Crack detection in a rotating shaft using artificial neural networks and PSD characterisation

A. A. Mohammed, R. D. Neilson, W. F. Deans. P. MacConnell
School of Engineering, University of Aberdeen, Kings College, Aberdeen, U.K.

Abstract

The diagnosis of cracks in rotating shafts using non-destructive techniques provides a route for avoiding catastrophic failure of these common components. This study measured the dynamic response of a full-scale rotating shaft with three different crack depths. A novel non-destructive system is developed and described. The system uses vertical vibration of the system measured over time and characterises its behaviour using elements of the power spectral density (PSD) gained from a fast Fourier transform of the time-history. The PSDs were used as an input into an artificial neural network (ANN) to detect the presence of cracks using changes in the spectral content of the vibration of the system. A novel method for reducing the amount of data input into the ANN is described. The Peak Position Component Method (PPCM) reduces data transfer by using statistical characterisation of the position of the peaks in the PSD. The peak positions represent a small fraction of the information contained in the total frequency range. The number of the PSD peaks used as input to the neural net is a small fraction of the total frequency range. The ANN was a supervised feed-forward network with Levenberg-Marquardt back-propagation algorithm acting on the PPCM results. The frequency spectrum for the three different crack lengths examined showed clear shifts in the peak positions of the PSD and the results clearly demonstrate the feasibility of using the new system to detect cracks in-service.

1. Introduction

Maintenance and inspection play an important role in the performance of mechanical systems and the selection of the right type of maintenance and inspection strategy extends operating life, improves availability and retains the system in its proper condition [1]. Strategies for mechanical equipment based on machine condition monitoring can provide significant economic and safety advantages [2]. The performance of rotating shafts is of particular interest to this study. Such shafts can be subjected to repeated bending and may develop cracks which can grow by fatigue causing catastrophic component failure. Identifying the development of a crack early is essential for safe operation. There have been many analytical, numerical and experimental studies on cracked shafts [3, 4, 5], however, there is little evidence that the behaviour of industrial scaled rotating shafts containing cracks has been examined experimentally.

In condition monitoring, relevant information is acquired from various sensors and analyzed to judge the condition of the machine's components. In rotating equipment, vibration and sound signals are directly related to the structural dynamics of the machine and contain abundant information about the condition of individual components. Extracting reliable features from sound and vibration signals is a common method for machine condition monitoring [6, 7, 8]. Vibration analysis is a particularly useful condition monitoring technique for fixed-plant rotating equipment due to the relatively fast data collection and interpretation compared to other available off-line techniques. Vibration data is collected as digitally sampled time domain signals and the development of transforms, such as the fast Fourier transform (FFT), have allowed the conversion of the time domain data into frequency spectra [9]. The use of digital signal processing (DSP) has allowed the implementation of filters and signal enhancing calculations to be performed on the vibration data for improved noise reduction and signature detection. This technology has enabled vibration analysis to be used for monitoring road vehicles, which inherently have a high noise component in the raw vibration data. Many of the DSP algorithms have been included in the data acquisition units, which feature time to frequency domain conversion using FFT, demodulated spectra acquisition, as well as coupling with a tachometer to allow the analysis of variable speed machinery. This approach contrasts to oil and wear debris analysis techniques, which often rely on extensive chemical analysis[10], and data interpretation by experienced/trained analysts, which take time and can be expensive.

Sinou and Lees [11] investigated the influence of transverse cracks. Changes to the shaft frequencies, as well as the harmonic component of the dynamical system response and the evolution of the orbits are the principal effects due to the presence of a crack in a rotating shaft. Darpe et al. [12] experimentally investigated the effect of bow on the nonlinear nature of the crack response. Patel and Darpe [13] numerically and experimentally investigated the rotor whirl characteristics of unbalance, crack and rotor-stator rub faults with a classical Jeffcott rotor model. Differences in the whirl nature of the lateral vibration response of these faults are proposed for fault detection when these faults exist alone as well as when in combination. Antoni and Randall [14] demonstrated how the spectral kurtosis can be used efficiently in the vibration-based condition monitoring of rotating machines to detect incipient faults even in the presence of strong masking noise while Sinou and Lees [11]

obtained some distinguishing features of a cracked rotor by examining the evolution of the first critical speed, associated amplitudes at the critical speed and half of the critical speed, and the resulting orbits during transient operation. Various types of nonlinear resonances of a Jeffcott rotor with a crack have been investigated both numerically and experimentally by Ishida and Inoue [15].

Despite the use of computers for manipulation of vibration data, the interpretation of the vibration spectra and diagnosis of machine faults has generally remained the job of highly trained experts. The difficulty of building a knowledge base from human experts and implementing the expert system for a broad range of possible faults is a common drawback of artificially intelligent systems [16]. Although artificially intelligent systems have been developed for vibration analysis, they have been developed for a particular machine component, including rolling element bearings [17] and transformers [18,19]. Generally, a large number of features are preferred for effective condition monitoring [20]. However, this affects the training cost and time, as well as the classification accuracy due to the presence of irrelevant or redundant features [20, 21]. Moreover, the sensitivity of the features may vary considerably under different operating conditions [21]. Therefore, it is beneficial but challenging to extract the effective features again from the original data for effective machine condition monitoring. The advantage of utilising artificial intelligence systems is in resolving the complex relations between the system inputs and the outputs without the need for a model of the system or an assumed relationship.

This study developed a method for crack diagnosis for a rotating shaft system comprising an electric motor, a load and a long shaft with different pre-crack depths at the mid-span using vibration measurement. The method includes data reduction of the vibration spectra to extract the most important features prior to the use of ANNs for the actual diagnosis, speeding up ANN training and increasing the robustness of the analysis.

