Experimental Monitoring of Eccentric Gears with different Mechanical Conditions

Mohamed Elforjani[a], David Mba[b], Bello Salihu[c]

[a]*SupervisoryEye, Bedford MK43 0JA, UK*
[b]*University of the Arts London, London WC1V 7EY, UK*
[c]*AAXAL, Milton Keynes MK10 9RG, UK*

m.elforjani@supervisoryeye.com or elforjani@gmail.com

Abstract

The use of Condition Monitoring (CM) for engineering design, and industrial gears, has dramatically changed the gear design and manufacturing process in a short space of time. As CM deployment grows rapidly, the feasibility of CM for the diagnosis and prognosis of most common gear failure modes has well been investigated and documented. However, this is not the case for the CM of industrial gear eccentricity. Previous published work, in particular, focused only on the use of simulation models as a basis to investigate the gear eccentricity. Simulations cannot always ease the detection of complex problems and complications that are experienced in real operations. For instance, excessive eccentricity can be considered as one of manufacturing errors that may severely lead to direct effect on the overall dynamic performance of gears. Yet, it produces very high modulated mesh frequency rates; thus, making the detection of faulty machine components very difficult or even impossible. With this in mind, this paper presents the first known attempt at the diagnosis and prognosis of experimental vibration datasets from five different eccentric gear conditions. Datasets were first analysed using Signal Intensity Estimator (SIE) method in time and frequency domains. Then, the data was subjected to additional processing for the classification of gearbox status. Observations from the results showed that the proposed techniques could successfully discriminate the "good" and "bad" gears.

Keywords: Condition Monitoring, Industrial Gearboxes, Gears Eccentricity, Simulation Models, Experimental Vibration Data, Signal Intensity Estimator.

1. Introduction

Gears have been utilized for various machining and industrial processes since the ancient era. Applications of industrial gears include wind turbines, paper mill plants, power

plants, iron industry, aerospace industry, automobiles, railways and trains, lifts and elevators etc. Unlike other power transmission drives, gears are compact and can carry more load. The assessment of condition monitoring (CM) feasibility for gears health integrity was determined. Off the shelf, research attempts at the use of different CM techniques for the diagnosis and prognosis of well-known gear failure modes such as tooth wear, cracked or broken gears etc., to name a few, were well documented [1 and 2]. However, some other failure modes were not experimentally investigated. For instance, procedures and methods for the evaluation of the impact of gear eccentricity on the fault detection and gear performance were instead limited to the use of simulation models. The gears are considered to be eccentric if they are spinning around a center of rotation that is offset from the theoretical center of rotation [3]. When the eccentricity far exceeds allowed range, it affects the mesh depth and the backlash required for a film lubrication between the teeth to prevent the overheating and the subsequent teeth damages. Yet, improper backlash and mesh depth may lead to heavy loading (high stress concentration) at the tip or root of the mating pair; thus, causing what-so-called tip and root interference. This prematurely tight mesh depth will break down the lubricant film and eventually produce definite signs of metal removal at the tips and roots, generally abrasion of the teeth, and may often show destructive scratch marks in the direction of rotation [4 and 5]. Further, excessive tooth loads result in root stresses higher than the endurance limit of the material. When the gears are loaded in this manner and subjected to enough repeated stress cycles, the gear tooth will fail. Overload may also result in a sudden misalignment; leading to a bearing seizure, failure of driven equipment, or failed or wiped gear bearings.

In this paper, only published work relevant to the study of gear's eccentricity is briefly highlighted. Several studies have analytically explored the effects of the eccentricity on the overall gear behaviour. For instance, Zhou et al. [6] discussed the impact of gear eccentricity errors on the mesh stiffness and calculated the time varying mesh stiffness for spur gears. Effects on dynamic transverse and torsional responses were also quantified. Chen [7] investigated the dynamics of a helical-geared rotor-bearing system 3D-dimensional motion due to rotating shaft deformation. Lagrange's equation was employed to drive the equations of motion. A dynamic model for a planetary gear train (PGT) was developed to simulate the impacts of gear eccentricity on mesh stiffness and overall dynamic performance [8]. It is worth to mention that the Potential Energy method was used to calculate the mesh stiffness. Influence of eccentricity errors on the dynamic characteristics of planetary gears were also numerically investigated by some other researchers [9 - 13]. A nonlinear dynamic model was proposed to investigate the impacts of the time variation of the pressure angle and gear backlash on spur gear structures [14]. The model was also used to study the effects of internal and external excitation, gear gravity and unbalance mass on the overall gear performance. It was concluded that the time variation of the pressure angle and the gear backlash may lead to a chaotic motion and change the tooth impact status. Another type of studies involved the investigation of the effect of eccentricity

on the leakage of circular arc involute circular arc gear pumps was undertaken by Zhou et al. [15]. Characteristics of the leakage were studied under different operating conditions such as outlet pressure and rotational speeds. It was shown that the leakage rate increases with the increase of the eccentricity. Suxiang et al. [16] presented the design principle with a new transmission mechanism for an eccentric modified gear to ensure a variation of gear tooth backlash within the allowable range.

Until recently, major studies in this area only considered the use of simulation models to alleviate the difficulty in the analysis of problems and complications of eccentric gears. A closer look at the review shows that there has been no known alternative work for the direct experimental investigation of the CM of gear eccentricity. The reason behind this can be attributed to some factors such as the simulation models are easy to run, require lower costs than the significant costs for experimental investigations, and the simplicity to handle the irreversibility of some processes. However, the experiments should be performed in a practical environment to ensure reproducible results, though it appears that these models could generally produce some promising results. Further, in the previous simulation work both failure modes (e.g. eccentricity and gear backlash) and faulty components were usually simulated at controllable operating conditions. If a specific machine fault can possibly be simulated, there is no way to ensure that all other failure modes in practice can appropriately be virtualized due to structural noise, extremely rapid variation in load and rotational speed and signals originating from head-to-head components. Yet, the collected data from gears, in real-world applications, very often, consists of extremely different modulation rates, and hence it needs efficient and workable CM tools that are not only capable of analyzing the data very thoroughly and accurately but also able to offer an immediate online response for sudden machines' abnormality. No published work exists to support the feasibility of the simulation models in analysing such high modulated data and identifying the offline and/or online machine faults (e.g. gear eccentricity). This problem has not received a simple solution so far and there exists no method that gives a satisfactory solution to the detection of eccentric gears. Hence, the aim here is to propose some CM techniques, which can be useful in a number of domains. The presented methods in this investigation are fundamentally different from the approaches proposed in the literature. In summary, the main contributions and the distinct advantage of this work include the following:

- This study is set out to explore the feasibility of CM tools for the detection of gear eccentricity by analyzing experimental vibration dataset (first known attempt).
- First known use of SIE technique, Wavelet Transform and Multi-Taper Spectral analysis for the detection of gear eccentricity.
- First known attempt at the analysis of eccentric gearboxes with multi-faulty components (e.g. bearings and gears).
- First known attempt at the classification of eccentric gearboxes with multi-faulty components (e.g. bearings and gears) using Machine Learning algorithms.

Nomenclature	
BPFI	Bearing Inner Race Frequency
BPFO	Bearing Outer Race Frequency
d_b	Bearing Ball Diameter
d_p	Bearing Pitch Diameter
f_r	Rotor Speed
n	Number of Rollers
Θ	Bearing Contact Angle
ENG	Signal Energy
ENT	Entropy
KU	Kurtosis
CF	Crest Factor
SIE	Signal Intensity Estimator in Time Domain
SIEF	Signal Intensity Estimator in Time Domain
WF	Wave Factor
RMS	Root Mean Square
EI	Energy Index
SPI	Spectral Power Intensity

2. Vibration Data and Instrumentations Setup

For the purpose of this research work, the 2009 challenge gearbox dataset (representative of generic industrial gearboxes) supplemented by the Prognostics and Health Management (PHM) Society is analysed [17]. Although some attempts have been made to previously analyze this dataset, it is somewhat surprising that the feasibility of CM tools for the detection of gear eccentricity has not been previously evaluated [18 and19]. The Endevco Data Acquisition System (specifications: 10mv/g Accel, +/- 1% error, Resonance > 45KHz) has synchronously acquired the data from the gearbox, shown in figures 1 and 2, at a sampling rate of 66.6667kHz (200/3 kHz) through two accelerometer channels attached to the retaining plates of the input and output shafts. In addition to high and low load conditions, different rotational speeds ranging from 30Hz to 50Hz were

applied. To control the rotational speed, a tachometer, generates ten pulses per one revolution, was employed. The gearbox consists of one input pinion with 32 teeth (installed on the input shaft), idler shaft with two gears (96 teeth and 48 teeth), and one output pinion with 80 teeth (installed on the output shaft), see figure 2. The reduction gear ratio from the input shaft to the output shaft is 5 to 1 (32/96*48/80).

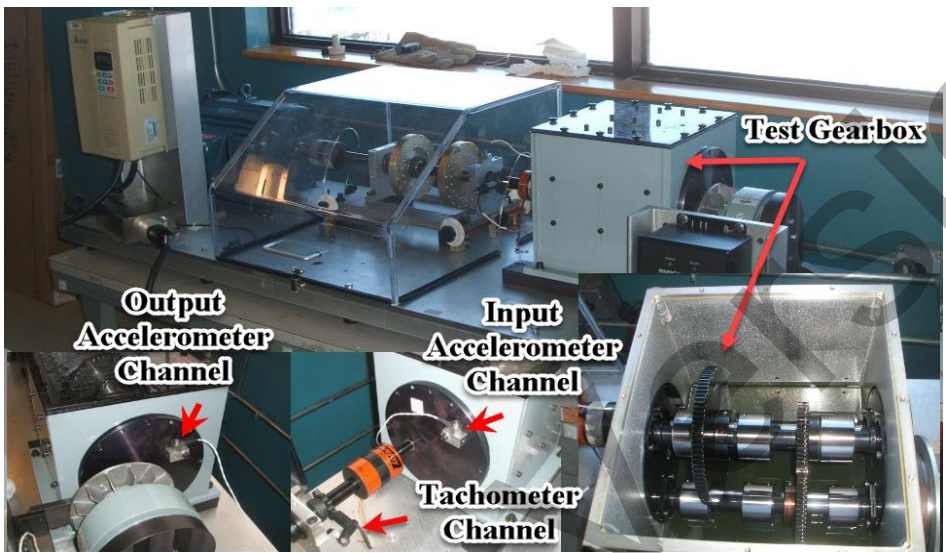

Fig. 1. Test rig with the test gearbox.

