

A new methodology for fault detection in rolling element bearings using singular spectrum analysis

Hussein Al Bugharbee¹, and Irina Trendafilova²

¹ Department of Mechanical engineering, the University of Wasit, Iraq

² Department of Mechanical and Aerospace Engineering, the University of Strathclyde, Glasgow G1 1XJ

Abstract. *This paper proposes a vibration-based methodology for fault detection in rolling element bearings, which is based on pure data analysis via singular spectrum method. The method suggests building a baseline space from feature vectors made of the signals measured in the healthy/baseline bearing condition. The feature vectors are made using the Euclidean norms of the first three PC's found for the signals measured. Then, the lagged version of any new signal corresponding to a new (possibly faulty) condition is projected onto this baseline feature space in order to assess its similarity to the baseline condition. The category of a new signal vector is determined based on the Mahalanobis distance (MD) of its feature vector to the baseline space. A validation of the methodology is suggested based on the results from an experimental test rig. The results obtained confirm the effective performance of the suggested methodology. It is made of simple steps and is easy to apply with a perspective to make it automatic and suitable for commercial applications.*

1 Introduction

Rolling element bearings (REB) constitute a major part of most rotating machinery. No matter how well and precisely the bearings are designed, how normally loaded, correctly assembled and properly lubricated they are, they fail due to material fatigue after a certain number of revolutions or operational time. The defects in REBs are considered as one of the main reasons for rotating machinery failures. The usual monitoring philosophy for REB's is vibration-based monitoring and it is based on their measured vibration response signals. During recent decades, a number of studies suggested different approaches for purposes of REB fault detection. Several of these focus on the review and the comparison of some of the approaches for REB's vibration-based health monitoring.

Most fault detection methods based on bearing vibration signal analysis usually include the extraction of certain features representative of the bearing state, which are used as means of comparison against the healthy state. There are many features suggested using a number of methods that are based on analysing the vibration signal in the different domains (*i.e* time, frequency and time-frequency) [1, 2]. Several challenges accompany the feature selection and extraction which include the sensitivity of the features to the change of the bearing condition and their noise perceptiveness.

The methodology proposed in this study is based on singular spectrum analysis (SSA) and is simple in structure and easy to apply. SSA is a time series analysis method that is primarily used to uncover the trend and

the periodic components of the original time series [3]. It has two main stages: decomposition and reconstruction. Only the decomposition stage is used in this study, whereas bearing vibration signals are decomposed into a new space made of a number of their principal components (PCs), which correspond to the directions of the biggest variability in the original data [4]. Thus most of the variance of the original data is preserved as the original signal is projected onto the space of some of its first PC's. Accordingly a certain amount of the information contained in the original signal is preserved in each PC in terms of its variance. In this study the norms of the first several PC's are used to build a reference space corresponding to the healthy bearing condition. Thus, the reference space preserves a certain amount of the information for the healthy state of the bearings depending on the number of PC's used to build it. From this perspective any new signal is projected onto the reference space in order to assess its similarity to the healthy/ baseline condition.

As the main goal of the current study targets only the detection of faults and, due to the limitation of the manuscript length, the capabilities of the method regarding location identification and fault severity estimation will not be discussed in details here. We'll merely mention that the same method can be used for purposes of further fault identification [5].

This paper offers a validation of the methodology using data measured on an experimental test rig, which was built specifically for the purpose at Strathclyde University.

* Corresponding author: hrazzaq@uowasit.edu.iq, hussainmcc@yahoo.com

The rest of the paper is organised into the following sections. The next section 2 presents the methodology for fault detection. Section 3 covers an experimental case study where the application of the developed methodology is demonstrated. Section 4 discusses the results obtained in the previous section 3. The last section 5 gives some general conclusions and directions for future research.

2. Methodology for fault detection in rolling element bearings.

The methodology suggested in this paper is based mainly on the decomposition stage of the singular spectrum analysis. It has two main steps: building a baseline/reference space and fault diagnosis. These two steps are detailed below.

