Expert Systems with Applications 38 (2011) 1876–1886

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Fault diagnosis of ball bearings using machine learning methods

P.K. Kankar, Satish C. Sharma, S.P. Harsha *

Mechanical and Industrial Engineering Department, Indian Institute of Technology Roorkee, India

ARTICLE INFO

Keywords:
Fault diagnosis
Artificial neural network
Support vector machine

ABSTRACT

Ball bearings faults are one of the main causes of breakdown of rotating machines. Thus, detection and diagnosis of mechanical faults in ball bearings is very crucial for the reliable operation. This study is focused on fault diagnosis of ball bearings using artificial neural network (ANN) and support vector machine (SVM). A test rig of high speed rotor supported on rolling bearings is used. The vibration response are obtained and analyzed for the various defects of ball bearings. The specific defects are considered as crack in outer race, inner race with rough surface and corrosion pitting in balls. Statistical methods are used to extract features and to reduce the dimensionality of original vibration features. A comparative experimental study of the effectiveness of ANN and SVM is carried out. The results show that the machine learning algorithms mentioned above can be used for automated diagnosis of bearing faults. It is also observed that the severe (chaotic) vibrations occur under bearings with rough inner race surface and ball with corrosion pitting.

1. Introduction

Condition monitoring of rotating machinery helps in early detection of faults and anticipation of problems in time, so as to prevent complete failure. Bearing vibration can generate noise and degrade the quality of a product line. Severe vibrations of bearings can even cause the entire system to function incorrectly and that results in downtime for the system and economic loss to the customer. Rolling bearings defects may be categorized as point or local defects and distributed defects. The vibrations are generated by geometrical imperfections on the individual bearing components and these imperfections are caused by irregularities during the manufacturing process as well as wear and tear. The various distributed defects are surface roughness, waviness, misaligned races, and off-size rolling elements. The local defects include cracks, corrosion pitting, brinnelling and spalls on the rolling surfaces. McFadden and Smith (1985, 1984) have developed the models for vibration produced by a single and multiple point defects on the inner race of the rolling element bearing under radial load based on high-frequency resonance technique. Prabhakar, Mohanty, and Sekhar (2002) have considered single and multiple point defects on inner race, outer race and the combination faults and used discrete wavelet transform (DWT) to detect these faults on bearings. Kankar, Harsha, Pradeep, and Sharma Satish (2009) have applied response surface methodology (RSM), to investigate the effects of various defects on the non-linear vibrations of rotor bearing system.

Various artificial intelligent (AI) techniques such as hidden Markov models (HMM) (Li, Wu, He, & Fulei, 2005), artificial neural networks (ANN) (Vyas & Satishkumar, 2001) and support vector machines (SVM) (Widodo & Yang, 2007; Yuan & Chu, 2007) have been used in the fault diagnosis of machines. Zhitong, Jiazhong, Hongpingn, Guoguang, and Ritchie (2003) have carried out fault detection of induction motor using SVM technique for detecting broken rotor bars. In their experiment, induction motor was experimented with no fault, one broken bar, two broken bars and three broken bars. They used stator current to obtain the signal and calculated the frequency spectrum for fault detection. Samanta (2004) has compared the performance of gear fault detection using ANN and SVM. The time-domain vibration signal of a rotating machine with normal and defective gears were processed for feature extraction. The results compare the effectiveness of both types of classifiers without and with GA-based selection of features and the classifier parameters. The main difference between ANNs and SVMs is in the principle of risk minimization (RM) (Gunn, 1998). In case of SVMs, structural risk minimization (SRM) principle is used for minimizing an upper bound on the expected risk whereas in ANNs, traditional empirical risk minimization (ERM) is used for minimizing the error on the training data.

This paper is mainly focused on bearing fault classification using two machine learning methods ANN and SVM as both can work with non-linear classifications. Vibration data are collected by piezoelectric accelerometers as a time domain signals for the healthy bearing and bearing with different faults. Defects are con-

* Corresponding author. Tel.: +91 9917489849/01332 286602; fax: +91 1332 285665.
E-mail address: spharsha@gmail.com (S.P. Harsha).

0957-4174/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.07.119

Nomenclature

$\lVert w\rVert^{-2}$	geometrical margin
C	error penalty
ξ_i	the distance between the margin and the examples x_i that lying on the wrong side of the margin
B	bias or threshold
λ_i	Lagrange multipliers
$W(\lambda)$	Lagrange function
x_{pi}	ith input of the pth input vector
net^h_{pi}	net input to hidden layer
o^h_{pj}	output of hidden layer
net^o_{pk}	net input to output layer
o^o_{pk}	output of output layer
E_p	sum of squares error
Δ^o_p	change in weight with respect to weight change
θ^h_j	bias for hidden layer
θ^o_j	bias for output layer
T_{pk}	desired output
w^o_{jk}	synaptic weight between hidden and output layer
T	number of iterative step
TP rate	true positive rate is the number of correctly classified fault divided by the total number of instances for that fault
FP rate	false positive rate is the number of incorrectly classified fault divided by the total number of instances other than considered fault

sidered as crack in outer race, inner race with rough surface and corrosion pitting in balls. The signals obtained are processed for machine condition diagnosis as shown in the flow chart Fig. 1. Features are extracted from the time domain signal by statistical method. These features are fed to a supervised attribute filter that can be used to select features. Selected features with the known output are used for training and testing of ANN and SVM.

2. Support vector machine

Support vector machine (SVM) is a supervised machine learning method based on the statistical learning theory. It is a useful method for classification and regression in small-sample cases such as fault diagnosis. Pattern recognition and classification using SVM is described in brief (Cristianini & Shawe-Taylor, 2000).

A simple case of two classes is considered, which can be separated by a linear classifier. Fig. 2 shows triangles and squares stand for these two classes of sample points. Hyper plane H is one of the separation planes that separate two classes. H_1 and H_2 (shown by dashed lines) are the planes those are parallel to H and pass through the sample points closest to H in these two classes. Margin is the distance between H_1 and H_2. The SVM tries to place a linear boundary between the two different classes H_1 and H_2, and orientate it in such a way that the margin is maximized, which results in least generalization error. The nearest data points that used to define the margin are called support vectors.

