



Gear and bearing fault classification under different load and speed by using Poincaré plot features and SVM

Rubén Medina¹ · Jean Carlo Macancela² · Pablo Lucero² · Diego Cabrera² · René-Vinicio Sánchez² · Mariela Cerrada²

Received: 12 May 2020 / Accepted: 9 November 2020 / Published online: 23 November 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

This paper describes two algorithms for feature extraction from the Poincaré plot which is constructed with the vibration signals measured in roller bearings and gearboxes. The extracted features are used for classifying 10 types of fault conditions in a gearbox and 7 types of fault conditions a roller bearings. Both vibration signal datasets were acquired at different loads and speeds. The feature extraction using Algorithm 1 performs the feature calculation from the Poincaré plot constructed with the raw vibration signals. In contrast, the Algorithm 2 requires an additional stage where the vibration signal is pre-processed for identifying the peaks of the signal. This peak sequence is equivalent to a non-uniform sub-sampling of the vibration signal that retains relevant information useful for fault classification. The fault classification is attained by using a multi-class Support Vector Machine. The proposed method is tested using the tenfold cross-validation. Results show that both algorithms could attain classification accuracies as high as 99.3% for the gearbox dataset and 100% for the roller bearings. The results are compared to other classification approaches performed on the same datasets by using other different features. The comparison shows that the approach in this paper has a performance as good as obtained by using well-known statistical features.

Keywords Poincaré plot · Support vector machine · Gearbox fault diagnosis · Bearing fault diagnosis · Rotating machinery · Vibration signals

Abbreviations

PHM	Prognostics and health management	LZC	Lempel–Ziv complexity
CD	Correlation dimension	SampEn	Sample entropy
LLE	Larger Lyapunov exponent	ApEn	Approximate entropy
		SVM	Support vector machines
		ANN	Artificial neural networks
		ECOC	Error-correcting output codes
		NI	National instruments
		A/D	Analog to digital converter
		HP	Horse power
		RBF	Radial basis function
		FNR	False negative rate
		FN	False negative
		TP	True positive
		FPR	False positive rate
		FP	False positive
		TN	True negative
		TPR	True positive rate
		TNR	True negative rate
		P	Precision
		SN	Sensitivity
		SP	Specificity

✉ Mariela Cerrada
mcerrada@ups.edu.ec

Rubén Medina
ruben.djmedina@ieee.org

Jean Carlo Macancela
jmacancelap@est.ups.edu.ec

Pablo Lucero
pluceroo@est.ups.edu.ec

Diego Cabrera
dcabrera@ups.edu.ec

René-Vinicio Sánchez
rsanchezl@ups.edu.ec

¹ CIBYTEL-Engineering School, Universidad de Los Andes, Mérida 5101, Venezuela

² GIDTEC, Universidad Politécnica Salesiana, 010105 Cuenca, Ecuador

ROC	Receiver operator curve
KNN	K nearest neighbors
RF	Random forest
MDSVC	Multi-modal deep support vector Classification
GDBM	Gaussian–Bernoulli deep Boltzmann machine

Introduction

Rotating machinery requires the development of accurate fault detection and diagnosis methods aiming at minimizing costs and avoiding accidents in industry. Rotating machinery includes several components susceptible of failure, however, bearings and gearboxes are the main location of faults (Moumene and Ouelaa 2012). The faults generation process is complex and related to the fabrication of each component and the operating conditions. Early detection of faults has motivated developing interesting research works to obtain accurate methods for detecting and classifying faults in rotatory machinery (Cerrada et al. 2016b; Li et al. 2018; Luo et al. 2020; Zhou et al. 2019). Gearboxes in rotatory machinery enable transmission of mechanical motion. They usually have complex configurations of gears usually connected by shafts and bearings. In such an equipment, the detection and classification of fault conditions can be performed through acquisition and analysis of vibration signals (Li et al. 2015b; Liu et al. 2016; Huang et al. 2018). A review of gearbox fault diagnosis using several features extracted from the vibration signal is presented in Sharma and Parey (2016). Vibration signal analysis is also useful for detection of faults in roller bearings (Li et al. 2015a). Roller bearings have several components where faults could appear: the cage, the inner and outer race, and the rolling elements. The faults in roller bearings are due to several factors such as defects in their construction, poor lubrication or overloading (Patil et al. 2008). Roller bearings defects lead to faults that could be detected by vibration signal analysis (Randall and Antoni 2011; Li et al. 2016a, b).

Fault diagnosis consists in the identification of a machine fault based on features usually extracted from measured signals (Wang et al. 2015). Although several signals can be sensed and processed for extracting useful features for feeding data-driven automatic classifiers, the vibration signal has been commonly used in many Prognostics and Health Management (PHM) applications (Cheng et al. 2010).

Several applications based on vibration signal analysis rely on the assumption that such signals are collected from linear and time-invariant systems where time-frequency methods (Sait and Sharaf-Eldeen 2011), as well as well-known time-statistical analysis are feasible and valid. However, rotating machinery are actually non-linear systems

where harmonics and even chaotic motions can occur as a response to harmonic excitation forces. Research works have been previously reported for verifying the non-linear nature of the vibration time series collected from gearboxes (Wang et al. 2003; Bajric et al. 2011), and bearings (Cui and Qian 2010; Liqin et al. 2008), as well as their self-similarity properties (Loutridis 2008). The non-linear vibration of gear systems is studied in Wang et al. (2003) by using a vibro-impact model of a gear pair. One of the methodological tools is non-linear systems with time-variant coefficients modeling mesh backlash and stiffness (Aherwar 2012). Both parameters contributes with a strong non-linear term to the dynamics equation showing the non-linear dynamic response. Other factors contributing to non-linear vibration are transmission errors and the friction between tooth mesh faces. Gearboxes actually involve multi-nonlinear elements contributing to a non-linear vibration. Kahraman and Blankenship (1997) several experiments were performed on geared systems showing a rich frequency spectrum of non-linear phenomena as well as chaotic motion. Such chaotic motion could lead to irregular operations and fatigue failure in rotating machinery (Szuminski and Kapitaniak 2012). The complex non-linear behavior of bevel gears is investigated in Motahar et al. (2016) by using models for three types of teeth modifications. The motion equations of the models are non-linear and time-varying. In these equations the backlash function and mesh stiffness are also considered. A genetic algorithm is used for finding the optimal profile modification and the dynamic behavior of the system is analyzed using several methodological tools including bifurcation diagrams as well as Poincaré maps (Kahraman 1992).

Dealing with non-linear vibration signal analysis requires application of appropriate tools such as non-linear dynamics and chaos theory (Soleimani and Khadem 2015). Within this set of tools, the correlation dimension (CD) has been used in Janjarasjitt et al. (2008) for diagnosis and prognosis of bearings in rotating machinery. The results showed that CD for a roller element close to failure was statistically different with respect to a new roller element. A modification of the original Grassberger-Procaccia algorithm was proposed. The algorithm was known as the partial correlation integral aiming at estimating the CD. Authors also found that the CD tended to increase as the life of the bearing was consumed. Other tools for non-linear rotating systems analysis are useful for detection and progression evaluation of faults (Sun 2012). In particular, the Largest Lyapunov Exponent (LLE) was used for studying incipient faults in the bearings of an induction motor, based on acquired data representing the stator currents. Authors performed the estimation of the LLE for a fault as it evolved with increasing damage. An increased change in LLE as the fault evolves to total damage of the bearing was found. Measurement of complexity in vibration signals has also been used for bearings elements health

evaluation (Yan and Gao 2004). Similarly, there are research works concerning Lempel-Ziv Complexity (LZC). An application of LZC on vibration signals of roller bearings was aimed at condition monitoring of rotating machine systems. Results showed the usefulness of complexity for machine health condition evaluation. Kedadouche et al. (2015), the comparison between LZC, Sample Entropy (SampEn) and Approximate Entropy (ApEn) calculated from vibration signals was reported aiming at detecting faults in gears. After comparison, combination of kurtosis with LZC was proposed by the authors for detecting faults in the early stages.

Typical tools for analyzing non-linear and chaotic dynamical systems are the phase space diagram and the Poincaré section, also known as Poincaré map. The phase space diagram is a tool for visualizing the behavior of a dynamical system by representing the trajectories such as a 3-D trajectory for a chaotic system (Jáuregui 2011). The Poincaré section is obtained from the phase space diagram by sampling points in the phase space at regular intervals. In the case of a 3-D phase space, the Poincaré section is obtained as the interception of a plane with the trajectories representing the phase space. The Poincaré map has been used extensively for visualizing non-linear, chaotic behavior of faults in gears and bearing elements (Szuminski and Kapitaniak 2012; Jáuregui 2011; Kahraman 1992; Kahraman and Blankenship 1997; Motahar et al. 2016; Rocha et al. 2010; Wang et al. 2003). However, this tool has been mainly used for visualization of the non-linear system behavior.

Vibration signal analysis using Poincaré maps has also been reported. The relation between phase space and the accurate calculation of Poincaré maps have been reported in Tucker (2002). Mevel and Guyader (2008) a model and an experiment were developed for investigating ball bearings motion and their chaotic features. The Poincaré maps were used by the authors for visualizing the phase space and their chaotic trajectories. Machine tool components and machining processes condition have been monitored by using information from several internal sensors as reported in Repo (2010). One of the tools used for performing the signal analysis was the Poincaré map. However, only qualitative analysis was stated in this research. Jáuregui (2011) a discussion about the analysis of phase diagrams for predicting non-linearities and transient responses in rotating machinery was presented. The author suggested the use of such a plots for faults detection. Fault characterization based on features extracted from Poincaré map was presented in Trendafilova and Manoach (2012). The damage index quantified the degree of damage based on length calculation of trajectories in the Poincaré map.

In applications of time series analysis to biomedical signals, rather than using the Poincaré map the researchers have used an approximation of this tool that is known as the Poincaré plot. Denoting the time series as $x(t)$ and a lag in time

as τ , this type of plot is a 2D graph where a lagged sample $x(t + \tau)$ is plotted against the sample $x(t)$. The Poincaré plot has been also known as scattergram or scatter plot. In biomedical signal processing a relevant application of Poincaré plots has been reported for electrocardiogram signal analysis. In this type of application, the R–R inter-peak intervals between consecutive heart beats are extracted from surface electrocardiogram signals. The extracted time series signal is known as Heart Rate Variability (HRV) and the Poincaré plot is an important tool that enables features extraction useful for cardiac diseases diagnosis (Hoshi et al. 2013).

Data-driven fault classification methods are commonly used for PHM. This type of methods includes two stages where the first one is devoted to feature extraction from the measured signals. The extracted features are used in the second stage for fault detection and classification. Several machine learning methods could be used for accomplishing this task such as Random Forests (RF), Support Vector Machines (SVM) (Yin and Hou 2016; Jedliński and Jonak 2015; Sánchez et al. 2017), and Artificial Neural Networks (ANN) (Pacheco et al. 2016).

A machine learning model for binary classification is SVM. This type of model is able to construct a hyperplane separating two classes such that the distance to the closest data point in each class is maximal (Luts et al. 2010). In contrast to the binary version of SVM, there are also multi-class versions of the SVMs where the decision function separating all classes are calculated simultaneously (Wang and Xue 2014). However, the preferred approach consists in splitting the original multi-class classification into a set of binary classification tasks. Their combination is performed by using Error-Correcting Output Codes (ECOC) (Escalera et al. 2010), to solve the multi-class classification task. In the ECOC method three symbols $\{-1, 0, +1\}$ are used for coding the results of binary classifiers where the zero symbol encodes a *don't care*. In the coding step, a codeword is assigned to each class. In the decoding stage, a test sample is given and the most similar class codeword is found. The ECOC framework includes several strategies such as one-versus-all or one-versus-one that extend the possibilities for solving the classification problem.

Medina et al. (2019) the Poincaré plot constructed with vibration signals of gearboxes has been used for extracting a set of features describing the symbolic dynamics of the underlying mechanical system. This set of features is an array with twelve elements that is used for training a SVM model. This method has been successfully applied for classification of faults in a gearbox.

In this paper, we propose the use of Poincaré plots of the vibration signal to extract a small set of features that are the inputs to a machine learning algorithm aiming at fault classification in gearboxes and bearings. The fault classification will be attained using a SVM model that performs

multi-class classification according to the ECOC approach (Wang and Xue 2014; Escalera et al. 2010). Classification of 10 different classes corresponding to the healthy class and 9 faulty classes is performed from the gearbox vibration signal dataset. On the other hand, an experiment for classification of 7 different classes in the vibration signal bearing dataset considering the healthy class and 6 faulty classes is also reported. Only three features derived from the Poincaré map are needed for classification. Two algorithms are proposed that have the advantage of being simple, fast and accurate for fault detection. Novelty of the proposed approach is extracting quantitative features from the Poincaré plot for attaining accurate fault classification rather than using this plot as a simple graphical qualitative tool. In this application, the delay τ used for constructing the Poincaré plot is related to a simplified time-delay embedding that is applied to chaotic time series modeling (Von Oertzen and Boker 2010). The high classification accuracy with only three features and low computational costs of the two algorithms suggest their application to continuous condition monitoring of rotating machinery (Bangalore and Tjernberg 2015; Goyal and Pabla 2016; Jedliński and Jonak 2015). Application in the context of continuous monitoring could be performed in two stages. The first stage would work using a SVM for online detection of the presence of faults in the gearbox (or bearings). Such fault detection could trigger a second stage where the multi-class SVM could identify the type of fault.

This paper is an extended version of the conference paper reported in ETCM 2017 (Medina et al. 2017). The work presented in the conference reported a method for gearbox fault detection (Algorithm 1) based on features extracted from vibration signals that were used for fault classification using a ECOC SVM. In addition to the work previously mentioned in this paper the following contributions are reported:

- Proposal of a new method (Algorithm 2 in Fig. 4) for feature extraction based on the time series of the sequence of peaks obtained from the raw vibration signal that is later used for plotting the Poincaré map and the subsequent feature extraction for classification.
- Additional tests and results were included for illustrating the application of both algorithms. Results concerning the classification of faults for roller bearings using a vibration signal dataset were also included.
- A more detailed description of algorithms, discussion and comparison of results are included as well as guidelines for selecting the optimal lag value (τ) using mutual information and the average auto-correlation sequence.

The rest of paper is organized to include the following parts: In “Methodology” section, the methodology for signal acquisition, feature extraction and fault classification is presented. A detailed discussion of the results is described

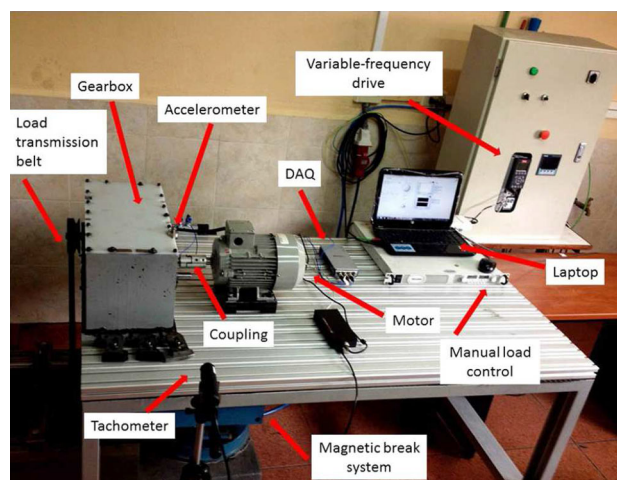


Fig. 1 Test-bed used for signal acquisition

in “Results” section, conclusion as well as future research is reported in “Conclusion” section.