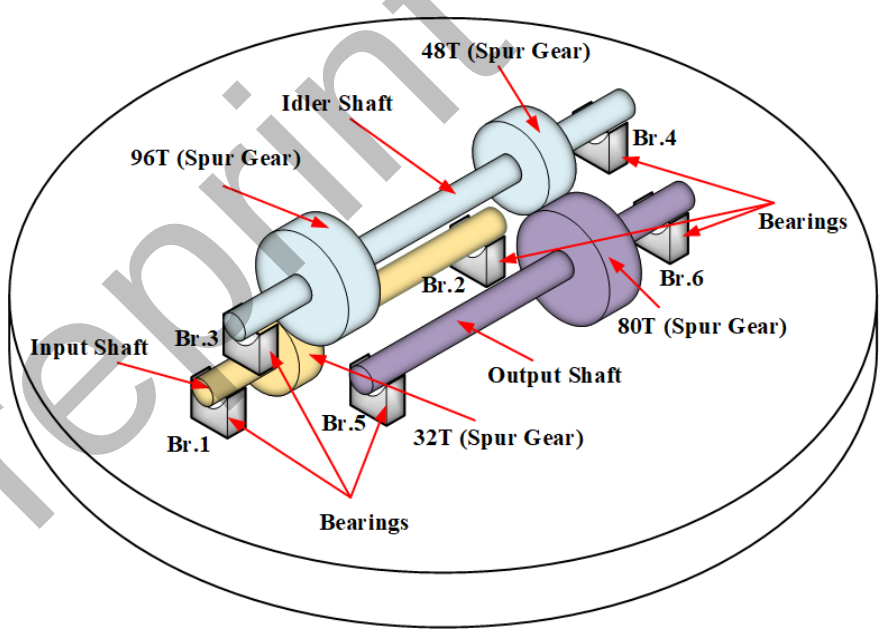

Fig. 2. Schematic diagram for the test gearbox.

The shafts were supported by six MB Manufacturing ER-10K bearings (specifications: 8 rolling elements (n) with a diameter (d_b) of 0.3125", 1.319" pitch diameter (d_p), and 0^o contact angle (Θ)). Prior testing, different failure modes were introduced to the mechanical components (shafts, bearings and gears) in the gearbox. Two types of gears (spur and helical) were experimented and it should however be noted, that only four different gear cases (spur type) were tested under the eccentricity conditions, presented in Table 1. Due to this, the primary focus of this work will only be on the analysis of the healthy gear and four gear cases under eccentricity conditions.

Table 1: Failure modes and mechanical conditions of gearbox components.

	Gears				Bearings					
Case	**32T**	**96T**	**48T**	**80T**	**Br.1**	**Br.2**	**Br.3**	**Br.4**	**Br.5**	**Br.6**
1	healthy	healthy	healthy	healthy	healthy	healthy	healthy	healthy	healthy	healthy
2	healthy	healthy	eccentric	healthy	healthy	healthy	healthy	healthy	healthy	healthy
3	chipped	healthy	eccentric	healthy	healthy	healthy	healthy	healthy	healthy	healthy
4	healthy	healthy	eccentric	broken	ball	healthy	healthy	healthy	healthy	healthy
5	healthy	healthy	eccentric	broken	inner	healthy	ball	healthy	outer	healthy

3. Data Analysis & Results Presentation

The vibration data, including Case 1 as a baseline, was analyzed using the Signal Intensity Estimator (SIE) technique. SIE is a well-established method and has successfully been applied to the condition monitoring of real-world applications [2, 20, 21 and 22]. It is basically a piecewise windowing method that attempts to analyse any CM data following a twostep process. The process incorporates the statistic chart of the signal intensity over a time period and its counterpart in signal spectrum as the frequency domain feature. Elforjani et al. [22] discussed the computational complexity and the mathematical correlations used to compute the SIE in very sufficient details. In general, the SIE method differs from the other tools because it uses the cumulative sum (CUSUM) to monitor any sequential sample [22]. Using CUSUM allows the total at any time interval to be determined without having to sum the entire sequence. Yet, if certain activities are not individually important, CUSUM can easily save on recording the sequence itself. Further, processing multiple samples of failure histories using CUSUM's results in greater sensitivity for detecting transient shifts or fluctuations in trends over time. At this point, it is very important to note that the SIE method is also completely different from the classic envelope approach, as its algorithm mainly involves splitting the entire signal into equally sized segments and then calculating the signal intensity from current and previous values using CUSUM method. When non-transient type signals are analyzed, identical statistically SIE charts will be noted and the ratio of SIE between any two adjacent segments will approximately approach a value of one. For the signals associated with transient

characteristics, this ratio will be greater than one [22].

To apply the SIE, the sum of the cumulative sum ($SCS_{segment}$) of a predefined segment (window) in a given time domain signal is normalized to the overall root mean square ($RMS_{overall}$) value of the same signal. As the rotational speed (RPM) is considered to be one of the causative factors of non-stationarity in the measured data, SIE algorithm makes use of it to determine the maximum analysis frequency (F_{max}), which is eventually used to calculate the size of the segments. The calculated F_{max} values are compared with published SIE Standard Frequencies (SSF's), which have been quantitatively and qualitatively validated experimentally to eventually ensure the reproducible results. SSF is very important factor in the SIE algorithm as it allows for a robust calculation of the most appropriate windows (segments) size and significantly helps to avoid any indefinite settings. Further, it significantly helps to alleviate the issue of what is called the losses in time localization and frequency localization [22].

3.1. Diagnostic Analysis of Eccentric Gears

The analysis involved the processing of vibration data using frequency domain techniques. Classical Fast Fourier Transform (FFT) was applied to the resulting SIE values. It is very obvious that the results completely tallied with the expectations of the feasibility of SIE in the identification of mechanical faults even if the analyzed data is very extremely modulated; thus, reinforcing the previous view of Elforjani et al. [21]. As expected, the resulting signal spectrum from Case 1 (healthy gears/baseline case) is insignificant and difficult to understand. In the case of normal/healthy gears, scattered low amplitude peaks may be present [23 and 24]; in Case 1 a vibration spectrum with broad frequencies was noted, see figure 3. In addition to this, it is very typical in such cases to see 1x RPM of the pinion and its first Gear Mesh Frequency (GMF).

Unlike Case 1, observations from the power spectrum of Case 2 show a clear indication of eccentric gear (48T). Evidence of this is a significant high amplitude in the sidebands, corresponded to 1Xg RPM (10Hz), with the harmonics around the GMF of this gear; this is a clear sign of a very severely eccentric gear [23 and 24]. Yet, another strong evidence of the eccentricity is the high amplitude in 1x, 2x and 3x GMF (480Hz, 960Hz and 1440Hz), see figure 4. Also, was noted that peaks at 1x and 2x RPM of the eccentric gear (Xg = 10Hz) and the pinion (Xp = 30 Hz) are present [23 and 24].

In Case 3, one tooth of the input pinion (32T) was pre-chipped and the eccentricity was introduced to the 48T gear. In addition to the appearance of 1x and 2x RPM (Xp = 30Hz) of the damaged pinion (32T gear), clear damage signs such as high amplitude of the first GMF and multiple sidebands around it, spaced at 1x RPM (Xp = 30Hz), were also observed in the obtained spectrum. One of the key findings from figure 5 was the high amplitude peak surrounded by some sidebands at 480Hz. This spike along with the presence of 1x and 2x RPM (Xg = 10Hz) represent clear signs of eccentric 48T gear [23 and 24]. The relatively low amplitude of the GMF of the 48T gear and its harmonics

comparing with the amplitude of the GMF of the 32T pinion can be attributed to the high friction as a result of the high load on the pinion teeth.

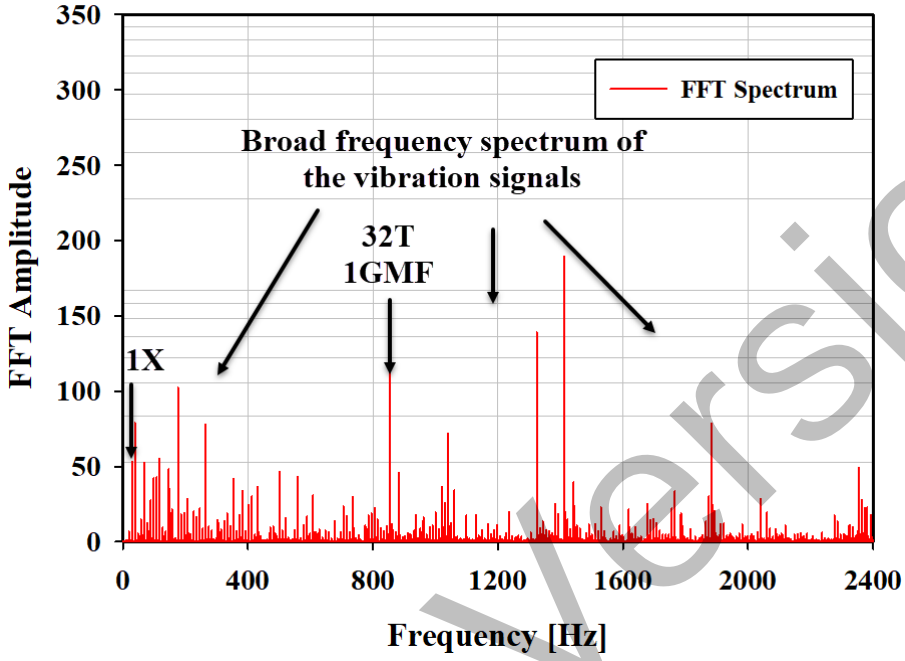

Fig. 3. Results of FFT (Case 1, 1X = 30Hz).

The last two cases (Case 4 & Case 5) are relatively different from the previous cases as the gearbox was experimented under different health conditions. In addition to the faulty gears, a set of experiments were also undertaken using different mechanical conditions for some of supporting bearings. It should be, however, telling that the same operating conditions (e.g. load & speed) and same processing were used for all tests. In Case 4, for instance, failure modes such as broken tooth, eccentricity and faulty ball element were introduced to the idler gear (48T), the output pinion (80T) and the supporting bearing, labelled Br.1, respectively. Yet, prior testing, mechanical conditions for supporting bearings, labelled Br.1 (faulty inner race), Br.3 (faulty ball element) and Br.5 (faulty outer race), were made in Case 5. It is, therefore, difficult to reliably diagnose/identify the faulty gearbox components (e.g. gears & bearings) under these conditions at the same time due to the presence of extremely different modulation rates in the collected data. Also, a number of other factors such as the selection of unreasonable maximum frequency of analysis, poor resolution of the spectrum window etc. may influence the overall analysis [21]. Hence, it must be noted that all these factors have substantial impact on the final results as they led to the fact that similar measurements were unworkably analyzed using a number of well-established CM tools that are commonly applied in online CM applications and diagnostic packages [21].

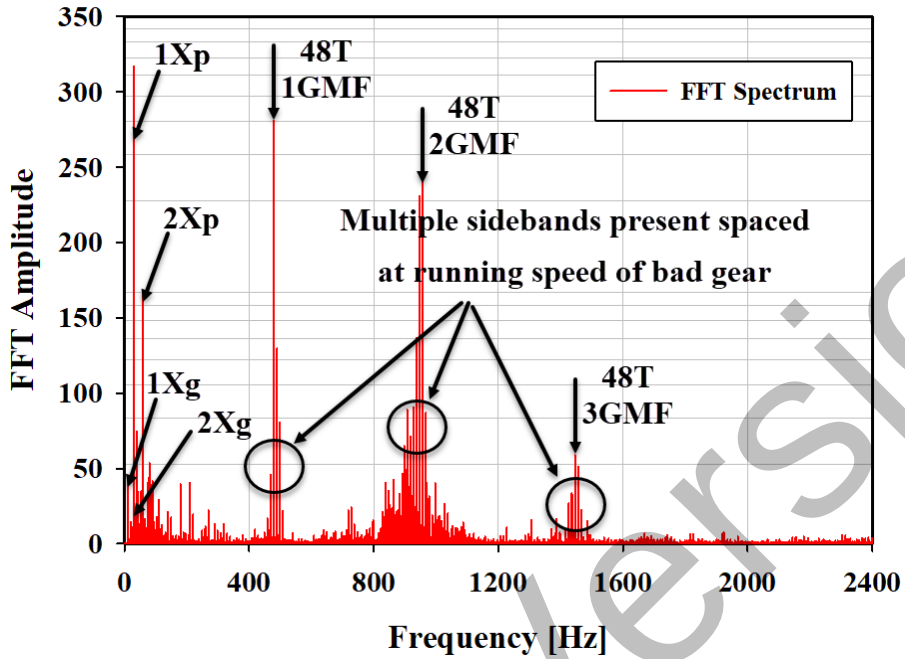

Fig. 4. Results of FFT (Case 2, 1Xp = 30Hz, 1Xg = 10Hz).

