Mechanical Systems and Signal Processing 25 (2011) 1765–1772

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/jnlabr/ymssp

# Evaluation of principal component analysis and neural network performance for bearing fault diagnosis from vibration signal processed by RS and DF analyses

E.P. de Moura [a,*], C.R. Souto [b], A.A. Silva [b], M.A.S. Irmão [c]

[a] Departamento de Engenharia Metalúrgica e de Materiais, Universidade Federal do Ceará, 60455-760 Fortaleza, CE, Brazil
[b] Departamento de Engenharia Mecânica, Universidade Federal de Campina Grande, 58109-970 Campina Grande, PB, Brazil
[c] Colegiado de Engenharia Mecânica, Universidade Federal do Vale do São Francisco, 48909-810 Juazeiro, BA, Brazil

ARTICLE INFO

*Article history:*
Received 30 June 2010
Received in revised form
26 November 2010
Accepted 30 November 2010
Available online 7 December 2010

*Keywords:*
Bearing
Fault diagnosis
Vibration analysis
Hurst analysis
Detrended-fluctuation analysis
Pattern recognition

ABSTRACT

In this work, signal processing and pattern recognition techniques are combined to diagnose the severity of bearing faults. The signals were pre-processed by detrended-fluctuation analysis (DFA) and rescaled-range analysis (RSA) techniques and investigated by neural networks and principal components analysis in a total of four schemes. Three different levels of bearing fault severities together with a standard no-fault class were studied and compared. Signals were acquired from bearings working under different frequency and load conditions. An evaluation of fault recognition efficiency was performed for each combination of signal processing and pattern recognition techniques. All four schemes of classification yielded reasonably good results and are thus shown to be promising for rolling bearing fault monitoring and diagnosing.



## 1. Introdution

Mechanical vibrations can be detected by displacement, velocity and acceleration sensors. Since these signals present random characteristics, one can treat them statistically and correlate the vibration signal characteristics to the behavior of the system under consideration.

The easiness on vibration signals handling enables the application of analysis techniques for predictive maintenance of industrial equipments, avoiding losses from production shut down, which is relatively expensive.

The application of vibration analysis on predictive maintenance can be extended to fault diagnosis of various machine components. Bearings require close attention during the inspection as malfunctioning of these elements causes increased machine vibrations, which can lead to failure.

Nowadays, there are several techniques used to detect defects in bearings. The spectral and envelope analyses are two of the most commonly applied techniques in the industry [1,2]. Spectral analysis consists of using Fourier transform on vibration signals collected from bearings in work regime, aiming to identify vibration frequencies for comparison to characteristic frequencies obtained from known defective conditions. The technique of the envelope is not so different from the previous one unless the signal collected is filtered by mathematical manipulation. Other more robust techniques are also used such as time–frequency techniques [3].

* Corresponding author. Tel.: +55 85 3366 9077; fax: +55 85 3366 9969.
*E-mail addresses:* elineudo@ufc.br (E.P. de Moura), almeida@dem.ufcg.edu.br (C.R. Souto), cicero@dem.ufcg.edu.br (A.A. Silva), marcos.silva@univasf.edu.br (M.A.S. Irmão).

0888-3270/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ymssp.2010.11.021

These techniques, however, rely on professional expertise on the subject and so, of course, skilled labor and the time to perform the analysis are required. In order to be cost effective, the current trend is to employ expert systems to diagnose and correct faulty systems based on the standards established for the defects. This research is aimed to be a contribution on the subject by the evaluation of two approaches: detrended-fluctuation analysis (DFA) and Hurst method or an RS analysis.

The DF analysis was developed to differentiate local patchiness and long-range correlations in DNA sequences [4]. The presence of long-range correlations – which reflects the presence of correlated noise in the sequence – is identified by studying the detrended-fluctuations in the sequence as a function of the size of the windows through which the sequence is examined. The first aim of the Hurst method was to study the Nile river and the problems related to water storage [5]. The RS method provides a reliable measure of some statistical aspects of time series records as shown by Mandelbrot [6] and Mandelbrot and Wallis [7].

Thus, it is likely that DFA and RSA are useful tools for vibration signals diagnosis by filtering out noise contributions from unimportant sources and focusing on the correlated noise which presumably comes from the roller bearing faults.

