear shaft operates at the same speed as shaft “B”. In this study, the motor was operating at a rotational speed of 900 rpm while the test gear shaft was running at 94 rpm for the different gear conditions. The vibrational signals of gearbox number 3 were recorded for 64 seconds for healthy, single crack tooth and multiple crack teeth conditions, and for 3.2 seconds for the missing tooth condition. Samples of the measured vibration for different gear conditions in the three directions are shown in **Figure 5**.

4. Recurrence Analysis and Discussion

An overview of the fault detection method used in this paper is summarized in **Figure 6** and is described as follows. Vibration data were recorded using a triaxial accelerometer in three space directions for different gear conditions such as healthy, single crack defect, multiple crack defect and missing tooth defect. Vibration data was divided into multiple segments in order to analyze it. Since the rotational speed of gears typically varies slightly during normal operation, the tachometer signal was used to divide the vibrational data into segments of revolutions. The tachometer provides the phase information of the shaft in the form of pulses synchronized to the rotation of the test gear. An example illustrates the segmentation process for the healthy gear data shown in **Figure 7**.

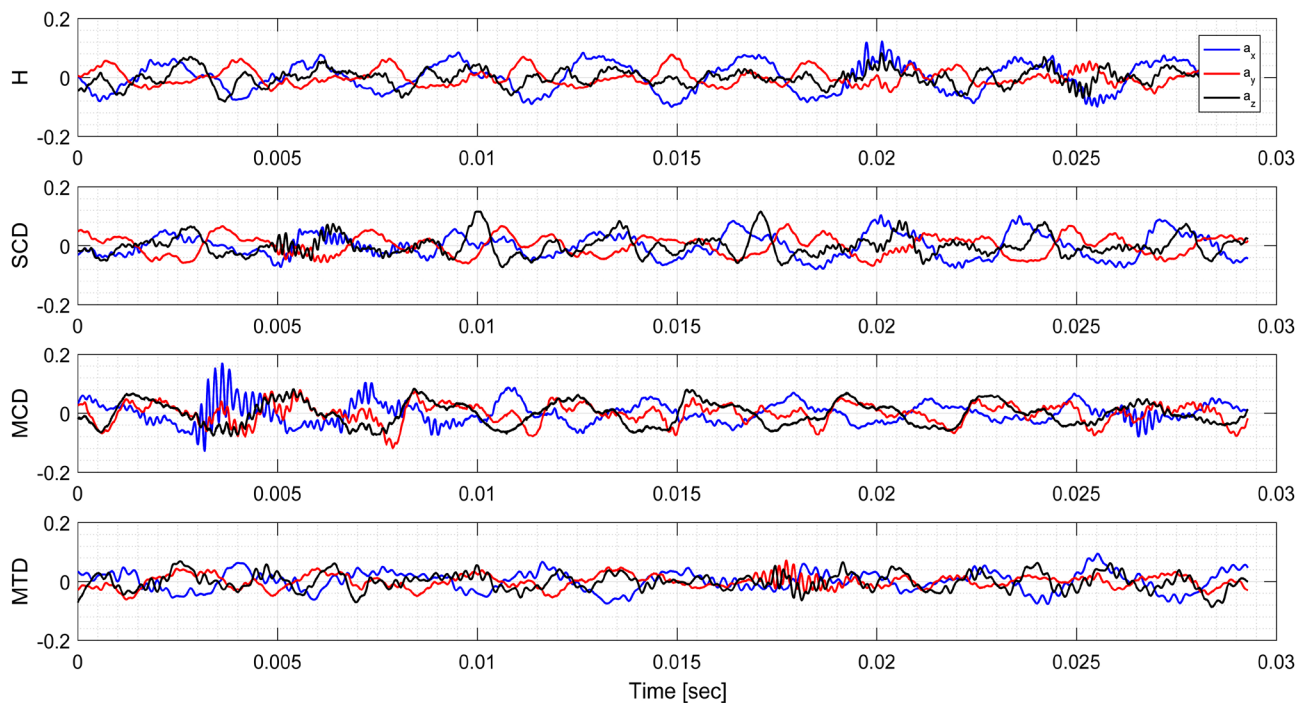


Figure 5. Samples of the gear vibration data for each configuration and direction.

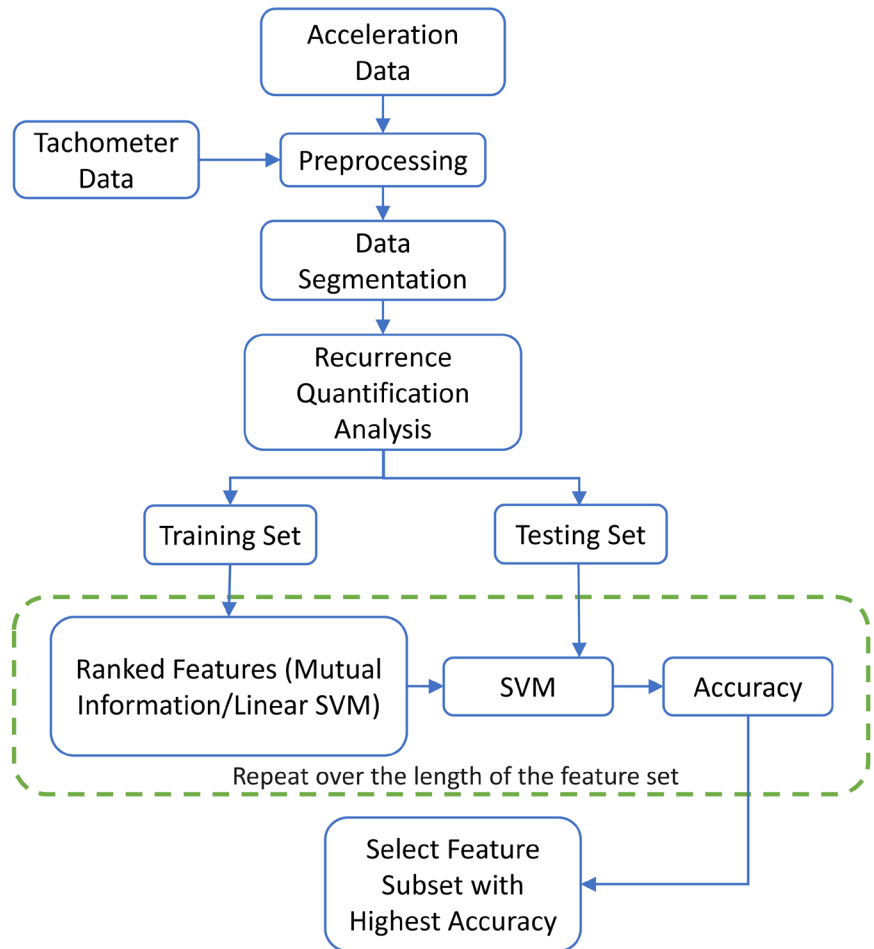


Figure 6. The process of the diagnostics method.