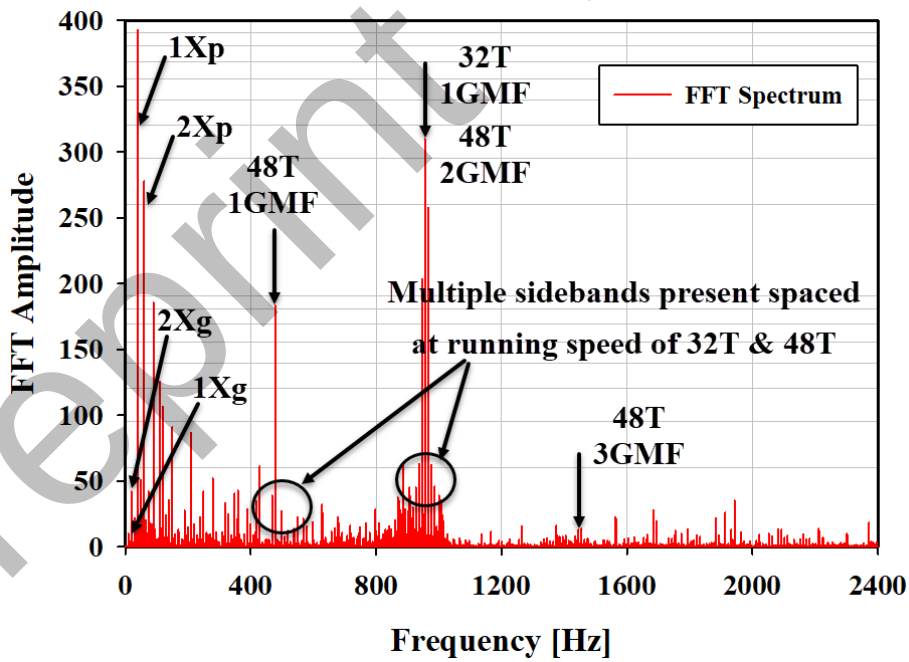

Fig. 5. Results of FFT (Case 3, 1Xp = 30Hz, 1Xg = 10Hz).

To analyze the measured data in Case 4 and Case 5, the characteristic frequencies of bearing components, presented in Table 2, were first determined using the following mathematical correlations [21]:

$$BPFO = \frac{n}{2} f_r \left(1 - \frac{d_b}{d_p}\cos\theta\right) \quad (1)$$

$$BPFI = \frac{n}{2} f_r \left(1 + \frac{d_b}{d_p}\cos\theta\right) \quad (2)$$

$$BSF = \frac{d_p}{2d_b} f_r \left(1 - \frac{d_b^2}{d_p^2}\cos\theta\right) \quad (3)$$

$$FTF = \frac{f_r}{2}\left(1 - \frac{d_b}{d_p}\cos\theta\right) \quad (4)$$

Table 2: Characteristic Frequencies of Bearing Components

Bearing Component	Forcing Frequency
Cage	FTF = 0.38*f_r
Ball	BSF = 1.99*f_r
Inner Race	BPFI = 4.99*f_r
Outer Race	BPFO = 3.05*f_r

Whereas f_r is the shaft rate (f_r = 30Hz for the input shaft, f_r = 10Hz for the idler shaft and f_r = 6Hz for the output shaft).

In addition to the analysis using the SIE method, the widespread known Spectral Kurtosis (SK), developed by Antoni [25], was also employed to process the data. Faulty machine components show significant high kurtosis levels; thus, give a rise to the impulsive responses. This will lead the acquired data to have varying statistical contents, in particular, if the machine is working under varying operating conditions. Thus, the idea behind the SK method is to locate the varying and/or non-stationary events in the frequency domain. The SK algorithm basically attempts to first calculate the kurtosis at different frequency bands and then identify the location of the maximum kurtosis. The frequency band of the maximum SK level will eventually be used to design a signal filter to extract the high-level impulsive signals. However, the method is an iterative and complicated computational process due to the difficulty in the investigation of the entire plane though several attempts were made to ease the overall calculation procedures [25]. Yet, the proper selection of the optimal size for the frequency bands has substantial impacts on the results.

For this particular investigation, figures 6 and 7 present the level of maximum

kurtosis and the carrier frequency for the data from Case 4 and Case 5 respectively. The kurtosis by nature was reported as a very noisy fault indicator. It is becoming increasingly levels when the incipient faults (faults are pronounced) produce short impulsive forces with very high excitation frequency but it drops to lower levels with low excitation frequency when these faults are physically becoming larger (well advanced faults) or even smoother (fully mature deformation) [21 and 26]. Thus, the relaxation period will be much longer than the fault repeatability period (e.g. bearings) and the adjacent impulse levels, therefore, will be overlapped and tend to produce Gaussian kurtosis values.

Observations from the resulting vibration spectrum of Case 4 and Case 5 are not any exceptional from this wide acknowledged view. Evidence of this is the very noisy resulting spectrum with very scattered and broad frequency spikes, shown in figures 8 and 9. It can obviously be noted that the SK could not resolve the noisy frequency peaks and it failed to demodulate the analysed vibration signals and, hence, it could not identify the faulty components (failure modes), though some signs of faulty bearings and gears could hardly be spotted, see figures 8 and 9.

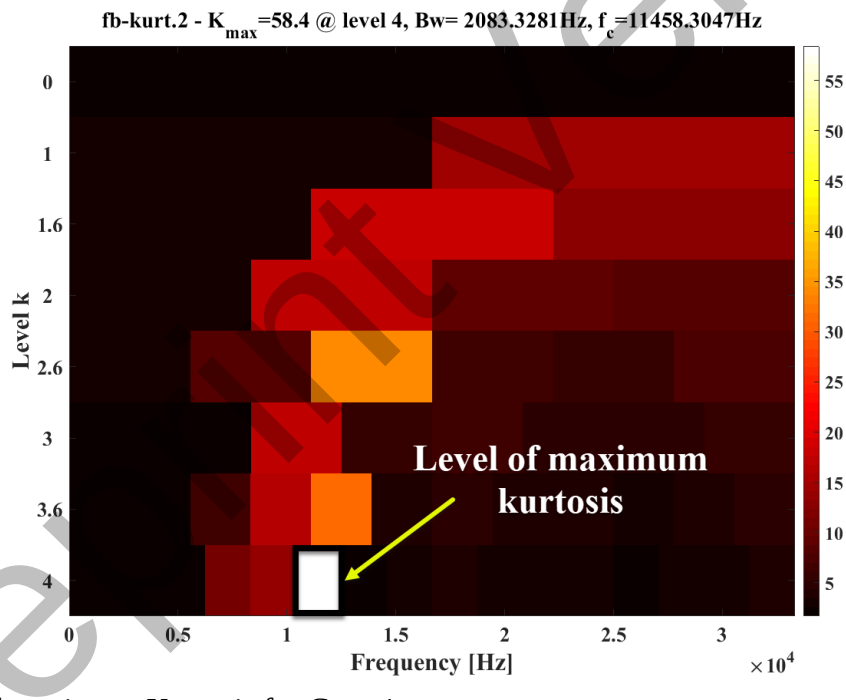

Fig. 6. Level of maximum Kurtosis for Case 4.

This can be attributed to the low excited frequencies, associated with the low kurtosis levels; these gearbox components (bearings and gears) were tested under fully mature fault conditions. Also, in addition to the contribution of high operational noise and signal interference, originating from the adjacent components, several introduced machine

faults will very likely modulate unnecessary carrier frequencies (e.g. gear mesh frequencies) that lead most of CM tools (e.g. SK) to unworkably observe the carrier frequencies of the faulty components under diagnosis [21]. Thus, it is difficult to draw any conclusions based on the observations from the SK spectrum and results of the study should be interpreted with caution.

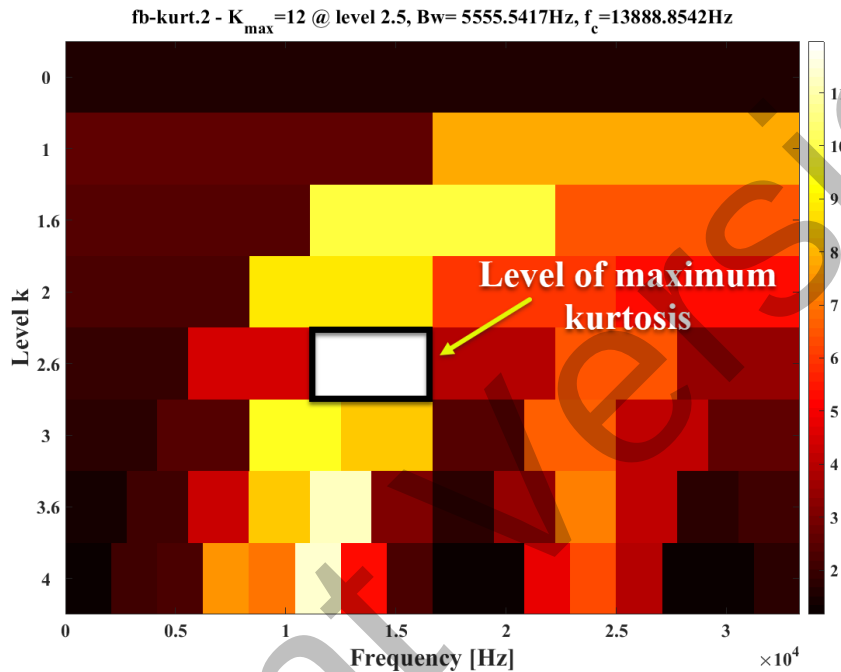

Fig. 7. Level of maximum Kurtosis for Case 5.

SIE algorithm, unlike SK, could produce very broad demodulated bands in the two cases and hence, the feasibility of SIE in handling such high extremely modulation rates in the vibration data from severely damaged machine components was here proven again. It can evidently be seen that the diagrams of SIE spectrum are dominated by cyclic frequencies of faulty bearings and gears, presented in figures 10 and 11. Symptoms of eccentric gear (48T) and/or gear installed on a bent shaft such as high amplitudes at 1GMF, 2GMF, and 3GMF can obviously be noted, see figure 10. Yet, the GMF's are circumjacent with multiple sidebands spaced at the running speed (1Xg = 10Hz) of the bad gear (48T).

The resulting SIE spectrum also shows multiple harmonics of this rotational speed (1Xg). These signs in the addition to the significant development of a peak at 240Hz, which is probably the excited resonance/natural frequency (GNF) of the 80T gear, are a clear indication of pitch diameter wear and/or broken tooth of the gear (here the 80T gear has a broken tooth). The reason behind the excitation of the GNF can be attributed to the high friction as a result of the broken tooth. Further, when the gear spectrum is analysed, it is

also useful to note that the gear is either eccentric or has a broken tooth and/or a pitch diameter wear is present, if a relatively higher amplitude peak at the 3GMF comparing to the 2GMF amplitude is noted [23 and 24]. However, this was not the case in the resulting spectrum for Case 4. The authors postulated that the high load on the gear teeth was the prime causal factor of this different level (higher 2GMF amplitude comparing to the 3GMF) in the amplitudes of the GMF's. Also, was observed that the SIE technique could locate the two frequency peaks, 60Hz (1BSF) and 120Hz (2BSF), which are corresponded to the faulty bearing ball (Br.1).