Recently, Moura et al [8] presented a combined application of detrended-fluctuation analysis (DFA) and principal component analysis (PCA) for gearbox fault diagnosis. It was shown that the variation of the detrended-fluctuations with the size of the time window is a signature of the type of gear fault and this signature can be pointed out by pattern-classification tools.

This paper is organized as follows: Section 2 is a description of the techniques employed for producing and capturing the vibration signals; Section 3 is a detailed presentation of the motivation and mathematics behind DF and RS methods; Section 4 depicts the results obtained by applying a combination of fluctuation methods (DF and RS) with techniques of pattern recognition (principal component analysis—PCA and artificial neural network—ANN) to the vibration signals and their discussion; Section 5 contains the conclusions.

## 2. Experimetal setup

Essays were performed on a bench composed of three-phase motor (0.37 kW at 1555 rpm) elastically connected to a shaft supported by two bearings. Defects were simulated by a notch on the farthest bearing from the motor. Three different severities of notch (0.15, 0.50 and 1.00 mm wide) were introduced at the outer race of the bearing. A no-fault class was also considered.

For each class of fault severity, vibration signals were obtained from two different rotation frequencies (25 and 60 Hz) and under two load conditions (1000 and 2000 N). The motor rotation was controlled by a frequency inverter. An external load was applied on the third bearing located halfway between the other two bearings using a lever connected to a hydraulic system and recorded by a manometer (Fig. 1).

For each combination of fault severity, frequency and load, 20 signals were captured, giving a total of 320 data sets. For data acquisition, PCB Model 353B03 accelerometer placed on the upper side of the test bearing, PCB Model 480E09 load amplifier and a National Instruments Model AT-MIO-16E-10 data acquisition system were used. Each signal was acquired with 4096 data points and a sampling rate of 20 kHz.

## 3. Detrended-fluctuations (DF) and Hurst (RS) methods

Detrended-fluctuation analysis and Hurst rescaled-range analysis were used in this work. Each of them is described as follows.

### 3.1. Detrended-fluctuation analysis (DFA)

The detrended-fluctuation analysis (DFA) [4] aims to improve the evaluation of correlations in a time series by eliminating linear trends from the data. The method consists in obtaining a new integrated series initially:

$$y_i = \sum_{j=1}^{i} [x_j - \langle x \rangle] \tag{1}$$

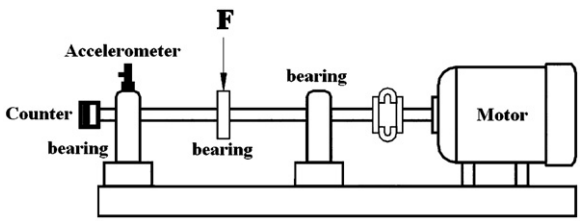


**Fig. 1.** Experimental setup of test bench.

With the average being taken over all $N$ points of the original series

$$\langle x\rangle = \frac{1}{N}\sum_{j=1}^{N} x_j \tag{2}$$

The series is broken into $N/\tau$ equal intervals of size $\tau$. A straight line is drawn through the points in each interval to obtain the local trend $h_i$. A detrended-variation function $\Delta_i$ is calculated by subtracting the value $h_i$ from the integrated data $y_i$ as follows:

$$\Delta_i = y_i - h_i \tag{3}$$

Finally, the root-mean-square fluctuation $F(\tau)$ for an interval $k$ is calculated as

$$F(\tau) = \sqrt{\frac{1}{\tau}\sum_i \Delta_i^2} \tag{4}$$

and averaged over all $N/\tau$ intervals. This process is repeated for several values of $\tau$ in order to construct the function $F(\tau)$ which, for a true fractal series, is presented in the form

$$F(\tau) \sim \tau^{\alpha} \tag{5}$$

where $\alpha$ is the scaling exponent.

### 3.2. Hurst method (RS)

The rescaled-range (RS) method was introduced by Hurst [5] as a tool for evaluating the persistence or antipersistence of a time series. The method consists of dividing the series into intervals of a given size $\tau$ and calculating the average ratio $R/S$ of the range (the difference between the maximum and minimum values of the series) to the standard deviation from each interval. The size $\tau$ is then varied and a curve of the rescaled range $R/S$ as a function of $\tau$ is obtained.