2. Experimental Methodology

2.1. Crack Rotor Test Rig

The test set up comprised a 50HP/37kW three phase electric motor connected to a 3 phase generator by means of a shaft 1450 mm long and 38 mm in diameter. The shaft is connected to the motor and generator using Fenner® TaperLock™ coupling as shown in Figure 1. These couplings rigidly couple the motor and the generator to the power transmission shaft. A variable resistive load was attached to the generator output. Three piezoelectric accelerometers (DJB type A20-J) were attached using stud mountings, one on top of the bearing housing of the generator nearest the transmission shaft (called Ch1), one on the motor shaft bearing housing nearest the transmission shaft (called Ch2) and the third in the centre of the guard covering the transmission shaft (called Ch3) as shown in Figure 2. A fourth accelerometer was connected to the charge amplifier and used as a control, to identify any faults with any of the other accelerometers or their wiring. The accelerometer positions were chosen to replicate the types of positions, which would be used in an industrial environment in which direct access to the shaft might not be possible. In such cases, bearing housing are typically instrumented. In the rig the bearings of the motor and generator are the main lateral support for the shaft. For this investigation the shaft cover was also selected to assess how measurement at a less well coupled position could be taken and still provide useful information on the shaft condition. The aim was to demonstrate a system which could diagnose a crack in a shaft from signals not obtained directly from the shaft itself.

The measured vibration signals from the accelerometers were fed to charge amplifiers. The voltage signals from the charge amplifiers were connected to a 12 bit analogue to digital converter comprising a dSPACE DS1102 DSP Controller Board, based on a Texas Instruments TMS320C31 floating-point processor. The software also provides data acquisition instruments to capture data from the model running on the real-time hardware. The DSP system is supported by Matlab-based software. An overview of the data acquisition system is shown in Figure 3. The sampled time domain signals were converted to the frequency domain using the PSD.
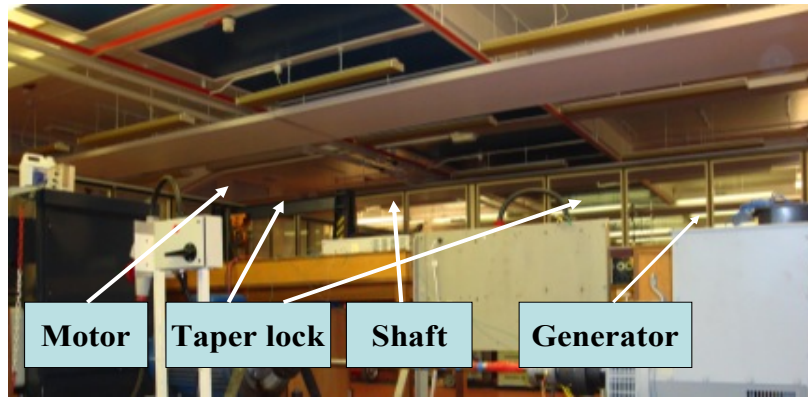
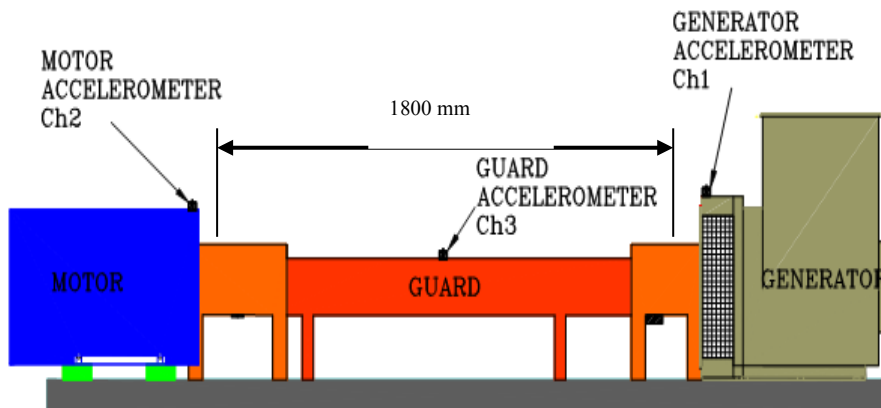Figure 1 Photo of the long shaft rig without shaft cover
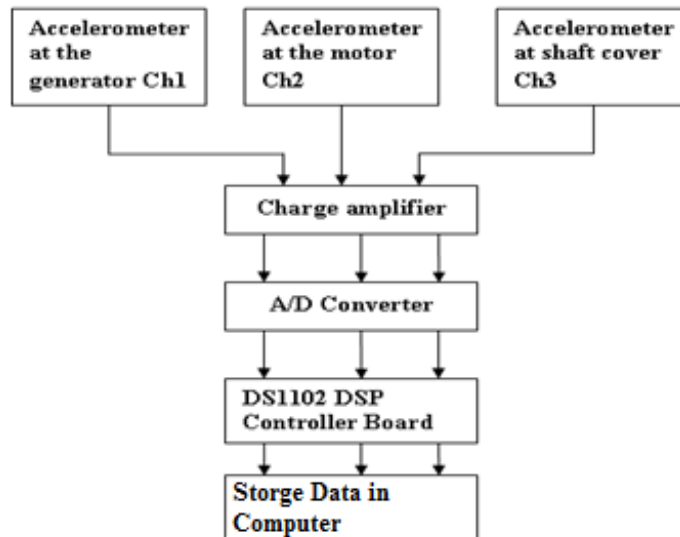

Figure 2 Vibration Sensor Locations

Figure 3Diagram for data acquisition system for vibration

2.2. Shaft defects

A series of shaft samples were tested with the experiments being carried out for increasing depth of cut. As the maximum possible loading of the system would not have caused a crack in an undamaged shaft, a crack was seeded in the shaft and the depth of the cut was increased incrementally to simulate the growth of the crack

as presented in Table 1. The experiments were carried out on two shafts; the first shaft without any defects was used as a reference. The second shaft had a 90° slot (Figure 4), made using a 3.2 mm wide milling cutter at 90°, midway along the shaft to simulate a crack.

Testing was begun with the reference shaft. This reference was used for all shafts and was used to identify any changes to the system from any cause. The reference shaft was mounted in the test rig after every test and compared with previous data to ensure consistency of the data. The three shafts with cuts were then tested. The testing procedure for every case was repeated three times. After each time reference shaft is mounted in the test rig.