Methodology

Measurement of vibration signals in a gearbox

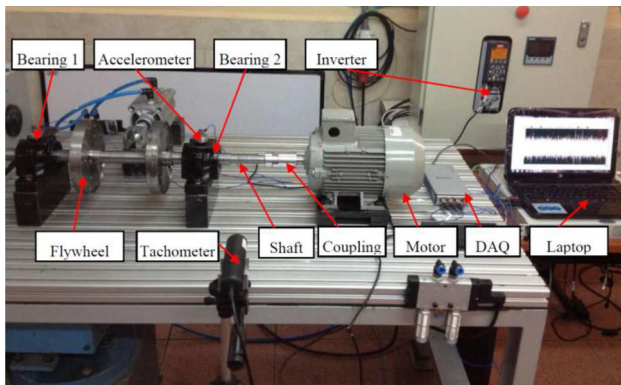
The signal acquisition was performed on a test-bed located in the vibration laboratory at the Salesian Polytechnic University in Cuenca-Ecuador. The acquisition system is shown in Fig. 1.

An induction motor of 1.1 kW that operates at 1650 rpm powered by a three-phase electric line of 220 Vac at 60 Hz is used for generating the rotating motion. A gearbox is connected to the motor using a shaft and the transmitted torque is transferred to a pulley that is used for driving a magnetic brake. Within the gearbox several gear faults can be configured. A power inverter with variable-frequency drive was used for generating 3 constants speeds of 480, 720 and 900 rpm and also 3 variable speeds in the ranges 720–1080 rpm, 300–720 rpm and 480–900 rpm. The load of the motor is controlled by the magnetic brake and three different loads (L) are configured by varying the voltage feed with respect to the maximum value: $L1$ considering 10% of the maximum feed, $L2$ with 50%, and $L3$ with 90%.

The signal acquisition was performed by using a National Instruments (NI) A/D conversion card including anti-aliasing filtering with a sample frequency of 50 kS/s. Each sample was recorded with a resolution of 24-bits. Two accelerometers IMI Sensor 603C01, 100 mV/g, were used for sensing the horizontal and vertical vibration motion. Each recorded signal had a length of 500,000 samples representing a duration time of 10s.

Table 1 Gear fault conditions

Label	Pinion	Gear
<i>P1</i>	Healthy	Healthy
<i>P2</i>	Tooth chafing	Healthy
<i>P3</i>	Tooth wear	Healthy
<i>P4</i>	25% tooth breakage,	Healthy
<i>P5</i>	50% tooth breakage	Healthy
<i>P6</i>	100% tooth breakage	Healthy
<i>P7</i>	Healthy	25% crack
<i>P8</i>	Healthy	100% crack
<i>P9</i>	Healthy	50% chafing
<i>P10</i>	25% tooth breakage	25%crack

**Fig. 2** Bearings vibration signal acquisition test-bed

The configuration of the gearbox considered ten different health conditions, including the healthy case (see Table 1). Class *P1* is the healthy (normal) condition, incipient faults are *P2* and *P3*. Moderate faults are denoted *P4*, *P5*, *P7*, *P9*. Two severe faults are denoted *P6* and *P8*, and finally a multi-fault is denoted *P10*. The case of incipient and moderate faults are important to be diagnosed in industrial applications because their detection allow to prevent further damage and accidents.

A total of 6 motor speeds and 5 repetitions were considered for each load $\{L1, L2, L3\}$. The number of acquired vibration signals for each fault was $6 \times 5 \times 3 = 90$ and considering the 10 faults a total of 900 vibration signals were recorded.

Measurement of vibration signals in roller bearings

The rolling element bearing test configuration for vibration signal acquisition was also constructed at the Salesian Polytechnic University in Cuenca, Ecuador. The test-bed allows configuring different health conditions of the bearings and measuring vibration signals. The experimental test-bed configuration is presented in Fig. 2.

Table 2 Fault conditions for the rollers bearings

Class	Bearing 1	Bearing 2
<i>P1</i>	Healthy	Healthy
<i>P2</i>	Inner race fault	Healthy
<i>P3</i>	Outer race fault	Healthy
<i>P4</i>	Ball fault	Healthy
<i>P5</i>	Inner race fault	Outer race fault
<i>P6</i>	Inner race fault	Ball fault
<i>P7</i>	Outer race fault	Ball fault

The test-bed included a 2.0 HP Siemens motor (3~ Vac) that was fed and controlled by using a Danfoss inverter with 1.5 kW. The motor was connected to a steel shaft with a diameter of 30 mm using a coupling. Two bearings 1207 EKTN9/C3, SKF (one at each end of the shaft) serve as shaft supporting. A tachometer (VLS5/T/LSR optical sensor, Compact) was used for monitoring the rotating speed of the motor. Similarly, the vibration signals were measured from each of the bearings with an accelerometer (ICP 353C03, PCB) mounted on each bearing housing (SNL 507-606, SKF). The load of the motor was implemented using flywheels on the shaft.

A data acquisition card (cDAQ-9234, NI) was used for measuring the vibration signals. The acquisition card was attached to the chassis (NI 9188) which was connected to a computer. The signal acquisition was performed by varying the motor speed considering three values: 480 rpm, 600 rpm and 900 rpm. Three different loads were considered by using zero (load *L1*), one (load *L2*) or two flywheels (load *L3*). Seven health bearing conditions in Table 2 were implemented, and each experiment was repeated 5 times. In total $3 \times 3 \times 7 \times 5 = 315$ signals were recorded using a sampling frequency of 50 kHz, over 20 seconds.

The Poincaré method

The Poincaré method is a methodological tool that can be used for analyzing the non-linear dynamics of systems that have a chaotic behavior (Alligood et al. 1997). Such type of dynamical systems can be represented by a set of first order differential equations and the solution for this set of equations describes an orbit or trajectory in the phase-space. This is illustrated in Fig. 3 where the phase-space trajectories for the 3D Lorenz attractor are shown intercepting a surface corresponding to a plane. The interception of such trajectories with the plane corresponds to the Poincaré map (Cerrada et al. 2020).

In practical applications, the differential equations modelling the system are unknown, and the Poincaré map cannot be obtained. In consequence, it is necessary to use approx-

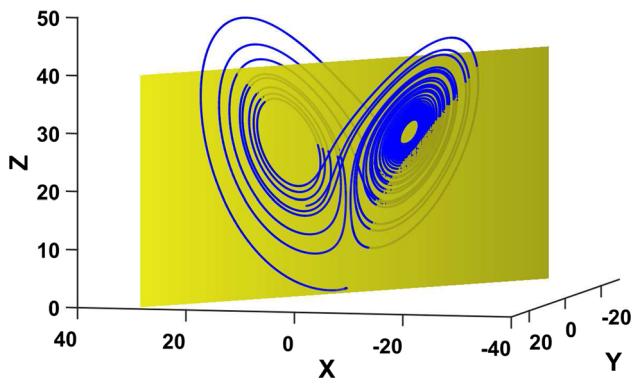


Fig. 3 Poincaré map for the 3D Lorenz attractor

imate solutions known as the Poincaré plot, from partial knowledge about the system. This plot is obtained in a simple manner by plotting the time series $x(t)$ of some variable of the system with respect to its lagged representation $x(t + \tau)$ (Alligood et al. 1997; Wu et al. 1998; Brennan et al. 2001).

The most common situation in mechanical systems consists in sensing variables such as the vibration, which is a time series $x(t)$, and the Poincaré plot is obtained by properly selecting the lag τ . Usually, this is selected by identifying the value for which the auto-correlation of $x(t + \tau)$ is zero or near zero. More details on the Poincaré method and Poincaré plot is presented in Cerrada et al. (2020). From this graphical representation, it is possible to extract features useful for fault classification as is described in the following sections.

Algorithms based on Poincaré plots

In this paper, two algorithms for fault classification using vibration signal analysis are proposed. The steps involved in each algorithm are shown in the diagrams presented in Fig. 4. Concerning the Algorithm 1, the features are calculated from the Poincaré plot as proposed in Medina et al. (2017). Such features are fed to a SVM for the fault classification.

The Algorithm 2 is an original contribution reported in this paper. The Algorithm 2 is inspired in the successful application of peaks detection and Poincaré plots for cardiac electrocardiographic signal analysis (Hoshi et al. 2013). According to Algorithm 2, rather than working with the raw time series, an additional step was considered. In this step, the peaks in the vibration signal are detected and the Poincaré plot is generated from the detected sequence of peaks. The feature extraction and classification stages are similar to Algorithm 1.

During the peak detection stage of Algorithm 2, the vibration signal is processed for searching the peaks separated by at least *lag* samples without any restriction concerning the sign of the peak. In consequence, the retained sequence of peaks includes positive and negative peaks. The relevance of

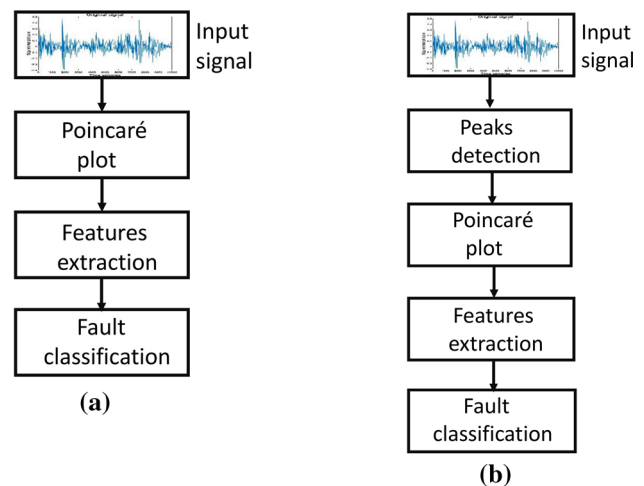


Fig. 4 **a** Algorithm 1 for fault classification using features extracted from the Poincaré plot, **b** Algorithm 2 includes an additional stage corresponding to the peaks detection

this approach relies on the fact that the peaks of the vibrations signal convey relevant information for fault detection as reported in Igba et al. (2016), Doguer and Strackeljan (2009) and Xia et al. (2012). The Algorithm 2 has the additional advantage of working with a shorter signal including only the peak, which provides a non-uniform sampled sequence that could be useful for extraction of other features for fault classification.

Feature extraction for algorithm 1

The shape of the Poincaré plot from the vibration signals of the gearbox dataset shows that their shape varies according to the type of fault. An illustration of this fact is shown in Fig. 5 where the Poincaré plots for two different fault classes are shown. Specifically, the plots corresponds to the healthy (normal) class *P1* and a class representing a combined fault *P10*. The Poincaré plot is considering all samples of the vibration signal. Construction of the Poincaré plot was performed with $\tau_{1_{map}} = 10$ (details about this selection will be given in “Criteria for the lag selection in Poincaré plots” section). There are apparent differences concerning the shapes of the plots for both classes. The healthy (normal) class *P1* has a shape that tends to be ellipsoidal. In contrast, the faulty class *P10* has a irregular shape with trajectories departing from the centre of the shape.

The first extracted feature from Poincaré plot corresponds to the standard deviation of points denoted *SD1*, measured as the dispersion of points with respect to the line-of-identity axis $y = x$. Denoting the sample of the vibration signal as $x(t)$ and the vector representing one vibration signal as \mathbf{x}_t , the calculation of the dispersion is based on fitting an ellipse shape to the cloud of points in the Poincaré plot constructed

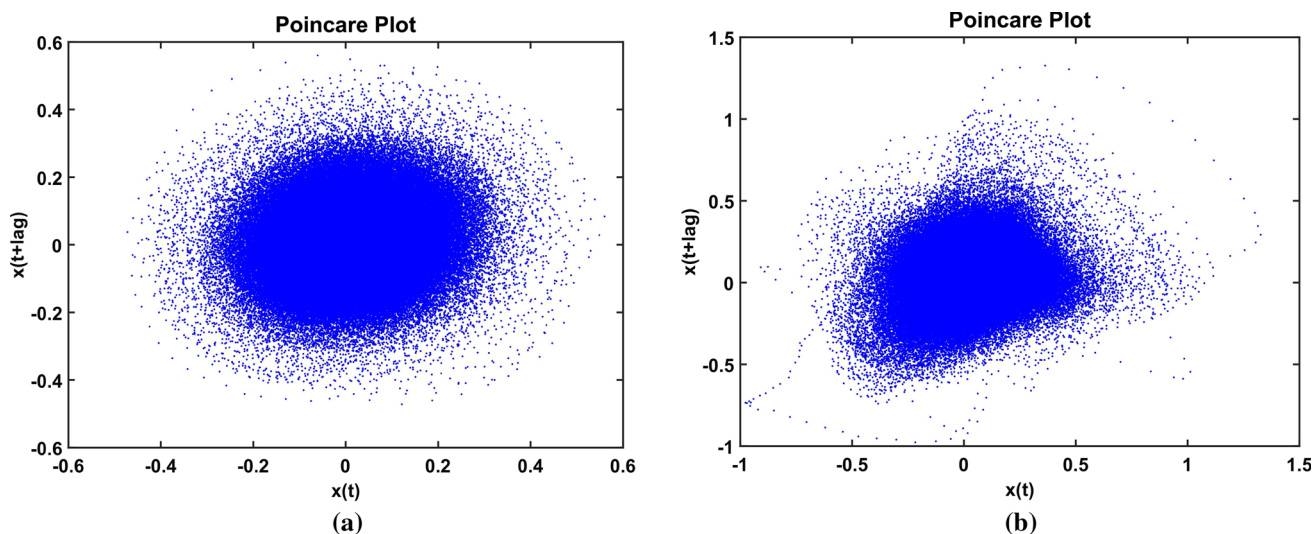


Fig. 5 Poincaré plots for vibration signals extracted from the gearbox dataset. **a** Poincaré plot for healthy case *P1*, **b** Poincaré plot for faulty class *P10*

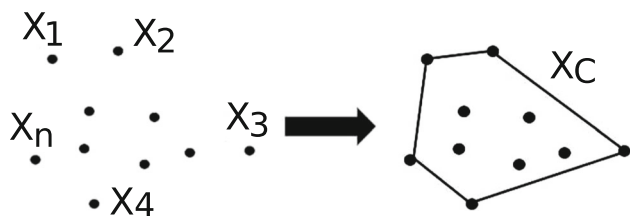


Fig. 6 The convex hull of the set of points is the polygon shown at the right enclosing all points

with a lag τ , and the equation is shown in Eq. (1).

$$SD1^2 = \frac{1}{2} var\{\mathbf{x}_t - \mathbf{x}_{t+\tau}\} = \frac{1}{2} SDSD^2 \tag{1}$$

where the standard deviation SDSD is calculated by considering lagged differences.

The standard deviation or dispersion of points *SD2* is the second feature calculated with respect to the axis $y = -x$ as shown Eq. (2).

$$SD2^2 = 2var\{\mathbf{x}_t\}^2 - \frac{1}{2} SDSD^2 \tag{2}$$

The area of the convex-hull is an additional feature that can be extracted from the Poincaré plot (De Berg et al. 2008). A set of bidimensional points \mathbf{X} is considered as convex when given any pair of points $r, s \in \mathbf{X}$, the line segment joining both points $rs \subset \mathbf{X}$. The convex-hull for a finite set of points $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$ is defined as the smallest convex set \mathbf{X}_C such that $\mathbf{X} \subset \mathbf{X}_C$. An illustration of a convex hull is shown in Fig. 6. The convex polygon enclosing all points is the convex hull. The area of the convex hull of the Poincaré plot is the feature for fault classification.