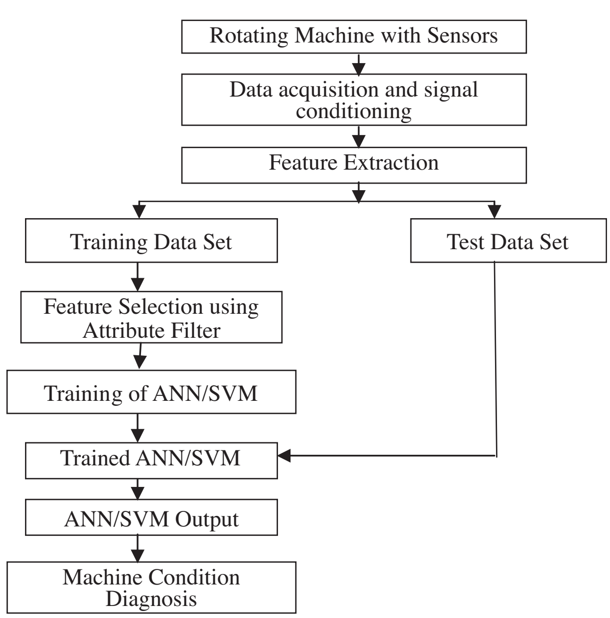

Fig. 1. Flow chart of bearing health diagnosis.

This is implemented by reducing it to a convex optimization problem: minimizing a quadratic function under linear inequality constraints (Cristianini & Shawe-Taylor, 2000). Consider a training sample set $\{(x_i, y_i)\}$; $i = 1$ to N, where N is total number of samples. It is wished to determine a separation plane among all linear separation planes that separates input samples into two classes with the smallest generalization error. Let us assume the samples can be classified into two classes namely triangle class and square class. Labels $y_i = -1$ and $y_i = +1$ are associated with triangle class and square class, respectively. Slack variables are considered ($\xi_i \geqslant 0$) for non-separable data. The hyperplane $f(x) = 0$ that separates the given data can be obtained as a solution to the following optimization problem

$$\text{Minimize} \quad \frac{1}{2}\lVert w\rVert^2 + C\sum_{i=1}^{N}\xi_i \tag{1}$$

$$\text{Subject to} \quad \begin{cases} y_i(w^T x_i + b) \geqslant 1 - \xi_i \\ \xi_i \geqslant 0, \quad i = 1, 2, \ldots, N \end{cases} \tag{2}$$

where C is a constant representing error penalty. Rewriting the above optimization problem in terms of Lagrange multipliers, leads to the problem

$$\text{Maximize} \quad W(\lambda) = \sum_{i=1}^{N}\lambda_i - \frac{1}{2}\sum_{i,j=1}^{N} y_i y_j \lambda_i \lambda_j (x_i \cdot x_j) \tag{3}$$

$$\text{Subject to} \quad \begin{cases} 0 \leqslant \lambda_i \leqslant C \\ \sum_{i=1}^{N}\lambda_i y_i = 0, \quad i = 1, 2, \ldots, N \end{cases} \tag{4}$$

The sequential minimal optimization (SMO) algorithm gives an efficient way of solving the dual problem arising from the derivation of the SVM. SMO decomposes the overall QP problem into QP sub-problems.

3. Artificial neural network

Artificial neural network (ANN) is an interconnected group of artificial neurons. These neurons use a mathematical or computational model for information processing. ANN is an adaptive system that changes its structure based on information that flows through the network (Zurada, 1999).

A single neuron consists of synapses, adder and activation function. Bias is an external parameter of neural network. Model of a neuron shown in Fig. 3 can be represented by following mathematical model

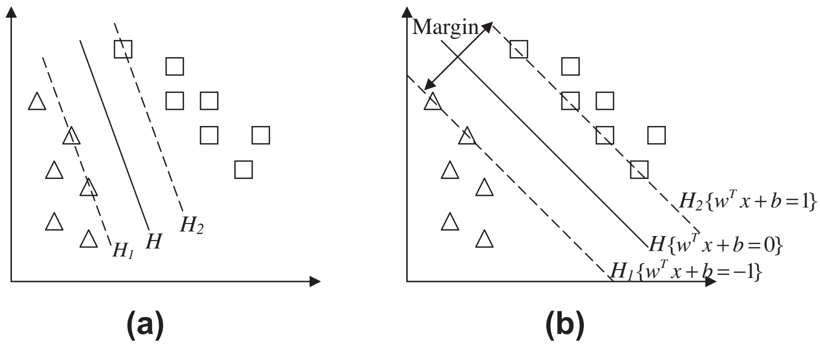

Fig. 2. Hyperplane classifying two classes: (a) small margin and (b) large margin.

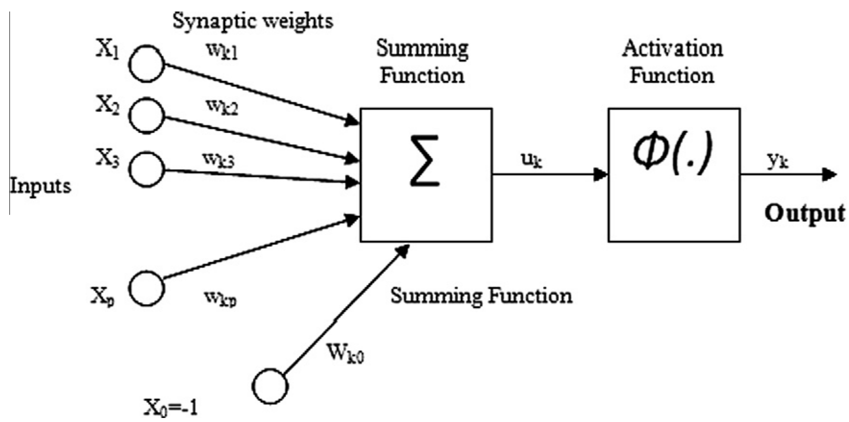

Fig. 3. Model of a single non-linear neuron.

$$y_k = \phi\left(\sum_{i=1}^{p} w_{ki}x_i + w_{k0}\right) \tag{5}$$

Input vector comprising of 'p' inputs multiplied by their respective synaptic weights, and sum off all weighted inputs. A threshold (bias) is used with constant input. Activation function converts output into a limited range output.