Table 1 Cut increase depth

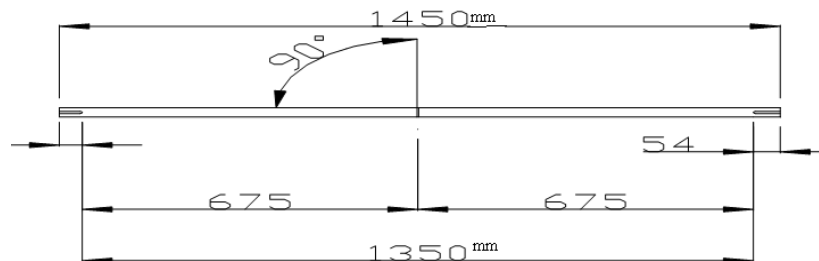| Slot depth (% of Shaft diameter)	Slot Depth (mm)
0	0
40	15.2
50	19
60	22.8

Figure 4 The 90° slot details on shaft

2.3 Experimental procedures

The aim of the work was to monitor the vibration on a loaded motor system with a long shaft and identify the changes to the vibration caused by the cut depth increasing. To ensure that the vibration data was taken during once the system had reached a stable state and steady state running, a prolonged test was carried out where the temperature was measured on the motor bearing and generator bearing. The results from the temperature test indicated that the test rig should be run for more than 4 hours for every defect condition to ensure a stable system. This ensured that all measurements were taken at the same running conditions and any changes in the vibration spectra were due to the fault condition and not to changes resulting starting from cold after a shutdown. The following test procedure was designed around this requirement.

The test consisted of three stages
1) The reference shaft, which was used to check the system behaviour had not been changed by any other causes, was mounted in the test rig and the frequency responses were measured. The shaft was run at 1550 rpm with a load of 20 KW for 4 hours until a steady state was achieved and then vibration measurements were taken at 30 minutes intervals. This data was used to assess the differences in the system behaviour between runs.
2) The cracked shaft was tested with different crack depths ranging between 0% and 60% of the shaft diameter as shown in Table 1. Between each stage of cut depth the reference shaft was mounted again and the vibration response evaluated to check the system behaviour had not been changed by other causes.
3) Step two was repeated three times to confirm the vibration response data was consistent.

In terms of processing the vibration data, the PSD was used.

3. Theoretical Development

3.1. Power Spectrum Definition and Estimation

One of the most commonly used approaches to investigating signal characteristics is the power spectrum. For a time signal, denoted by *x(n)* the Fourier transform is given by *X(k)*. Fourier analysis is perhaps the most useful and well used tool for vibration analysis today. Its ability to efficiently decompose a signal to constituent harmonic components has proved powerful in gaining a greater understanding of machine operation and many results have been developed to identify faults in terms of Fourier components [22, 23]. The second order spectra known as the power spectral density, P(f), is defined by the following equation:

$$P(f) = E[X(f)X^*(f)] \qquad (1)$$

Where *E[.]* represents the expectation operator and the term X^* is the complex conjugate of the DFT defined above. The power spectral density is the first member of the class of higher order spectra [24]. In common with the Fourier transform, the PSD is a function of the frequency f, but contains only magnitude information and considers each frequency component to be independent of all the others [24, 25]. The PSD analysis used resulted in a 1024 point spectrum. Although this could be used directly for diagnosis it was decided to use an automated method to extract the most relevant data from the spectrum and apply this to an artificial neural network for diagnosis.

3.2. The Peak Position Component Method (PPCM)

This work proposes the peak position component as a new method for data extraction. Peak position component analysis can be viewed as a statistical technique for achieving a dimensionality reduction. Using the higher peaks as the original set of variables in an n-dimensional space is transformed in to new set of variables, the so called peaks position component (PPC), in a P-dimensional space such that p < n. Here the basic theory for PSD data will be developed, though generalization to any other quantity is straightforward.

Using all available PSD data, let matrix $[Q(\omega)]_{mxn}$ which has m rows of PSDs, each with n frequency points. A typical element is denoted by $q_{ij}(\omega)$, indices *i* and *j* indicating the position in the matrix:

$$Q = \begin{bmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{m1} & \cdots & q_{mn} \end{bmatrix} \qquad (2)$$

where *i,* = 1, 2, 3, . . . , *m*, *j* = 1, 2, 3, . . . , *n*. The mean response of jth column, i.e. the mean of all of the jth components of the PSDs, can be defined as follows. The mean value for the jth column of the matrix Q is $\overline{q}_j$, where

$$\overline{q}_j = \frac{1}{m}\sum_{i=1}^{m} q_{ij} \qquad (3)$$

The corresponding standard deviation (σ_j) for the jth column of the matrix is:

$$\sigma_j = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}(q_{ij} - \overline{q}_j)} \qquad (4)$$

To clear the noise for the spectra matrix the standard deviation for each column is subtracted from the corresponding column of the spectral matrix to give a new matrix *D* with individual elements d_{ij} where:

$$d_{ij} = q_{ij} - \sigma_j \qquad (5)$$

The next stage is to find the highest peak in the matrix D. First the mean value for matrix D is calculated:

$$\overline{d}_j = \frac{1}{n} \sum_{i=1}^{n} d_{ij} \tag{6}$$

This provides a new array, W, comprising n elements

$$W = \begin{bmatrix} \overline{d}_1 & \cdots & \overline{d}_n \end{bmatrix} \tag{7}$$

The highest peaks contained within the array W are identified and their positions computed. The position of the ith maximum peak in the spectra is denoted p_i and these make up an array P_H:

$$P_{H_P} = \begin{bmatrix} p_1, p_2, & \cdots & , p_n \end{bmatrix} \tag{8}$$

The final matrix used for the analysis for every case, $ANda$, is created by extracting from the matrix D the columns designated by the array PH. This will give only the data which has the highest peaks from the matrix D.