Examples of convex-hull calculated from Poincaré plots of vibration signals are shown in Fig. 7. The Poincaré plots of the healthy class *P1* and the faulty class *P10* are shown with the convex hull represented as a red polygon. The shape of both polygons are clearly different and the area is a feature useful for fault classification.

Feature extraction for algorithm 2

Algorithm 2 differs from Algorithm 1 mainly in the pre-processing stage which requires the additional step of the peaks detection, as shown in Fig. 4. The algorithm 2 includes the following stages:

1. The vibration signal is pre-processed by calculating its absolute value $\|x(t)\|$. Peaks separated by at least τ_{peak} samples are retained. The resultant time series is denoted $x_p(t)$. This *peak signal* is an array with smaller size than the original signal. An example of this pre-processing is shown in Fig. 8 using a small fragment of the vibration signal. The sequence of peaks are shown in red.
2. After obtaining the sequence of peaks, the next stage is constructing the Poincaré plot using a lag parameter $\tau_{2map} = 1$. Each point in the Poincaré plot represents the location defined by two consecutive peaks from $x_p(t)$. The Poincaré plot is composed by four clouds of points, one cloud for each quadrant. The sign of the pair of consecutive peaks define the quadrant where each point is located. This is illustrated in Fig. 9.
3. The Poincaré plot for $x_p(t)$ is analyzed for extracting the features stated in Algorithm 1: the area of convex-hull, SD1, and SD2.

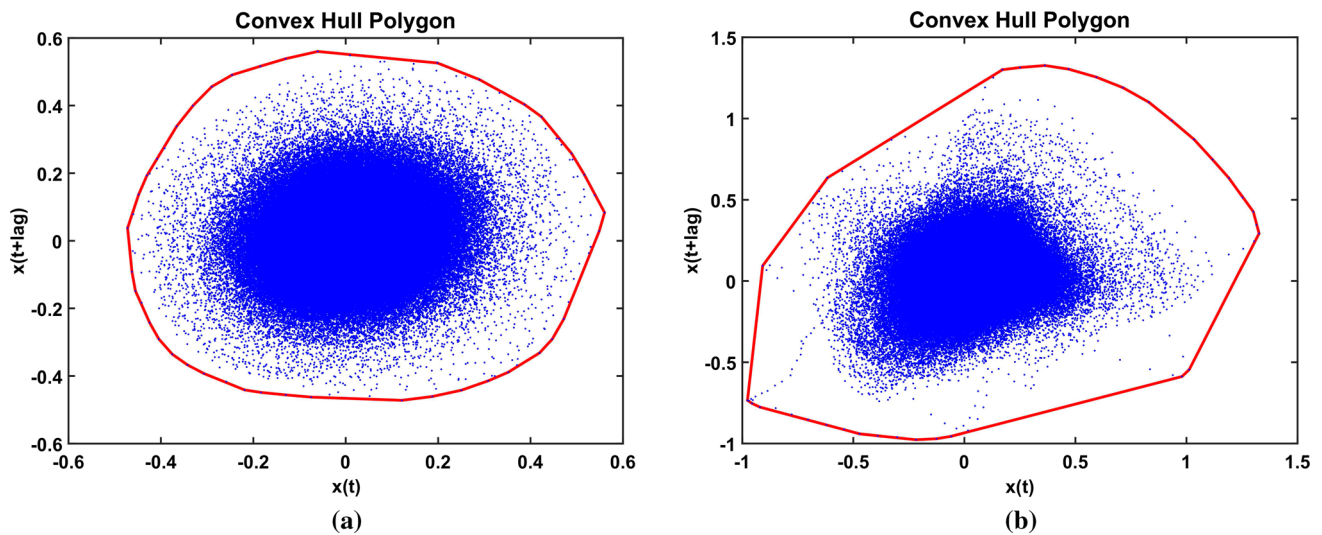


Fig. 7 Convex hull (in red) of Poincaré plots from a vibration signal of the gearbox dataset. **a** Convex hull for the healthy class *P1*, **b** Convex hull for the faulty class *P10* (Color figure online)

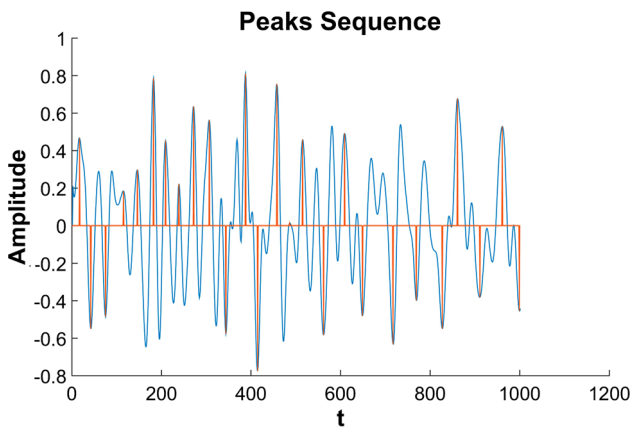


Fig. 8 Peaks sequence (in red) overlaid on a fragment of the original vibration signal (in blue) (Color figure online)

The Poincaré plots for two signals from the gearbox dataset are shown in Fig. 9. The Poincaré plot for the healthy class *P1* is shown in Fig. 9 while the plot for the severe fault *P6* is shown in Fig. 9b. The shape of both plots is different. In the case of the severe fault the clouds of points have a star shape. However, the clouds of the normal class is smaller and round.

Criteria for the lag selection in Poincaré plots

The Poincaré plot represents the intersection between a plane and the trajectories of the phase-space of a possible chaotic dynamical system (Daw et al. 2003). Within this context, the reconstruction of the phase-space dynamics is attained by using time-delay embedding. This implies the construction of the state vector $\mathbf{X}(t) = [x(t), x(t + \tau), \dots, x(t + (D - 1)\tau)]$,

where $x(t)$ is a time series signal. The phase-state dimension is denoted as D (also known as fractal dimension) and the time delay or lag is denoted as τ . The phase-space dynamics is reconstructed by plotting each of the points $\mathbf{X}(t)$ in the D dimensional space ($D = 2$ in our application). When the measured signal is considered as an infinite amount of noiseless data, an arbitrary selection of the lag τ is feasible (Takens 1981), however, when the available data is finite and noisy, the lag could be selected such that the samples in $\mathbf{X}(t)$ are independent or uncorrelated. This is accomplished by selecting the lag in two ways, from the available time series: the value for which the plot Autocorrelation vs Lag has zero crossing (normally the first zero crossing), or the value for which the plot Mutual Information vs Lag has a local minimum (Fraser and Swinney 1986).

In the approach of this paper, the goal of constructing the Poincaré plot is the extraction of features that enable the classification of the 10 different faulty conditions for the gearbox dataset and seven faulty conditions for the roller bearings dataset. This classification is feasible when the feature vectors forms proper cluster structures for each class on the feature space. Based on this fact, the search for the lag value to obtain Poincaré plots from uncorrelated samples could be a useful criterion for trying a solution to this problem.

In this research, the gearbox dataset contains 900 vibration signals. A representation of the autocorrelation for this dataset is obtained by averaging the autocorrelation of each of the signals included in the dataset. The averaged autocorrelation for this dataset is shown in Fig. 10. The candidate values for the lag would be located in this plot around the lagged samples with τ values around 10, 30 and 60. With

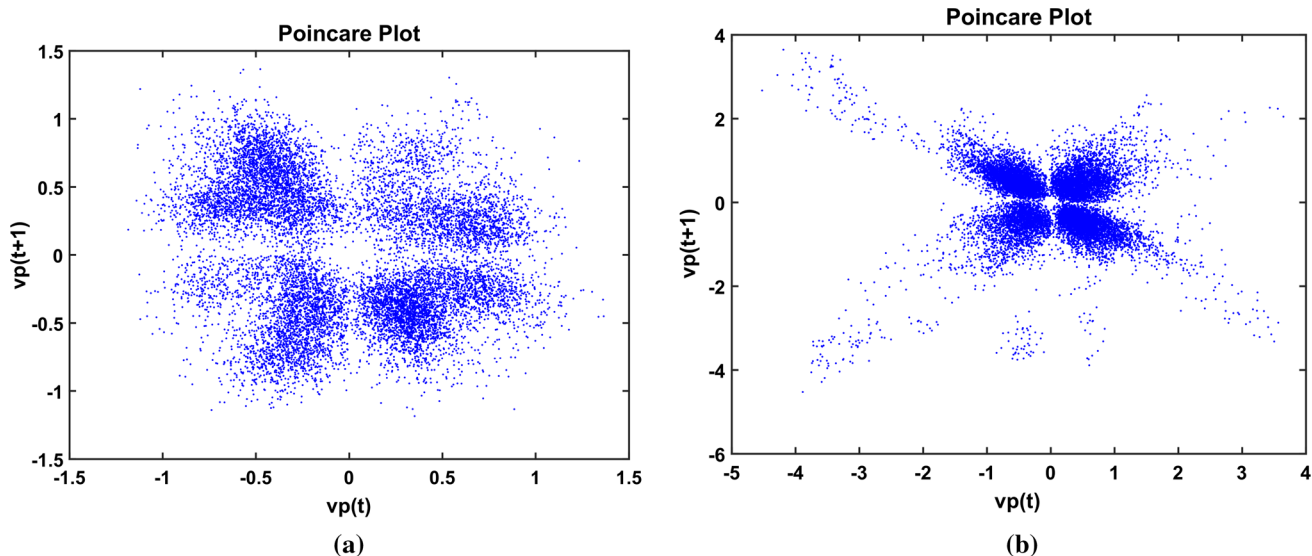


Fig. 9 Poincaré plots for *peaks signals* $x_p(t)$ extracted from the vibration signal gearbox dataset **a** Poincaré plot for healthy class *P1*, **b** Poincaré plot obtained for the faulty class *P6* corresponding to a severe fault

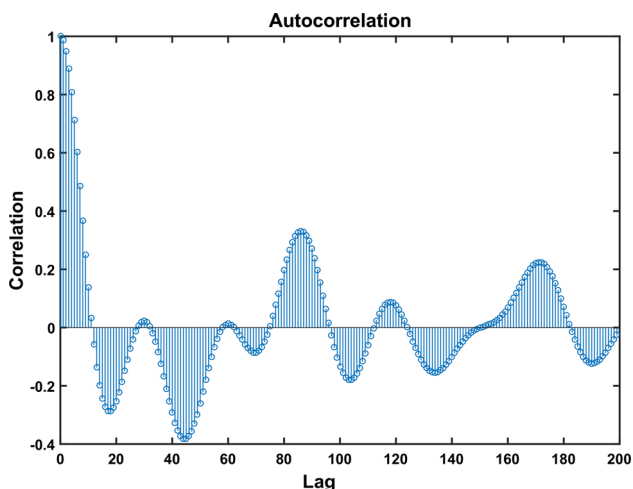


Fig. 10 Averaged autocorrelation considering the 900 vibration signals in the gearbox dataset

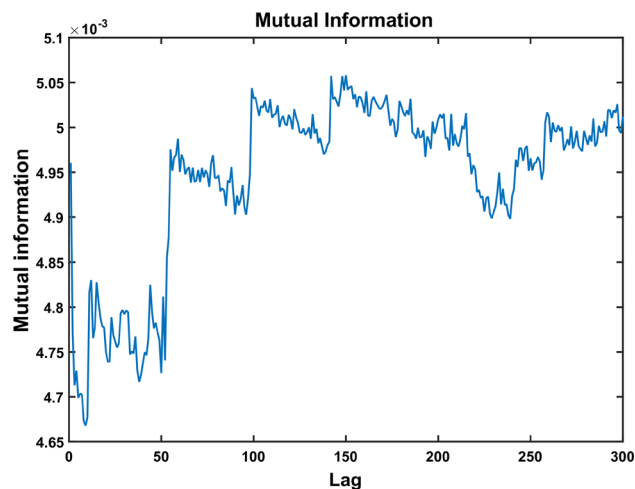


Fig. 11 Averaged mutual information considering the 900 vibration signals in the gearbox dataset

these lags, the correlation between samples of vibration signal should be low.

Alternatively, the mutual information for a pair of random variables quantifies their mutual dependence. This represents the *amount of information* that can be obtained about the first variable based on the other one. Mutual information for discrete random variables $x(t)$ and $x(t + \tau)$ is defined as:

$$I(\tau) = \sum_{i \in x(t)} \sum_{j \in x(t+\tau)} p_{i,j}(\tau) \log_2 \frac{p_{i,j}(\tau)}{p_i p_j} \tag{3}$$

where $p_{i,j}(\tau)$ is the joint probability function for variables $x(t)$ and $x(t + \tau)$ and p_i, p_j are the marginal probability distribution functions for both random variables respectively.

The mutual information is calculated for each of the 900 vibration signals in the gearbox dataset and then averaged. Results obtained by varying the lag τ between 1 and 300 is shown in Fig. 11

Lower values of the mutual information plot are obtained for lag values lower than 50 where several local minimums are present. An empirical procedure is used for selecting the lag for constructing the Poincaré plot. Several lag values between 1 and 50 were tried. The selected lag values lead to a local minimum of the mutual information plot and for the lag intervals where the autocorrelation value is low. For each of the selected values, the Poincaré based features were extracted and the accuracy of classification using multi-class SVM was calculated. The selected lag was the value pro-

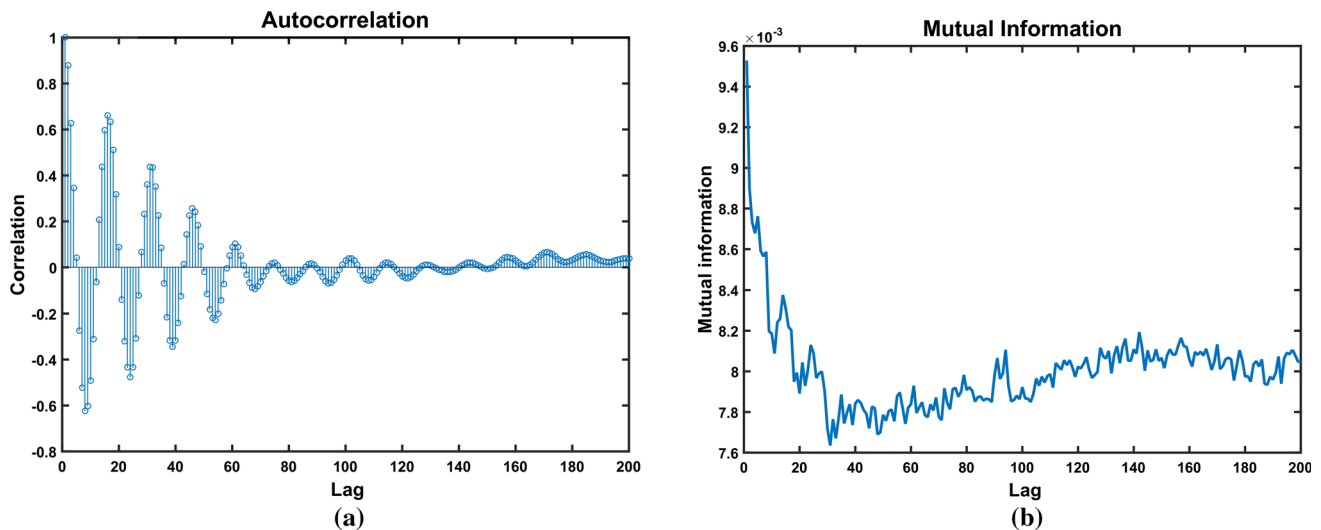


Fig. 12 **a** Averaged autocorrelation considering the 315 vibration signals in the bearings dataset, **b** averaged mutual information considering the 315 vibration signals in the bearings dataset

viding the maximum classification accuracy. Concerning the Algorithm 1, the lag value was $\tau 1_{map} = 24$.