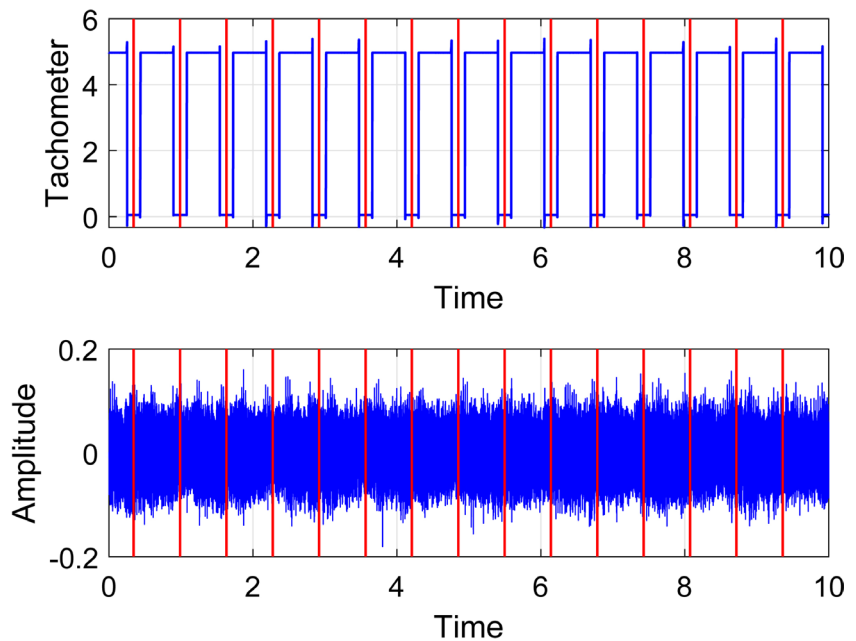


Figure 7. Segmentation preprocess (top plot: tachometer signal; bottom plot: vibrational signal for healthy gear condition).

The red vertical lines mark the locations of the pulses in the tachometer signal and their corresponding locations in the acceleration signal.

The segment $x_k(i)$, for $i=1,2,\dots,N$ is the vibration data of revolution k of a total number of K revolutions. The total number of data segments that were obtained is 341 segments including 98 data segments each for healthy, single crack defect, and multiple crack defect conditions while 47 data segments were obtained for the missing tooth defect condition.

In order to obtain the recurrence matrix and plots, we need to reconstruct the state space from the acceleration time series. As discussed earlier, the vibrational acceleration of the test gearbox in x , y and z directions is considered as the state vector as indicated in Equation (2). By application of the method explained in section 2, the RP of the state vector for each data segmentation is constructed. Samples of the RPs of the same state vector length (6000 points) for different gear conditions are presented in **Figure 8**. This visualization represents points in time series that are visited more than once.

By inspecting the RPs, the following observations are made. First, diagonal segments of varied lengths that are parallel to the main diagonal (line of identity LOI) are the main pattern of the plots. A diagonal line occurs when the trajectory visits the same region of the phase space at different times. This is also true for diagonal segments that are orthogonal to the LOI but with reversed time sequence. These orthogonal segments appear in the multiple crack (state points within the range of 5000 - 6000) and missing tooth (randomly distributed) gear conditions. Second, small vertical segments are observed in some recurrence plots, *i.e.*, the gear with the missing tooth condition. This indicates that the state at that location changes slowly or does not change. Third, white bands with different widths appear in the RPs at different locations *i.e.*, in the missing tooth condition (state points within the range of 4500 - 5000). These white bands appear to form a rectangle in the center of the healthy RP and an upper right edge of a rectangle in the single crack RP, which usually develops when some states are rare to occur. Finally, the missing tooth condition is easily distinguished from the other gear conditions. It has a higher density of points compared to the other gear conditions. Other than that, the plots consist of complicated patterns which are hard to interpret. Hence, quantitative measures are necessary to obtain a more objective way to investigate the plots.

Figure 9 presents a comparison between the computed RQA parameters for all four gear conditions. The calculated RQA for the four cases are first normalized to make the comparison coherent. A measure with a zero value implies that the quantity is close to the average of the corresponding quantities for the four gear conditions. A positive measure indicates a magnitude higher than the average and vice versa for a negative measure. It is evident from the results of the analysis that RQA parameters can easily distinguish the differences and detect the alterations in the signal, even though the difference of the recurrence plots for the four cases is challenging to detect by the naked eye. For