The steps are repeated for each inspection trying to detect a defect. The array $ANda$ is the input data fed to the artificial neural network. The choice of potential features is found to have a significant impact by choosing the largest peaks to reduce the size of the spectrum matrix. Figure 5 shows the schematic diagram of the PPCM method. To avoid the peaks being too close together and to prevent the same peak being picked multiple times, these condition were built into the search algorithm as shown in Figure 5.
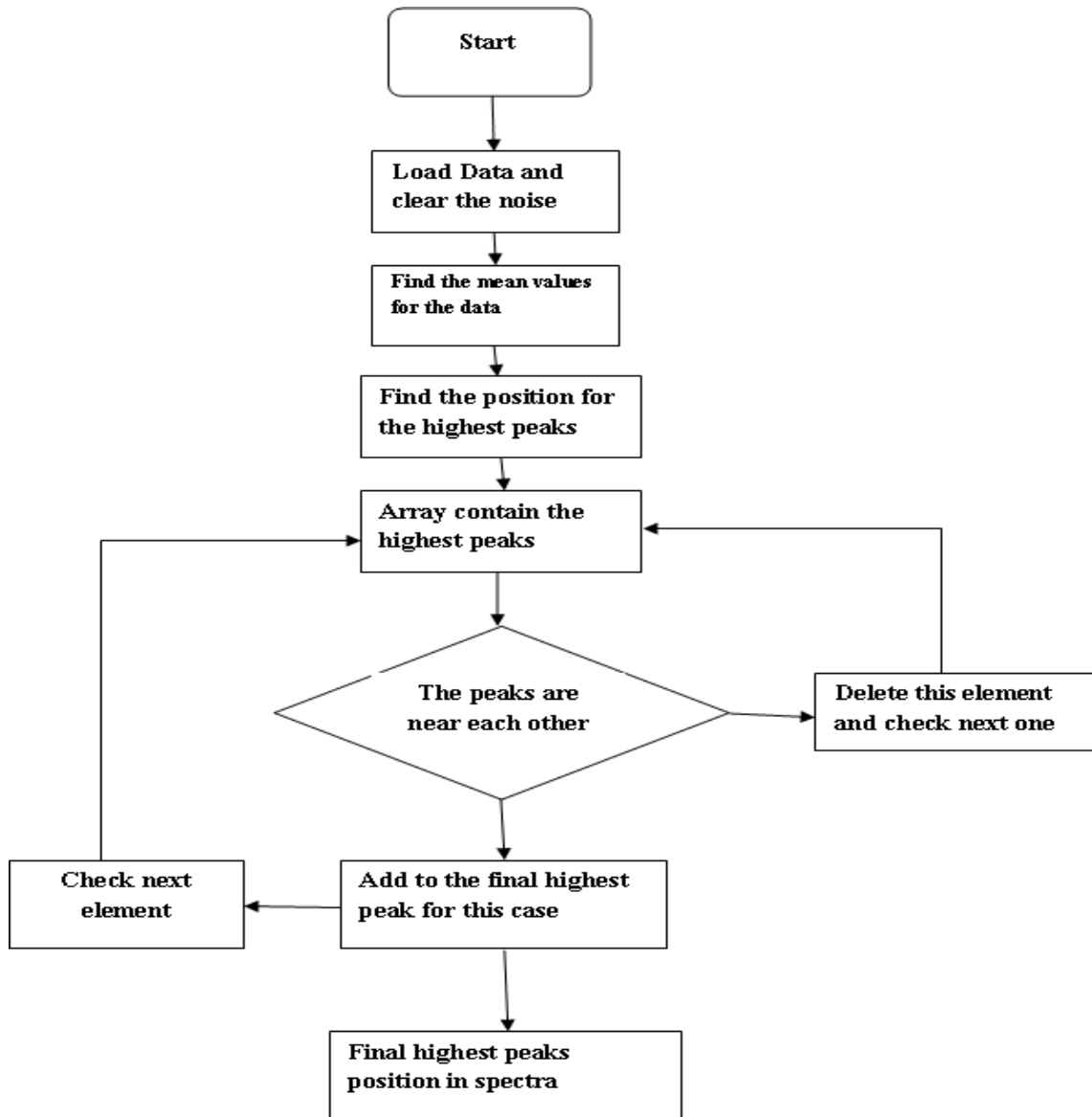
Figure 5 Flowchart of the peaks position component method (PPCM).

3.3. Artificial Neural Network

Artificial neural networks provide a general, non-linear parameterised mapping between a set of inputs and a set of outputs. A network with three layers of weights and sigmoidal activation functions can approximate any smooth mapping and such a type will also be used here. A typical multi-layer perceptron (MLP) supervised feed-forward multi-layer neural network, known as a BP neural network, is used in this study [26, 27].

The MLP, used in this study was a four layer scheme as shown in Figure 6. Each of these layers has nodes (neurons) which contain a non-linear activation function. One of the most commonly used activation functions is the sigmoid function [27]. $x_i(n)$ represents the input to the network, f_i and f_k represent the output of the two hidden layers and $y_i(n)$ represents the output of the final layer of the neural network. The connecting weights between the input to the first hidden layer, first to second hidden layer and the second hidden layer to the output layers are represented by respectively [28].
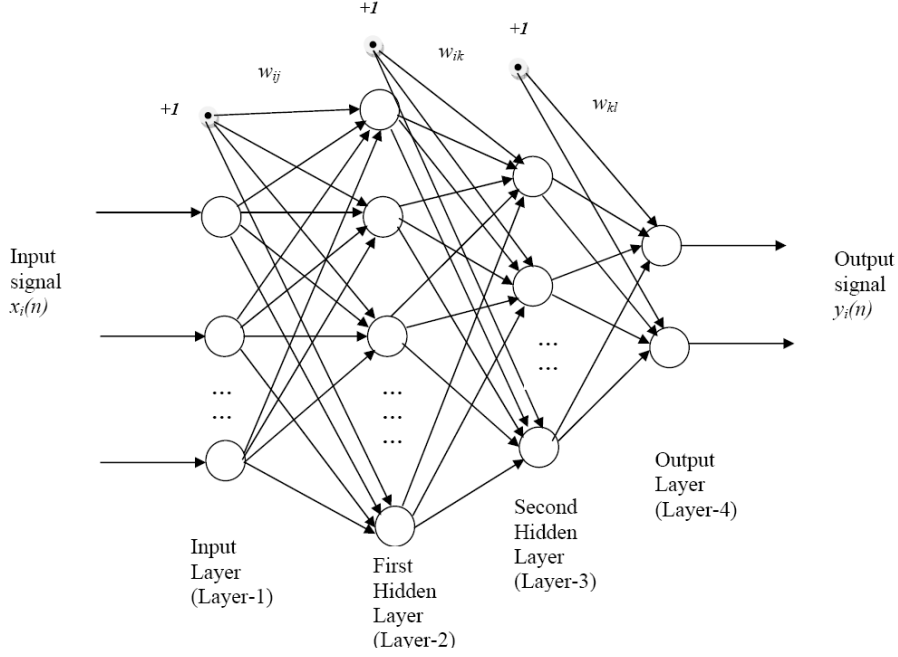
Figure 6 Structure of multilayer perceptron networks