In the Algorithm 2, the τ_{peak} was selected by using the same method as in Algorithm 1 based on the averaged inter-correlation, and by considering also the averaged mutual information. The selected values were $\tau_{peak} = 30$ for obtaining the sequence of peaks, and $\tau 2_{map} = 1$ for constructing the Poincaré plot.

The same procedure was applied for selecting the lag for the bearings dataset. Figure 12a shows the average autocorrelation with the 315 vibration signals in the dataset. The locations of the zero crossing for this function are on $\{5, 10, 20, 25, 35, 43\}$. The average mutual information for all vibration signals of the bearings dataset is shown in Fig. 12b. The global minimum for the averaged mutual information as a function of the lag is located at 30 with several local minimums in the neighborhood approximately between 15 and 100. In this neighborhood, there are several zero crossing for the autocorrelation function. The selected lag value for Algorithm 1 was $\tau 1_{map} = 25$ which is close to the lag value for obtaining the global minimum of averaged mutual information function.

The impact of lag selection on classification accuracy is further explored in “Results” section.

Cluster structure from features generated by using algorithm 1 for the gearbox dataset

The cluster structure was analyzed for the gearbox dataset considering each load separately. The set of three features was calculated for each signal under different load and fault classes. After calculation, we perform a comparison between the clusters of healthy class $P1$ and all the faulty conditions.

The small set of features provides complementary information about the fault condition able to form cluster structures in the feature space that could be separated from each other through single hyper-planes.

Concerning the case of load $L1$, the cluster structure representing an incipient fault $P3$ (see Table 1 including the list of gearbox faults) is shown in Fig. 13a. In this Figure, the green dots are used for representing the healthy class and red dots for the faulty class. The healthy class could be separated from the faulty class by using a single surface. The case of a moderate fault $P5$ is shown in Fig. 13b. The clusters could also be separated each others for this type of faults.

A case representing the severe fault $P8$ compared to the healthy class is shown in Fig. 13c. The cluster for $P8$ could be separated from the healthy cluster. Finally, class $P10$ corresponding to a combined fault is shown in Fig. 13d compared to healthy class $P1$. Separation of clusters for this type of fault is feasible by using machine learning models.

Faulty classes also generates cluster structures that are distinguished from each others. A simple example is the comparison between the multi-fault class $P10$ and the severe fault $P8$. The cluster structure is shown in Fig. 14. The clusters generated could be separated in the feature space and both classes could be classified by using the appropriate classification algorithm.

Cluster structure from features generated by using algorithm 2 for the gearbox dataset

The analysis of the Algorithm 2 applied on signals from the gearbox dataset was also performed by considering each load separately and the same features considered by the Algorithm 1.

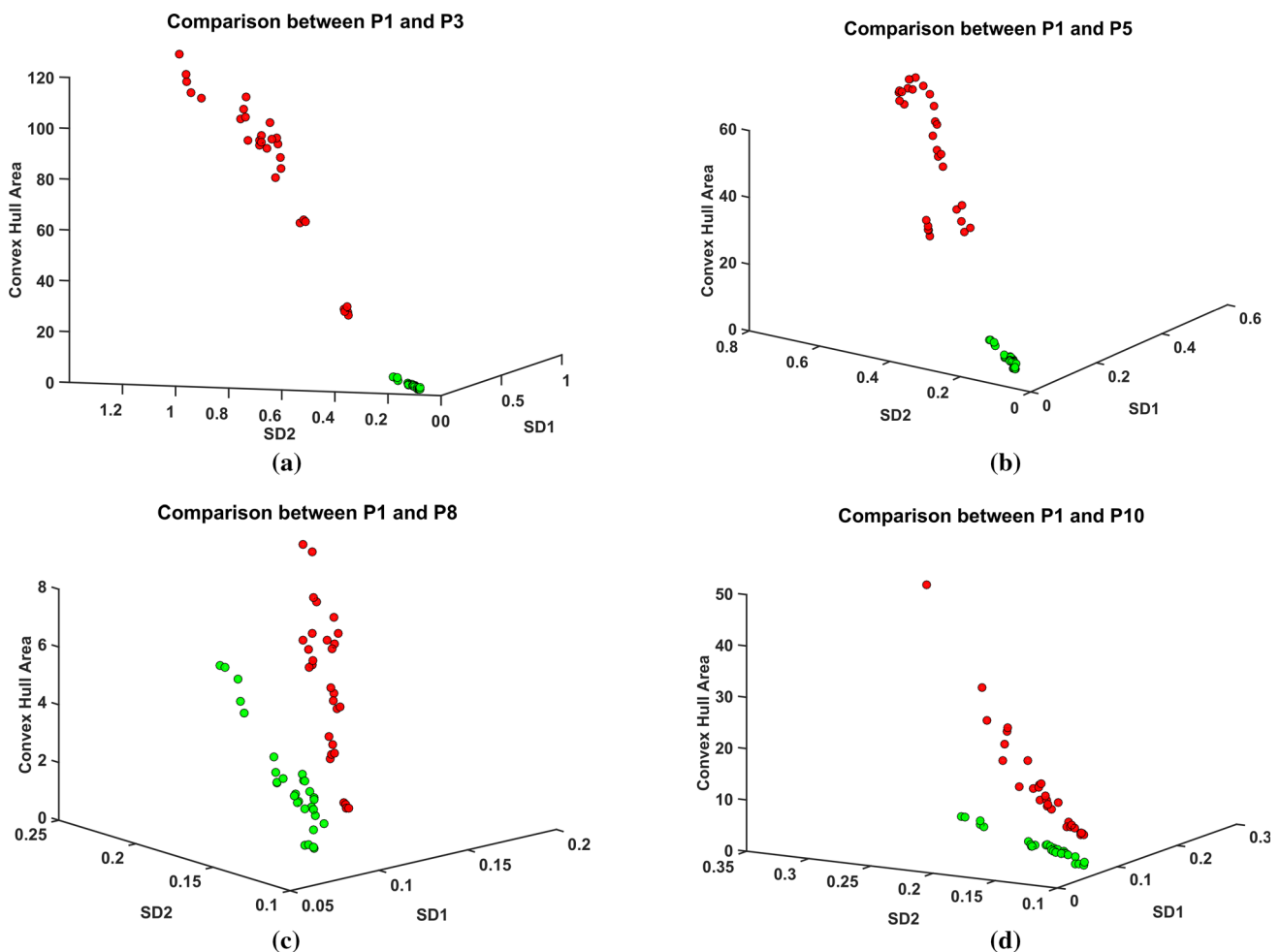


Fig. 13 Results of the Algorithm 1 using signals from the gearbox dataset: cluster structure for load $L1$. The faulty class is represented with red dots and the healthy class $P1$ is represented with green dots. **a** The healthy class tends to a cluster structure that is separated from

the cluster defined by faulty class $P3$. **b** The faulty class $P5$ is also separated from the healthy class. **c** Comparison between $P1$ and $P8$. **d** Classes $P1$ and faulty class $P10$ are also clusters feasible for separation

Concerning the case of load $L1$, the cluster representing an incipient fault $P3$ is shown in Fig. 15a. The points associated to the faulty class are shown in red while the healthy class is in green dots. Separation between both classes is feasible by using a single surface. The case of a moderate fault $P5$ is shown in Fig. 15b. The clusters could also be clearly separated for this type of fault and their classification is feasible.

Comparison between the healthy class and the severe fault denoted $P8$ is shown in Fig. 15c. The points representing $P8$ tend to form a cluster that could be separated from the healthy class by using a machine learning algorithm. Finally, a combined fault corresponding to class $P10$ is compared to normal class $P1$ and it is shown in Fig. 15d. The cluster for this type of fault could also be separated from the healthy class $P1$ by using machine learning techniques.

Comparison between faulty classes also generates cluster structures enabling separation between them. An example of this fact is attained by comparing the multi-fault denoted $P10$ with the severe fault denoted $P8$. The comparison between the cluster generated by fault $P10$ with respect to fault $P8$ is shown in Fig. 16. In this case, the clusters could be separated and both faulty classes could be detected by using the appropriate classification algorithm.

Cluster structure from features generated by using algorithm 1 for the bearing dataset

The gear fault patterns in this experiment occupied larger areas of damage producing greater changes in the vibration signals. In contrast, during incipient bearing fault the damaged surface is smaller and the vibration signal changes are weak (Hiroaki and Nader 2012). In this experiment, the

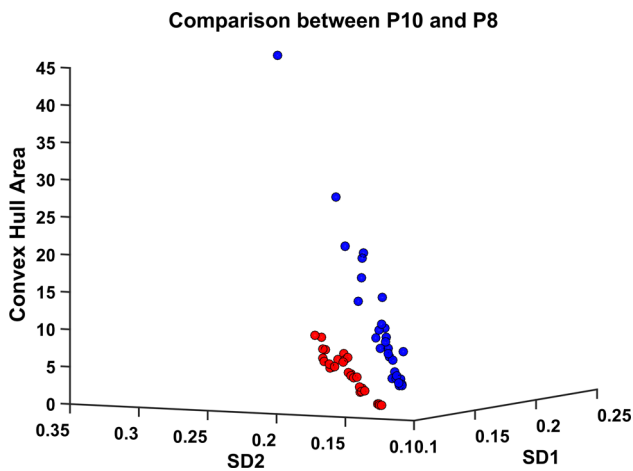


Fig. 14 Comparison between the cluster structures for classes $P10$ (blue) and $P8$ (red) for load $L1$, by using the Algorithm 1 with signals extracted from the gearbox dataset. Each faulty class forms a cluster in the feature space amenable for classification (Color figure online)

three features were extracted considering the vibration signals measured by the accelerometer 1 for load $L2$ and a comparison between healthy class $P1$ and other faulty classes is shown in Fig. 17 (see Table 2 for a description about the simulated faults). The healthy class $P1$ is shown in green, while the faulty classes are shown in red. Comparison of cluster structures generated by classes $P1$ and $P2$ are shown in Fig. 17a. The cluster structures of class $P1$ and $P3$ are shown in Fig. 17b. The classes $P1$ and $P5$ are shown in Fig. 17c. Similarly, the comparison between classes $P1$ and $P7$ is shown in Fig. 17d. In all cases, the healthy class can be separated from the faulty class by using a single surface. In each of the plots in Fig. 17 the dots corresponding to each of the classes tend to create three small clusters that corresponds to changes in speed. The separation between the three clusters for the faulty class is larger and their location is also clearly separated from the clusters associated to the healthy class.

Cluster structure from features generated by using algorithm 2 for the bearing dataset

In a second experiment, the Algorithm 2 was used for extracting three features from the vibration signal measured by the accelerometer 1 for load $L2$. The healthy class $P1$ is compared to the other faulty classes as shown in Fig. 18. Like the comparison obtained for the bearing dataset using the algorithm 1, the healthy class is represented in green dots while the faulty classes are represented in red dots. In Fig. 18a, the cluster for the healthy class $P1$ and faulty class $P2$ are compared. Figure 18b shows the comparison of healthy class to faulty class $P3$. The healthy class is compared to faulty class $P5$ in Fig. 18c. Similarly, the cluster structures of $P1$ and $P7$

are presented in Fig. 18d. In all the cases, separation between the healthy class and the faulty class is feasible by using the appropriate machine learning technique. In this comparison, the load is $L2$, however three different speed values are considered. The changes in speed are represented by using this set of features as three small clusters that are clearly visible in Fig. 18b. However, each of the sub-clusters representing the healthy class is clearly separated from the corresponding sub-cluster for the faulty class. The comparison between classes $P2$, $P5$ and $P7$ shows that the sub-clusters for the healthy class are grouped in a compact cluster while the points for the faulty classes tend to create several sub-clusters. In all cases, however, there is a clear separation between the healthy and faulty classes.

Fault classification

Several alternatives for multi-class classification are available (Aly 2005). However, SVM is a highly robust and successful classification algorithm (Borges 1998). The basic SVM performs binary classification and extensions intended for multi-class classification are also available (Aly 2005; Escalera et al. 2010; Goyal et al. 2020; Wang and Gan 2017).

In this experiment of gearbox fault classification, ten classes are considered while for the roller bearings dataset seven classes are represented. In consequence, a multi-class classification technique was applied using Matlab multi-class ECOC Support Vector Machines (Escalera et al. 2010). The multi-class SVM was trained using a Radial Basis Function (RBF). Even when the linear kernel could be useful for performing binary classification, in this case we have a problem of multi-class classification where the separation surface could become very complex requiring a more flexible kernel. The best results were obtained by using the RBF kernel compared to the linear and the polynomial kernel (Ali and Smith-Miles 2006).

Evaluation of classification accuracy is necessary to assess the classifier performance. This evaluation is performed by subdividing the available data into two subsets: one for training and the other for testing the trained model (Blum et al. 1999). Three alternatives are usually considered (Kohavi 1995): *hold-out*, the *k-fold* cross-validation method and the *Leave-one-out* technique. The preferred method is *k-fold* cross-validation as it provides lower variance. Selection of k is in general dependent on the data, however, $k = 10$ is commonly used (Anguita et al. 2012).

The *k-fold* cross-validation is applied by subdividing randomly the input data into k sub-samples that are mutually exclusive and have approximately the same size. The procedure consists in training the model with $k - 1$ sub-samples and then testing it by using the sub-sample excluded during the training phase. Repetition is performed k times using as testing data one of sub-samples and training with the rest.