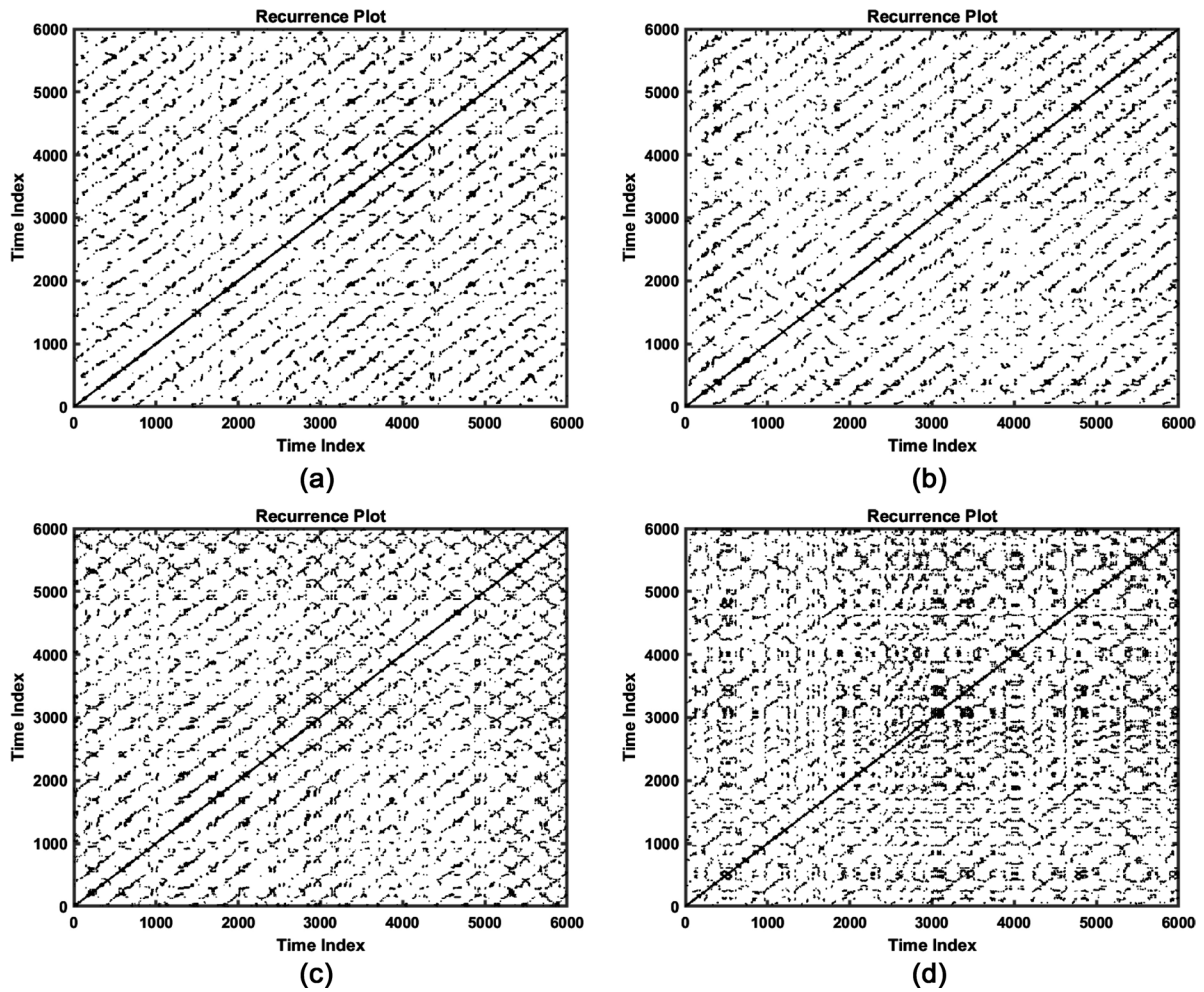


Figure 8. Recurrence plots for different gear conditions. (a) Healthy gear; (b) Gear with a single crack defect; (c) Gear with multiple crack defect; (d) Gear with a missing tooth.

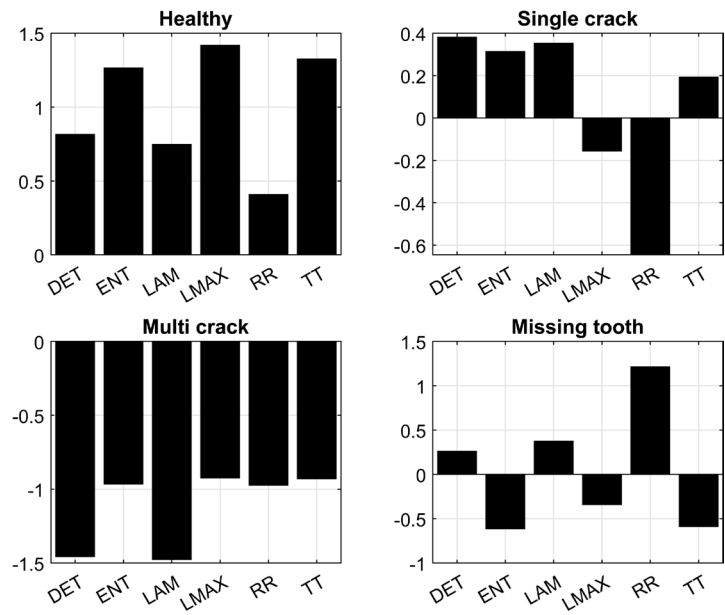


Figure 9. Recurrence quantification analysis for different gear conditions.

instance, the following can be observed. The healthy gear condition has positive RQA parameters relative to other gear conditions, which indicates higher RQA parameters than the average of all four gear conditions. On the other hand, the multiple crack condition has negative RQA parameters implying lower RQA parameters than average. A closer inspection of the figure reveals interesting characteristics of the individual RQA parameters. First, RR for the missing tooth condition has the highest magnitude compared to other gear conditions. This supports the results discussed previously when the RP was analyzed (missing tooth RP is the most dense). Second, inspecting the trapping time and the entropy leads to the same conclusion; the lowest values of the TT and $ENTR$ correspond to a gear system with multiple cracks and the highest values corresponds to a healthy gear system.

5. Fault Classification and SVM

In the previous section, we have illustrated how the condition of the gear system can be detected using the RQA parameters. We were able to measure and represent these influences by quantitative criteria. In contrast to this, the diagnostics procedure is the inverse, where the system status will be predicted using its dynamic response. In order to do that, support vector machine will be used as a classifier.

Support vector machine (SVM) is a supervised learning model used for classification and regression. This classification technique exploits training data in order to define a hyperplane that maximizes the separation distance between different classes in the feature space [31]. Support vectors are the data points that lie closest to the hyperplane or the decision surface. The optimum hyperplane can be determined by solving the following optimization problem.

$$\begin{aligned} \max Q(\alpha) &= \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j \gamma_i \gamma_j K(\mathbf{x}_i, \mathbf{x}_j) \\ \text{subject to } & 0 \leq \alpha_i \leq C \text{ and } \sum_{i=1}^N \alpha_i \gamma_i = 0 \end{aligned} \quad (10)$$

where α_i is the Lagrangian multiplier, \mathbf{x} and γ are the training feature set and the desired output set, respectively. N is number of training samples and $K(\mathbf{x}_i, \mathbf{x}_j)$ is the kernel function which maps the input features to a higher dimension space in order to change the features representation and to capture nonlinear patterns. For this study, two kernel functions are investigated: linear kernel, which can be expressed as:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$$

and radial basis (Gaussian) kernel, defined as the following:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\zeta \|\mathbf{x}_i - \mathbf{x}_j\|^2\right)$$

where $\zeta > 0$ is the kernel scale which determines the width of the kernel function. A small ζ value defines a kernel function with a wide width. Finally C is

the box constraint that controls the values of the Lagrangian multipliers. A higher value indicates a higher misclassification cost, leading to a more strict separation.

After solving the optimization problem and obtaining the optimal parameters α_i , a new test example can be classified using:

$$g(\mathbf{x}) = \sum_{i \in \sigma} \alpha_i \gamma_i K(\mathbf{x}, \mathbf{x}_i) + b \quad (11)$$

where σ denotes the set of indices of the support vectors. Note that only the support vectors play a role in making predictions for new data points. This stems from the fact that $\alpha_i = 0$ for data points that are not corresponding to support vectors. The bias b in Equation (11) can be calculated as follows:

$$b = \frac{1}{N_\nu} \sum_{n \in \nu} \gamma_n - \sum_{m \in \sigma} \alpha_m \gamma_m K(\mathbf{x}_n, \mathbf{x}_m) \quad (12)$$

where ν denotes the set of indices of the points that have $0 \leq \alpha_n \leq C$.