3.1. Back propagation (BP) algorithm

Intelligence of neural network lies in the weights between neurons. Back propagation (BP) algorithm as shown in Fig. 4, is most extensively used learning algorithm (Simon, 2005; Zurada, 1999). BP algorithm consists of two passes – forward pass and backward pass.

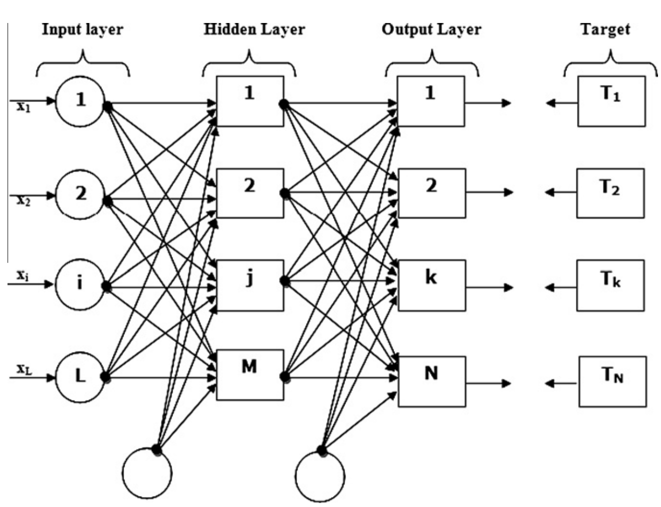

Fig. 4. Back propagation algorithm in multilayer neural network.

(a) Forward pass:
The pth input vector

$$x_p = (x_{p1} \cdots x_{pi} \cdots x_{pL}) \tag{6}$$

Net input to the hidden layer

$$net^h_{pi} = \sum_{i=1}^{L} w^h_{ij}x_{pi} + \theta^h_j \tag{7}$$

Net input of hidden layer is passed through the activation function of hidden layer. So, output obtained from hidden layer

$$o^h_{pj} = f^h_i(net^h_{pj}) \tag{8}$$

Output of the hidden layer becomes the input to the next hidden layer or output layer. Different synaptic weights are applied to each output of hidden layer which with bias term constitute the input to the output layer

$$net^o_{pk} = \sum_{j=1}^{M} w^o_{jk}o^h_{pi} + \theta^o_j \tag{9}$$

After obtaining the net input to the out layer activation function of output layer is applied. Activation function for hidden layer may be different from output layer. Output of the output layer

$$o^o_{pk} = f^h_k(net^o_{pk}) \tag{10}$$

Thus we obtain the output based on the current synaptic weights. But initially the synaptic weights are random. So an error signal is generated based on the output obtained ad the desired output. Most commonly the error signal is taken as mean square error. So mean square error signal

$$E_p = \frac{1}{2}\sum_{k=1}^{N}(T_{pk} - o^o_{pk})^2 \tag{11}$$

(b) Backward pass:
According to error signal changes in the weights are made. Change of an output layer weight

$$\Delta^o_p w = -\frac{\partial E_p}{\partial w^o_{jk}} = \delta^o_{pk} o^h_{pi} \tag{12}$$

where $\delta^o_{pk} = (T_{pk} - o^o_{pk})$.

Similarly changes in synaptic weights of hidden layer can be evaluated (Simon, 2005). As this process of weight change is iterative so a learning rule is to be applied to make process smooth. Gradient decent learning rule is widely used. Using modified gradient decent learning rule with momentum

$$w^o_{jk}(t+1) = w^o_{jk}(t) - \eta\Delta^o_p w + \alpha w^o_{jk}(t-1) \tag{13}$$

The terms 'η' and 'α' are introduced to make the learning process smooth and to ensure that the weight changes take place in the same direction.

4. Experimental setup and data acquisition

Experimental tests are carried on a test rig to generate training and test data. The rig is connected to a data acquisition system through proper instrumentation. A variety of faults are simulated on the bearing at various speeds. Various parameters of bearing used for the study are listed in Table 1. Accelerometers are used for picking up the vibration signals from various stations on the rig. Signatures for healthy bearings operation establish the baseline data. This baseline data can then be used for comparison with signatures obtained under faulty conditions. After collecting the vibration signals, salient features are extracted and compiled to form a feature vector which is fed to ANN/SVM to train it.

A variety of faults are simulated on the rig at various speed and using no loader, one loader and two loaders. Training/test data are generated using bearing with various faults and without fault. The following faults are introduced in the bearing:

(a) Outer race with crack and rough surface (Fig. 5).
(b) Inner race with rough surface (Fig. 6).
(c) Ball with corrosion pitting (Fig. 7).
(d) Combination of above faults.

The following cases are considered for acquiring training data from the data acquisition system:

(a) Healthy bearings (HB).
(b) Bearing with outer race crack (BORC).
(c) Bearing with rough inner race surface (BRIR).
(d) Ball with corrosion pitting (BCP).
(e) Combined bearing component defects (CBD).

Table 1
Parameters of bearing used for experiment.

Parameter	Value
Bearing specification	6205
Outer race diameter	52 mm
Inner race diameter	25 mm
Ball diameter	7.94 mm
Ball number	10
Contact angle	0°
Clearance	C3 Type

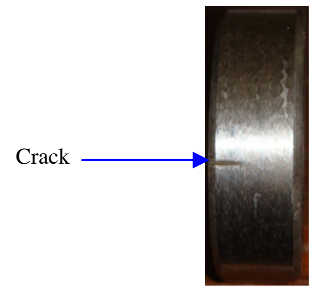

Fig. 5. Outer race with crack.