2.1 Building the reference space

To build a reference space, a number L of discretized vibration signals measured on healthy bearing (i.e $x_i=[x(1), x(2), \dots, x(n)]$) are arranged into a matrix \mathbf{X}_{ref} as shown in equation (1):

$$\mathbf{X}_{ref} = \begin{bmatrix} x_1(1) & x_1(2) & x_1(3) & \dots & x_1(n) \\ x_2(1) & x_2(2) & x_2(3) & \dots & x_2(n) \\ x_3(1) & x_3(2) & x_3(3) & \dots & x_3(n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_L(1) & x_L(2) & x_L(3) & \dots & x_L(n) \end{bmatrix} \quad (1)$$

Where L is the window size of the signal decomposition. Then the covariance matrix of the matrix \mathbf{X}_{ref} is obtained as shown in equation (2).

$$\mathbf{C}_X = \frac{\mathbf{X}_{ref} \mathbf{X}_{ref}^T}{L} \quad (2)$$

The letter ‘‘T’’ denotes the transpose. Then the covariance matrix is subjected to singular value decomposition to extract its eigenvectors \mathbf{U}_i and eigen values λ_i according to the equation (3).

$$\mathbf{C}_X \mathbf{U}_i = \lambda_i \mathbf{U}_i \quad (3)$$

Each eigenvalue λ_i represents the percentage of the variance of the original signal contained in the direction of \mathbf{U}_i . In this study the first three eigenvectors are used to build the reference space where all the lagged versions of others signals are projected as it will be described in the next section.

2.2. Fault assessment methodology

The fault assessment methodology has two main stages: feature extraction and a fault detection process. The fault detection process involves the setting of a threshold to distinguish between features corresponding to the healthy and the faulty states and the classification process.

2.2.1 Feature extraction

The signals measured on the healthy bearings are initially divided into a training and a testing sample. While all other signals obtained from the faulty bearing are used as a testing sample. The signals from the training sample are transformed into their lagged version (see equation (4)) and then multiplied (i.e projected) by the eigenvectors \mathbf{U}_i obtained in equation (3). This is presented in equation (5).

$$\mathbf{X} = \begin{bmatrix} x(1) & x(2) & x(3) & \dots & x(K) \\ x(2) & x(3) & x(4) & \dots & x(K+1) \\ x(3) & x(4) & x(5) & \dots & x(K+2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x(L) & x(L+1) & x(L+2) & \dots & x(n) \end{bmatrix} \quad (4)$$

where $K=n-L+1$ and L is the pre-selected window size of the signal decomposition.

Projecting the j^{th} signal onto a number of eigen vectors results in a set of PC’s, $i=1,2, \dots, L$ as detailed in equation (5) below.

$$PC_{ij}(m) = \sum_{k=1}^L X^T(m+k-1) * U_i(m) \quad (5)$$

where PC_{ij} is the i^{th} projection corresponding to j^{th} observation, $i=1,2, \dots, L$, $j=1,2, \dots, [please\ complete,]$, and $m=1,2, \dots, K$ are the components of the PC. Then the Euclidean norms of the first Q PCs are calculated according to equation (6).

$$f_{ij} = \sum_{m=1}^K (PC_{ij}(m))^2 \quad (6)$$

Where $i = 1,2, \dots, Q$ and $j = 1,2,3 \dots k$. The notation Q refers to the number of PC’s used in the analysis and in this case $Q=3$. The notation k refers to the number of the signals used in the analysis. The above features are arranged to form k feature vectors as shown in equation (7). Each FV has $Q = 3$. components.

$$fv_j = [f_{1j} \ f_{2j} \ f_{3j}] \quad (7)$$

Accordingly the feature vector space is defined. In the present study, only the first three features have been used in the analysis. Thus, the FVs dimension is three

and it can be represented as a point into the three dimensional space.