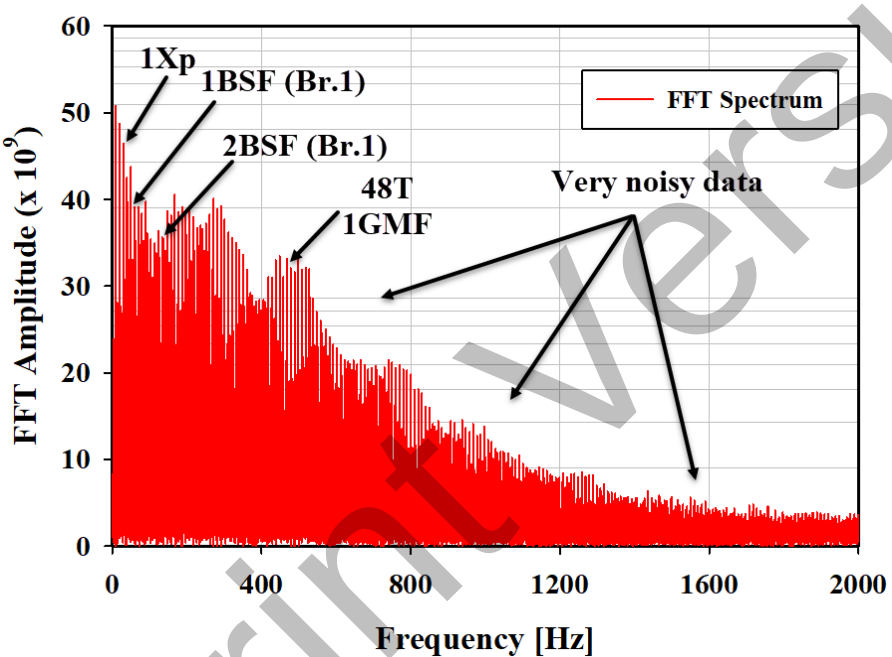

Fig. 8. Results of SK (Case 4, 1Xp = 30Hz).

As presented in Table 1, various types of mechanical conditions were introduced to the gearbox components in Case 5. The results from the analysis of this case show that some interesting amplitude modulations on the spectrum are unusual and should closely be investigated. In the spectrum of figure 11, peaks at 1GMF, 2GMF, and 3GMF of the 32T gear (input pinion) can clearly be noted. The significantly high amplitude of the 2GMF (1920Hz) of this gear comparing to the 1GMF (960Hz) and the 3GMF (2880Hz) is a clear evidence of a misaligned gear [23 and 24]. Another cause of these significant peaks at the GMF and its harmonics may be referred to the gear wear. Yet, important multiple sidebands around these GMF's are also common signs of the broken tooth and/or the presence of pitch diameter wear, see figure 11. Further, the authors believe that the high friction and high load on the teeth have excited the GNF peak of this gear, which was eventually convoluted

with the modulated 1GMF's (480Hz) of 48T and 80T gears.

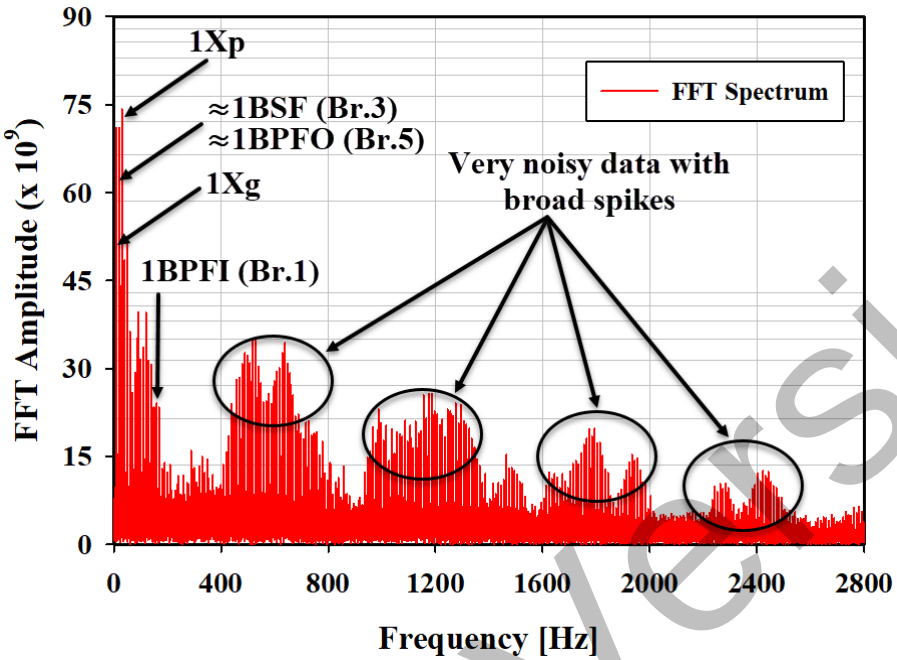

Fig. 9. Results of SK (Case 5, 1Xp = 30Hz, 1Xg = 10Hz).

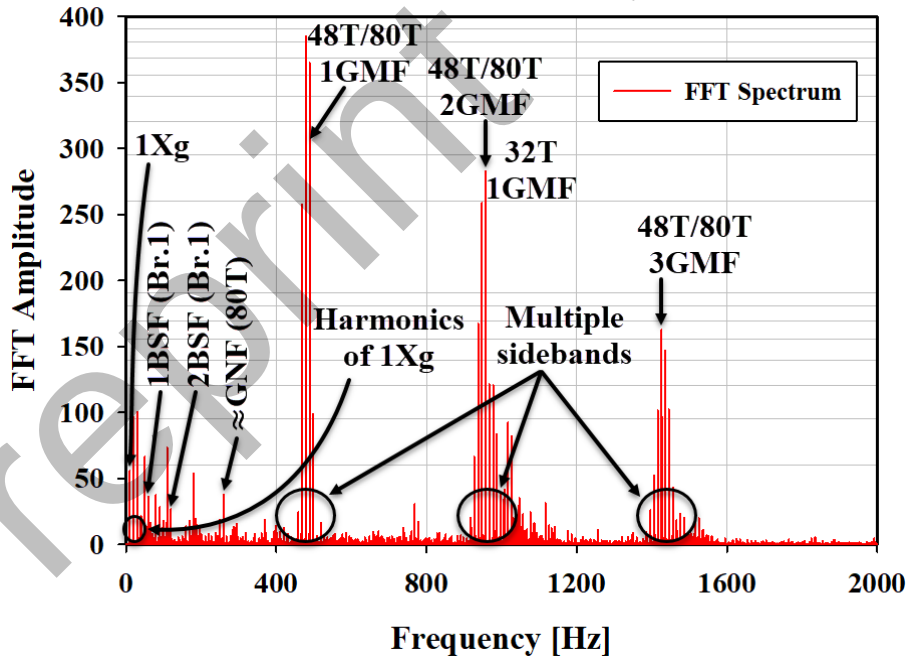

Fig. 10. Results of SIE (Case 4, 1Xp = 30Hz, 1Xg = 10Hz).

The amplitude modulation corresponded to the eccentric 48T gear is also present. Evidence of this is the appearance of significant peaks surrounded by multiple sidebands at 1GMF, 2GMF, and 3GMF of this gear. However, it must be telling that the significant high amplitude in the 1GMF (480Hz) can also be related to the broken 80T gear, whose natural frequency may also have been excited at 240Hz, see figure 11. Observations from figure 11 also display an uncommon increase in the 1GMF (480Hz) and 3GMF (1440Hz) spikes; this is another clear symptom of an eccentric gear and/or a gear installed on a bent shaft, though it also suggests that what-so-called improper gear backlash phenomenon has occurred. Yet, circumstantial evidence supports the presence of severe eccentricity was the modulation of the idler shaft speed (1Xg = 10Hz) and its harmonics [23 and 24].

As the acquired data has high modulation rates and the resulting SIE spectrum is dominated by significant GMFs' peaks, it was, thus, expected that the information of bearings' faults will be superimposed by the gear meshing signals; thus, generating what-so-called high order cyclostationary process (signals with varying properties e.g. statistical variance). Although this case exhibited the same features, the resulting SIE could have revealed additional important findings. For instance, evidence of the presence of 1BPFI for Br.1, 1BSF for Br.3, and 1BPFO for Br.5 can obviously be observed, see figure 11.

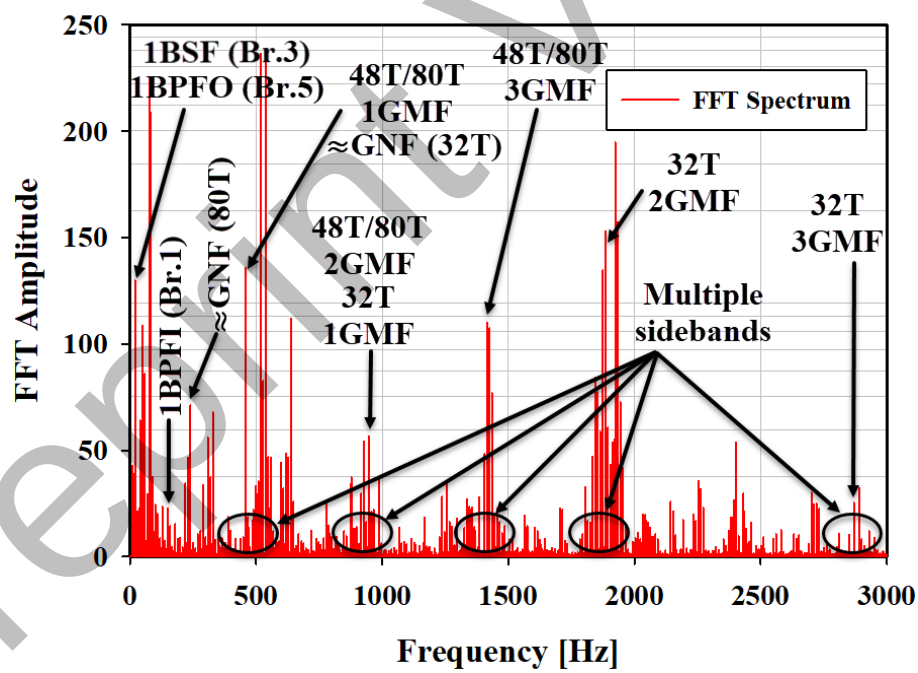

Fig. 11. Results of SIE (Case 5, 1Xp = 30Hz, 1Xg = 10Hz).

Resulting SIE from Case 5 was further processed using Continuous Wavelet Transform (CWT). The Morlet wavelets with Wavelet Parameter (WP) value of 8 were employed. As seen in figure 12, the integrated power of SIE offers good frequency localization, and these findings are in line with previous results observed in figure 11, where the presence of intense peaks, corresponded to faulty components in the gearbox, can clearly be noted.