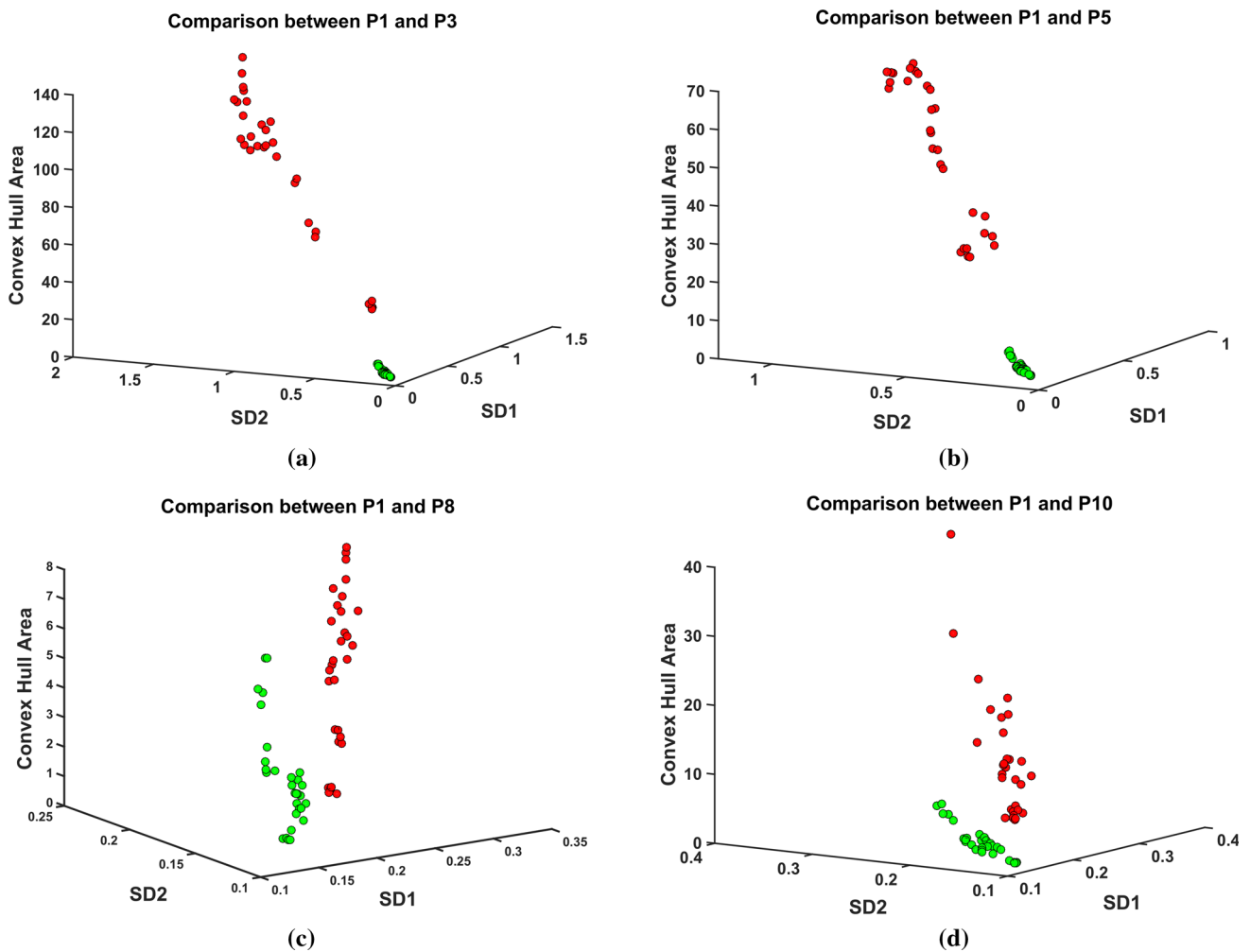


Fig. 15 Results of the Algorithm 2 using signals from the gearbox dataset: cluster structure for load *L1*. **a** Comparison between healthy class *P1* and faulty class *P3*. **b** Comparison between classes *P1* and

P5. **c** Comparison between classes *P1* and *P8*. **d** Comparison between *P1* and *P10* (Color figure online)

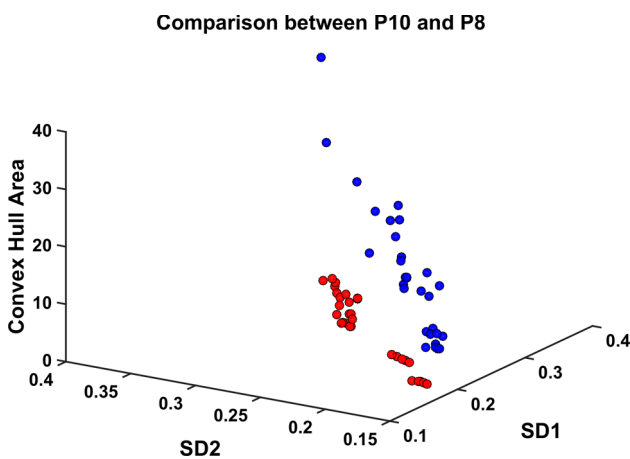


Fig. 16 Comparison between the clusters of classes *P10* (blue) and *P8* (red) for load *L1* with signals extracted from the gearbox dataset, by using the Algorithm 2. Each faulty class forms a cluster in the feature space amenable for classification

Evaluation of the classification results relies on several metrics (Hossin and Sulaiman 2015). Such metrics are defined on the elements of the confusion matrix (Sokolova and Lapalme 2009). The accuracy as well as the area under the Receiver Operator Curve (ROC) are important metrics for assessing the classifier performance.

Fault classification of gearbox

Each of algorithms proposed were validated on four classification experiments. The first experiment was performed by considering the total number of 900 signals in the dataset. The set of signals was partitioned at random into 10 equal sized sub-sets (10-fold) with 90 signals. A single sub-set is used for testing the trained model. The model is trained by using the other 9 sub-sets of data. This process is repeated 10 times, each time considering a different sub-set of data for testing

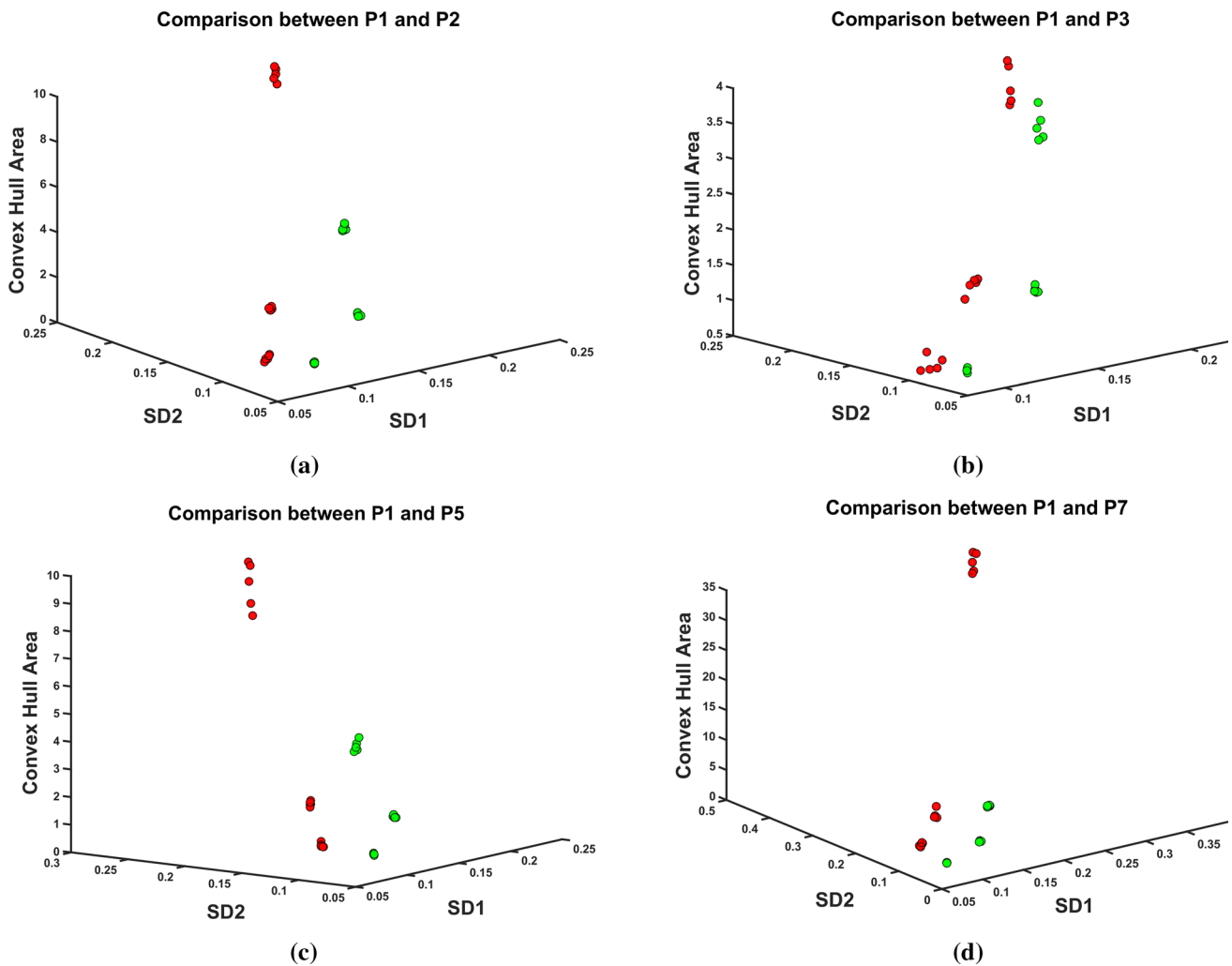


Fig. 17 Results by using the Algorithm 1 applied on vibration signals measured with accelerometer 1 in the bearing dataset. The acquisition was performed by considering load L_2 . The cluster structure for points of the healthy class P_1 is compared to other faulty classes. Green dots are used for points of the healthy class and red dots for the faulty classes.

a Comparison between healthy class and the faulty class P_2 . **b** Cluster structures for the healthy class and faulty class P_3 . **c** Clusters for healthy class and the faulty class P_5 . **d** Representation of the healthy class and the faulty class P_7 . All faulty classes tends to form clusters in the feature space that are amenable for classification

and the rest of data for training. This cross-validation provides 10 trained models. The accuracy of the cross-validation process is the averaged accuracy obtained by the 10 models. The 10-fold cross-validation was also used for three experiments at constant loads L_1 , L_2 and L_3 respectively. The multi-class SVM was trained using a Radial Basis Function (RBF) with $\sigma = 0.10$ in all experiments of gearbox faults classification.

Fault classification of roller bearings

In this case, four classification experiments were considered for validating each of the proposed algorithms. The multi-class classification of the set of seven bearings health conditions was performed by the ECOC Support Vector

Machines algorithm. The roller bearings dataset was composed by 45 signals for each of the 7 bearings health conditions for a total of 315 recorded signals. The SVM model was validated using tenfold cross-validation with a Radial Basis Function (RBF) with $\sigma = 0.07$.

Similarly to the gearbox dataset, for the bearings dataset the 10-fold cross-validation was also performed during three faults classification experiments performed at constant loads L_1 , L_2 and L_3 . The Radial Basis Function (RBF) with $\sigma = 0.18$ was used in the experiments at constant load.

Evaluation of classifier performance

The standard method for evaluating the classifier performance consists in calculating the confusion matrix obtained

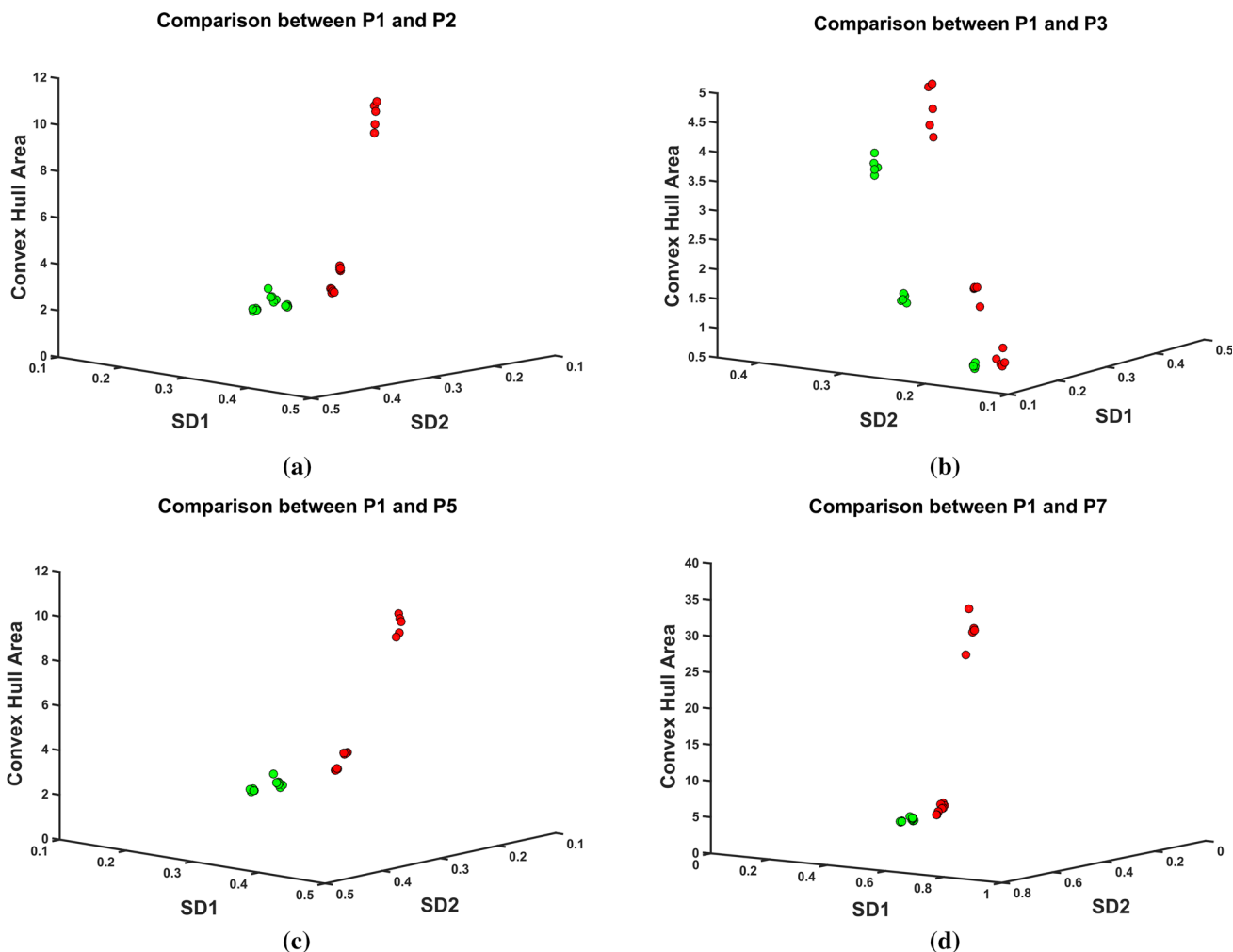


Fig. 18 Results by using the Algorithm 2 applied on vibration signals measured with accelerometer 1 in the bearing dataset. The load of the mechanical system is $L2$ and the features were extracted using the Algorithm 2. The comparison is performed between the healthy class and the faulty classes. In all cases both clusters of points could be separated

using a single surface. **a** Comparison between healthy class and the faulty class $P2$. **b** Clusters for healthy class and the faulty class $P3$. **c** Cluster for the normal class and the faulty class $P5$. **d** Comparison between cluster structures for class $P1$ and the faulty class $P7$

from the classifier using data from the test set (Sokolova and Lapalme 2009; Tharwat 2018). Using the information included in the confusion matrix enables calculation of performance metrics for the classifier as well as the Receiver Operator Curve (ROC) (Landgrebe and Duin 2008). In the case of a multi-class classifier considering K classes C_k with $k \in \{1, 2, \dots, K\}$, the confusion matrix has dimension $K \times K$ where each of the columns represents the target class denoted C_k and each of the rows represents the estimated class obtained as the output of the classifier. This output class is denoted \hat{C}_k . Examples of the confusion matrix are shown in Fig. 21. The numbers along the main diagonal represents the number of examples correctly classified for each of the classes. These values are known as true positives for each of the classes and they are denoted as TP_k . Concerning the mis-

classified samples, the false negatives for class k are denoted FN_k and it represents the samples that were not recognized as belonging to the class k . This quantity is obtained by adding along each column the cells that are different from the diagonal cell. Similarly, the false positives for class k are denoted as FP_k and it represents the samples that were incorrectly assigned to class k . Their calculation is performed by adding along the rows, the cells that are different from the diagonal cell. The true negatives for class k are denoted TN_k and it represents the samples with target labels different from k that were correctly classified as belonging to an output class different from k . Their calculation is attained by performing the addition of cells in a sub-matrix excluding the row and column k . Some of the performance metrics are presented in Table 3.