2.2..2 Fault detection

The fault detection in this case is regarded as a classification problem where two classes are introduced – the class of signals from healthy bearings and the class of signals from faulty bearings. Thus each signal from the testing sample can be assigned to one of these classes. The classification process has two steps: setting a threshold and comparison to of the features calculated for a signal under test to the threshold.

For setting a threshold, a feature matrix ($\mathbf{F}_{baseline}$) is first made by arranging all the training FVs in rows. Then the Mahalonobis distance (D) of each of the training FVs to the $\mathbf{F}_{baseline}$ is calculated as in equation (8).

$$D_i = \sqrt{(\mathbf{fv}_i - \mathbf{E}_{baseline}) \cdot \mathbf{S}^{-1} \cdot (\mathbf{fv}_i - \mathbf{E}_{baseline})^T} \quad (8)$$

Where $\mathbf{E}_{baseline}$ is the mean of the rows of $\mathbf{F}_{baseline}$ and \mathbf{S}^{-1} is the inverse of the covariance matrix of $\mathbf{F}_{baseline}$ which can be calculated as per equations (2) and (3) applied for $\mathbf{F}_{baseline}$.

The threshold (Thr_{ref}) is eventually selected as the maximum value of the distances D_i calculated for all the training FV's (9)

$$Thr_{baseline} = \max(D_i) \quad (9)$$

where $i=1,2,..N$ and N is the number of training FV's used.

Once threshold is set, the distance of the testing FVs to $\mathbf{F}_{baseline}$ is calculated as per equation (8). Then each new testing FV is assigned to one of the classes-healthy or faulty -by comparing it to $Thr_{baseline}$ as detailed in equation (10).

$$\left. \begin{array}{l} D_i > Thr_{baseline} \quad \text{a } FV_i \text{ is assigned to F class} \\ D_i \leq Thr_{baseline} \quad \text{a } FV_i \text{ is assigned to H class} \end{array} \right\} \quad (10)$$

3. Methodology verification

3.1 Case study and experimental setup

The experimental setup for this case study, which is built at the Department of Mechanical and Aerospace Engineering at Strathclyde University consists of 1 HP shunt DC motor, bearing assembly and a mechanical loading system, and is presented in Figure 1.



Fig. 1. Bearing test rig at Strathclyde university.

The vibration acceleration data is collected from the test bearing housing (in this case a SKF deep groove 6308 ball bearing) is mounted inside. A pinion-toothed belt mechanism is used to transmit the torque from the motor to the bearing assembly. A fault diameter of 0.05 inch was introduced by an electrical discharge machine on the inner race, on a ball and on the outer race for different bearings. The bearing vibration acceleration datasets is collected at a sampling rate of 12 kHz for different speeds (*ie* 250 r/min, 750 r/min and 1250 r/min) using a magnetic-based accelerometer.

3.2 Data collection and signal analysis

The data used in the analysis is shown in Table 1, where (H) refers to a healthy bearing condition, IRF refers to a bearing with an inner race fault, BF refers to a bearing with a ball fault and ORF refers to a bearing with an outer race fault.

Table 1. Signals used in the verification

Case No	Motor speed	Signal category	No of signals
1	250	Healthy,IRF,BF,ORF	320 (4X80)
2	750	Healthy,IRF,BF,ORF	320 (4X80)
3	1250	Healthy,IRF,BF,ORF	320 (4X80)

The signals are segmented into 320 equal non-overlapping signals (*ie* 4 signal classes \times 80 signals per class, each one containing 2048 data points). Forty of the healthy signals are used as training sample and the remaining 280 (*ie* 40 healthy and 240 faulty) signals are used as a testing sample.

Figure 2 presents a signal measured on a healthy bearing at a speed of 1250 r/min. The x -axis represents the number of data points while the vertical y -axis is the acceleration of the signal vibration.