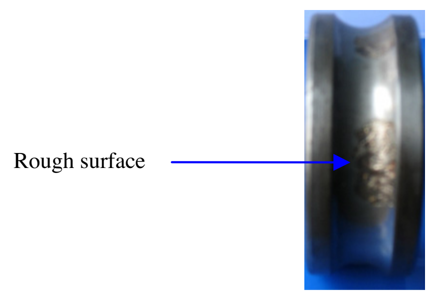

Fig. 6. Inner race with rough surface.

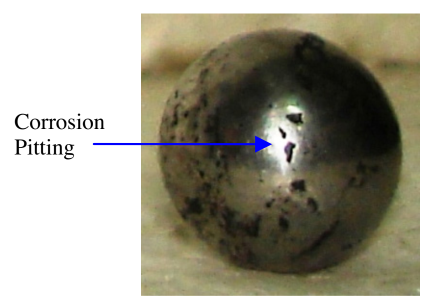

Fig. 7. Ball with corrosion pitting.

Time responses are obtained at various speeds and with different load conditions with considerations of all cases in a phased manner. The time responses at 1000, 1500 and 2000 rpm with no loader and two loader conditions are shown in Figs. 8–11, respectively.

5. ANN/SVM training and testing

5.1. Feature extraction and selection

A wide set of features are calculated from the vibration signals using statistics. These statistical features are explained below.

(a) **Range:** Range refers to the difference between maximum and minimum value of a signal.
(b) **Mean value:** Average value of a signal is termed as mean value.
(c) **Standard deviation:** Standard deviation is measure of energy content in the vibration signal

$$\text{Standard deviation} = \sqrt{\frac{n\sum x^2 - (\sum x)^2}{n(n-1)}} \tag{14}$$

(d) **Skewness:** Skewness is a measure of symmetry, or more precisely, the lack of symmetry

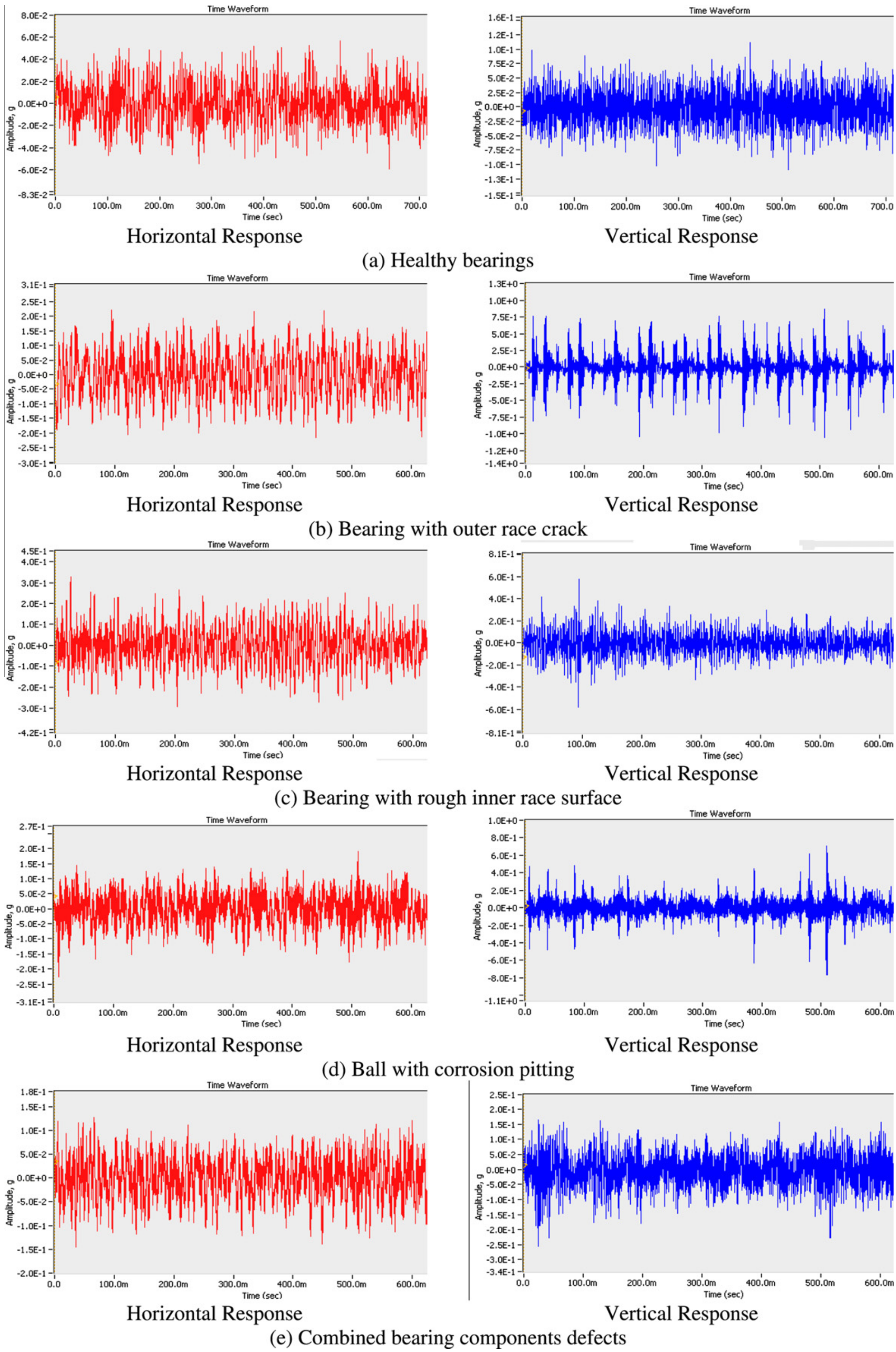

Fig. 8. Vibration signals for various bearing conditions at 1000 rpm with no loader.