Originally, SVM was formulated for binary classification problems [32]. SVM applicability was then expanded to multi-class classification problems by performing a series of binary classification [33] [34] using various coding schemes such as one-versus-all (OVA) and one-versus-one (OVO). In this paper, OVO approach was used because of its ability to train quickly [35]. OVO performs $k(k-1)/2$ individual binary classifiers to evaluate all possible pairwise classifiers. When testing a new example, each classifier will contribute with one vote and the class with the most votes will be assigned as the predicted class.

The RQA parameters, which represent the feature vector, are used as input to the SVM classifier. For each set of data, six parameters are used in the feature set for training. Various SVMs were trained using 60% of the data samples (202 cases) and a cross validation algorithm with five folds. 40% of the data samples (139 cases) were used for testing the classifier. To find an optimal hyperplane, an optimization problem needs to be solved, as mentioned above. However, another optimization problem is encountered in tuning some of the SVM parameters such as box constraint and kernel scale. To solve this problem, the k-fold cross validation loss was minimized by searching on a given range of each optimized parameter. Box constraint was optimized for the two SVMs (linear and Gaussian). Meanwhile, kernel scale parameter was optimized for Gaussian SVMs.

The fault classification process is divided into three parts: 1) Anomaly detection 2) Defect classification 3) The effect of ranked RQA parameters. The detailed description of the above mentioned procedures is provided in the sequel.

5.1. Anomaly Detection

Anomaly detection is a technique used to identify abnormal classes or irregular behavior from what is defined as a normal standard. This has a lot of practical advantages in system condition monitoring applications such as the ability to

distinguish defective classes from healthy. In this part of the study, two-class (normal or fault) detection was performed where all gear defects were grouped in one class (D) and the healthy gear condition was used to define the second class (H). Two SVMs were optimized (Linear and Gaussian) and trained using all RQA parameters. The effectiveness of the classification models for the training data is presented by means of confusion matrix plots. The test confusion matrices of the linear and Gaussian SVMs are shown in **Table 1** and **Table 2**.

In the confusion matrix, the diagonal cells show the number and percentage of correct classifications by the trained classifier, while the off diagonal cells represent the misclassified predictions. For example, in **Table 1**, the linear SVM correctly predicted 38 cases as a healthy gear condition and 97 cases as a defective condition. It also misclassified 4 cases out of the 139 total test samples. On the other hand, the trained Gaussian SVM was slightly superior compared to the linear SVM and misclassified 3 cases as defective conditions.

Analyzing the confusion matrix is an important step in building a classification model. It gives strong clues as to where the classification model is going wrong. However, the number of misclassifications is not adequate to evaluate the performance of the SVM classifier. Taking only one performance metric is sometimes misleading. For example, a classifier with a relatively low misclassification rate might predict some of the classes fairly accurately but performs poorly for other classes, which might be critical to a certain application. Thus, to describe the performance of the classifier for the fault detection application, some metric rates are calculated from the confusion matrix such as precision (P), recall (R) and overall accuracy.

For each class, precision measures the rate of correct predictions out of all

Table 1. Confusion matrix of the linear SVM classifier for anomaly detection.

| | | Test | | Recall | Precision |
|--------|---|------------------|-------------|------------------------|-----------|
| | | Confusion Matrix | | | |
| Output | H | 38 27.3% | 2 1.4% | 95.0% | 95.0% |
| | D | 2 1.4% | 97 69.8% | 98.0% | 98.0% |
| | | H | D | Overall Accuracy 97.1% | |
| | | Target | | | |

Table 2. Confusion matrix of the Gaussian SVM classifier for anomaly detection.

| | | Test | | Recall | Precision |
|--------|---|------------------|-------------|------------------------|-----------|
| | | Confusion Matrix | | | |
| Output | H | 38 27.3% | 1 0.7% | 95.0% | 97.4% |
| | D | 2 1.4% | 98 70.5% | 99.0% | 98.0% |
| | | H | D | Overall Accuracy 97.8% | |
| | | Target | | | |

predictions that were made by the classifier. In other words, when the classifier predicts a class, precision indicates how often the prediction is correct. Recall measures the correctly predicted rate of the actual samples for a given class. If the classifier has high recall and low precision for a certain class, this means that the classifier is biased to that class. A high precision and low recall classifier for a given class indicates that the classifier is too conservative. When the classifier predicts that given class, it is usually correct but it is highly unlikely to predict it because of its low recall rate. Ideally, a classifier with high recall and high precision is what we seek. Finally, the overall accuracy of the classifier is the rate of the correct prediction.

For a binary confusion matrix, healthy condition is designated as negative class and defective condition is designated as positive class. True positive (TP) is correctly classified as a defective condition, false positive (FP) is incorrectly classified as a defective condition. In contrast, true negative (TN) is correctly classified as healthy condition and false negative (FN) is incorrectly classified as healthy condition. For example in the confusion matrix (**Table 1**) $TN = 38$, $TP = 97$, $FN = 2$, and $FP = 2$.

The generalized formulas of false negative FN_i , false positive FP_i , true negative TN_i , true positive TP_i , precision P_i , recall R_i and overall accuracy for a given class i are presented below. Note that M_{ij} is the element (i, j) in the confusion matrix. This generalization is presented here so it can be applied later for multiple class confusion matrix problems.

$$FN_i = \sum_{j=1(j \neq i)} M_{ji} \tag{13}$$

$$FP_i = \sum_{j=1(j \neq i)} M_{ij} \tag{14}$$

$$TN_i = \sum_{j=1(j \neq i)} \sum_{k=1(k \neq i)} M_{jk} \tag{15}$$

$$TP_i = M_{ii} \tag{16}$$

$$R_i = \frac{TP_i}{TP_i + FN_i} \tag{17}$$

$$P_i = \frac{TP_i}{TP_i + FP_i} \tag{18}$$

$$\text{Overall Accuracy} = \frac{TP_{\text{all}}}{\text{Total test samples}} \tag{19}$$

where $TP_{\text{all}} = \sum_{j=1} TP_j$ is the total number of true positives and represents the summation of the confusion matrix diagonal. Now, by inspecting the confusion matrices in **Table 1** and **Table 2**, the recall and precision are calculated for each gear condition and the overall accuracy is calculated for each classifier. For linear SVM, the healthy condition has 95.0% for both recall and precision whereas the defective condition has 98.0% for both recall and precision. The overall accuracy of the linear SVM classifier is 97.1%. In contrast, Gaussian SVM

has better precision in predicting the healthy condition (97.4%) and a better recall rate in predicting the defective condition (99.0%). The overall accuracy of the Gaussian SVM is 97.8%.

This result is remarkable for several reasons. First, employing RQA parameters as SVM features can provide valuable information in characterizing the dynamics of various gear faults. This helps to discriminate healthy gear condition from defective conditions. Second, both SVMs have high recall and high precision for each gear condition, which indicates an accurate prediction for detecting the health status and identifying any abnormality including the three defective gear conditions (SCD, MCD, and MTD). Finally, no *a priori* knowledge of the system was included in the features. This implies that the RQA approach can be conveniently applied to diverse dynamical systems in an automated process, with minimal need for adaptation and reliance on expert knowledge about the system.