Mathematically, the RS method is defined in the following way: given an interval of size $\tau$, whose left end is located at point $i_0$, the average of the time series $x$ of the interval is calculated

$$\langle x\rangle_\tau = \frac{1}{\tau}\sum_{t=i_0}^{i_0+\tau-1} x_t \tag{6}$$

An accumulated deviation from the mean is defined as

$$X(t,\tau) = \sum_{u=1}^{t}\{x_u - \langle x\rangle_\tau\} \tag{7}$$

A range is extracted from the above operations

$$R(\tau) = \max X(t,\tau) - \min X(t,\tau) \tag{8}$$

And the corresponding standard deviation

$$S = \sqrt{\frac{1}{\tau}\sum_{t=1}^{\tau}\{x(t) - \langle x\rangle_\tau\}^2} \tag{9}$$

Finally, the rescaled range $R(\tau)/S(\tau)$ is obtained and its average is determined over all intervals.

For a surface with true fractal features, the rescaled range should satisfy the scaling form

$$RS(\tau) = \frac{R(\tau)}{S(\tau)} \sim \tau^{H}, \tag{10}$$

where $H$ is the Hurst exponent.

As can be noticed, the DF and RS methods can be interpreted as functions $F(\tau)$, which compresses the information contained in the original signal in a small number of components. In this particular study, the vibration signal comprising of 4096 points was transformed into a curve $F(\tau)$ of 37 values. Curves $F(\tau)$ represent the various conditions of severity of defects in roller bearing, which may be used in conjunction with statistical tools aimed at pattern classification.

## 4. Pattern recognition

### 4.1. Principal component analysis

Given a set of column vector $\{\mathbf{x}_i\}$ generated during Hurst analysis or detrended-fluctuation analysis, the principal component analysis (PCA) works by projecting the vectors onto the directions defined by the eigenvectors of the group-covariance matrix $\mathbf{S}$

defined as

$$\mathbf{S} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i - \mathbf{m})(\mathbf{x}_i - \mathbf{m})^{\mathbf{T}} \tag{11}$$

in which $\mathbf{m}$ is the average vector

$$\mathbf{m} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i \tag{12}$$

and $\mathbf{T}$ denotes the vector transpose.

The projection along the direction of the eigenvector corresponding to the largest eigenvalue of $S$ is the first principal component and accounts for the largest amount of variation in the original vectors. The remaining principal components are arranged in decreasing order of the corresponding eigenvalues. It follows that the projection on the second largest eigenvector related to the second largest eigenvalue gives the second principal component [9].

### 4.2. Artificial neural network

In this study, each input of the neural network is represented by a vector $\{\mathbf{x}_i\}$ with dimension 37 (one signal with 37 points) or, geometrically, by one point in a space of dimension 37, called space of inputs.