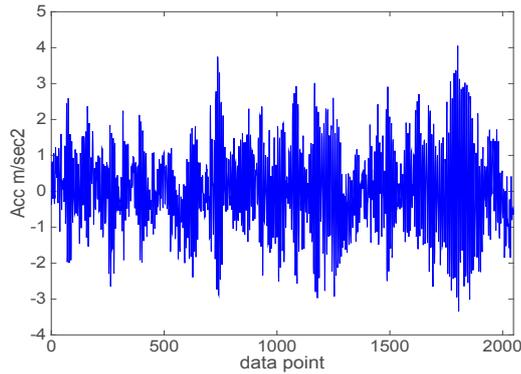


Fig. 2. An example of healthy bearing vibration signal

4. Results and discussion

As an illustration Figure 3 shows a 3D visualisation of the 40 FVs corresponding to the baseline-training sample, which are used to form the baseline features.

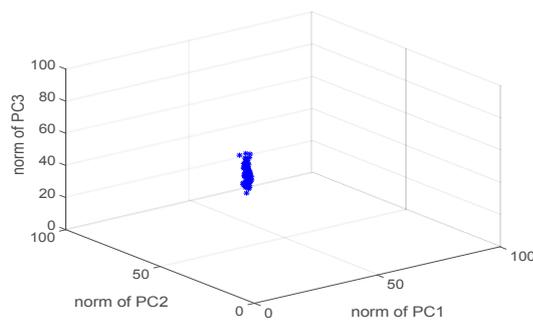


Fig. 3. 3D visualization of the training sample of signals measured on healthy bearings

In this study, only the first three PCs are used to form the feature vectors. The feature vectors are formed using the norms of these first three PC's.

The next Figure 4 shows a 3D visualization of the features corresponding to all the signals from the different bearing conditions.

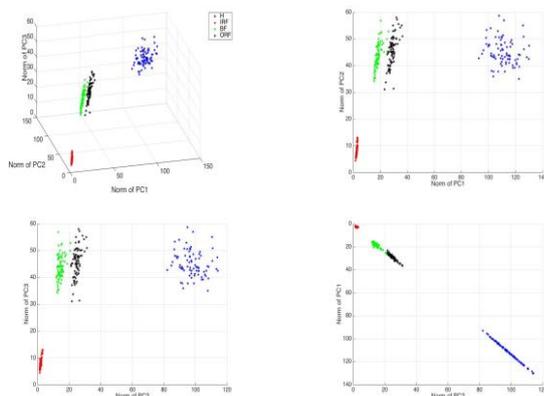


Fig. 4. 3D and 2D visualizations of FV's from different bearing conditions.

The blue dots refer to healthy FVs, red dots refer to IRF FVs, green dots refer to BF FVs and the black dots refer to ORF FVs. It can be seen that the performed signal analysis has resulted in clustering the signals from the different fault categories and in terms of the resulting features the fault categories clearly distinguishable. From the figures on Figure 4 it can be seen that there is separability not only between the baseline and non-baseline FVs but also among the different fault categories.

The next Figure 5 shows the Mahalanobis distances (MDs) of both the training and the testing FVs measured to the baseline testing FVs category.

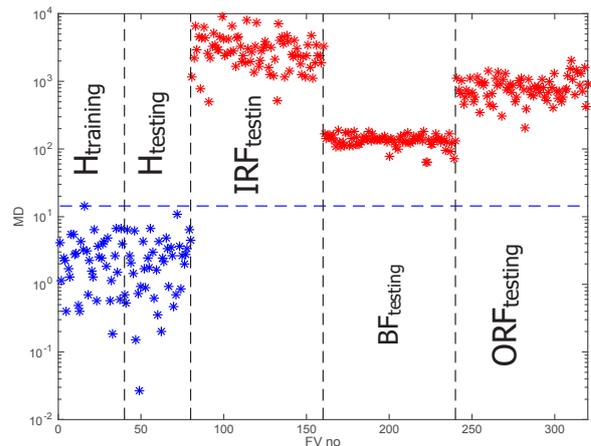


Fig. 5. MD's for the training and the testing samples corresponding to different fault states

As the MD is used as a fault indicator in this study, it is compared to the predetermined threshold constructed as detailed in equation (9), & 2.2.2 .