Table 3 Performance metrics for a multi-class classifier

Metrics	Formula
Accuracy	$ACC_k = \frac{TP_k + TN_k}{TP_k + TN_k + FP_k + FN_k}$
Error rate	$ERR_k = \frac{FP_k + FN_k}{TP_k + TN_k + FP_k + FN_k}$
Sensitivity	$SN_k = \frac{TP_k}{TP_k + FN_k}$
Specificity	$SP_k = \frac{TN_k}{TN_k + FP_k}$
False negative rate	$FNR_k = \frac{FN_k}{FN_k + TP_k}$
False positive rate	$FPR_k = \frac{FP_k}{FP_k + TN_k}$
Precision	$P_k = \frac{TP_k}{TP_k + FP_k}$
F1-score	$F1 - S = \frac{2P * SN}{P + SN}$

Results

Fault classification results for the algorithm 1

Gearbox fault classification

The results concerning all vibration signals in the gearbox dataset attained a classification accuracy of 91.8 %. In this experiment, the lag value was set as $\tau_{1map} = 24$. The Receiver Operator Curve (ROC) for results obtained during the classification is shown in Fig. 19a. The highest accuracy of classification is attained by classes $P1$, $P2$, $P3$ and $P5$ which are healthy and incipient faults. The accuracy of faults $P6$, $P9$ and $P4$ is in the middle and finally, the lower accuracy is attained by faults $P7$, $P8$ and $P10$.

Results for the experiments concerning the gearbox fault classification at constant load attained classification accuracies up to 95.7% (load $L1$), 96.0% (load $L2$) and 99.3% (load $L3$). In each experiment, the lag value was also set as $\tau_{1map} = 24$. The accuracy obtained at load $L3$ is the highest with respect to the other loads. The confusion matrix obtained at load $L3$ using the Algorithm 1 is shown in Fig. 21a. The lower right cell presents the average accuracy corresponding to 99.3% and the average error of 0.7%. The bottom row of the confusion matrix reports the specificity and false negative rate for each of the classes while the rightmost column reports the sensitivity and false positive rate. The cells of the matrix includes the number of examples obtained as result of the cross-validation process. These values are used for calculating several performance metrics as explained in “Evaluation of classifier performance” section.

Results concerning the performance metrics obtained from the confusion matrix obtained in the experiment at load $L3$ are presented in Table 4. The second column represents the false negative rate. Classes $P8$ and $P10$ have a FNR of 3.3%, the rest of classes have a value of 0.0%. The third column shows the FPR with values of 3.2% for classes $P4$ and $P5$, while the rest of classes have a value of 0.0%. The

sensitivity is presented in the fourth column with values of 96.8% for classes $P4$ and $P5$, while the rest of classes have a value of 100%. The specificity for classes $P8$ and $P10$ is 99.6% while the rest of classes have a value of 100%.

The resultant ROC for load $L3$ is shown in Fig. 20a. The performance of the classifier shows two groups: classes $P8$ and $P10$ have the lowest performance of 96.8% while the rest of classes attain a value of 100%.

Bearings fault classification

The first classification experiment on bearings faults considered all the 315 signals in the dataset. The Algorithm 1 was used with the lag parameter set as $lag_{1map} = 25$. The Radial Basis Function (RBF) parameter was set as $\sigma = 0.07$ for training the multi-class SVM. The total classification accuracy was 91.4%. Results concerning the ROC are shown in Fig. 19c.

The second fault bearing classification experiment was performed at constant load by using features from the Algorithm 1. The lag parameter was set as $\tau_{1map} = 25$. A Radial Basis Function (RBF) with $\sigma = 0.18$ was used for training the multi-class SVM. The total accuracy obtained was 93.3% (load $L1$), 98.1% (load $L2$) and 98.1% (load $L3$). Results in terms of the confusion matrix are shown in Fig. 22a for load $L2$. The classes $P1$ and $P7$ attained the higher FNR value of 6.7% and the lowest specificity of 93.3%. The specificity for the rest of classes is 100%. Classes $P4$ and $P6$ attained the higher FPR value of 6.3% and the lowest sensitivity with a value of 93.8%. The the rest of classes attained a sensitivity of 100%. Results concerning ROC are shown in Fig. 20c considering the load $L2$. Classes $P4$ and $P6$ attained the lower performance of 93.3% while rest of classes attained a performance of 100%.

Fault classification results for the algorithm 2

Gearbox fault classification

The classification accuracy is 91.8% for the first experiment considering all signals in the gearbox dataset. The parameter corresponding to the lag value is set as $\tau_{peak} = 10$. This value is different from the optimal value used for the case of constant load (this selection will be further discussed in “Accuracy as function of the lag” section). The Receiver Operator Curve for this experiment shows that classification accuracies higher than 95% are attained by classes $P1$, $P3$, $P4$, $P5$ as shown in Fig. 19b. The lowest accuracy of classification is attained by classes $P6$, $P7$, $P8$, $P9$ and $P10$ with classification accuracies between 80% and 90%.

The classification of gearbox faults for three experiments performed at constant load, provided classification accuracies of 99.3% ($L3$), 94.3% ($L2$) and 92.3% ($L1$). The lag

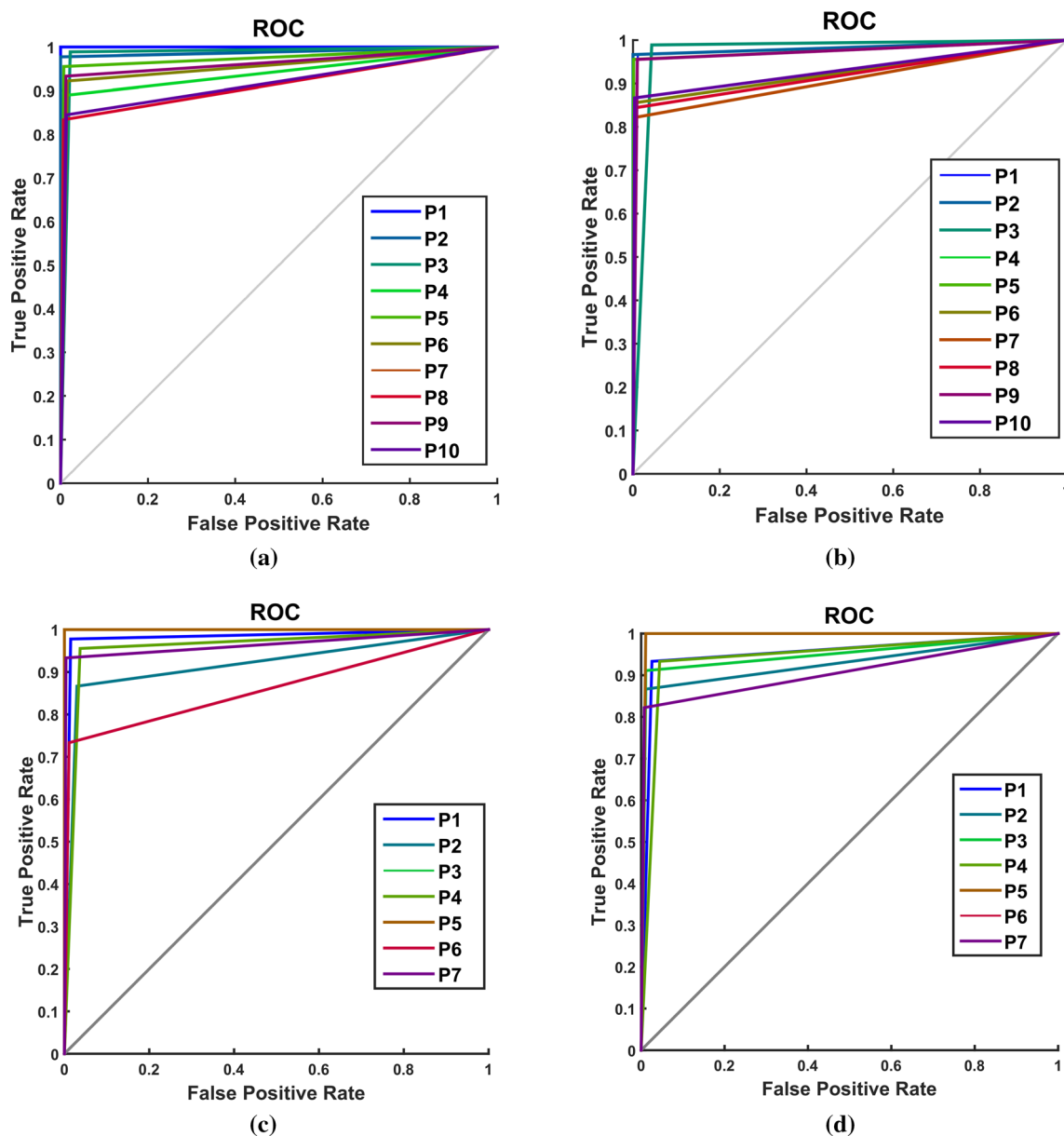


Fig. 19 Receiver Operator Characteristic (ROC) by using SVM classifier. **a** ROC for the first experiment of faults classification using features from Algorithm 1. All signals in the gearbox dataset were considered. **b** ROC for the experiment using features extracted from the Algorithm 2. All signals in the gearbox dataset were considered. **c** Results concern-

ing the ROC by using all signals in the bearings dataset. The features extracted from the Algorithm 1 were used for training and validating the SVM. **d** ROC from the SVM model trained and validated using features extracted from Algorithm 2 using all signals in the bearing dataset

value was set to $\tau_{peak} = 30$. The results of the confusion matrix for load $L3$ are shown in Fig. 21b. In Table 5, the metrics calculated from the confusion matrix are presented. The second column represents the false negative rate with higher values attained by $P3$ and $P8$ with values of 3.3%. The false positive rate is shown in third column with higher values attained by class $P6$ with a value of 6.3%. The sensitivity is shown in the third column with a lower value of 94.8% attained by class $P6$. The rest of classes have the high-

est value of 100%. The specificity is shown in the fifth column with a lowest value for classes $P3$ and $P8$ with 99.6%. The rest of classes attain a value of 100%.

Results concerning the ROC for the SVM classifier trained with features extracted by using the Algorithm 2 from vibration signals of the gearbox at load $L3$ are shown in Fig. 20b. The lowest performance of the classifier corresponds to classes $P3$ an $P8$ with an accuracy of 99.6%. The accuracy for the rest of classes is 100%.

Table 4 Performance metrics obtained with Algorithm 1 for load $L3$ in the gearbox dataset

Class	FNR	FPR	Sensitivity	Specificity
$P1$	0.000	0.000	1.000	1.000
$P2$	0.000	0.000	1.000	1.000
$P3$	0.000	0.000	1.000	1.000
$P4$	0.000	0.032	0.968	1.000
$P5$	0.000	0.032	0.968	1.000
$P6$	0.000	0.000	1.000	1.000
$P7$	0.000	0.000	1.000	1.000
$P8$	0.033	0.000	1.000	0.996
$P9$	0.000	0.000	1.000	1.000
$P10$	0.033	0.000	1.000	0.996

Bearing fault classification

The first classification experiment of roller bearing faults considered all 315 signals in the dataset. The Algorithm 2 was used with the lag parameter set as $\tau_{1map} = 25$. The SVM classification model was trained using a Radial Basis Function (RBF) with $\sigma = 0.09$. The total classification accuracy was 89.8%. The lowest classification accuracy is attained by $P7$ while the highest accuracy is obtained by classes $P6$ and $P5$. Results expressed in terms of ROC are shown in Fig. 19d.

Three additional bearing fault classification experiments were performed at constant load with features by using the Algorithm 2. The lag parameter was set as $\tau_{1map} = 25$. A Radial Basis Function (RBF) with $\sigma = 0.15$ was used for training SVM model. A total accuracy of 87.6% was obtained with $L1$, the accuracy for load $L2$ was 100.0% while for $L3$ the accuracy was 94.3%. Results concerning the confusion matrix using signals at load $L2$ are shown in Fig. 22b. The values of specificity are 100% for all classes while the FNR is 0.0%. Similarly, the sensitivity is 100% and the FPR is 0.0% for all classes. The ROC for this experiment considering load $L2$ is shown in Fig. 20d. In this experiment, the Algorithm 2 enables a perfect classification with only three features.

Accuracy as function of the lag

The effect of the lag selection on the accuracy of the gearbox faults classification was investigated. The accuracy is estimated and plotted as a function of the lag for the Algorithm 1. The analysis concerning τ_{1map} was performed considering all vibration signals as well as for loads $L1$, $L2$ and $L3$. Results are shown in Fig. 23a. The highest accuracy is attained for load $L3$. There is a global maximum at $\tau_{1map} = 24$ with an accuracy of 99.3% and two local maximums at $\tau_{1map} = 45$ and $\tau_{1map} = 62$ with accuracies of 98%. The maximum accuracy attained for load $L2$ is 97.3%

at $\tau_{1map} = 40$ and for load $L1$ the maximum accuracy is 96.3% at $\tau_{1map} = 45$. The accuracy of classification considering all vibration signals is lower with a maximum of 91.8% at $\tau_{1map} = 24$. In general, the higher accuracy for this algorithm is obtained considering values for the lag between 20 and 40 samples.

The lag selection analysis was also performed for Algorithm 2 (τ_{2peak}) considering all gearbox vibration signals as well as for loads $L1$, $L2$ and $L3$. Results are shown in Fig. 23b. The highest accuracy is attained for load $L3$. There is a global maximum at $\tau_{2peak} = 30$ with an accuracy of 99.3% and two local maximums at $\tau_{2peak} = 10$ with an accuracy of 98% and $\tau_{2peak} = 65$ with an accuracy of 98.7%. The maximum accuracy attained for load $L2$ is 97% at $\tau_{2peak} = 35$ and at $\tau_{2peak} = 65$. The load $L1$ attained a maximum accuracy of 95.0% at $\tau_{2map} = 15$. The accuracy of classification considering all vibration signals is lower with a maximum of 91.8% at $\tau_{2peak} = 10$. In general the higher accuracy for this algorithm is obtained considering values for the lag between 5 and 35 samples.

An experiment concerning the effect of the lag selection on the accuracy of the bearing fault classification was also performed. The variation of the accuracy of classification for the Algorithm 1 as a function of the lag (τ_{1map}) was studied for the bearing dataset considering all signals as well as for loads $L1$, $L2$ and $L3$. Results of this experiment are shown in Fig. 23c. Higher accuracies are obtained for loads $L2$ and $L3$ with values of 99% and 98.1% respectively. The classification accuracy is lower considering all vibration signals in the roller bearings dataset. However, the accuracy is higher than 90%. The larger classification accuracies are obtained for lag values between 15 and 80 samples. The variation of the accuracy on bearing faults classification was also investigated for the Algorithm 2. When the features for fault classification are extracted by using the Algorithm 2, the resultant accuracy is shown in Fig. 23d. The highest accuracy of 100% is attained by the Algorithm 2 considering load $L2$ when $\tau_{2peak} = 25$. The maximum accuracy considering the load $L3$ is 97.14% when $\tau_{2peak} = 35$. Algorithm 2 attains an accuracy of classification of 90.5% considering the load $L1$ when $\tau_{2peak} = 35$. The maximum accuracy with all signals in the bearing dataset is 89.5% when $\tau_{2peak} = 35$. In general, the best classification results with Algorithm 2 are considering a lag value between 15 and 40 samples when dealing with faults in the bearing dataset.