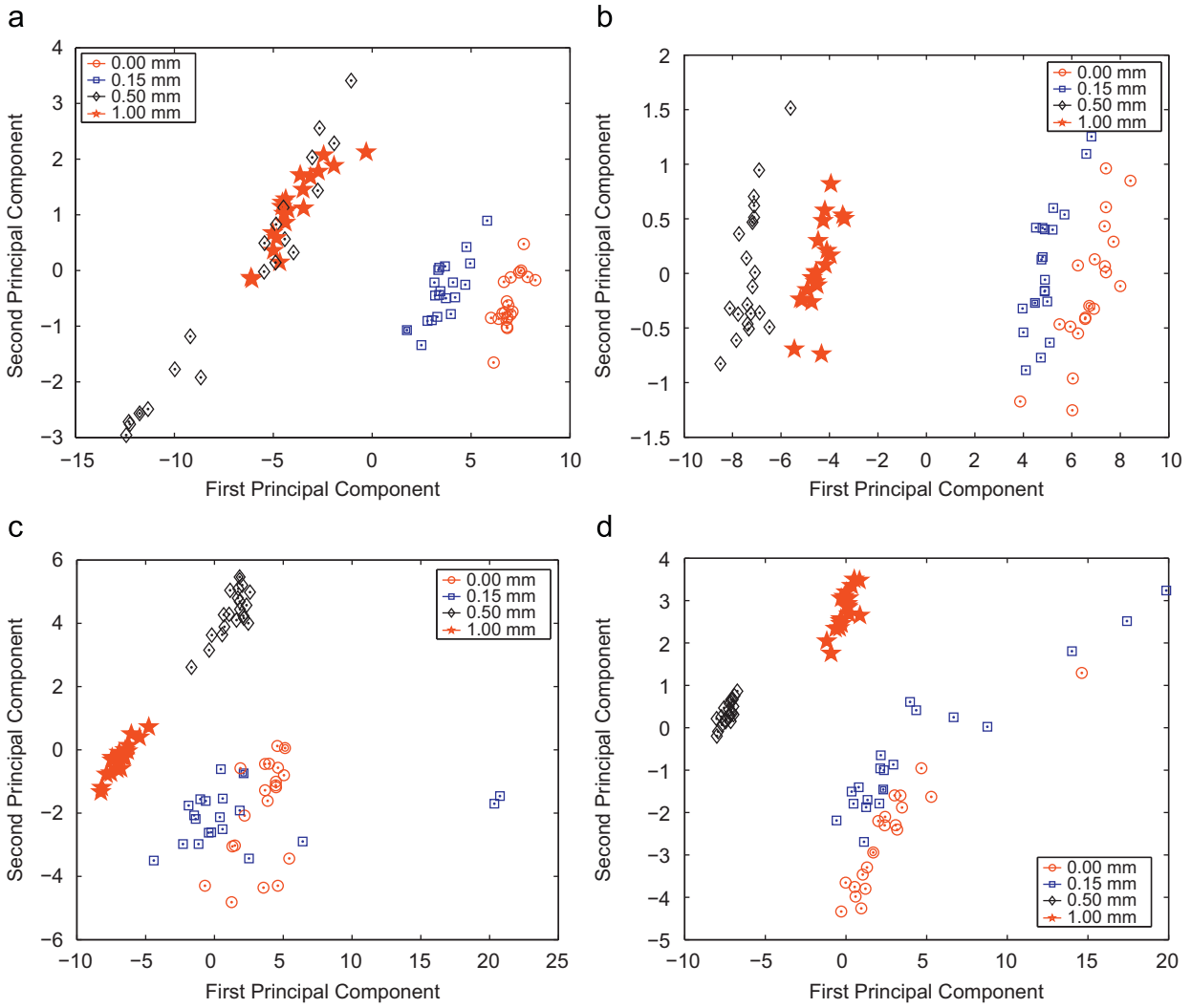


**Fig. 2.** Projection of the vector corresponding to DFA curves of vibration signals from four frequency and load conditions along the plane defined by the first two principal components: (a) 25 Hz and 1000 N, (b) 25 Hz and 2000 N, (c) 60 Hz and 1000 N and (d) 60 Hz and 2000 N.

The nonlinear classifiers were implemented by a neural network with only one hidden layer totally connected. For the output layer, one discriminator should be determined to classify each of the classes. This is obtained through the neural network with one output neuron for each class, i.e., in this work, a total of four neurons of the output layer were used. The number of neurons of the hidden layer was calculated as an average from number of class and number of attributes (size of input vector). So, twenty neurons were used in the hidden layer. These best discriminators were determined by the adjustment of synaptic weights during training based on standard output with the backpropagation learning rule [10,11].

The training-test pairs were made up by random choice. The training was carried out at 1000 epochs with an alpha learning rate and a moment equal to 0.1 and 0.75, respectively [10]. Logistic sigmoid activation functions were used for all the neurons. The output neuron of the last layer that was greater than zero corresponds to the selected class.

## 5. Results and discussion

Two classifiers were used for this work. The first one is based on PCA. With this classifier, the vector $y_i$ is assigned to the class whose average vector is closer. The second one was implemented by an artificial neural network, as described in Section 4.2.

With the purpose to compare the efficiency of PCA to the one of an ANN for bearing fault diagnosis, vibration signals acquired from bearings with only one class of fault, but different severities were evaluated with respect to frequency and load conditions.