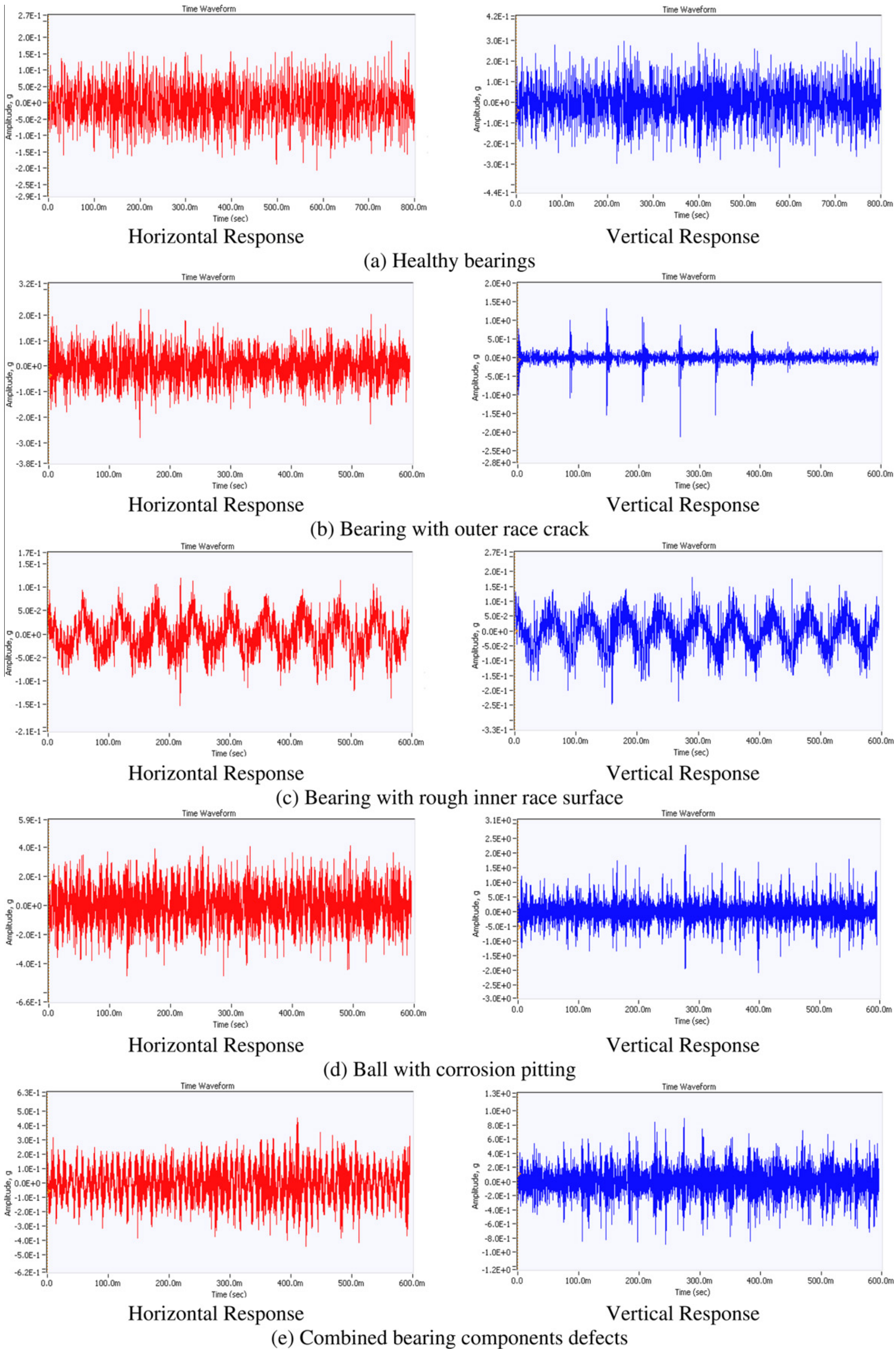

Fig. 9. Vibration signals for various bearing conditions at 2000 rpm with no loader.