5.2. Defect Classification

In this subsection, we studied the effectiveness of RQA parameters as features to classify different gear conditions including healthy gear and defective gears with single crack, multiple crack and missing tooth conditions. Two SVMs (linear and Gaussian) were trained as indicated previously by using optimal parameters then tested on a new set.

The test confusion matrices for the linear and Gaussian SVMs are shown in **Table 3** and **Table 4**, respectively. Out of 139 test samples, 4 cases were misclassified, which corresponds to 2.9% of all test samples. The misclassified cases are divided equally between the healthy and the single crack gear conditions. Moreover, the classifier is successful in identifying multiple crack and missing tooth conditions with zero misclassifications. Similarly, the Gaussian SVM performed well, with 5 misclassifications in total, in which 4 healthy cases were incorrectly classified as a single crack condition. This corresponds to 2.9% of all samples. In addition, one case of the single crack defect was misclassified as a healthy gear condition. Furthermore, multiple crack and missing tooth conditions were correctly classified with zero misclassifications.

Table 3. Confusion matrix of the linear SVM classifier.

| | | Test Confusion Matrix | | | | Recall | Precision |
|--------------|-----|-----------------------|-------------|-------------|-------------|------------------------|-----------|
| Output Class | H | 38 27.3% | 2 1.4% | 0 0.0% | 0 0.0% | 95.0% | 95.0% |
| | SCD | 2 1.4% | 38 27.3% | 0 0.0% | 0 0.0% | 95.0% | 95.0% |
| | MCD | 0 0.0% | 0 0.0% | 40 28.8% | 0 0.0% | 100.0% | 100.0% |
| | MTD | 0 0.0% | 0 0.0% | 0 0.0% | 19 13.7% | 100.0% | 100.0% |
| | | H | SDC | MCD | MTD | Overall Accuracy 97.1% | |
| | | Target Class | | | | | |

Table 4. Confusion matrix of the Gaussian SVM classifier.

| | | Test Confusion Matrix | | | | Recall | Precision |
|--------------|-----|-----------------------|-------------|-------------|-------------|------------------------|-----------|
| | | H | SCD | MCD | MTD | | |
| Output Class | H | 36 25.9% | 1 0.7% | 0 0.0% | 0 0.0% | 90.0% | 97.3% |
| | SCD | 4 2.9% | 39 28.1% | 0 0.0% | 0 0.0% | 97.5% | 90.7% |
| | MCD | 0 0.0% | 0 0.0% | 40 28.8% | 0 0.0% | 100.0% | 100.0% |
| | MTD | 0 0.0% | 0 0.0% | 0 0.0% | 19 13.7% | 100.0% | 100.0% |
| | | H | SCD | MCD | MTD | Overall Accuracy 96.4% | |
| | | Target Class | | | | | |

Performance metrics such as recall, precision and overall accuracy were calculated for the linear and Gaussian SVMs. First, linear SVM shows a more balanced precision/recall trade-off for healthy and single crack defect conditions. Meanwhile, Gaussian SVM is more biased to the single crack condition over the healthy condition. Second, both linear and Gaussian SVMs have virtually perfect recall and precision for multiple crack and missing tooth conditions. Finally, the overall accuracies of linear SVM (97.1%) and Gaussian SVM (96.4%) are fairly close. The linear SVM is chosen to continue further analysis in the next subsection.

In general, RQA parameters and optimized SVMs seem to achieve significant results in identifying various gear conditions in a mock-up of a helicopter gearbox. Outstanding performance was achieved with 100% accuracy, 100% recall and 100% precision in detecting multiple crack and missing tooth conditions. The most challenging problem was distinguishing between healthy and single crack defect conditions where RQA parameters were capable of capturing the differences. A balanced classifier with 95.0% recall and 95.0% precision was achieved. In summary, the classifier is extraordinarily effective in predicting all gear conditions using the RQA parameters as features.

5.3. The Effect of Ranked RQA Parameters

As previously indicated, different kinds of information were extracted from the RPs using the RQA parameters. For each gear condition, some of the RQA parameters are more correlated and informative than others. A robust and accurate prediction is achieved by extracting an efficient set of features that can characterize the system response in a unique way. The feature set should also provide as much information as possible about the intrinsic dynamics of the system. This motivates ranking and selecting a subset of relevant features to maximize the correlation between the extracted features and the predicted classes. Furthermore, this aids in understanding the contribution of each ranked feature to the classification process and to which gear condition the feature information contributes. To do so, mutual information is used as a feature ranking technique. This technique was developed in our past work [36] [37] [38]

to diagnose rolling bearings with various health conditions.

The features are ranked based on the mutual information content between the feature subset and the gear condition using a greedy search algorithm [39]. The feature with the highest mutual information will be assigned to another set. Along with that, an evaluation measure (prediction accuracy) that scores the different feature subsets is determined. This process will be repeated until the whole feature set is covered. The new set that corresponds to the highest mutual information is the ranked feature set. The feature subset with the highest prediction accuracy is the selected feature subset.

Figure 10 depicts the mutual information relationship with the ranked feature set. As illustrated by the figure, *LAM* parameter has the highest mutual information with the four gear conditions under study, then followed by *ENT*, *RR*, *LMAX*, *TT* and lastly, *DET*. This important result is summarized in **Table 5**. To test these results, a linear SVM is trained using a subset of the ranked feature set. The first subset is the highest ranked feature. Then, in an iterative fashion, a new feature will be added from the ranked feature set to form a new subset. For clarification, **Table 6** lists the feature subset indices and their corresponding features. Recall and precision were calculated for each gear condition to examine the effect of the new added feature. The overall accuracy and the F_1 score are calculated for each subset. F_1 score is the harmonic mean of recall and precision, which compares both measures in one metric and is calculated as follows:

$$F_1 \text{ score} = \frac{2P_i R_i}{P_i + R_i} \quad (20)$$

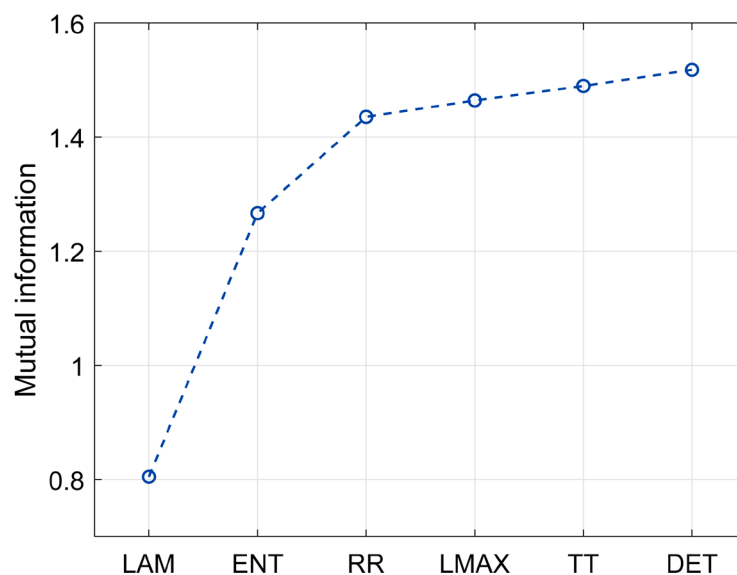


Figure 10. Ranked features with their corresponding mutual information contribution.