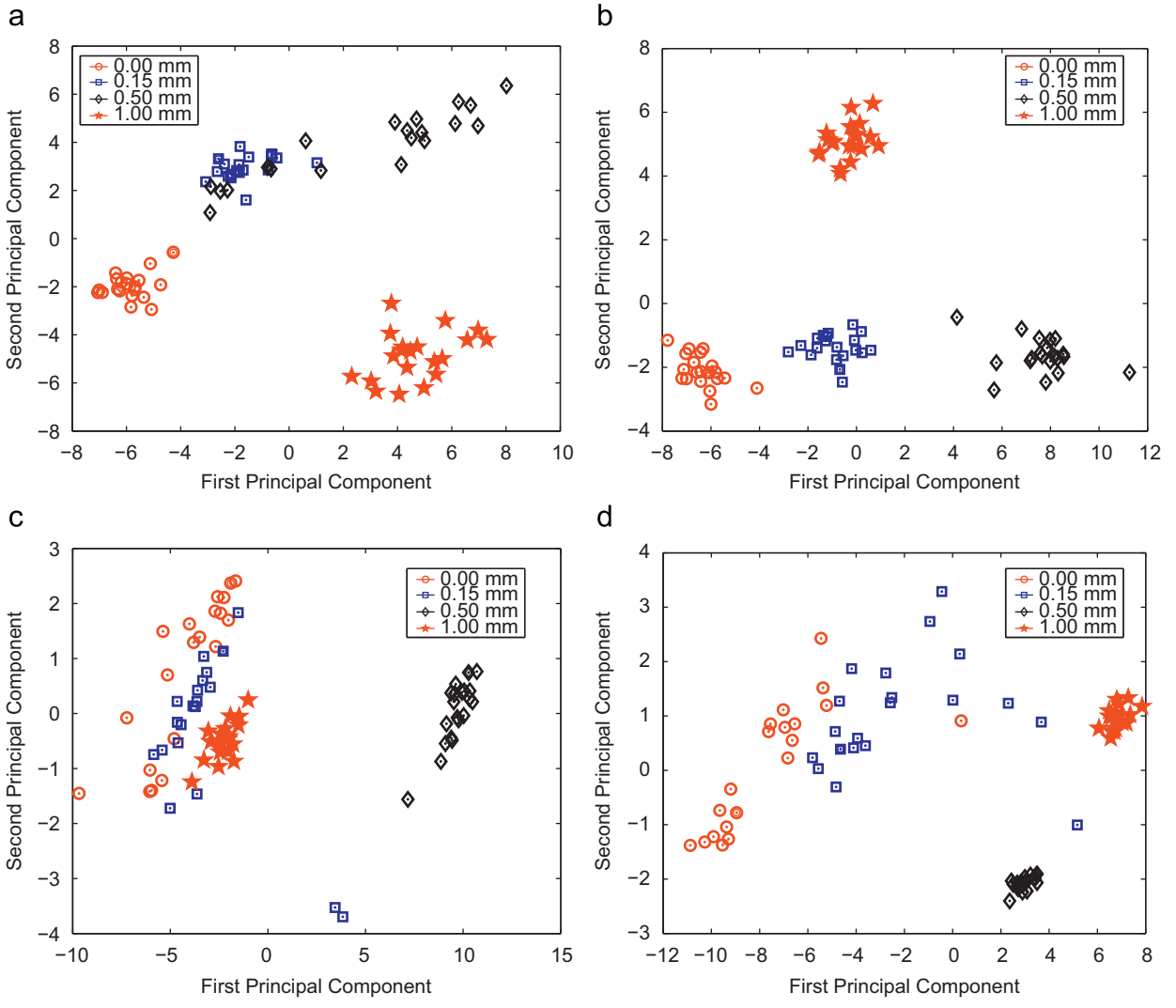


**Fig. 3.** Projection of the vector corresponding to RSA curves of vibration signals from four frequency and load conditions along the plane defined by the first two principal components: (a) 25 Hz and 1000 N, (b) 25 Hz and 2000 N, (c) 60 Hz and 1000 N and (d) 60 Hz and 2000 N.

The results obtained by applying PCA to the vectors yielded by DFA and RSA on vibration signals acquired from bearings working under different frequencies and load conditions are portrayed in Figs. 2 and 3, respectively.

As introduced in Section 4.1, Figs. 2 and 3 were obtained by the projection of vectors along the direction of the eigenvectors (principal components) corresponding to two largest eigenvalues of covariance matrix of input vectors [9].

As described earlier, during classification based on PCA, the vector $y_i$ is assigned to the class, whose average vector is closer. By calculating the average percentages of fault severities which are correctly classified (average success rate), it is possible to make a quantitative evaluation of the classification approach.

For PCA, with $n$ possible classes, the fully transformed vectors have $(n-1)$ relevant components [10]. Thus, only the first three principal components were used for the nearest-class-mean rule.