If P_1 is the number of neurons in the first hidden layer, each element of the output vector of the first hidden layer may be calculated as:

$$f_j = \varphi_j\left[\sum_{i=1}^{N} w_{ij}x_i(n) + b_j\right], j = 1,2, \dots P_1 \tag{9}$$

where b_j is the threshold to the neurons of the first hidden layer, N is the number of inputs and φ is the nonlinear activation function in the first hidden layer, and in this study the sigmoid function was chosen. The activation function used is given by:

$$f(netinput) = \frac{1}{1+e^{-netinput}} \tag{10}$$

The time index n was dropped to make the equations simpler. If P_2 is the number of neurons in the second hidden layer, the output of this layer could be represented as, f_k:

$$f_k = \varphi_k\left[\sum_{j=1}^{P_1} w_{jk}f_j + b_k\right], k = 1,2, \dots P_2 \tag{11}$$

where, b_k is the threshold to the neurons of the second hidden layer

The output of the final output layer can be calculated as

$$y_i(n) = \varphi_l\left[\sum_{k=1}^{P_2} w_{kl}f_k + b_l\right], l = 1,2, \dots P_3 \tag{12}$$

where, b_l is the threshold to the neuron of the final layer and P_3 is the number of neurons in the output layer.

The output of the MLP may then be expressed as

$$y_i(n) = \varphi_l\left[\sum_{k=1}^{P_2} w_{kl}\left(\varphi_k\left[\sum_{j=1}^{P_1} w_{jk}\left(\varphi_j\left[\sum_{i=1}^{N} w_{ij}x_i(n) + b_j\right]\right) + b_k\right]\right) + b_l\right] \tag{13}$$

To implement an ANN the system needs trained. Back propagation training of an MLP [29,26] can be summarised in the following eight steps:
1. Random initialisation of the weights in order to start with the training.

2. The input pattern is given to the input layer by assigning each value of the pattern to each neuron of the input layer.

3. The network passes the input values to the first hidden layer. In each of the neurons of the hidden layer a value is computed by doing the sum over the products of the output of each neuron in the input layer by the weights of the connection between this neuron of the input layer and the neuron of the hidden layer.

4. The value computed is now passed through a non-linear activation function, giving an output for each of the hidden layer neurons.

5. Step 3 and 4 are repeated for each additional hidden layer.

6. Once the output layer of the neural networks is reached and the non-linear activation function has been applied, the values of the output layer are considered as the output pattern of the network.

7. An error value is calculated, by comparing the actual output pattern with the desired target. If the error satisfies the specifications the training is stopped.

8. If the error is not low enough, the weights are modified using a training algorithm such as the Delta Rule algorithm, Levenberg-Marquardt algorithm, Quasi-Newton algorithm, etc [29, 30, 31]. Steps 2-7 are repeated until the error satisfies the requirements or a maximum number of iterations is achieved.

Once the training is complete and the results are validated, the neural network is ready to be used in a mode dealing with the unseen data. Once the neural network has then been tested and validated as successful for previously unseen data, it can be used as a black box system that only cares about the inputs and the outputs of the system.

When the ANN training is completed the performance of the ANN can be validated by the success rate (*SR*) of how many unseen patterns are correctly classified. The success rate is defined as follows

$$SR = \frac{number\ of\ un\ seen\ patterns\ correctly\ classified}{number\ of\ unseen\ patterns} \times 100\% \qquad\qquad 15$$

The application of an ANN to condition monitoring of rotor shafts is now presented below. For this application, the multi layer perceptron supervised network was used due to the fact that the desired outputs were well known. The back-propagation algorithm chosen for this study was the Levemberg Marquart. This has been used extensively and has given good results in other studies [28-32].

In this study the data was divided into 2 groups for each fault condition all of which were processed according to the procedure documented in section 3.2 to provide the arrays *ANda*. The first group of data was used to train the ANN. The second group was used for testing using "unseen" data. The network was trained using the fault condition data as inputs along with the corresponding fault conditions as the known outputs.

4. Experimental results

The experimental results are divided into three parts. The first sets of experiments results are a presentation of the averaged spectra for the healthy shaft (reference shaft). The second part comprises a comparison between the un-cracked shaft and shaft with a 40, 50 or 60% pre-crack using Ch1, Ch2 and Ch3. The third part of the results presents application of the PPCM and neural networks to the vibration results