Algorithms performance comparison

A comparison in terms of the confusion matrix is shown in Figs. 21 and 22. Figure 21 reports the confusion matrices obtained by using the Algorithm 1 and Algorithm 2 for the gearbox dataset at load $L3$. Both confusion matrices were obtained during the cross-validation process and reports the

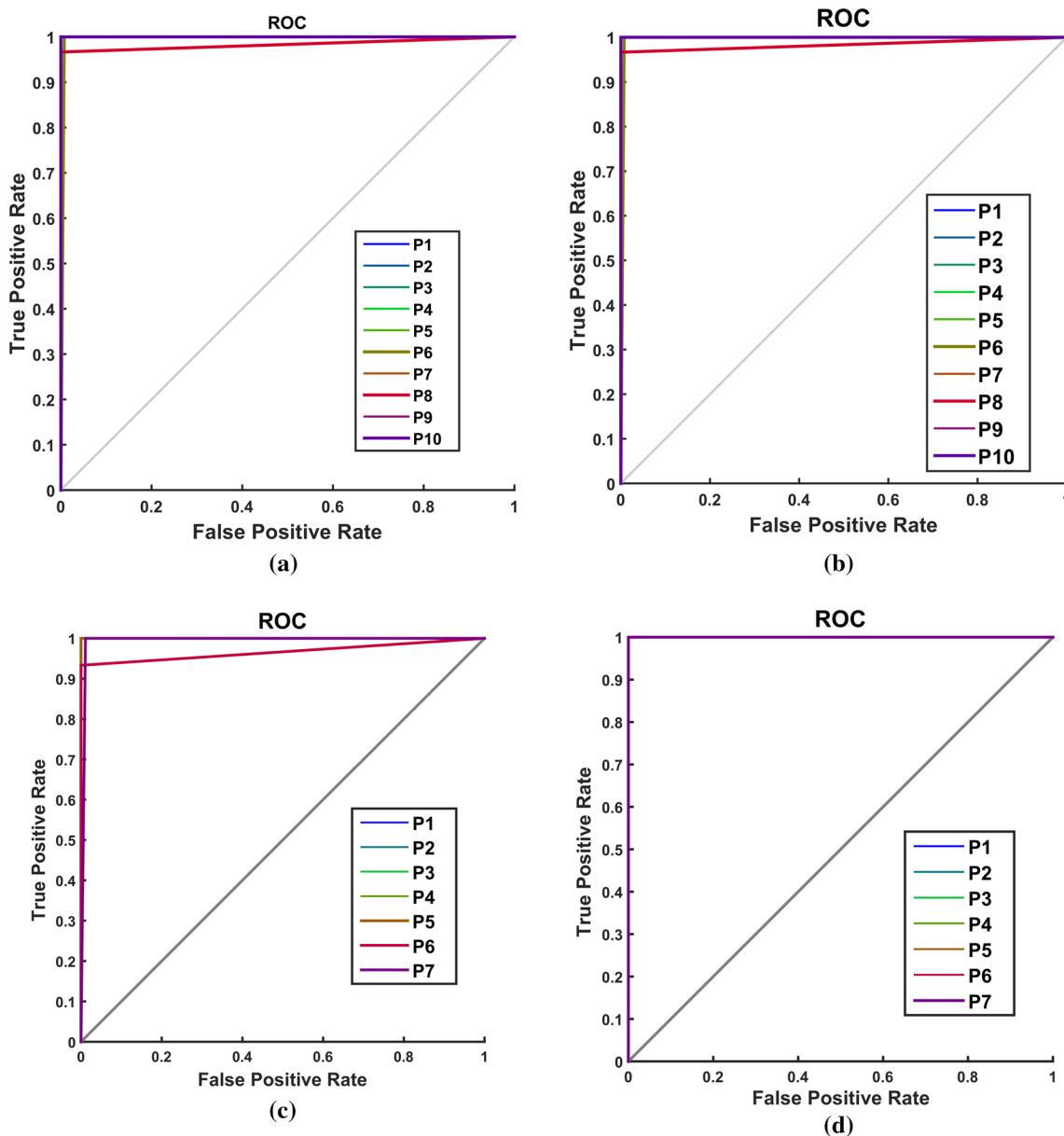


Fig. 20 Results concerning the ROC for the SVM model using vibration signals acquired at a constant load. **a** ROC for the model trained with features extracted by using the Algorithm 1. The vibration signals correspond to the gearbox dataset at load *L3*. The average accuracy is 99.3%. **b** ROC for the model trained with features extracted by using the Algorithm 2. The vibration signals correspond to the gearbox dataset at

load *L3*. The average accuracy is 99.3%. **c** ROC by using signals in the roller bearings dataset at load *L2* and features extracted by using the Algorithm 1. The average accuracy is 98.1%. **d** ROC by using signals in the bearings dataset at load *L2*, and features extracted by using the Algorithm 2. The average accuracy is 100%

average accuracy as well as other metrics calculated including FNR, FPR, sensitivity and specificity. These results were discussed in “Results” section. In terms of the average accuracy the results are similar for both algorithms corresponding to 99.3%. Figure 22 shows the confusion matrices comparing Algorithm 1 and Algorithm 2 for the bearing dataset at load *L2*. The average accuracy attained by algorithm 1 is 98.1% while the Algorithm 2 has an average accuracy of 100%.

Results concerning the classification accuracy for gearbox faults is presented in Table 6, when the classification is performed based on all vibration signals. The results are almost similar for both algorithms as the result provided by Algorithm 1 is 91.33% while the result provided by Algorithm 2 is 91.78%. Other results presented are the average error of 8.67%, the average specificity of 99.04%, the average precision of 91.96%. The F1-score is 93.10% while the average

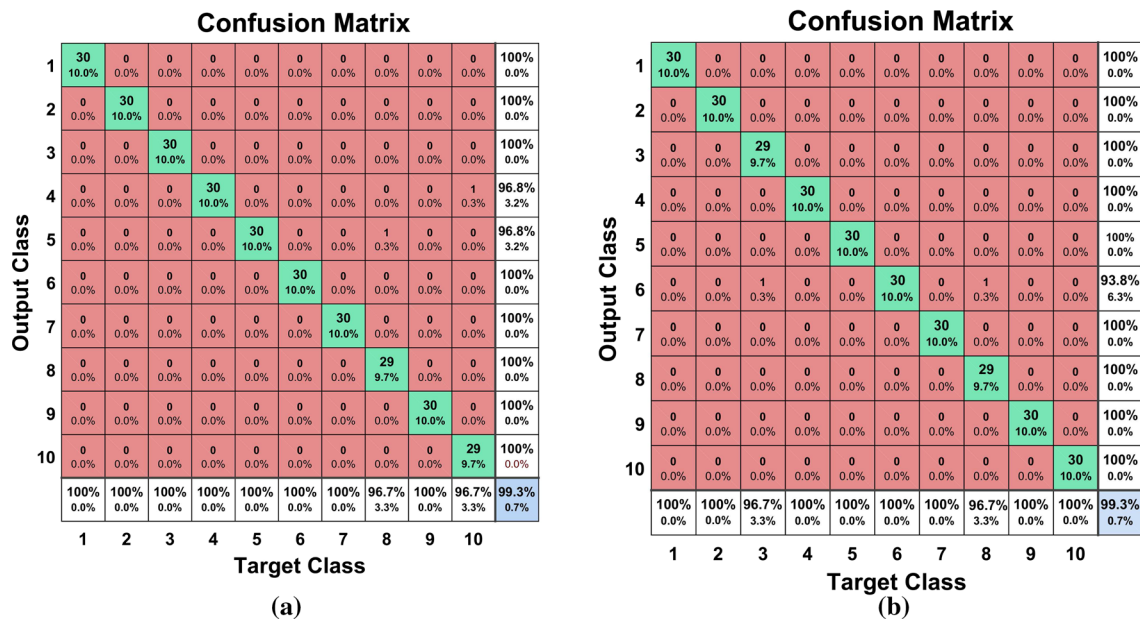


Fig. 21 **a** Results for gearbox faults classification using features extracted by using the Algorithm 1 at load *L3*. The accuracy is 99.3%. **b** Results for gearbox faults classification at load *L3* by using features from the Algorithm 2. The classification accuracy is 99.3%

Table 5 Performance metrics obtained with Algorithm 2 with signals at load *L3* in the gearbox dataset

Class	FNR	FPR	Sensitivity	Specificity
<i>P1</i>	0.000	0.000	1.000	1.000
<i>P2</i>	0.000	0.000	1.000	1.000
<i>P3</i>	0.033	0.000	1.000	0.996
<i>P4</i>	0.000	0.000	1.000	1.000
<i>P5</i>	0.000	0.000	1.000	1.000
<i>P6</i>	0.000	0.063	0.948	1.000
<i>P7</i>	0.000	0.000	1.000	1.000
<i>P8</i>	0.033	0.000	1.000	0.996
<i>P9</i>	0.000	0.000	1.000	1.000
<i>P10</i>	0.000	0.000	1.000	1.000

FPR is only 0.96%. The best results of the Algorithm 1 are attained at load *L3* corresponding to an accuracy of 99.33%, the average error of 0.67%, an average specificity of 99.93% and an average FPR of 0.07%. The results obtained using the Algorithm 2 for the case concerning the fault classification using all the signals is an average accuracy of 91.78%, which is very close to the results provided by Algorithm 1. The maximum average accuracy for the Algorithm 2 is similar to the accuracy obtained by Algorithm 1, however there are differences concerning the accuracy of classification for each of the classes as shown in Fig. 21. Accuracy for the Algorithm 1 at loads *L1* and *L2* are 95.67% and 96.00% respectively while the accuracy with Algorithm 2 for these loads are 92.33% and 94.33% respectively.

The fault classification results for bearing are shown in Table 7. The accuracy attained by Algorithm 1 (91.43%) is slightly higher with respect to the accuracy of 89.84% provided by the Algorithm 2. Results for the bearing dataset at a constant load are highest for load *L2* and Algorithm 2, which attains an accuracy of 100% while the highest accuracy for the Algorithm 1 is 98.1%.

Results concerning the accuracy of the Poincaré based classification methods applied to the gearbox dataset are close to results attained by other methods using the same dataset. However, the proposed approach performs the classification by extracting only three features.

One of the methods that uses the same gearbox dataset was reported in Sánchez et al. (2018). In this research, features in the time domain are used for classification. According to the ranking of features, at least four features are required for overcoming 91% of accuracy for both K Nearest Neighbors (KNN) and Random Forests (RF) classifiers, while in this work, with only three features the accuracy is higher than 91%. Other method using the same gearbox dataset has been reported in Cerrada et al. (2016a). In that research the classification was performed based on a subset 12% of 811 features. Several machine learning techniques were compared and a classification accuracy of 98% was attained by using Random Forest classifiers. The gearbox dataset was also analyzed in Li et al. (2015b). In that research, the machine learning technique was Multimodal Deep Support Vector Classification (MDSVC) and the trained model provided a classification accuracy of 97.08%. Additional research was reported in Pacheco et al. (2016) where a set of multimodal homolo-

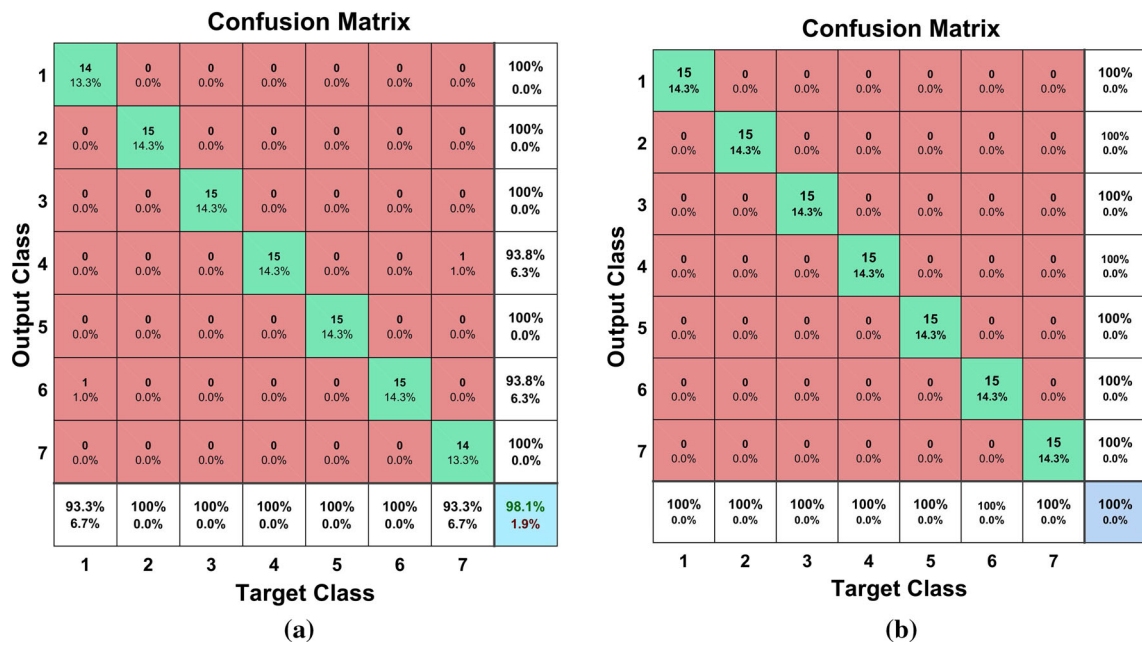


Fig. 22 a Results of the confusion matrix by using Algorithm 1 for load *L2*. In this experiment the features are extracted from the bearing dataset. The total accuracy is 98.1% **b** Results concerning the fault

classification with features by using the Algorithm 2 at load *L2*. The features were extracted from vibration signals from the bearing dataset. The total accuracy is 100.0%

Table 6 Gearbox classification accuracy for the proposed algorithms (in percent)

Algorithm	Load	Accuracy	Error	Specificity	Precision	F1-Score	FPR
Algorithm 1	All	91.33	8.67	99.04	91.96	93.10	0.96
	<i>L1</i>	95.67	4.33	99.52	95.92	95.64	0.48
	<i>L2</i>	96.00	4.00	99.56	96.64	96.13	0.44
	<i>L3</i>	99.33	0.67	99.93	99.35	99.33	0.07
Algorithm 2	All	91.78	8.22	99.09	92.70	91.90	0.91
	<i>L1</i>	92.33	7.67	99.15	92.62	92.15	0.85
	<i>L2</i>	94.33	5.67	99.37	94.59	94.35	0.63
	<i>L3</i>	99.33	0.67	99.93	99.38	99.33	0.07

Table 7 Roller bearings classification accuracy for the proposed algorithms (in percent)

Algorithm	Load	Accuracy	Error	Specificity	Precision	F1-Score	FPR
Algorithm 1	All	91.43	8.57	98.57	91.83	91.36	1.43
	<i>L1</i>	93.33	6.67	98.89	94.15	93.10	1.11
	<i>L2</i>	98.10	1.90	99.68	98.21	98.09	0.32
	<i>L3</i>	98.10	1.90	99.68	98.21	98.09	0.32
Algorithm 2	All	89.84	10.16	98.31	90.43	89.85	1.69
	<i>L1</i>	87.62	12.38	97.94	88.05	87.34	2.06
	<i>L2</i>	100.00	0.00	100.00	100.00	100.00	0.00
	<i>L3</i>	94.29	5.71	99.05	94.81	91.39	0.95