Tables 1 and 2 show the average success rate from PCA of the DFA and RSA curves of vibration signals acquired from bearings under load of 1000 and 2000 N, respectively. It is noticeable that the increase in load produces a better classification of the severity of the defect for both RSA and DFA. This was already expected from the PCA plots (Figs. 2(b) and 3(b)). It can also be verified that, in a general way, the RSA presents better performance than DFA.

**Table 1**
Percentage of vibration signals acquired from bearings working under load of 1000 N, which were correctly classified by applying the nearest-class-mean rule to the results of a combined approach of PCA and DFA/RSA.

| | DFA/25 Hz | DFA/60 Hz | RSA/25 Hz | RSA/60 Hz |
|---|---|---|---|---|
| Classified as 0.00 mm | 100 | 70 | 100 | 85 |
| Classified as 0.15 mm | 95 | 80 | 100 | 35 |
| Classified as 0.50 mm | 50 | 100 | 70 | 100 |
| Classified as 1.00 mm | 90 | 100 | 100 | 90 |
| **Average success rate** | **83.75** | **87.50** | **92.50** | **77.50** |

**Table 2**
Percentage of vibration signals acquired from bearings working under load of 2000 N, which were correctly classified by applying the nearest-class-mean rule to the results of a combined approach of PCA and DFA/RSA.

| | DFA/25 Hz | DFA/60 Hz | RSA/25 Hz | RSA/60 Hz |
|---|---|---|---|---|
| Classified as 0.00 mm | 90 | 85 | 100 | 95 |
| Classified as 0.15 mm | 90 | 35 | 100 | 70 |
| Classified as 0.50 mm | 95 | 100 | 100 | 100 |
| Classified as 1.00 mm | 100 | 100 | 100 | 100 |
| **Average success rate** | **93.75** | **80.00** | **100.00** | **91.25** |

**Table 3**
Average percentage calculated over 100 training sets derived from RSA and DFA of vibration signals acquired from bearings working under load of 1000 N, which were correctly classified by the neural network.

| | DFA/25 Hz | DFA/60 Hz | RSA/25 Hz | RSA/60 Hz |
|---|---|---|---|---|
| Classified as 0.00 mm | 100 | 99.56 | 100 | 100 |
| Classified as 0.15 mm | 100 | 100 | 100 | 100 |
| Classified as 0.50 mm | 100 | 100 | 100 | 100 |
| Classified as 1.00 mm | 100 | 100 | 100 | 100 |
| **Average success rate** | **100** | **99.89** | **100** | **100** |

**Table 4**
Average percentage calculated over 100 training sets derived from RSA and DFA of vibration signals acquired from bearings working under load of 2000 N, which were correctly classified by the neural network.

| | DFA/25 Hz | DFA/60 Hz | RSA/25 Hz | RSA/60 Hz |
|---|---|---|---|---|
| Classified as 0.00 mm | 100 | 98.81 | 100 | 100 |
| Classified as 0.15 mm | 100 | 99.50 | 100 | 100 |
| Classified as 0.50 mm | 100 | 100 | 100 | 100 |
| Classified as 1.00 mm | 100 | 100 | 100 | 100 |
| **Average success rate** | **100** | **99.58** | **100** | **100** |

**Table 5**
Average percentage calculated over 100 test sets derived from RSA and DFA of the vibration signals acquired from bearings working under load of 1000 N, which were correctly classified by the neural network.

| | DFA/25 Hz | DFA/60 Hz | RSA/25 Hz | RSA/60 Hz |
|---|---|---|---|---|
| Classified as 0.00 mm | 100 | 77.75 | 100 | 99.00 |
| Classified as 0.15 mm | 99.5 | 80.25 | 100 | 98.50 |
| Classified as 0.50 mm | 100 | 99.75 | 100 | 100 |
| Classified as 1.00 mm | 100 | 100 | 100 | 100 |
| **Average success rate** | **99.88** | **89.44** | **100** | **99.38** |

**Table 6**
Average percentage calculated over 100 test sets derived from RS and DF analyses of the vibration signals acquired from bearings working under the load of 2000 N, which are correctly classified by the neural network.

| | DFA/25 Hz | DFA/60 Hz | RSA/25 Hz | RSA/60 Hz |
|---|---|---|---|---|
| Classified as 0.00 mm | 100 | 90 | 100 | 94.25 |
| Classified as 0.15 mm | 99.75 | 84.75 | 100 | 90.75 |
| Classified as 0.50 mm | 100 | 100 | 100 | 100 |
| Classified as 1.00 mm | 100 | 100 | 100 | 100 |
| **Average success rate** | **99.94** | **93.69** | **100** | **96.25** |

A second classifier was implemented by an ANN as described in Section 4.2. Supervised learning was employed during this classifier training. In order to make a quantitative quality assessment of the discrimination obtained by neural network, the input data (vectors yielded from DFA and RSA) were divided into a training set and a testing set for a given group of frequency and load conditions.