4.1 Reference situation

The reference was used for two things, firstly to check if there were any changes in the system response for reasons other than the crack, and then secondly to use it to compare with the shafts with defects. The shaft with a simulated crack was then examined with the vibration responses are measured for different defect depths. The measured reference situation is reported in Figure 7 for the un-cracked shaft and shows the peaks that are going to be considered for comparing the un-crack and pre-crack shafts for Ch1, Ch2 and Ch3. As shown in Figure 7 Ch1 has three frequency ranges containing peaks. The first two peaks are from 300 to 400 Hz. There is then a peak between 800 to 1000 Hz, and the last peak is in the range 1000 to 1200 Hz. For Ch2 there are two ranges that are going to use for comparing between un-cracked and cracked shaft. These ranges are 325 to 365 Hz and 1100 to 1200 Hz. The Ch3 vibration represents a measure of the vibration of the test rig, because it is not connected directly with any source of vibration (motor or load). The vibration of Ch3 gives a general picture of how the vibration changes when the state of the shaft changes. It should be noted that these frequencies do not correspond to the shaft rotational speed. The peaks appear to be excitation of the natural frequencies of the shaft excited by small random excitation components. The crack in the shaft changes the structural properties of the shaft. The change in the structural stiffness will alter the frequency spectrum produced by the shaft when operating or disturbed. Thus, if the shaft vibrations can be sensed, then it is possible that the presence of a defect can be inferred due to a change in the frequency spectrum produced by the shaft.
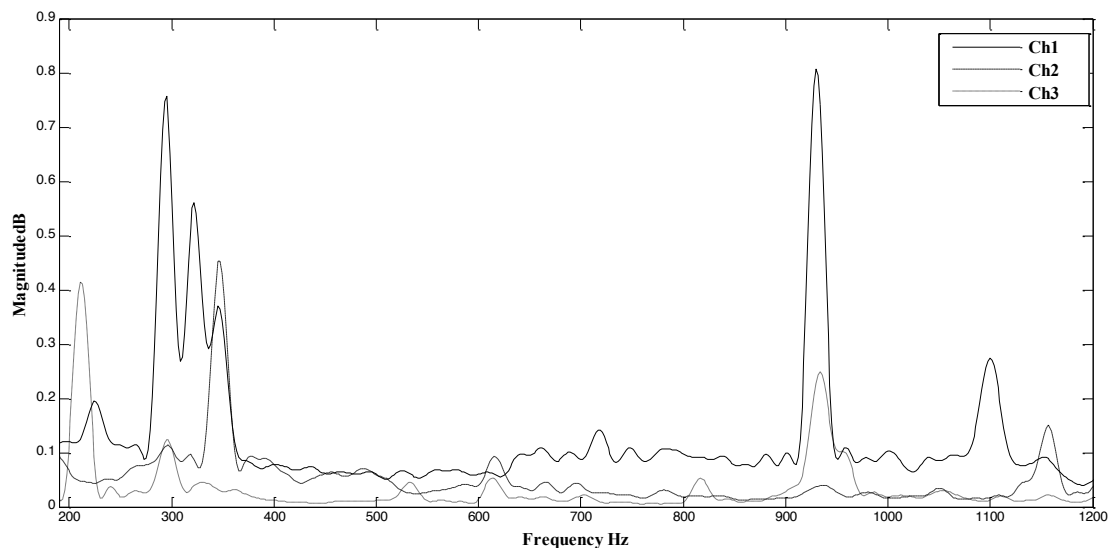

Figure 7 Vibration spectra of Ch1, Ch2 and ch3 with un-cut shaft (Reference situation)

4.2 Comparison between un-cracked shaft 40, 50 and 60% pre-cracked shaft using Ch1, Ch2 and Ch3

Representative results for the frequency shifts in the system response for the experiments with the cracked shaft are presented in Figure 8. It can be seen from Table 2 and Figure 8b that for Ch1 the peak in frequency range 800 to 1000 Hz reduces in frequency from the reference case (un cracked shaft) by 70.3 Hz for the 40% cracked shaft shifts by, 85.9 Hz for the 50% pre-cracked shaft and 107 Hz for the 60% pre-cracked. For Ch2 there are two ranges from 325 to 365 Hz and 1100 to 1200 Hz are used for comparing between un-cracked and pre-cracked shaft. The results in Table 2 and Figure 8 show that Ch2 follows the same trend as Ch1 with the largest shift from the reference to the 40% cut and progressive reduction in frequency as the cut depth increases. The vibration of Ch3 gives a similar although somewhat less clearly defined trend in the ranges 190 to 230Hz and 500 to 550 Hz even although this measurement position is not so closely coupled to the shaft as the bearings of the motor and generator which support the shaft.

Although quite large cut depths were used in the experiments, the data for Ch1 in the region 800 to 1000 Hz has a large shift for the 40% cut. Plotting this frequency as a function of cut depth shows an approximately linear

shift from which a frequency shift of 5 Hz would correspond to a 3% cut depth change. A 5 Hz shift is substantially larger than the 1.8Hz resolution of the spectra.
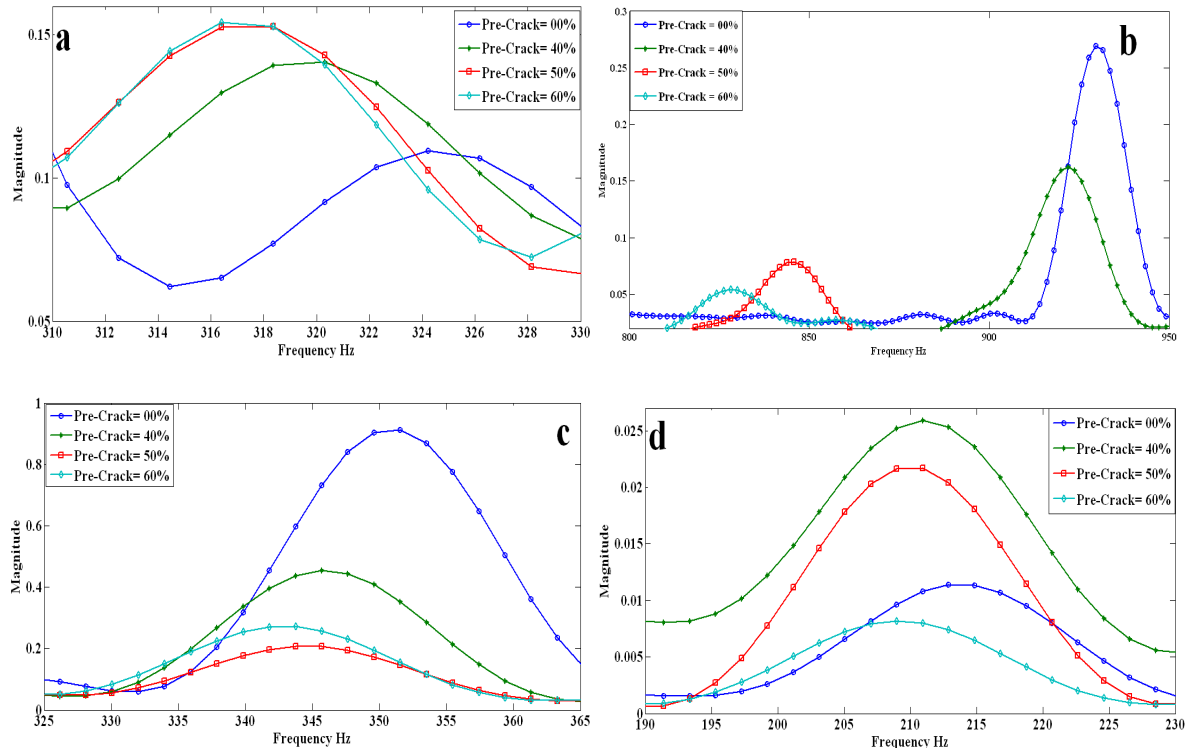