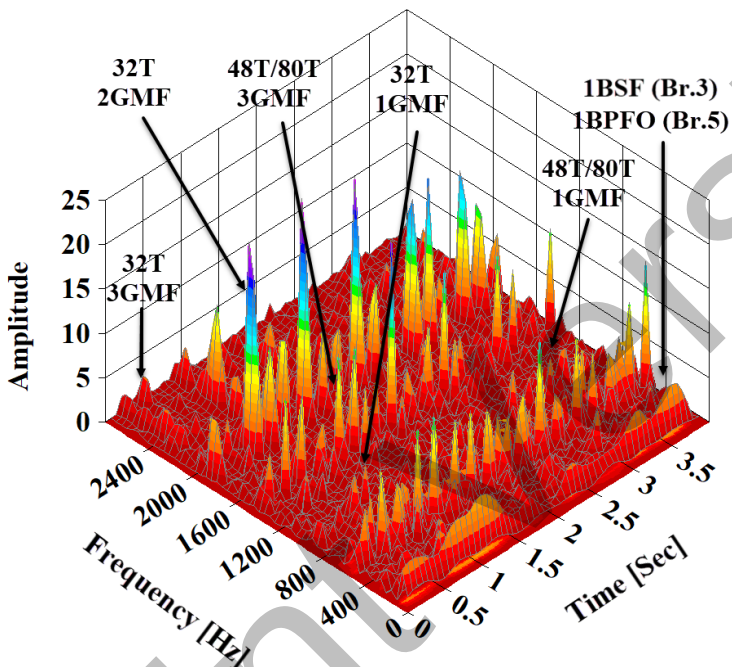

Fig. 12. Results of CWT (Case 5, 1Xp = 30Hz, 1Xg = 10Hz).

Thus far the observations have shown that the findings of this experiment using SIE are completely compatible with those reported in standard machine diagnostic references [23 and 24]. The next phase of analysis involved the use of Fourier Multi-Taper Spectral (FMS) as the counterpart of SIE in frequency domain. FMS was first developed to be an improved spectral tool for analysing what-so-called sleep staging EEG (electroencephalogram) [27, 28 and 29]. It is basically a time frequency decomposition method that is based on something called the Discrete Prolate Spheroidal Sequences (DPSS) or DPSS tapers. Selection of the shapes of these DPSS tapers is very important as it helps to remove the false power from the side lobes and concentrate it in the main lobe; thus, reducing the broadband bias caused by the side lobes. Yet, these shapes should be orthogonal to one another so that they produce estimated projections of uncorrelated data. This will simply help to average the resulting tapers together and eventually reduce the variance. To estimate the Multi-Taper spectrum, a single taper spectrum is first determined

for each piece of the tapered data and then the mean single taper spectrum is computed; sufficient details were well documented in [27, 28 and 29].

The FMS method was applied to the SIE values calculated from Case 5 and the results are presented in a stacked format with two different spectrums. The bottom plot represents the FFT average of the tapers' FFT's, while the top graph is the F-value spectrum. The F-value (F-Ratio) is the ratio between the variance between the segments to the variance within the segments. The use of F-ratio plot could ease the isolation of the central frequencies. It should be telling that the number of tapers was set at 7 and the bias width was selected to be 4. This configuration, shown in figure 13, has provided a good mean for closely monitoring of each computed frequency. Unlike the previous analysis, for instance, the second forcing frequency (2BPFI) for the faulty inner race of the bearing (Br.1) could easily be identified using the F-ratio diagram.

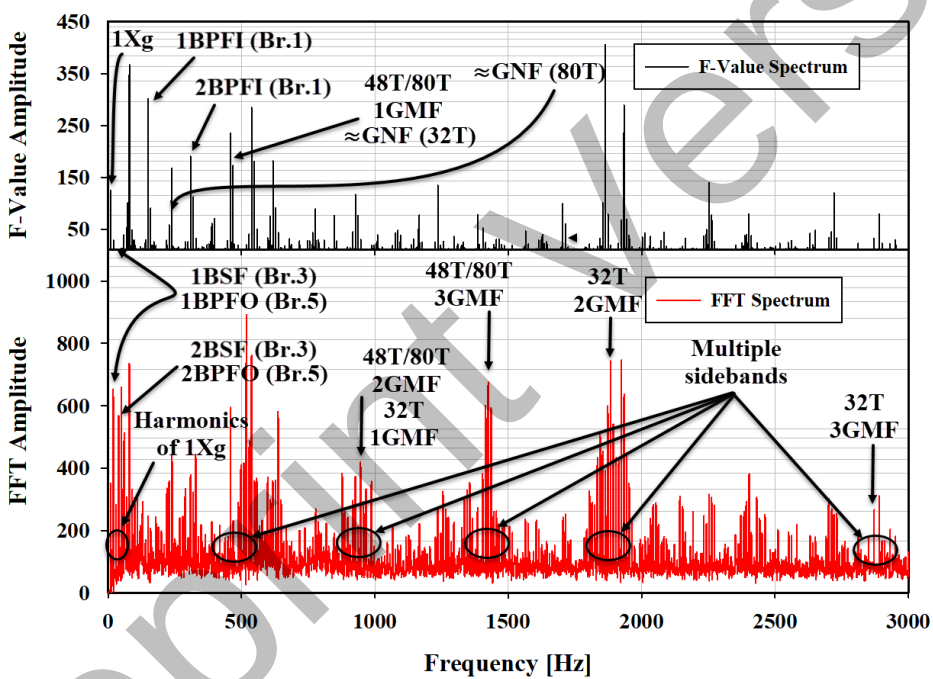

Fig. 13. Results of FMS (Case 5, 1Xp = 30Hz, 1Xg = 10Hz).

3.2. Classification of Eccentric Gears

This research also investigated the possibility of using broader approaches for the machine prognostic. By obtaining adequate extracted information from the provided data, viable Machine Learning (ML) algorithms can simply be devised for the potential of eccentric gears classification. ML is a study of computer algorithms that automatically find patterns in data to improve and predict the overall performance. Finding patterns is usually

based on previous data or through trial and error processes. Generally, ML algorithms adjust some internal values to better map the outputs to the input features. Input features are very often broken down into one or more features that describe general characteristics of the data, while outputs can be specific values or just more favourable outcomes. This process is implemented during the training phase, where ML models with training algorithms are fed with a bunch of data [30, 31 and 32].

Different types of learning employ different structures of algorithms and different learning mechanisms. Yet, some types are better suited to solving certain kinds of problems. Generally, three learning mechanisms including Supervised Learning, Unsupervised Learning and Reinforcement Learning are widely used. In supervised learning, labelled data is fed into the ML model, which eventually tries to match outputs to inputs. This can be done by instructing the model which output should get with each input and accordingly the model will attempt to achieve as many accurate predictions as possible. Applications of this learning mechanism involve the classification or regression problems. Unlike the supervised learning, unsupervised models are provided with unlabelled data to derive its own patterns and understand the overall data distribution. This means that the models are not trained to achieve certain targets but they rather instead tend to cluster the related features in the data. Examples of unsupervised learning may include Principle Component Analysis (PCA) and clustering or association problems. The third type of ML is the reinforcement learning, where the models are essentially supplied with the environment, actions, and reward system and the models and/or agents are accordingly required to obtain the highest rewards. Each action yields positive or negative reward and any model not only learns which actions to choose but also keeps a track of state, and environment that eventually lead to maximize the reward. Applications of this learning type can be found in state-action-reward algorithms such as in games and Markov decision process. A good overview of existing ML approaches, including the most commonly used methods, the analysis and theory of algorithms and computational complexity are given in [30, 31 and 32].

Supervised ML approach has been used in this research work. In this approach, several processes including the extraction of the statistical features from data and the application of unsupervised PCA (unsupervised dimensionality reduction technique) have been implemented. Eleven statistical features such as peak value, crest factor, kurtosis, entropy, etc., to name a few, were first extracted and the unsupervised PCA was then applied to summarize and visualize the most important information contained in a multivariate dataset. It should be noted that the mathematical correlations used to extract these statistical features were presented in the work by Elfojani [2]. One of the difficulties inherent in multivariate analysis is the problem of visualizing and determining the relationship between variables in a given data that has many variables. Some variables are often correlated and hence, more than one variable may measure the same driving principle that governs the behaviour of the system [30, 31 and 32]. This type of correlation is a clear

indication of the redundancy in the data, which can be omitted by replacing a group of the correlated variables with a single new variable using the PCA method. The method creates a set of Principle Components (PC's), each of which is a linear combination of the original variables and is orthogonal to the other PC's; thus, leading to no redundant information. PCA assumes that the directions with the largest variances are the most important and what-so-called the Eigenvalues are used to determine the amount of variance retained by each PC [30, 31 and 32]. The extracted features from the provided data were further analysed by the incorporation of this concept and resulting PC's were assessed in multiple stages using eigenvalues method and the analysis of statistical correlation to determine the required number of PC's.

Here prior to PCA, attempts were made to investigate the correlation between the features. From the assessment of correlation analysis, it is notable that a close correlation exists between some of the extracted features. A graphical description of the resulting correlation is presented in figure 14, where a significant correlation, for instance, between RMS and Peak, Peak and ENG, and KU and CF can obviously be noted.

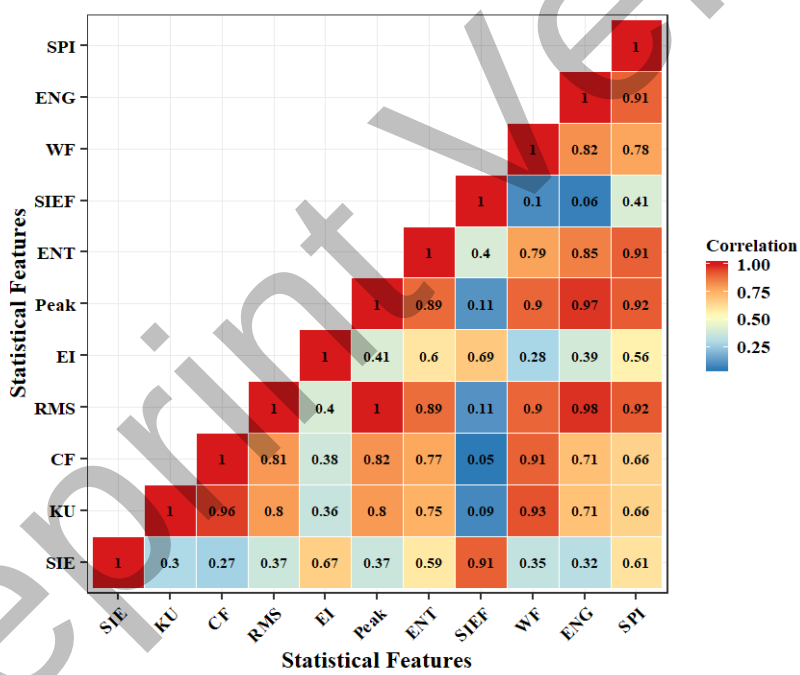

Fig. 14. Correlation Matrix.

The observed prevalence of correlation between the extracted features indicating a possible redundancy and this however may explain the need to reduce the dimensionality of the data. Following to the correlation analysis, the data was subjected to additional processing to obtain the eigenvalues as well as the proportion of variation explained by

each eigenvalue. The results are interesting and help to justify the selection of the required number of PC's. Results illustrated in Table 3 show that the first principle component (PC1) with an eigenvalue of 7.1121 and divided by 11 equals (0.6765) can provide an information of 64.65% from the total variation in the data, while PC4 and PC11 will only give 5.51% and 0.0178% respectively.