Table 5. Ranked features.

| Feature rank | 1 | 2 | 3 | 4 | 5 | 6 |
|--------------|------------|------------|-----------|-------------|-----------|------------|
| Feature | <i>LAM</i> | <i>ENT</i> | <i>RR</i> | <i>LMAX</i> | <i>TT</i> | <i>DET</i> |

Table 6. Ranked feature subsets and their corresponding features.

| Feature subset | Features |
|----------------|------------------------------------|
| 1 | <i>LAM</i> |
| 2 | <i>LAM, ENT</i> |
| 3 | <i>LAM, ENT, RR</i> |
| 4 | <i>LAM, ENT, RR, LMAX</i> |
| 5 | <i>LAM, ENT, RR, LMAX, TT</i> |
| 6 | <i>LAM, ENT, RR, LMAX, TT, DET</i> |

where R_i and P_i denote the overall recall and precision which can be calculated by averaging all the individual values of recall and precision.

Figure 11 presents various performance metrics for each feature subset: a) and b) show the change of recall and precision for each gear condition at each feature subset. Meanwhile, c) presents the overall accuracy and F_1 score for each feature subset. A detailed discussion of the effect of adding each of the ranked RQA parameters to the feature subset is presented below.

- The effect of laminarity (*LAM*)

Laminarity provides a high correlation to the multiple crack condition with high recall and precision rates. Detection of the healthy condition is in an acceptable range but it does not give decisive information. Furthermore, laminarity's information content regarding the single crack condition can be considered unreliable. It has low precision, which makes the classifier biased to that condition. Finally, laminarity does not help in detecting the missing tooth condition because of its poor recall rate, which makes the classifier too conservative to make any missing tooth predictions. Even though the overall accuracy is 75.0%, this represents an example of relying only on the accuracy of the classifier. The classifier failed in predicting an entire gear condition, *i.e.*, the missing tooth condition. Moreover, F_1 score can not be determined due to the lack of the precision metric for the missing tooth condition (no prediction was made).

- The effect of entropy (*ENT*)

Entropy adds an evident value for detecting the missing tooth condition. For example, 100.0% precision and 78.0% recall was achieved for the missing tooth condition. Also, entropy highly contributes to the prediction accuracy of the single crack condition by increasing the precision. This leads to a reduction in the classifier's bias toward the single crack condition. Adding the entropy feature has enhanced the classifier's ability to distinguish both the healthy condition and the multiple crack condition. The quality of detection has improved due to the information relevance contributed by the entropy with 92.0% overall accuracy and 92.0% F_1 score.

- The effect of recurrence rate (*RR*)

The recurrence rate's effect is significant due to a notable improvement in detecting the multiple crack and missing tooth conditions. The quality of

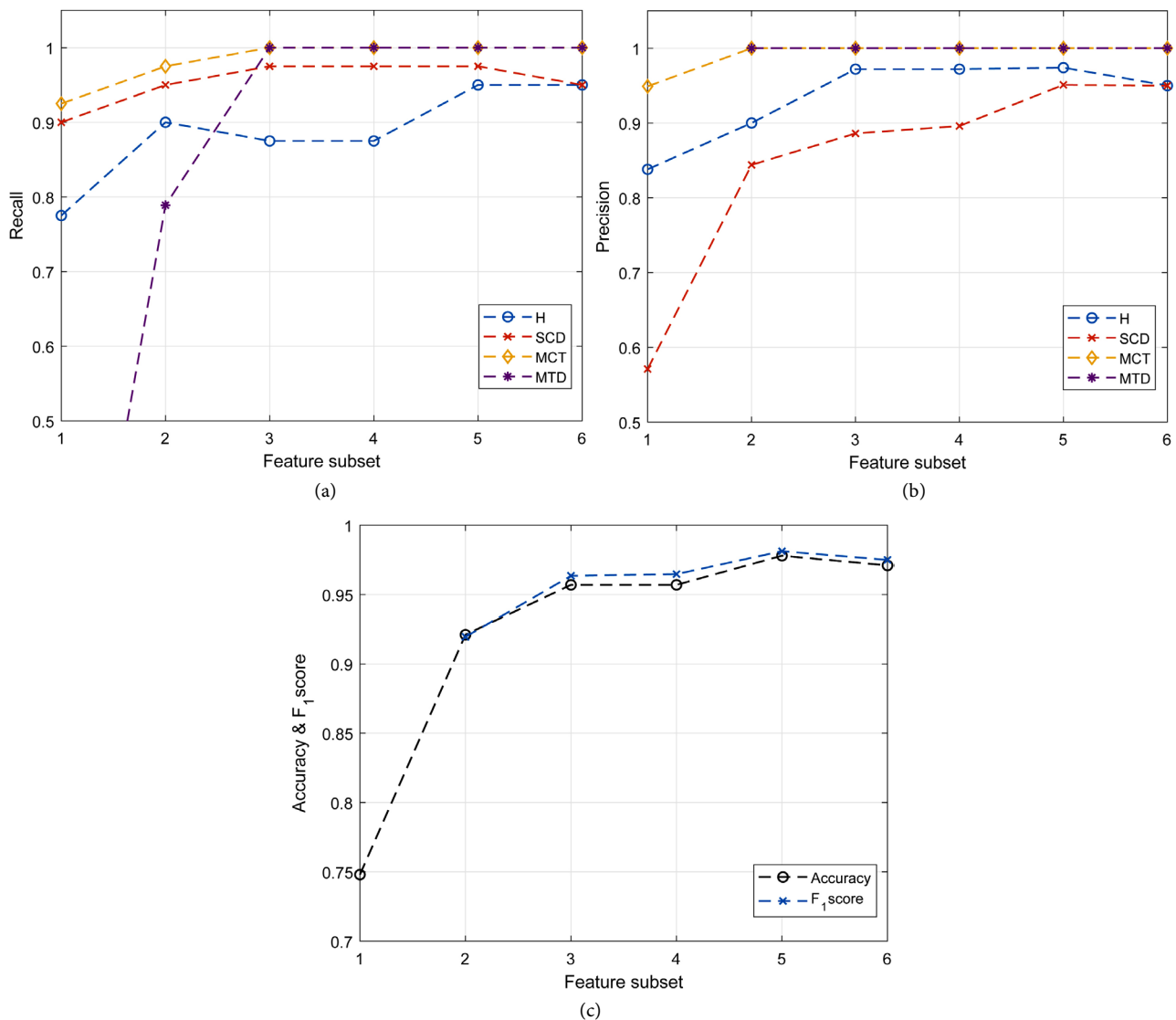


Figure 11. Performance metrics for ranked feature subsets. (a) Recall for each gear condition; (b) Precision for each gear condition; (c) Accuracy and F_1 score.

classification of the multiple crack and the missing tooth conditions is virtually perfect. It can be seen that the recall of the healthy condition decreased while the precision increased. This is reflected in the improvement in detection of the single crack condition.