Figure 5 is divided into 5 zones corresponding to the different bearing conditions and also separating the testing and the training MD's, using the vertical dashed lines. The H_{training} Zone corresponds to training sample of healthy bearings signals (H), the H_{testing} zone is made of the testing sample for healthy signals. The other three zones (IRF_{testing} , BF_{testing} & ORF_{testing}) represent the testing FVs corresponding to IRF, BF and ORF classes respectively. The horizontal dashed line represents the value of the threshold calculated. The separation between the baseline and non-baseline FVs is very clear and it results in very accurate classification and fault detection. As can be seen from figure 5 all the testing feature vectors (100%) are correctly classified.

The performance of the methodology using the datasets of this case study is shown in Table 2. The first column shows the type of dataset as corresponding to the revolutions of the motor at which the data is collected. The second, the fourth and the fifth columns show the correct classification rate of the training and testing FVs for the healthy and the faulty signals, respectively.

Table 2. Percentages of correct classification rates of signals used in the analysis

Speed (rev/min)	% correct classification		
	H training	H testing	F testing
250	100%	100%	100%
750	100%	100%	100%
1250	100%	100%	100%

From Table 2, it can be seen the all the 320 FVs are correctly classified to their correct real classes.

5. Conclusions and future work.

The present study suggests a rather simple and easy to apply but accurate method for fault detection in rolling element bearings based on singular spectrum analysis. The suggested method is simple in application as it solely uses the signals from the healthy bearing state to build the reference space. As such, it does not require any previous measurements corresponding to faulty/ anomalous conditions. A number of signals measured in the baseline/healthy condition can be used to build the reference space. SSA is used only for the purposes of building this reference space. The new signals, which are to be classified, are not subjected to SSA; their lagged versions are simply projected onto the first principal directions of the reference space. As such, the transformations applied to the measured signals are minimal and very simple. The classification rule, which is based on a threshold of the Mahalanobis distance, is also a robust and simple one.

As a result of this, the method holds considerable potential for automatization and practical implementation.

It should also be noted that the method is rather general and can be applied to any measured signals regardless of their stationarity. The 3D visualisation shows that not only baseline and non-baseline signal categories can be distinguished but in some cases the method can be used to separate different faulty categories. This remains a subject of further research and as such the interpretation of the PCs/the principal directions will be very helpful. Based on the results obtained, it has been shown for the case studies considered that SSA is capable of extracting essential information regarding the presence of faults in signals where different fault locations are detected at different motor speeds. As the primary goal of the current study is limited to the detection of faults, further investigations are targeted in terms of expanding the method for fault qualification and quantification purposes and also for online monitoring of the bearing condition.

References

1 A. Sharma, M. Amarnath, and P. Kankar, "Feature extraction and fault severity classification in ball

bearings," *Journal of Vibration and Control*, vol. **22**, pp. 176-192, (2016).
 2 B. Nayana and P. Geethanjali, "Analysis of Statistical Time Domain Features Effectiveness in Identification of Bearing Faults from Vibration Signal," *IEEE Sensors Journal*, (2017).
 3 H. Hassani, "Singular spectrum analysis: methodology and comparison," (2007).
 4 H. Al-Bugharbee and I. Trendafilova, "A fault diagnosis methodology for rolling element bearings based on advanced signal pretreatment and autoregressive modelling," *Journal of Sound and Vibration*, vol. **369**, pp. 246-265, (2016).
 5 H. R. Al-Bugharbee, "Data-driven methodologies for bearing vibration analysis and vibration based fault diagnosis," *PhD thesis*, University of Strathclyde, 2016.