It is necessary to understand that the cumulative variance percentage is determined by adding the successive proportions of the variation. This will help the analysts to only consider the PC's that account for a certain fraction of the total variance; this is only a rule of thumb method as there is no a global accepted method used to decide how many PC's are enough to represent the data. Indeed, this will highly depend on the specific field of application and the specific dataset. In practice, only first few PC's that can explain from 70% to 90% of the variation are considered [30, 31 and 32]. Possible contribution by every individual feature to the first four PC's was also investigated. A closer examination for the visual representation of the results, presented in figure 15, reveals that the extracted features contributed significantly to the first component (PC1) comparing to the other PC's. Apparent advantage in utilizing these various analyses is that the outcome has eventually led to the conclusion that selection from 2 up to 4 PC's may be sufficient for optimal classification results. However, the evaluation of overall performance of ML classification algorithms including a comparison between results from the use of 2 up to 6 PC's, was also discussed later in the paper.

Table 3: Eigenvalues, Variance Percent, and Cumulative Variance Percent

Principle Component	Eigenvalue	Variance %	Cumulative Variance %
PC 1	7.1121	64.6551	64.6551
PC 2	1.6223	14.7482	79.4033
PC 3	1.0953	9.9572	89.3605
PC 4	0.6064	5.5125	94.8730
PC 5	0.2377	2.1614	97.0344
PC 6	0.1423	1.2935	98.3279
PC 7	0.0955	0.8685	99.1963
PC 8	0.0497	0.4519	99.6483
PC 9	0.0219	0.1995	99.8478
PC 10	0.0148	0.1344	99.9822
PC 11	0.0020	0.0178	100.0000

Unlike the high dimensionality data, data with only 2 PC's (PC1 = Fault Indicator 1 and PC2 = Fault Indicator 2) can easily be visualized. Figure 16 illustrates the distribution of healthy and faulty observations (classes) from the test gearbox. Interpretation of this

distribution is difficult as the observations show broadly similar traits. Healthy and faulty classes either adhered and/or overlapped each other. This may be attributed to identical statistical contents or the stochastic nature of the data from gearboxes, see the above machine diagnostic section. Hence, the selection of the most efficient ML classification methods represents a big challenge.

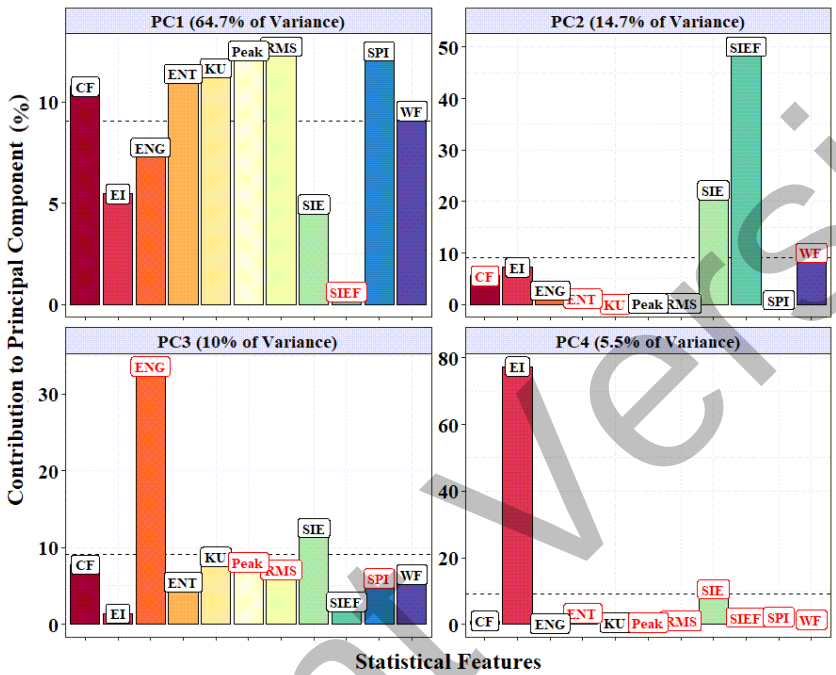

Fig. 15. Contribution of Features to PC's.

To tackle this problem, Random Forest (RF) and Support Vector Machine (SVM) classifiers have been proposed for gears prognostic analysis. RF is a supervised Ensemble of decision trees technique that is widely used for both regression and classification problems on large datasets. The most interesting aspect of this technique is the use of bagging method to construct a large number of relatively uncorrelated tress during the training process for the final prediction. The advantage of having many uncorrelated models is that the trees will protect each other from their individual errors. This structure, in turn, improves the accuracy and stability of the model by reducing the variance without increasing the bias and/or the noise. Each tree model is made up of a series of nodes and branches and splits the data at the root into subsets based on different features. At the end of the process, each individual tree will provide its prediction and classes with the most votes will represent the model prediction. It is worth pointing out that RF can also be used to identify the most significant variables in a given dataset from many input features. SVM models are based on the statistical learning theory and, more specifically, on the idea of

maximizing margins. SVM model plots each data item as a point in n-dimensional space, where each feature has the coordinate as the value. Observations for which the assigned class is correct will have positive and large margins (margins close to one), whilst small margins (margins close to zero) are a symptom of instability. Generally, SVM models attempt to find an appropriate line that splits data between two distinct groups, which is eventually used to classify new points.

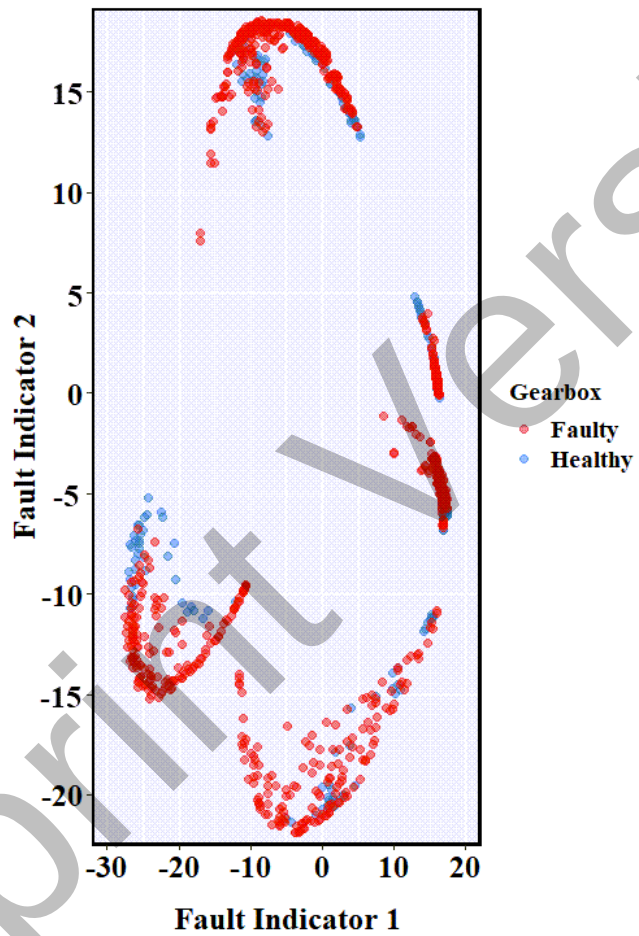

Fig. 16. Data distribution with 2 PC's (PC1 = Fault Indicator 1, PC2 = Fault Indicator 2).

A random forest model with 500 trees and SVM model with SVM-Kernel function, and cost value of 15 were trained to classify bad and good gearboxes. It is here important to make a note that the most relevant individual parameters for the proposed models were kept changing until low level of the Misclassification Rate (MCR) was achieved. For the evaluation of the impact of increasing the number of PC's in the dataset on the overall classification outcomes, it would be worth testing how well the proposed ML models

perform in different conditions. Hence, in the first test, the dataset has the all instances (observations), with two numeric attributes (PC1 denoted as Fault Indicator 1 and PC2 denoted as Fault Indicator 2) and two classes (Healthy and Faulty), while a number of PC's ranging from 3 to 6 as numeric attributes were used for test2, test3, test4 and test5 respectively. To maintain a structure proportional to that of the original dataset, 60% of the observations were randomly selected for the training set, while the total of 1200 observations constituted the test set. This division method has eased the validation of the trained models and helped to determine the error in test sets. The schematic diagram, presented in figure 17, walks the readers through various basic steps of the classification process.

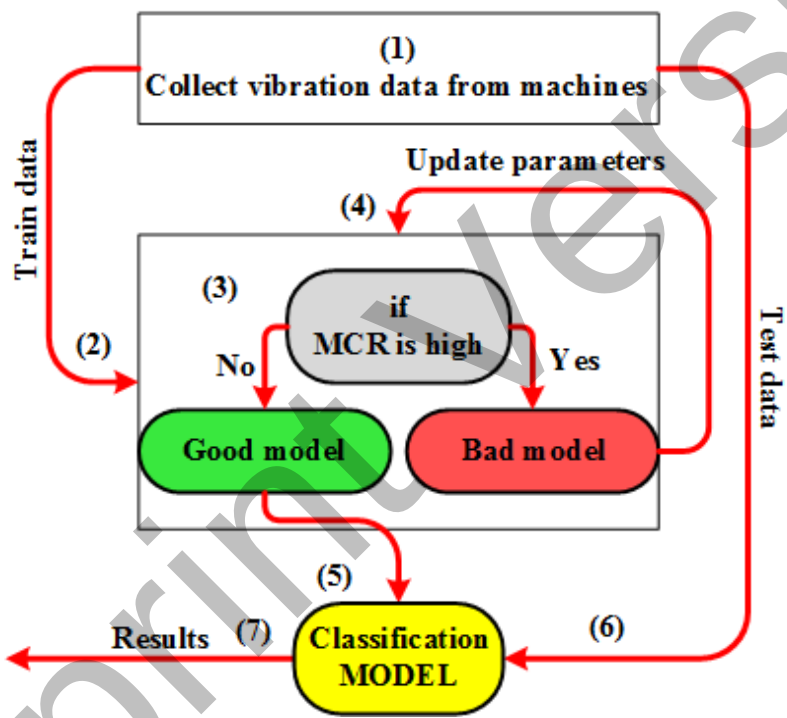

Fig. 17. Schematic Chart for Gearboxes Classification.

The results demonstrate the adequacy of the RF and SVM algorithms as both models, for instance with 2 PC's, could achieve a perfect fit to the training set and classify the test set at very high accuracy. The models provided flexibility to handle very complex issues found in the separation of the classes. It must be stressed that the visualisation of the results was only made for the dataset with two attributes as the visualisation for high dimensionality dataset is difficult to make and even impossible. The performance of the proposed techniques using two PC's is shown in figures 18 and 19. While RF interestingly tried to split the plane into healthy and faulty segments, the SVM classifier efficiently

capsulized each individual healthy observation. However, slightly better results were achieved when using a higher number of PC's. Evidence for this is the evaluation of the classification prediction results using two convenient methods known as Confusion Matrix (CM) and Receiver Operating Characteristic (ROC) curve.