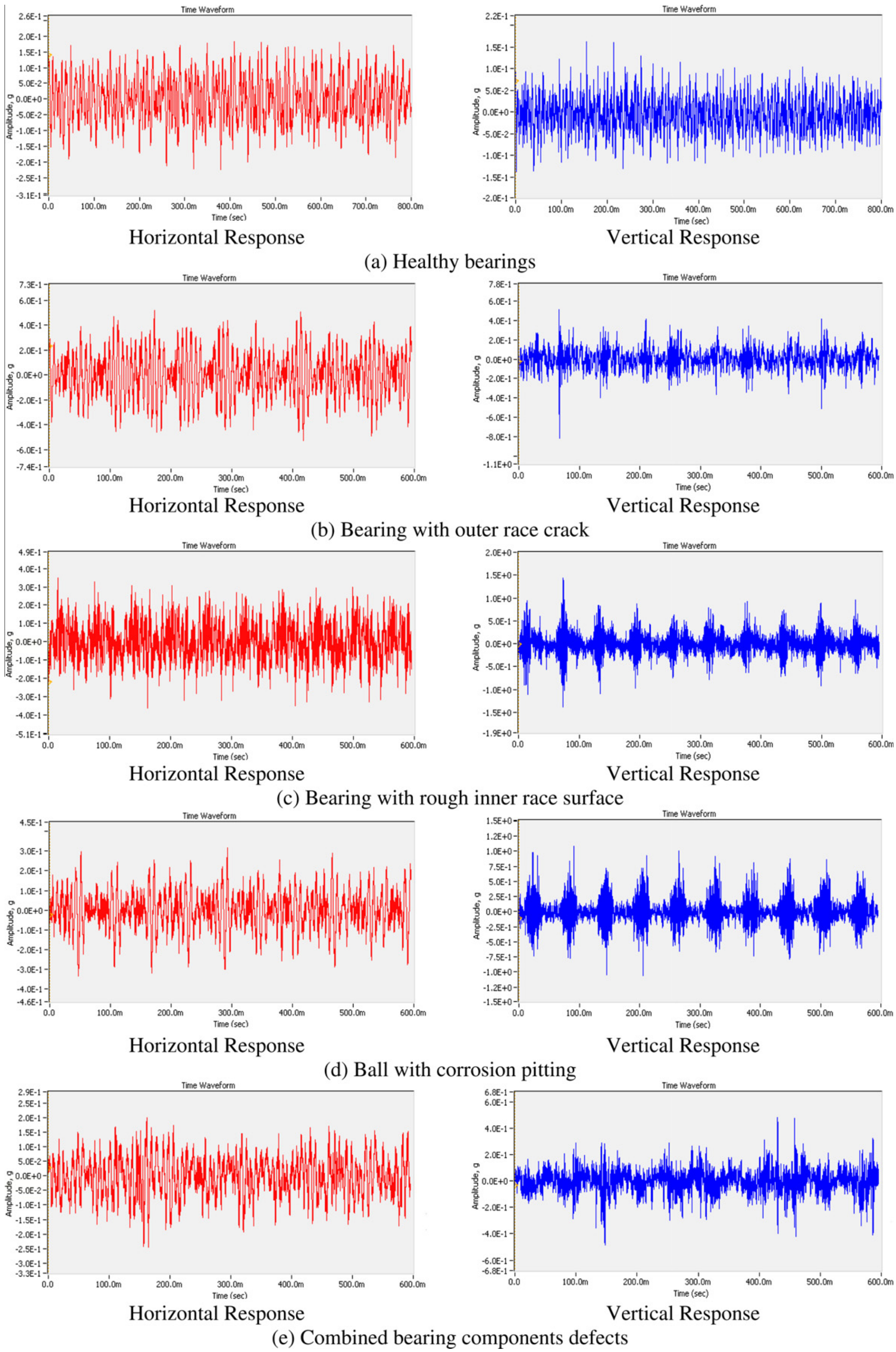

Fig. 10. Vibration signals for various bearing conditions at 1000 rpm with two loader.

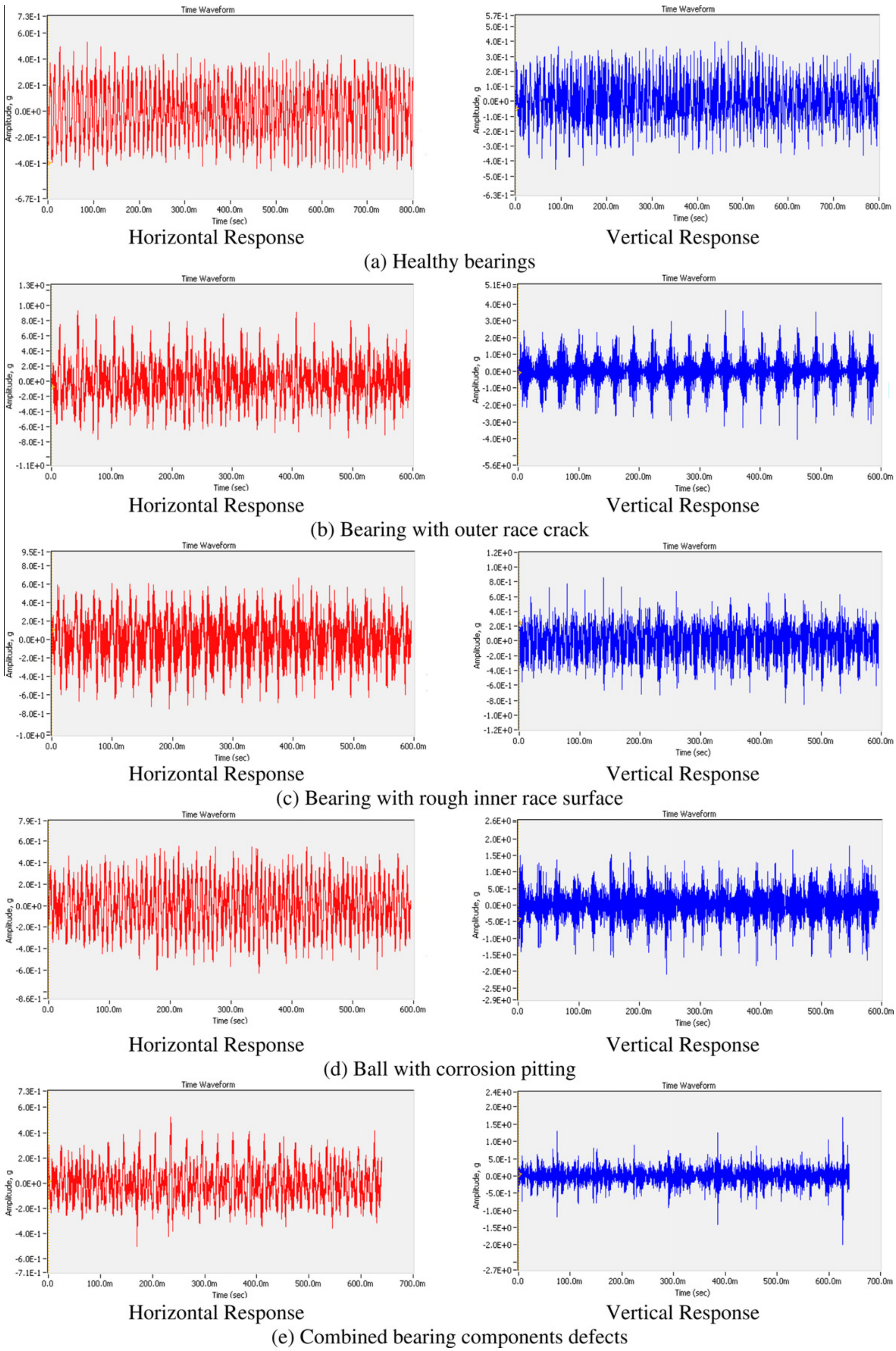

Fig. 11. Vibration signals for various bearing conditions at 2000 rpm with two loader.

$$\text{Skewness} = \frac{n}{(n-1)(n-2)} \sum \left(\frac{x_i - \bar{x}}{s} \right)^3 \quad (15)$$

(e) **Kurtosis:** Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution.

$$\text{Kurtosis} = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left(\frac{x_i - \bar{x}}{s} \right)^4 \right] - 3\frac{(n-1)^2}{(n-2)(n-3)} \quad (16)$$

(f) **Crest factor:** The crest factor is equal to the peak amplitude of a waveform divided by the RMS value.
Where n is the sample size and s is the standard deviation.