Figure 8 Vibration spectra of Ch1, Ch2 and Ch3 for different cut depths (a) Ch1 frequency range of 310 to 330 Hz (b) Ch1 frequency range of 800 to 950 Hz(c) Ch2 frequency range of 325 to 365 Hz (d) Ch3 frequency range of 190 to 230 Hz

Table 2 Frequency peaks of 90° Pre-crack for different pre-crack depths

	Frequency for different 90° pre-crack (Hz)			
	Depth(% of shaft diameter)			
Frequency range Hz	0%	40%	50%	60%
Ch1 290 to 310	296.9	293	291	291
Ch1 310 to 330	324.2	320	318.4	316.4
Ch1 800 to 960	988.3	918	898.4	880.9
Ch2 325 to 365	351.6	347.7	345.7	343.8
Ch2 1120 to 1170	1170	1154	1148	1145
Ch3 190 to 230	214.8	210.9	210.9	209
Ch3 500 to 550	531.3	525.4	527.3	525.4

It should be noted that the frequencies identified are well away from the rotational speed of the shaft which is 1550rpm (25.334Hz) and the peaks are generally not harmonics of the shaft speed. Analysis was undertaken to assess if these peaks corresponded to any of the classical bearing fault frequencies but they do not. In addition such frequencies are shaft speed dependent and so would not show a shift other than by a change in shaft speed.

4.3 Applying PPCM and Neural networks to vibration results

The PPC data reduction method reduces the no of inputs from 1024 to 5 for Ch1, 3 for Ch2, and 2 for Ch3. Although not a complex technique, the PPC method does provide effect data reduction and has picks out

peaks which are fault sensitive and would not normally be examined using the traditional selection of the harmonics of the shaft. This the allowed ANNs to be implemented for condition as follow.

1000 files of vibration data were generated for every shaft for Ch1, Ch2 and Ch3. Half of these files were used to train the neural networks; the other half of the data set was used to test the networks performance with each case using 500 training patterns for each case of the condition. Those cases are

1. Normal shaft
2. Faulty shaft with 40% cut
3. Faulty shaft with 50% cut
4. Faulty shaft with 60% cut.

The MATLAB Neural Network Toolbox was used for this investigation. A two layered feed-forward neural network trained using the back propagation algorithm was employed to perform shaft condition diagnosis.
Three separate neural networks were trained, one based on the vibration obtained from load side Ch1, the second one based on the vibration obtained from motor side Ch2 and third one based on the vibration obtained from shaft cover Ch3. The three neural networks are required to predict the shaft condition under normal working condition.

Since the multilayer perceptron (MLP) type neural network belongs to the class of supervised learning networks, a 'teacher' is required in order to achieve the required performance. The expected targets were vectors defined by the four different patterns T1= [0 0],   T2= [0 1], T3=[1 0] and T4=[1 1] for the normal shaft,40% defect shaft, 50% defect shaft, 60% defect shaft respectively. For the training of the networks, the sum-squared goal was 10^{-10} and the number of epochs was limited to a maximum of 1000.

After successfully training the ANN networks, Ch1, Ch2 and Ch3 were then used to diagnose the shaft states for unseen patterns. The results show that they are very effective in distinguishing the four shaft condition states. The results obtained are shown in Table 2.

The vibration based MLP for Ch1 and Ch3 distinguished the four kinds of shaft condition with a 100% success rate. The vibration based MLP for Ch2 also identified the four separate conditions with more than 99.7% success rate. In this investigation it has been found that the ANN identifies all type of faults with a high success rate.

Table 2 Success rate of network diagnose for 500 unseen test patterns

Networks name	Success rate (%)			
	Cut 0%	Cut 40%	Cut 50%	Cut 60%
Ch1	100	100	100	100
Ch2	99.4	99	99.3	99.7
Ch3	100	100	100	100

As well as the raw data collected from the rig, the ANN was also tested with reduced data sets interpolated between fault conditions to assess the results of such inputs with the ANN used as an estimator rather than the categorisation used earlier. Although not quite as successful as when identifying the actual fault conditions, the interpolated data could still be diagnosed.

5. Conclusion

When using the frequency domain (PSD) approach, the shaft vibration measured at the bearings and on the shaft cover provides a very clear indication of shaft condition for the 40, 50, and 60% cut shafts for all cases. The peaks in the frequency ranges from 0 to 1200 Hz typically shifted by between 2 and 85 Hz from the reference case depending on the frequency range being examined. Comparison between the 90° cut and 45° cut shafts show that the trends in the results from both cases were in agreement and in some cases the values of frequency shift within particular frequency ranges were similar.