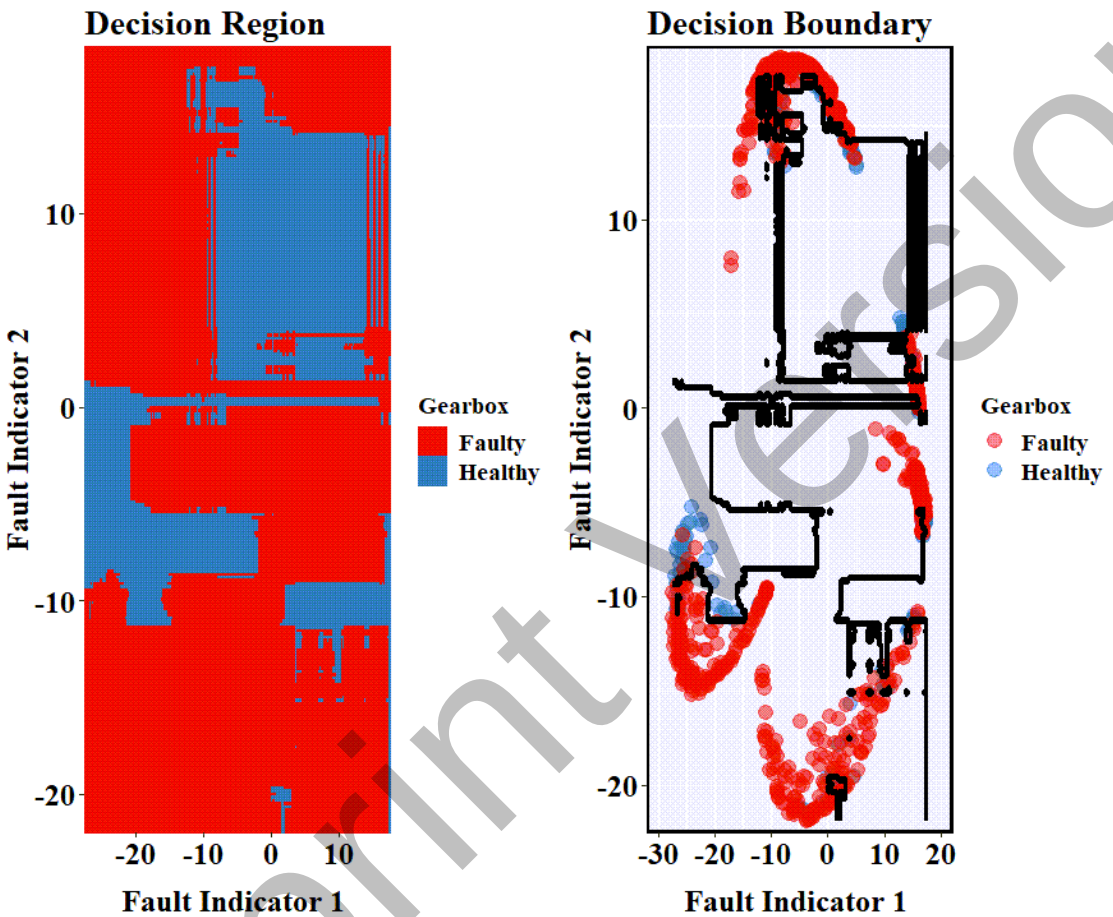

Fig. 18. Results of RF Classification Based on Gearboxes Status (PC = 2).

The confusion matrix is a handy table that makes a presentation of the accuracy of a model with two or more classes. The cells of the table are the number of predictions made by ML algorithms. The predictions are presented on the x-direction and the accuracy outcomes are vertically listed (y-direction). The diagonal contains the number of samples assigned to the right classes, while all the other elements represent the misclassifications. The ROC curve is another valuable tool used to compare different classifiers that can assign a score to their predictions. In general, this score can be interpreted as a probability, so it is bounded between 0 and 1.

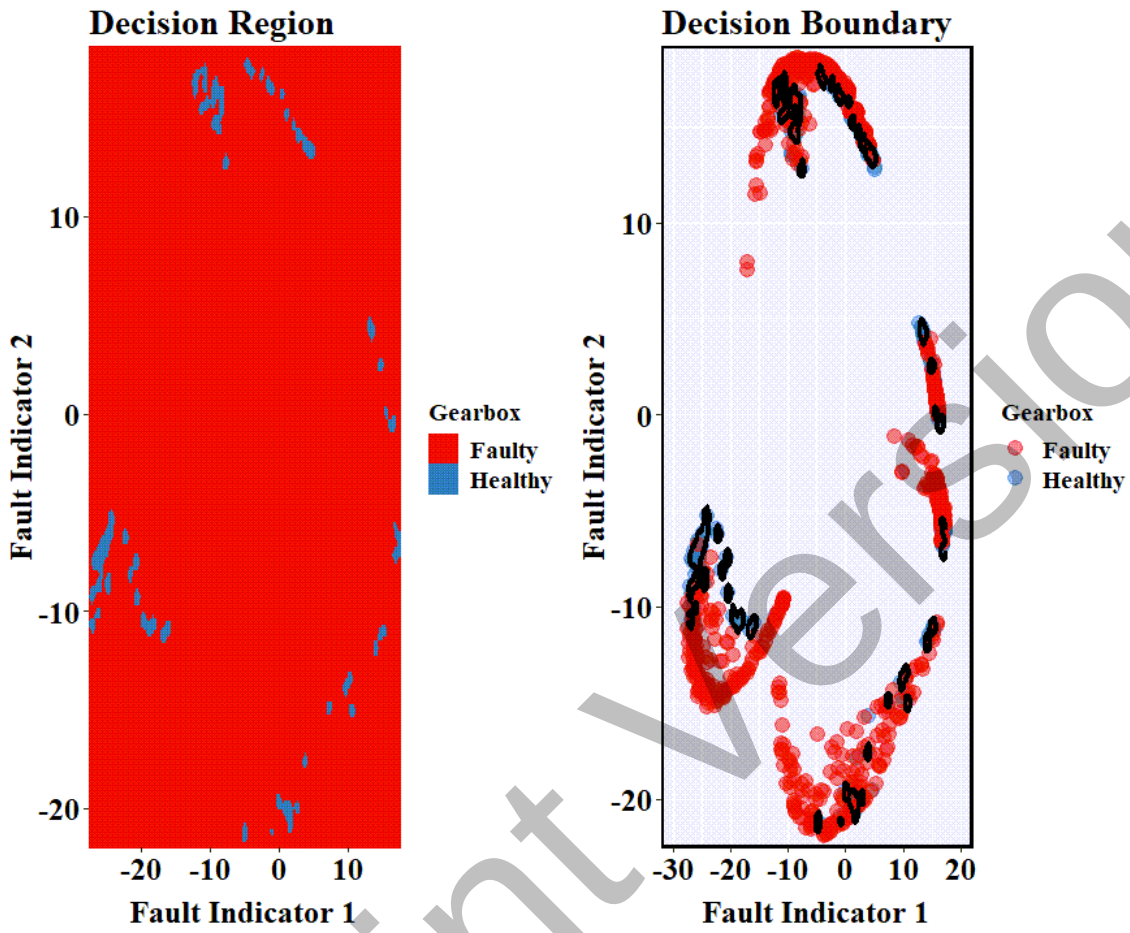

Fig. 19. Results of SVM Classification Based on Gearboxes Status (PC = 2).

The ROC is essentially a graphical representation of contrast between True Positive Rate (TPR) also known as Sensitivity (y-axis) and the increasing False Positive Rate (1 - FPR) also known as (1 – Specificity) (x-axis) at various thresholds. The solid oblique line in the graph represents a perfectly random classifier, so all the curves above this threshold (closer to left top corner) indicate more accurate performance. This means the ideal classifier will show a ROC curve with FPR = 0 and TPR = 1; the perfect curve is split into the segments (0, 0) – (0, 1) and (0, 1) – (1, 1). Area under ROC Curve (or AUC for short) is also another measure of the model accuracy. The AUC represents a model's ability to discriminate between positive and negative classes. AUC values are bounded between 0 (worst performance) and 1.0 (best performance) and hence, an area of 1.0 represents a model that made all predictions perfectly, while a value of 0.5 (which is the area of one of the two right triangles) corresponds to a model that is as good as random.

The resulting ROC curve is the following plot, presented in figure 20, and as confirmed by the AUC, ROC curve for different models with different number of PC's shows almost perfect performance. The results are interesting and help to justify the fact that good agreement was found when comparing results from ROC curve against published data such as the accuracy and the MCR, shown in Table 4. There is an important note to bear in mind that the results in Table 4 do not constitute a thorough evaluation for ML models with high number of PC's. With performing direct comparisons between the ML model with 4 PC's and the ML models with high number of PC's (5 PC's and 6 PC's), it can be considered that there is a little and insignificant improvement to distinguish between the overall performance, see figure 20.

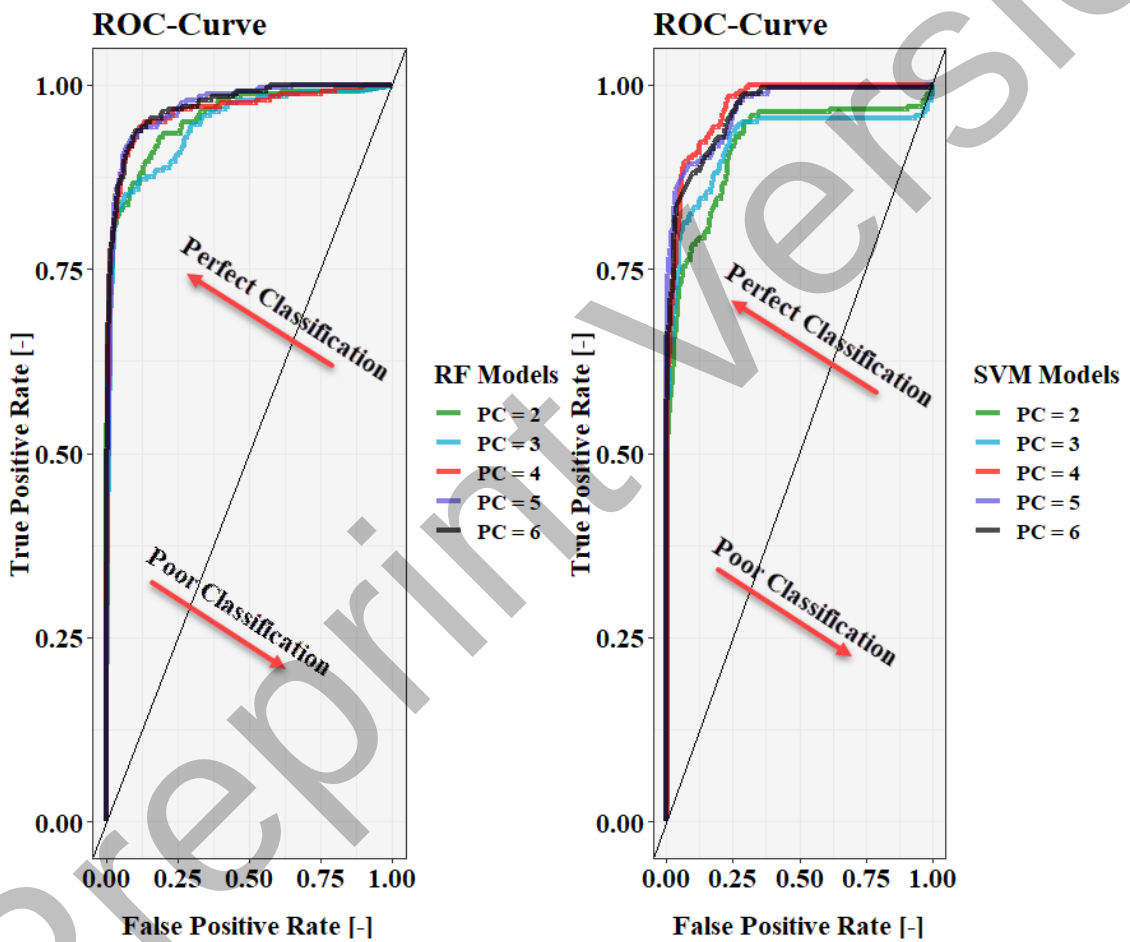

Fig. 20. ROC Curve for the Proposed ML Models.