- The effect of the longest diagonal line ($LMAX$)

No change is noticed on the performance metrics by adding the $LMAX$ parameter. However, changing the order and trying different RQA parameters affects the performance negatively. This supports the order of features that was determined by the mutual information ranking technique. Additionally, features work together in a nonlinear fashion, where a certain feature might not give enough information but combining it with another can add more value.

- The effect of trapping time (TT)

Trapping time has a positive effect on the healthy condition that is

represented by an increase the recall rate. It also has a positive effect on the single crack condition represented by an increase the prediction precision. This is reflected in the overall accuracy and the F_1 score, which reaches the highest values thus far with 97.8% and 98.0%, respectively.

- The effect of determinism (*DET*)

In this particular configuration, determinism gave no information. Moreover, the recall and the precision of the healthy and the single crack conditions are decreased.

The optimal subset is subset number 5, which contains all RQA parameters except determinism, which adds no additional information for classifying the gear conditions in the study. Laminarity provides valuable information about the multiple crack condition. The entropy parameter correlates to the missing tooth condition. Only by using three features *LAM*, *ENT* and *RR*, the accuracy of 95.6% is achieved (0.16% from using all features).

6. Conclusion

In this paper, a mock-up of a helicopter gear box system was studied in order to detect and identify various gear faults. Four gear conditions including healthy, single crack, multiple cracks, and missing tooth were investigated under a constant operating condition. The RP method, which is based on visualizing high-dimensional dynamical systems using a two-dimensional plot, was applied. The RQA parameters were then used as an input into two SVMs. The fault classification process was divided into three parts: 1) Anomaly detection, 2) Defect classification and 3) The effect of ranked RQA parameters. Results indicate that RQA parameters provide valuable information in characterizing the dynamics of various gear faults in order to discriminate the healthy gear condition from defective conditions. Also, an outstanding performance was achieved using RQA parameters to identify various gear conditions with 100% accuracy, 100% recall and 100% precision in detecting multiple cracks and missing tooth conditions. In general, the classifier is extraordinarily effective in predicting all gear conditions using the RQA parameters. Finally, mutual information is used to rank the extracted features. An optimal feature subset was determined using *LAM*, *ENT*, *RR*, *LMAX*, and *TT*. A correlation between the RQA parameters used in the study and the different gear conditions was discussed.

Acknowledgements

This work is supported by the US Office of Naval Research under the grant ONR N00014-15-1-2311 with Capt. Lynn Petersen as the Program Manager. We deeply appreciate this support and are humbled by ONR's enthusiastic recognition of the importance of this research.

References

- [1] Bajric, R., Zuber, N. and Isic, S. (2013) Recent Advances in Vibration Signal

- Processing Techniques for Gear Fault Detection—A Review. *Applied Mechanics and Materials*, **430**, 78-83. <https://doi.org/10.4028/www.scientific.net/AMM.430.78>
- [2] Hussain, S. and Gabbar, H. (2011) A Novel Method for Real Time Gear Fault Detection Based on Pulse Shape Analysis. *Mechanical Systems and Signal Processing*, **25**, 1287-1298. <https://doi.org/10.1016/j.ymsp.2010.11.013>
- [3] Wang, W.Q., Ismail, F. and Farid Golnaraghi, M. (2001) Assessment of Gear Damage Monitoring Techniques Using Vibration Measurements. *Mechanical Systems and Signal Processing*, **15**, 905-922. <https://doi.org/10.1006/mssp.2001.1392>
- [4] Dalpiaz, G., Rivola, A. and Rubini, R. (2000) Effectiveness and Sensitivity of Vibration Processing Techniques for Local Fault Detection in Gears. *Mechanical Systems and Signal Processing*, **14**, 387-412. <https://doi.org/10.1006/mssp.1999.1294>
- [5] Randall, R.B. (1982) A New Method of Modelling Gear Faults. *ASME Journal of Mechanical Design*, **104**, 259-267. <https://doi.org/10.1115/1.3256334>
- [6] Chad, E.F. (1998) Synchronous Sampling Sideband Orders from Helical Planetary Gear Sets. Master's Thesis, Blacksburg.
- [7] Jardine, A.K.S., Lin, D. and Banjevic, D. (2006) A Review on Machinery Diagnostics and Prognostics Implementing Condition-Based Maintenance. *Mechanical Systems and Signal Processing*, **20**, 1483-1510. <https://doi.org/10.1016/j.ymsp.2005.09.012>
- [8] Serridge, M. (1990) Ten Crucial Concepts behind Trustworthy Fault Detection in Machine Condition Monitoring. Springer, Berlin, 729-740. https://doi.org/10.1007/978-94-009-1914-3_42
- [9] Sankaravelu, A., Noah, S.T. and Burger, C.P. (1994) Bifurcation and Chaos in Ball Bearings. *ASME Applied Mechanics Division Publications*, **192**, 313.
- [10] Mevel, B. and Guyader, J. (1993) Routes to Chaos in Ball Bearings. *Journal of Sound and Vibration*, **162**, 471-487. <https://doi.org/10.1006/jsvi.1993.1134>
- [11] Kappaganthu, K. and Nataraj, C. (2011) Nonlinear Modeling and Analysis of a Rolling Element Bearing with a Clearance. *Communications in Nonlinear Science and Numerical Simulation*, **16**, 4134-4145. <https://doi.org/10.1016/j.cnsns.2011.02.001>
- [12] Wang, W.J. and McFadden, P.D. (1996) Application of Wavelets to Gearbox Vibration Signals for Fault Detection. *Journal of Sound Vibration*, **192**, 927-939. <https://doi.org/10.1006/jsvi.1996.0226>
- [13] Peng, Z. and Chu, F. (2004) Application of the Wavelet Transform in Machine Condition Monitoring and Fault Diagnostics: A Review with Bibliography. *Mechanical Systems and Signal Processing*, **18**, 199-221. [https://doi.org/10.1016/S0888-3270\(03\)00075-X](https://doi.org/10.1016/S0888-3270(03)00075-X)
- [14] Combet, F. and Gelman, L. (2007) An Automated Methodology for Performing Time Synchronous Averaging of a Gearbox Signal without Speed Sensor. *Mechanical Systems and Signal Processing*, **21**, 2590-2606. <https://doi.org/10.1016/j.ymsp.2006.12.006>
- [15] Yuan, J., He, Z. and Zi, Y. (2010) Gear Fault Detection Using Customized Multi-wavelet Lifting Schemes. *Mechanical Systems and Signal Processing*, **24**, 1509-1528. <https://doi.org/10.1016/j.ymsp.2009.11.003>
- [16] Mohammed, O.D., Rantatalo, M., Aidanp, J.-O. and Kumar, U. (2013) Vibration Signal Analysis for Gear Fault Diagnosis with Various Crack Progression Scenarios. *Mechanical Systems and Signal Processing*, **41**, 176-195. <https://doi.org/10.1016/j.ymsp.2009.11.003>
- [17] Li, B., Zhang, X. and Wu, J. (2017) New Procedure for Gear Fault Detection and