The peak position component (PPC) analysis is a powerful tool for reducing the size of measured frequency response data. Such a reduction has clear benefits over data reduction via point selection or modal analysis. Even when a relatively small number of inputs are used, the PPC-compressed PSD appear to retain most of the original information. This feature is likely to be suited to other signal processing applications. The use of PPC reduced the number of inputs to the ANN but retained the prominent features for defect detection.

In this paper the application of the multi layer perceptron (MLP) based system to the fault diagnosis of rotor shaft system has been presented. For the three channels of the long rotor shaft application presented in this paper, the vibration power spectra has been given to the ANN, using the PPC which has been able without help to find the features of each fault in the case of the particular situations. In order to interpret the power spectra signatures, an ANN was developed which has been trained to perform diagnosis of the condition of the shaft.
Three networks have been used, the networks using different inputs. All of those networks with different inputs size give good results. Results presented have also shown that the approach developed can be successfully and reliably applied in real time.

6. ACKNOWLEDGEMENTS

7. References

[1] L. Swanson, Linking maintenance strategies to performance, International Journal of production Economics 70 (2001) 237–244.
[2] Y. Bar-Cohen, Emerging NDE technologies and challenges at the beginning of the 3rd millennium - Part I., Material Evaluation 58 (1) (2000) 17–30.
[3] A. D. Dimarogonas, Vibration of cracked structures: A state of the art review, Engineering Fracture Mechanics 55 (5) (1996) 831–857.
[4] D. Dimarogonas, C. A. Papadopoulos, Vibration of cracked shafts in bending, Journal of Sound and Vibration 91 (4) (1983) 583–593.
[5] O. S. Jun, O. S. Eun, Modelling and vibration analysis of a simple rotor with a breathing crack, Journal of Sound and Vibration 155 (2) (1992) 273–290.
[6] G. Dalpiaz, A. Rivola, R. Rubini, Effectiveness and sensitivity of vibration processing techniques for local fault detection in gears, Mechanical Systems and Signal Processing 14 (3) (2000) 387–412.
[7] P. Samuel, D. Pines, A review of vibration-based techniques for helicopter transmission diagnostics, Journal of Sound and Vibration 282 (1- 2) (2005) 475–508.
[8] U. Benko, J. Petrovˇciˇc, J. Juriˇci´c, J. Tavˇcar, J. Rejec, A. Stefanovska, Fault diagnosis of a vacuum cleaner motor by means of sound analysis, Journal of Sound and Vibration 276 (3-5) (2004) 781–806.
[9] Z. Peng, F. Chu, Application of the wavelet transform in machine condition monitoring and fault diagnostics: A review with bibliography, Mechanical Systems and Signal Processing 18 2 (2) (2004) 199–221.
[10] L. Toms, Machinery oil analysis - Methods, automation and benefits, 2nd Edition, Coastal Skills Training, Virginia, 1998.
[11] J.-J. Sinou, A. Lees, A nonlinear study of a cracked rotor, European Journal of Mechanics A/Solids 26 (1) (2007) 152–170.
[12] A. Darpe, K. Gupta, A. Chawla, Dynamics of bowed rotor with transverse surface crack, Journal of Sound and Vibration 296 (4-5) (2006) 429–445.
[13] T. Patel, A. Darpe, Vibration response of a cracked rotor in presence of rotor-stator rub, Journal of Sound and Vibration 317 (3-5) (2008) 841–865.
[14] J. Antoni, R. Randall, The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines, European Journal of Mechanics A/Solids 20 (2) (2006) 308–331.
[15] Y. Ishida, T. Inoue, Detection of a rotor crack using a harmonic excitation and nonlinear vibration analysis, Journal of Vibration and Acoustics128 (6) (2006) 741–749.
[16] E. Rich, K. Knight, Artificial Intelligence, 2nd Edition, McGraw Hill, U.S.A., 1991.
[17] B. Li, M. Chow, Y. Tipsuwan, J. Hung, Neural network based motor rolling bearing fault diagnosis, IEEE Transactions on Industrial Electronics 47 (5) (2000) 1060–1069.
[18] T. W. S.M. Islam, G. Ledwich, A novel fuzzy logic approach to transformer fault diagnosis, IEEE Transactions on Dielectrics and Electrical Insulation 7 (2) (2000) 177–186.

[19] Y. L. Z. Wang, P. Griffin, A combined ann and expert system tool for transformer fault diagnosis, IEEE Transactions on Power Delivery 13 (4) (1998) 1224–1229.

[20] R. G. A. Malhi, Pca-based feature selection scheme for machine defect classification, IEEE Transaction on Instrumentation and Measurement 53 (6) (2004) 1517–1525.

[21] A.Widodo, B.-S. Yang, T. Han, Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors, Expert Systems with Applications 32 (2) (2007) 299– 312.

[22] P. Tavner, J. Penman, Condition Monitoring of Electrical Machines, Research Studies Press Ltd, England, 1987.

[23] R. Barron, Engineering condition monitoring: Practices methods and application, Addison-Wesley Longman, 1996.

[24] P. D. MacFadden, J. Smith, Model for the vibration produced by a signal point defect in rolling elements bearing, Journal of Sound and Vibration 96 (1) (1984) 69–82.

[25] S. Kay, Spectrum analysis a modern perspective, Proceeding of the IEEE 69 (11) (1981) 1380–1414.

[26] M. T. Hagan, M. B. Menhaj, Training feedward networks with the marquardt algorithm, IEEE Transactions on Neural Networks 5 (6) (1994) 989–993.

[27] A. Carling, Introducing Neural Networks, Sigma Press, U.K., 1992.

[28] S. Haykin, Neural Networks: A Comprehensive Foundation, Pearson `Edition Asia, 2002.

[29] R. P. Lippman, An introduction to computing with neural nets, IEEE ASSP Magazine April (1987) 4–22.

[30] P. J. Werbos, Beyond regression: new tools for prediction and analysis in the behavioral sciences, Ph.D. thesis, Harvard University (1974).

[31] K. Swingler, Applying Neural Networks, Academic Press Limited, U.K., 1996.