Table 4: Evaluation of the Proposed ML Classification Models

	No. of Principle Components																	
	PC = 2						PC = 3						PC = 4					
Classification Models	RF			SVM			RF			SVM			RF			SVM		
Confusion Matrix		F	H		F	H		F	H		F	H		F	H		F	H
	F	921	39	F	934	26	F	912	48	F	938	22	F	927	33	F	935	25
	H	42	198	H	83	157	H	40	200	H	79	161	H	43	197	H	65	175
True Classification	1119			1091			1112			1099			1124			1110		
False Classification	81			109			88			101			76			90		
Accuracy	93.25%			90.92%			92.66%			91.58%			93.66%			92.50%		
MCR	06.75%			09.08%			07.34%			08.42%			06.34%			07.50%		
AUC	95.46%			91.53%			94.32%			92.08%			96.29%			96.78%		

These observations also agree and are in keeping with the results reported in the above principle component analysis, where it was shown that ML models with 4 PC's can explain around 95% of the total variation (acceptably large variation percentage or sufficient information) in the health integrity of the test gearboxes. Generally, the overall performance is rather optimal because all AUC values are close to 1.0 (FPR << TPR) and yet, all models recorded a high classification accuracy above 90% with low overall MCR values (MCR < 10%), see Table 4.

4. Conclusion

As per the earlier published work in the literature, this study is the first of its kind as it presented the first known attempt at the experimental evaluation of the feasibility of CM for eccentric industrial gears. Acquired vibration data from a representative gearbox with eccentric gears and different mechanical conditions, introduced prior testing, was successfully analysed. It was shown that the characterizing of the abnormality of faulty machine components using SK technique is very difficult and even impossible due to excessive eccentricity and high modulation rates in the analysed data. However, this work provided another proof that the SIE approach can outperform other state-of-the-art CM algorithms such as SK in the most difficult cases. The results demonstrate the potential superiority of SIE algorithm for the fault detection in very high modulated measurements. These results also agree well with the existing CM standard references for machine fault monitoring. Yet, another conclusion was reached is that applying FMS to the SIE values could offer a better overall picture of time-frequency space than the use of classical FFT.

It was found that the use of PCA method will produce good quality results in the dimensionality reduction and data visualization. Despite the fact that even slightly better classification results are achieved when using a data with higher number of PC's, results showed that the selection of a limited number of PC's (e.g. 4 PC's), by which between 80% to 95% of the total variation in the machine data can be explained, will perform with sufficient reliability when used in real time machine classifications. Though it is difficult to reliably classify the data from gearboxes due to some practical difficulties such as the likelihood of the statistical contents, RF and SVM models could demonstrate a high consistency with the predefined different gearbox status.

Finally, this study can be considered as the first investigative step and it is, thus, more similar experimental investigations must be undertaken as they will ease not only a further improving upon the available CM tools but also developing new approaches that may eventually lead to adopt a general evaluation method.

5. References

1. M. Elforjani, D. Mba, A. Muhammad, and A. Sire, "Condition Monitoring of Worm Gears", *Applied Acoustics*, Vol. 73, DOI https://doi.org/10.1016/j.apacoust.2012.03.008, no. 8, pp. 859-863 2012.
2. M. Elforjani, "Diagnosis and Prognosis of Real World Wind Turbine Gears", *Renewable Energy*, Vol. 147, DOI https://doi.org/10.1016/j.renene.2019.09.109, no. 1, pp. 1676-1693, 2020.
3. J. C. J. Correa, and A. L. Guzman, *Mechanical Vibrations and Condition Monitoring*, 1st Edition, Academic Press, Mar. 2020.
4. A. S. Kale, "Bending Fatigue Failure in Gear Tooth", *International Journal of Engineering Research & Technology (IJERT)*, Vol. 2, no. 6, 2013.
5. D. Jelaska, *Gears and Gear Drives*, United Kingdom: John Wiley & Sons Ltd, Sept. 2012.
6. X. He, X. Zhou, Z. Xue, Y. Hou, Q. Liu and R. Wang, "Effects of Gear Eccentricity on Time-Varying Mesh Stiffness and Dynamic Behavior of a Two-Stage Gear System", *Journal of Mechanical Science and Technology*, Vol. 33, DOI https://doi.org/10.1007/s12206-019-0203-7, no. 3, pp 1019–1032, 2019.
7. Y. C. Chen, "Time-Varying Dynamic Analysis of a Helical-Geared Rotor-Bearing System with Three-Dimensional Motion Due to Shaft Deformation", *Applied Sciences*, Vol. 10, DOI https://doi.org/10.3390/app10041542, no. 4, pp 1542, 2020.
8. Z. Cao, Y. Shao, M. Rao, and W. Yu, "Effects of The Gear Eccentricities on the Dynamic Performance of a Planetary Gear Set", *Nonlinear Dynamics*, Vol., 91, DOI https://doi.org/10.1007/s11071-017-3738-0, no. 1, pp 1–15, 2018.
9. G. Cheon, and R. G. Parker, "Influence of Manufacturing Errors on the Dynamic Characteristics of Planetary Gear Systems", *KSME International Journal*, Vol. 18, DOI https://doi.org/10.1007/BF02983645, no. 4, pp. 606–621, 2004.
10. W. Kim, H. H. Yoo, and J. Chung, "Dynamic Analysis for a Pair of Spur Gears with Translational Motion Due to Bearing Deformation", Journal of Sound and Vibration, Vol. 329, DOI https://doi.org/10.1016/j.jsv.2010.04.026, no. 21, pp. 4409-4421, 2010.
11. W. Yu, C. K. Mechefske, and M. Timusk, The Dynamic Coupling Behaviour of a Cylindrical Geared Rotor System Subjected to Gear Eccentricities, *Mechanism and Machine Theory*, Vol. 107, DOI https://doi.org/10.1016/j.mechmachtheory.2016.09.017, pp. 105-122, 2017.

12. Z. Chen, and Y. Shao, "Dynamic Features of Planetary Gear Train with Tooth Errors", *Proceedings of The Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, Vol. 229, DOI https://doi.org/10.1177/0954406214549503, no. 10, pp. 1769–1781, 2015.
13. M. Inalpolat, and A. Kahraman "A Dynamic Model to Predict Modulation Sidebands of a Planetary Gear Set Having Manufacturing Errors", *Journal of Sound and Vibration*, Vol. 329, DOI https://doi.org/10.1016/j.jsv.2009.09.022, no. 4, 2010.
14. Y. Yi, K. Huang, Y. Xiong, and M. Sang, "Nonlinear Dynamic Modelling and Analysis for a Spur Gear System with Time-Varying Pressure Angle and Gear Backlash", *Mechanical Systems and Signal Processing*, Vol. 132, DOI https://doi.org/10.1016/j.ymssp.2019.06.013, pp. 18-34, 2019.
15. Y. Zhou, and M. Hao, "The Study of Leakage of Circular Arc–Involute–Circular Arc Gear Pump", *Advances in Mechanical Engineering*, Vol. 9, DOI https://doi.org/10.1177/1687814017720082, no. 9, 2017.
16. F. Suxiang, H. Shulim, "Design and Application of the Eccentric Modified Gears", *Atlantis Press*, Vol. 12, DOI https://doi.org/10.2991/mems.2012.153, pp. 585-589, 2012.
17. PHM, *2009 Data Challenge*. [Online]. Available: http://www.phmsociety.org/competition/09.
18. P. Boškoski, and A. Urevc, "Bearing Fault Detection with Application to PHM Data Challenge", *International Journal of Prognostics and Health Management*, Vol. 2, no. 1 2011.
19. H. Al-Atat, D. Siegel, and J. Lee, "A Systematic Methodology for Gearbox Health Assessment and Fault Classification", *International Journal of Prognostics and Health Management*, Vol. 2, no. 1, 2011.
20. M. Elforjani, S. Shanbr, and E. Bechhoefer, "Detection of Faulty High Speed Wind Turbine Bearing Using Signal Intensity Estimator Technique", *Wind Energy*, Vol. 21, DOI: https://doi.org/10.1002/we.2144, no. 1, pp. 53-69, 2018.
21. M. Elforjani, and E. Bechhoefer, "Analysis of Extremely Modulated Faulty Wind Turbine Data Using Spectral Kurtosis and Signal Intensity Estimator", Renewable Energy, Vol. 127, DOI: https://doi.org/10.1016/j.renene.2018.04.014, pp. 258-268, 2018.
22. M. Elforjani, D. Mba (2021). The Use of Signal Intensity Estimator for Monitoring Real World Non-Stationary Data, Book title: Applied Condition Monitoring, Springer Book Series: https://www.springer.com/series/13418
23. MOBIUS INSTITUTE, Vibration Analysis Definitions. [Online]. Available: https://www.mobiusinstitute.com/site2/item.asp?LinkID=2001
24. M. Nakhaeinejad, and D. O. Bukowitz, Practical Vibration Analysis of Machinery: Case Studies, *CreateSpace Independent Publishing Platform*, 2011.

25. J. Antoni, “The Spectral Kurtosis: A Useful Tool for Characterizing Non-Stationary Signals”, *Mechanical Systems and Signal Processing*, Vol. 20, DOI: https://doi.org/10.1016/j.ymssp.2004.09.001, no. 2, pp. 282-307, 2006.
26. M. Elforjani, “Estimation of Remaining Useful Life of Slow Speed Bearings Using Acoustic Emission Signals”, *Journal of Nondestructive Evaluation*, Vol. 35, DOI: https://doi.org/10.1007/s10921-016-0378-0, no. 4, 2016.
27. B. Babadi and E. N. Brown, "A Review of Multitaper Spectral Analysis", *IEEE Transactions on Biomedical Engineering*, Vol. 61, DOI: https://doi.org/10.1109/TBME.2014.2311996, no. 5, pp. 1555-1564, 2014.
28. P. Das, and B. Babadi, “Multitaper Spectral Analysis of Neuronal Spiking Activity Driven by Latent Stationary Processes”, *Signal Processing*, Vol. 170, DOI: https://doi.org/10.1016/j.sigpro.2019.107429, 2020.
29. J. M. Lees, and J. Park, “Multiple-taper spectral analysis: A stand-alone C-subroutine”, *Computers & Geosciences*, Vol. 21, DOI: https://doi.org/10.1016/0098-3004(94)00067-5, no. 2, pp. 199-236, 1995.
30. M. Mohri, A. Rostamizadeh, A. Talwalkar, and F. Bach, *Foundations of Machine Learning (Adaptive Computation and Machine Learning Series)*, The MIT Press, Aug. 2012.
31. G. Bonaccorso, *Machine Learning Algorithms: A Reference Guide to Popular Algorithms for Data Science and Machine Learning*, Packt Publishing, Jul. 2017.
32. J. Watt, R. Borhani, and A. K. Katsaggelos, *Machine Learning Refined, Foundations, Algorithms, and Applications*, 1st Edition, Cambridge University Press, Nov. 2016.