- Diagnosis Using Instantaneous Angular Speed. *Mechanical Systems and Signal Processing*, **85**, 415-428. <https://doi.org/10.1016/j.ymssp.2016.08.036>
- [18] Mohammed, O.D. and Rantatalo, M. (2016) Dynamic Response and Time-Frequency Analysis for Gear Tooth Crack Detection. *Mechanical Systems and Signal Processing*, **6667**, 612-624. <https://doi.org/10.1016/j.ymssp.2015.05.015>
- [19] Hong, L., Qu, Y., Dhupia, J.S., Sheng, S., Tan, Y. and Zhou, Z. (2017) A Novel Vibration-Based Fault Diagnostic Algorithm for Gearboxes under Speed Fluctuations without Rotational Speed Measurement. *Mechanical Systems and Signal Processing*, **94**, 14-32. <https://doi.org/10.1016/j.ymssp.2017.02.024>
- [20] Salem, A.-A. (2012) Condition Monitoring of Gear Systems Using Vibration Analysis. PhD Thesis, University of Huddersfield, Huddersfield.
- [21] Kwuimy, C.K., Kankar, P., Chen, Y., Chaudhry, Z. and Nataraj, C. (2015) Development of Recurrence Analysis for Fault Discrimination in Gears. *Proceedings of the ASME IDETC/CIE*, V008T13A079.
- [22] Vachtsevanos, G., Lewis, F., Roemer, M., Hess, A. and Wu, B. (2006) Intelligent Fault Diagnosis and Prognosis for Engineering Systems. Wiley, Hoboken. <https://doi.org/10.1002/9780470117842>
- [23] Randall, R.B. (2011) Vibration-Based Condition Monitoring. John Wiley & Sons, Ltd., Hoboken. <https://doi.org/10.1002/9780470977668>
- [24] Peng, F., Yu, D. and Luo, J. (2011) Sparse Signal Decomposition Method Based on Multi-Scale Chirplet and Its Application to The Fault Diagnosis of Gearboxes. *Mechanical Systems and Signal Processing*, **25**, 549-557. <https://doi.org/10.1016/j.ymssp.2010.06.004>
- [25] Syta, A., Jonak, J., Jedlinski, L. and Litak, G. (2012) Failure Diagnosis of a Gear Box by Recurrences. *Journal of Vibration and Acoustics*, **134**, Article ID: 041006. <https://doi.org/10.1115/1.4005846>
- [26] Mohamad, T.H., Samadani, M. and Nataraj, C. (2018) Rolling Element Bearing Diagnostics Using Extended Phase Space Topology. *Journal of Vibration and Acoustics*.
- [27] Mohamad, T.H., Kwuimy, C.K. and Nataraj, C. (2017) Discrimination of Multiple Faults in Bearings Using Density-Based Orthogonal Functions of the Time Response. *Proceedings of the ASME IDETC/CIE*, V008T12A035.
- [28] Mohamad, T.H. and Nataraj, C. (2017) Gear Fault Diagnostics Using Extended Phase Space Topology. *Annual Conference of the Prognostics and Health Management Society*.
- [29] Samadani, M., Mohamad, T.H. and Nataraj, C. (2016) Feature Extraction for Bearing Diagnostics Based on the Characterization of Orbit Plots with Orthogonal Functions. *Proceedings of the ASME IDETC/CIE*, Charlotte, 21-24 August 2016, V008T10A026. <https://doi.org/10.1115/DETC2016-60438>
- [30] Zbilut, J.P., Thomasson, N. and Webber, C.L. (2002) Recurrence Quantification Analysis as a Tool for Nonlinear Exploration of Nonstationary Cardiac Signals. *Medical Engineering & Physics*, **24**, 53-60. [https://doi.org/10.1016/S1350-4533\(01\)00112-6](https://doi.org/10.1016/S1350-4533(01)00112-6)
- [31] Bishop, C. (2006) Pattern Recognition and Machine Learning. Springer-Verlag, New York.
- [32] Cortes, C. and Vapnik, V. (1995) Machine Learning. *Support Vector Networks*, **20**, 273-297.
- [33] Vapnik, V. (1998) Statistical Learning Theory. Wiley, New York.

- [34] Schölkopf, B., Burges, C.J. and Smola, A.J. (1999) *Advances in Kernel Methods: Support Vector Learning*. MIT Press, Cambridge.
- [35] Wang, Z. and Xue, X. (2014) Multi-Class Support Vector Machine. In: *Support Vector Machines Applications*, Springer, Berlin, 23-48.
https://doi.org/10.1007/978-3-319-02300-7_2
- [36] Kappaganthu, K., Nataraj, C. and Samanta, B. (2011) Feature Selection for Bearing Fault Detection Based on Mutual Information. In: *IUTAM Symposium on Emerging Trends in Rotor Dynamics*, Springer, Berlin, 523-533.
https://doi.org/10.1007/978-94-007-0020-8_44
- [37] Kappaganthu, K. and Nataraj, C. (2011) Feature Selection for Fault Detection in Rolling Element Bearings Using Mutual Information. *Journal of Vibration and Acoustics*, **133**, Article ID: 061001. <https://doi.org/10.1115/1.4003400>
- [38] Kappaganthu, K. and Nataraj, C. (2011) Mutual Information Based Feature Selection from Data Driven and Model Based Techniques for Fault Detection in Rolling Element Bearings. *Proceedings of the ASME IDETC/CIE*, Washington DC, 28-31 August 2011, 941-953. <https://doi.org/10.1115/DETC2011-47822>
- [39] Duda, R.O., Hart, P. and Stork, D. (2001) *Pattern Classification*. Wiley Interscience, New York.