A sample training vector is shown in Table 2. Six statistical features are used each for horizontal and vertical response. Total 73 instances and 14 features are used for the study including statistical features, speed and number of loader used. To select the most appropriate features for making decisions by machine learning algorithms, these features are fed to a supervised attribute filter of WEKA software. Attribute filter uses a density-based cluster to generate cluster membership values; filtered instances are composed of these values plus the class attribute (if set in the input data) (Witten Ian, 2005).

5.2. Training and testing

Features selected from attribute filter with the known output are used for training and testing of ANN and SVM. The numbers of instances present in the input data are 73, instances are the number of data points in the input. The numbers of attributes present are 15 they include the statistical features as well as speed and number of loader.

In a multi-class prediction, the result on a test set is often displayed as a two dimensional confusion matrix with a row and column for each class. Each matrix element shows the number of test examples for which the actual class is the row and the predicted class is the column. Results correspond to large numbers down the main diagonal and small, ideally zero, off-diagonal elements gives accurate prediction.

After selecting fault as an attribute for class, classification is started and the classifier output consists of confusion matrix, detailed accuracy by class and evaluation of the success of the numeric prediction.

6. Results

ANN/SVM training and classification of faults is carried out using WEKA software. Training vectors are already compiled and are put as an input. The defects considered in the study are classified using ANN/SVM as shown in Table 3. Total 73 cases are considered for testing, which have 15, 14, 15, 14, 15 cases of ball with corrosion pitting, bearing with rough inner race surface, combined bearing component defects, healthy bearings and bearing with outer race crack respectively. From Table 3, we infer that ANN has correctly predicted 13, 10, 5, 14 and 10 cases, while SVM has classified 14, 10, 6, 14 and 10 cases correctly for ball with corrosion pitting, bearing with rough inner race surface, combined bearing component defects, healthy bearings and bearing with outer race cracks of bearing, respectively.

The vibration response at rotor speed 1000 and 2000 rpm with different defects conditions without loader is shown in Figs. 8 and 9, respectively. The five different bearing conditions are used as (a)

Table 2
Sample input vector for ANN/SVM.

Features														
1	2	3	4	5	6	7	8	9	10	11	12	13	14	Class
Amplitudes of features														
0.050998	−2.40E−05	0.005026	−0.019	0.419	4.898499	0.152438	1.98E−05	0.009818	−0.059	5.439	7.644003	1000	0	BCP
0.065839	−7.26E−05	0.006524	−0.042	0.429	4.715232	0.310888	−4.01E−04	0.016188	0.007	5.916	10.27636	1500	0	BCP
0.096399	−1.29E−05	0.011106	−0.025	0.078	4.314111	0.516477	−3.66E−04	0.030425	0.012	4.277	7.462858	2000	0	BCP
0.039209	−1.55E−04	0.004487	0.165	0.323	4.738265	0.038363	−3.39E−04	0.003414	0.047	0.312	5.893028	1000	0	BRIR
0.066436	−4.38E−05	0.008023	−0.038	0.066	4.081314	0.115677	−3.25E−04	0.009022	−0.054	0.68	6.381561	1500	0	BRIR
0.092168	−6.50E−05	0.01226	−0.061	−0.212	3.801125	0.128522	−3.40E−04	0.011543	−0.154	0.749	5.893816	2000	0	BRIR
0.036292	−4.03E−04	0.004529	−0.071	−0.189	3.708944	0.072709	−3.54E−04	0.006545	−0.101	0.125	4.220662	1000	1	CBD
0.067394	−4.09E−04	0.006937	−0.048	0.314	4.943423	0.173509	−4.49E−05	0.012398	−0.005	1.215	5.885301	1500	1	CBD
0.141367	−8.73E−05	0.018522	0.096	−0.522	3.828913	0.345842	−4.23E−04	0.026837	−0.238	1.582	6.541074	2000	1	CBD
0.027867	−2.91E−04	0.004136	0.018	−0.482	3.448514	0.02979	−2.57E−04	0.003311	0.102	0.319	4.522176	1000	1	HB
0.034905	−3.32E−04	0.004827	−0.034	−0.035	3.301827	0.070386	−2.65E−04	0.008588	−0.037	0.249	3.764996	1500	1	HB
0.056728	−3.38E−04	0.008671	0.104	−0.37	3.359181	0.067476	−2.68E−04	0.008985	−0.115	−0.061	3.522973	2000	1	HB
0.110407	−4.26E−05	0.018144	−0.088	−0.449	2.988353	0.161674	−7.63E−06	0.00959	0.102	3.488	7.800044	1000	2	BORC
0.12809	2.87E−04	0.016378	−0.145	−0.007	3.779761	0.521847	−3.40E−04	0.031014	0.13	6.05	9.587711	1500	2	BORC
0.177638	−4.24E−05	0.022025	−0.039	0.185	4.233465	1.03E+00	1.81E−05	0.061017	0.013	5.433	9.4605	2000	2	BORC

Table 3
Confusion matrix.

BCP		BRIR		CBD		HB		BORC		Classified as
ANN	SVM	ANN	SVM	ANN	SVM	ANN	SVM	ANN	SVM	
13	14	1	0	1	1	0	0	0	0	BCP
0	0	10	10	0	0	3	3	1	1	BRIR
5	6	1	1	5	6	2	2	2	0	CBD
0	0	0	0	0	0	14	14	0	0	HB
2	2	1	1	1	1	1	1	10	10	BORC

Table 4
Detailed accuracy by class.