For each work condition, 80% of vectors (64 vectors) were used for training and the remaining 20% vectors (16 vectors) for testing. The average success rate was calculated over 100 random choices from training and testing sets.

Tables 3 and 4 show the results obtained from the training data and Tables 5 and 6 show the results obtained from the test data. It can be seen that in Tables 3–6 the overall result for the training data is better than the overall result for the test data. This is so because it is easier to classify a data that was supplied to the network during the training process.

The sets derived from RSA and DFA of vibration signals were also divided according to the load applied to bearings. Tables 3 and 5 show the average percentage obtained for sets derived from vibration signals acquired from bearings working under load of 1000 N. The results obtained from vibration signal acquired from bearings under 2000 N are presented in Tables 4 and 6.

Comparing the average success rate as a function of external applied load (Table 3 versus Table 4, and Table 5 versus Table 6), it can be noticed that it is easier to distinguish the four classes of fault severities for bearing working under higher load. Improvement on classification due to more intense load was previously observed in geared-systems [8].

Tables 3 and 4 also show that the result of classification obtained from RS curves is better than the one obtained from DF curves for a same work condition. It was also perceived that the neural network classifier produces better results than the PCA in all cases. However, the computational cost of neural network is disadvantageous with respect to the classifier based on PCA.

## 6. Conclusions

In this investigation, two pattern recognition techniques (PCA and ANN) combined with two fluctuation analysis techniques (DF and RS) were used in an attempt to discriminate fault severities from vibration signals acquired from roller bearings under different conditions of frequencies and load.

The combinations of techniques employed made it possible to recognize patterns and distinguish vibration signals acquired from bearings with four different fault severities. The classifier based on principal component analysis presents performance slightly inferior to the one implemented by neural network, but with an inferior computational cost. Again, it should be considered that the principal component analysis is an unsupervised method, whereas the classifier implemented by neural network was trained by supervised learning.

Besides, the method employed presents the extra advantage of a significant data compression. Therefore, from the obtained results, it can be concluded that these techniques seem to be promising for monitoring and diagnosing bearing fault.

## Acknowledgments

The authors would like to kindly thank the financial support of the Brazilian agencies CAPES and CNPq.

## References

[1] R.B. Randall, A new method of modeling gear faults, Journal of Mechanical Design 104 (1982) 259–267.
[2] P.D. McFadden, J.D. Smith, Model for the vibration produced by single point defect in a rolling element bearing, Journal of Sound and Vibration 96 (1984) 69–81.
[3] L.R. Padovese, Hybrid time–frequency methods for nonstationary mechanical signals, Mechanical Systems and Signal Processing 18 (2004) 1047–1064.
[4] C.K. Peng, V. Buldyrev, S. Havlin, M. Simmons, H.E. Stanley, A.L. Goldberger, Mosaic organization of DNA nucleotides, Physical Review E 49 (1994) 1685–1689.
[5] H.E. Hurst, Long-term storage capacity of reservoirs, Transactions of the American Society of Civil Engineers 116 (1951) 770–799.
[6] B.B. Mandelbrot, The Fractal Geometry of Nature, W.H. Freeman & Company, New York, 1983.
[7] B.B. Mandelbrot, J.R. Wallis, Some long-run properties of geophysical records, Water Resources Research 5 (1969) 321–340.
[8] E.P. de Moura, A.P. Vieira, M.A.S. Irmão, A.A. Silva, Applications of detrended-fluctuation analysis to gearbox fault diagnosis, Mechanical Systems and Signal Processing 23 (2009) 682–689.
[9] A.R. Webb, Statistical Pattern Recognition, second ed., Wiley, West Sussex, 2002.
[10] S. Haykin, Neural Networks, a Comprehensive Foundation, Macmillian College Publishing, New York, 1994.
[11] P.D. Wasserman, Neural Computing Theory and Practice, Van Nostrand Reinthold, New York, 1989.