TP rate		FP rate		Precision		Recall		F – measure		Class
ANN	SVM	ANN	SVM	ANN	SVM	ANN	SVM	ANN	SVM	
0.867	0.933	0.121	0.138	0.65	0.636	0.867	0.933	0.743	0.757	BCP
0.714	0.714	0.051	0.034	0.769	0.833	0.714	0.714	0.741	0.769	BRIR
0.333	0.4	0.034	0.034	0.714	0.75	0.333	0.4	0.455	0.522	CBD
1	1	0.102	0.102	0.7	0.7	1	1	0.824	0.824	HB
0.667	0.667	0.052	0.017	0.769	0.909	0.667	0.667	0.714	0.769	BORC

Table 5
Evaluation of the success of the numeric prediction.

Parameters	Values (ANN)		Values (SVM)	
Correctly classified instances	52	71.2329%	54	73.9726%
Incorrectly Classified Instances	21	28.7671%	19	26.0274%
Kappa statistic	0.6407		0.6749	
Mean absolute error	0.1561		0.2586	
Root mean squared error	0.3003		0.3444	
Relative absolute error	48.6647%		80.6103%	
Root relative squared error	74.8483%		85.8576%	

bearings are defect free (b) bearings are with outer race crack (c) bearings with rough inner race surface (d) ball with corrosion pitting (e) Combined bearing components defects. From the obtained response, it can analyzed that with healthy bearings the periodic responses are observed with low magnitude as shown in Figs. 8 and 9(a). When the crack has been introduced in the outer race of bearings, the time response is of aperiodic in nature with type-I intermittency (with vertical response) as shown in Figs. 8 and 9(b). The more severe vibrations are appeared in the spectra with defect in inner race and balls as shown in Figs. 8 and 9(c) and (d). These are also of aperiodic in nature. The combined defects has more impact on vibration spectra and response shows chaos and beat like structure as shown in Figs. 8 and 9(e). The vibration responses are irregular for healthy bearings under loader condition as shows in Figs. 10 and 11(a). When the crack has been introduced in the outer race of bearings, the time response is of aperiodic in nature with type-II intermittency which give a clear indication of mixed behaviour as shown in Fig. 10 and 11(b). The more severe (chaotic) vibrations are appeared in the spectra with defect in inner race and balls as shown in Figs. 10 and 11(c) and (d). These are highly aperiodic in nature. The combined defects has significant impact on vibration spectra and response shows chaos and beat like structure as shown in Figs. 10 and 11(e).

The detailed accuracy of each class is shown in Table 4. Table 4 gives us information about TP rate, FP rate, precision, recall and F-measure values for the 5 classes by using ANN and SVM. The values of various measures of correctly classification of faults are given in Table 5.

7. Conclusions

This study presents a procedure for detection of bearing fault by classifying them using two machine learning methods, namely, ANNs and SVMs. Features are extracted from time-domain vibration signals using statistical techniques. Procedure incorporates most appropriate features selection by a filtering algorithm, which uses a density-based cluster to generate cluster membership values. The roles of different vibration signals, obtained with or without loader and at various speeds, are investigated. The time responses observed for different fault condition of bearing shows that severe (chaotic) vibrations occur under bearings with rough inner race surface and ball with corrosion pitting. The effect of combined defect has also significant on vibration of a rotor bearing system. It is also observed that the classification accuracy for SVM is better than of ANN. Both the machine learning methods give less accuracy to correctly predict the bearing condition with combined bearing component fault, though results obtained from SVM are slightly better, it may be because of small training data are taken for this study assuming that in practical situation less historical data is available. The results show the potential application of machine learning algorithm for developing a knowledge base system which can be useful for early diagnosis of defect for applying condition based maintenance to prevent catastrophic failure and reduce operating cost.

References

Cristianini, N., & Shawe-Taylor, N. J. (2000). *An introduction to support vector machines*. Cambridge: Cambridge University Press.

Gunn, S. R. (1998). Support vector machines for classification and regression. Technical Report, University of Southampton, Department of Electrical and Computer Science.

Kankar, P. K., Harsha, S. P., Pradeep, K., & Sharma Satish, C. (2009). Fault diagnosis of a rotor bearing system using response surface method. *European Journal of Mechanics A/Solids, 28*, 841–857.

Li, Z., Wu, Z., He, Y., & Fulei, C. (2005). Hidden Markov model-based fault diagnostics method in speedup and speeddown process for rotating machinery. *Mechanical Systems and Signal Processing, 19*(2), 329–333.

McFadden, P. D., & Smith, J. D. (1984). Model for the vibration produced by a single point defect in a rolling element bearing. *Journal of Sound and Vibration, 96*(1), 69–82.

McFadden, P. D., & Smith, J. D. (1985). The vibration produced by multiple point defects in a rolling element bearing. *Journal of Sound and Vibration, 98*(2), 263–273.

Prabhakar, S., Mohanty, A. R., & Sekhar, A. S. (2002). Application of discrete wavelet transform for detection of ball bearing race faults. *Tribology International, 35*, 793–800.

Samanta, B. (2004). Gear fault detection using artificial neural networks and support vector machines with genetic algorithms. *Mechanical Systems and Signal Processing, 18*(3), 625–644.

Simon, H. (2005). *Neural networks – A comprehensive foundation*. San Francisco: Pearson Education.

Vyas, N. S., & Satishkumar, D. (2001). Artificial neural network design for fault identification in a rotor bearing system. *Mechanism and Machine Theory, 36*, 157–175.

Widodo, A., & Yang, B.-S. (2007). Review on support vector machine in machine condition monitoring and fault diagnosis. *Mechanical Systems and Signal Processing, 21*, 2560–2574.

Witten Ian, H., & Eibe, F. (2005). *Data mining: Practical machine learning tools and techniques* (2nd ed.). Berlin: Morgan Kaufmann.

Yuan, S., & Chu, F. (2007). Fault diagnosis based on support vector machines with parameter optimization by artificial immunization algorithm. *Mechanical Systems and Signal Processing, 21*, 1318–1330.

Zhitong, C., Jiazhong, F., Hongpingn, C., Guoguang, H., & Ritchie, E. (2003). Support vector machine used to diagnose the fault of rotor broken bars of induction motors. In *Sixth international conference on electrical machines and system (ICEMS)* (pp. 891–894).

Zurada, M. J. (1999). *Introduction to artificial neural systems*. Delhi: Jaico Publishing House.