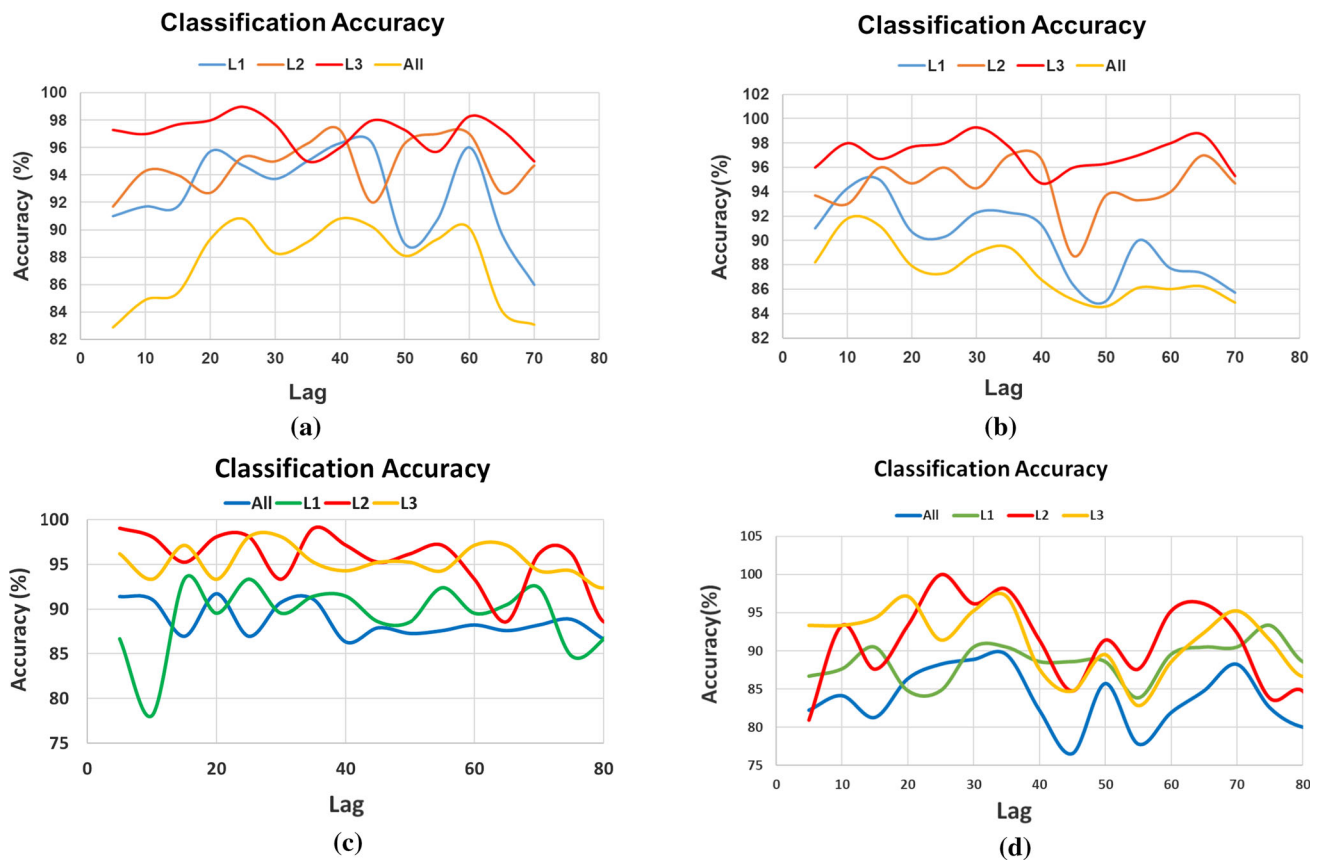


Fig. 23 **a** Accuracy of classification by using features extracted from Algorithm 1 as a function of the lag for the gearbox dataset considering all signals as well as loads $L1$, $L2$ and $L3$. **b** Variation of the accuracy as a function of the lag between peaks for the Algorithm 2 considering

the vibration signal in the gearbox dataset. **c** Variation of the accuracy of classification for the Algorithm 1 as a function of the lag considering vibration signals in the bearing dataset. **d** Variation of the accuracy on bearing faults classification using the Algorithm 2

gous features were used. In that research, the set of features extracted from the vibration signals included time, frequency and time-frequency features. A set of 330 features were selected from a total 817 features and several algorithms were compared considering genetic algorithms, entropy based algorithms, linear discriminant, principal components, random forests and non-negative matrix factorization. The best results were obtained by using Random Forest algorithms. The method using symbolic dynamics reported in Medina et al. (2019) performed the faults classification using the same gearbox dataset with a set of features corresponding to an array with twelve elements. The highest accuracy was 99.78% by considering all signals in the dataset at constant load to obtain accuracies up to 100.0%. This method also extracts the features from the Poincaré plot and the main difference concerning the results is related to the case considering all signals in the dataset. The accuracy attained with the method proposed in this paper is lower, however, only three features are considered and the accuracy is higher than 91%.

Concerning the bearings fault classification reported in Li et al. (2016b), the authors report a method for fault classification by using a Gaussian-Bernoulli deep Boltzmann machine (GDBM). Such a method is able to learn from statistical features extracted from vibration signals in time-domain, frequency-domain and time-frequency domain. The reported average accuracy for the bearing dataset is 79.98% with a maximum value of 91.75%. This accuracy is close to the value obtained by the algorithms reported in this document considering all signals and using only three features. In contrast, we have shown that using signals acquired at load $L3$ the average accuracy can be as high as 100% by considering only three features.

Conclusions

In this paper, we proposed a new feature extraction approach using Poincaré plots from vibration signals. The proposed method is useful for gearbox and roller bearing fault classification. The cluster structure generated in the feature space

was analyzed showing that such structure is amenable for classification tasks. Two algorithms were presented for feature extraction. Both extraction methods rely on using only three features for performing the fault classification. The features extracted describe the shape of the Poincaré plot. Two of the features are the classical $SD1$ and $SD2$ dispersion measures and the third feature is the area of the convex hull shape of the Poincaré plot.

The proposed algorithms attained a highest accuracy of 99.3 % when the fault classification is performed by using the gearbox dataset. Similarly, for the roller bearing dataset the highest classification accuracy is 100%. In the previous cases, the vibration signals are considered at a constant load.

The case concerning the feature extraction using all vibration signals in the datasets provided the following results: the accuracy of Algorithm 1 is higher than 91.3% for both gearbox and roller bearing fault classification. In this case, the Algorithm 2 has an accuracy of 91.8% for classification of gearbox faults and 89.8% for classification of roller bearing faults. The accuracy of both algorithms is higher than 90 % for lag values selected according to the procedure explained in “Criteria for the lag selection in Poincaré plots” section. The Algorithm 2 has a higher computational cost than the Algorithm 1. Such computational cost is necessary for extracting the signal peaks from the input vibration signal. The peaks extraction stage is equivalent to perform a non-uniform subsampling of the input signal. This intermediate result represented by a new time series, retains relevant information concerning the faults. In consequence, further research is needed aimed at proposing other feature extraction methods. In that sense, Algorithm 2 open new research avenues that could be explored for gearbox and roller bearing fault classification.

Both algorithms provide features that enables excellent classification accuracy by using multi-class support vector machines. The accuracy of classification is dependent on the load. Concerning the gearbox fault classification experiment, the highest classification accuracy is attained with the highest load corresponding to $L3$ when the Algorithm 1 is used. The highest accuracy for the roller bearing fault classification experiment is attained using the load $L2$. The accuracy of fault classification considering all the vibration signals is lower, however, both algorithms attain an accuracy higher than 89.84% with only three features for fault classification in gearboxes or roller bearings. It is expected that combination with other features could allow increasing the classification accuracy considering all vibration signals independent from the load.

Comparison with respect to other more complex classification methods tested with the same databases shows that the algorithms proposed in this paper attain better classification results considering only three features.

The feature extraction method that we are proposing is simple and accurate. Their application could be extended for extracting features from other types of signals recorded from roller bearings and gearboxes as well as other types of machines such as reciprocating compressors.

The Poincaré plot is susceptible of providing other features that could be useful for rotating machinery fault classification. The future research is aimed at studying the extraction of other features and their possible incorporation within the fault classification procedures.

Fault severity classification in gearboxes (Sun et al. 2016) is also important for assessing their structural health. Application of the proposed algorithms for fault severity classification within a data fusion framework is part of the current research work performed at our laboratory.

References

- Aherwar, A. (2012). An investigation on gearbox fault detection using vibration analysis techniques: A review. *Australian Journal of Mechanical Engineering*, 10(2), 169–183.
- Ali, S., & Smith-Miles, K. A. (2006). A meta-learning approach to automatic kernel selection for support vector machines. *Neurocomputing*, 70(1–3), 173–186.
- Alligood, K. T., Sauer, T. D., & Yorke, J. A. (1997). *Chaos: An introduction to dynamical systems*. Berlin: Springer.
- Aly, M. (2005). Survey on multiclass classification methods. *Neural Networks*, 19, 1–9.
- Anguita, D., Ghelardoni, L., Ghio, A., Oneto, L., & Ridella, S. (2012). The ‘k’ in k-fold cross validation. In *2012 proceedings of the European symposium on artificial neural networks, computational intelligence and machine learning* (pp. 25–27).
- Bajric, R., Sprecic, D., & Zuber, N. (2011). Review of vibration signal processing techniques towards gear pairs damage identification. *International Journal of Engineering and Technology*, 11(4), 124–128.
- Bangalore, P., & Tjernberg, L. B. (2015). An artificial neural network approach for early fault detection of gearbox bearings. *IEEE Transactions on Smart Grid*, 6(2), 980–987.
- Blum, A., Kalai, A., Langford, J. (1999). Beating the hold-out: Bounds for k-fold and progressive cross-validation. In *Proceedings of the twelfth annual conference on computational learning theory* (pp. 203–208). ACM.
- Brennan, M., Palaniswami, M., & Kamen, P. (2001). Do existing measures of poicare plot geometry reflect nonlinear features of heart rate variability? *IEEE Transactions on Biomedical Engineering*, 48(11), 1342–1347.
- Burges, C. J. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2), 121–167.
- Cerrada, M., Macancela, J. C., Cabrera, D., Estupiñan, E., Sánchez, R. V., & Medina, R. (2020). Reciprocating compressor multi-fault classification using symbolic dynamics and complex correlation measure. *Applied Sciences*, 10(7), 2512.
- Cerrada, M., Sánchez, R. V., Pacheco, F., Cabrera, D., Zurita, G., & Li, C. (2016a). Hierarchical feature selection based on relative dependency for gear fault diagnosis. *Applied Intelligence*, 44(3), 687–703.
- Cerrada, M., Zurita, G., Cabrera, D., Sánchez, R. V., Artés, M., & Li, C. (2016b). Fault diagnosis in spur gears based on genetic algorithm

- and random forest. *Mechanical Systems and Signal Processing*, 70, 87–103.
- Cheng, S., Azarian, M. H., & Pecht, M. G. (2010). Sensor systems for prognostics and health management. *Sensors*, 10(6), 5774–5797.
- Cui, L., Qian, Z. (2010). Study on dynamic properties of roller bearing with nonlinear vibration. In *2010 International conference on mechanic automation and control engineering (MACE)* (pp. 2723–2726). IEEE .
- Daw, C. S., Finney, C. E. A., & Tracy, E. R. (2003). A review of symbolic analysis of experimental data. *Review of Scientific Instruments*, 74(2), 915–930.
- De Berg, M., Cheong, O., Van Kreveld, M., & Overmars, M. (2008). *Computational geometry: Introduction*. Berlin: Springer.
- Doguer, T., Strackeljjan, J. (2009). Vibration analysis using time domain methods for the detection of small roller bearing defects. In *SIRM 2009-8th international conference on vibrations in rotating machines* (pp. 23–25).
- Escalera, S., Pujol, O., & Radeva, P. (2010). On the decoding process in ternary error-correcting output codes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(1), 120–134.
- Fraser, A. M., & Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical Review A*, 33(2), 1134.
- Goyal, D., & Pabla, B. (2016). The vibration monitoring methods and signal processing techniques for structural health monitoring: A review. *Archives of Computational Methods in Engineering*, 23(4), 585–594.
- Goyal, D., Choudhary, A., Pabla, B., & Dhami, S. (2020). Support vector machines based non-contact fault diagnosis system for bearings. *Journal of Intelligent Manufacturing*, 31, 1275–1289.
- Hiroaki, E., & Nader, S. (2012). Gearbox simulation models with gear and bearing faults. In *Mechanical engineering*. InTech.
- Hoshi, R. A., Pastre, C. M., Vanderlei, L. C. M., & Godoy, M. F. (2013). Poincaré plot indexes of heart rate variability: Relationships with other nonlinear variables. *Autonomic Neuroscience*, 177(2), 271–274.
- Hossin, M., & Sulaiman, M. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining and Knowledge Management Process*, 5(2), 1–11. <https://doi.org/10.5121/ijdkp.2015.5201>.
- Huang, W., Kong, F., & Zhao, X. (2018). Spur bevel gearbox fault diagnosis using wavelet packet transform and rough set theory. *Journal of Intelligent Manufacturing*, 29(6), 1257–1271.
- Igba, J., Alemzadeh, K., Durugbo, C., & Eiriksson, E. T. (2016). Analysing rms and peak values of vibration signals for condition monitoring of wind turbine gearboxes. *Renewable Energy*, 91, 90–106.
- Janjarasjitt, S., Ocak, H., & Loparo, K. (2008). Bearing condition diagnosis and prognosis using applied nonlinear dynamical analysis of machine vibration signal. *Journal of Sound and Vibration*, 317(1), 112–126.
- Jáuregui, J. C. (2011). Phase diagram analysis for predicting nonlinearities and transient responses. In *Recent advances in vibrations analysis*. InTech.
- Jedliński, Ł., & Jonak, J. (2015). Early fault detection in gearboxes based on support vector machines and multilayer perceptron with a continuous wavelet transform. *Applied Soft Computing*, 30, 636–641.
- Kahraman, A. (1992). On the response of a preloaded mechanical oscillator with a clearance: Period-doubling and chaos. *Nonlinear Dynamics*, 3(3), 183–198.
- Kahraman, A., & Blankenship, G. W. (1997). Experiments on nonlinear dynamic behavior of an oscillator with clearance and periodically time-varying parameters. *Journal of Applied Mechanics*, 64(1), 217–226.
- Kedadouche, M., Thomas, M., Tahan, A., & Guilbault, R. (2015). Nonlinear parameters for monitoring gear: Comparison between lempel-ziv, approximate entropy, and sample entropy complexity. *Shock and Vibration*. <https://doi.org/10.1155/2015/959380>.
- Kohavi, R., et al. (1995). *A study of cross-validation and bootstrap for accuracy estimation and model selection* (Vol. 14, pp. 1137–1145). Montreal: Ijcai.
- Landgrebe, T. C., & Duin, R. P. (2008). Efficient multiclass roc approximation by decomposition via confusion matrix perturbation analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5), 810–822.
- Li, B., Zhang, X., & Wu, T. (2018). Measurement of instantaneous angular displacement fluctuation and its applications on gearbox fault detection. *ISA Transactions*, 74, 245–260.
- Li, C., Liang, M., & Wang, T. (2015a). Criterion fusion for spectral segmentation and its application to optimal demodulation of bearing vibration signals. *Mechanical Systems and Signal Processing*, 64, 132–148.
- Li, C., Sanchez, R. V., Zurita, G., Cerrada, M., Cabrera, D., & Vásquez, R. E. (2015b). Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis. *Neurocomputing*, 168, 119–127.
- Li, C., de Oliveira, J. V., Cerrada, M., Pacheco, F., Cabrera, D., Sanchez, V., et al. (2016a). Observer-biased bearing condition monitoring: From fault detection to multi-fault classification. *Engineering Applications of Artificial Intelligence*, 50, 287–301.
- Li, C., Sánchez, R. V., Zurita, G., Cerrada, M., & Cabrera, D. (2016b). Fault diagnosis for rotating machinery using vibration measurement deep statistical feature learning. *Sensors*, 16(6), 895.
- Liqin, W., Li, C., Dezhi, Z., & Le, G. (2008). Nonlinear dynamics behaviors of a rotor roller bearing system with radial clearances and waviness considered. *Chinese Journal of Aeronautics*, 21(1), 86–96.
- Liu, Z., He, Z., Guo, W., & Tang, Z. (2016). A hybrid fault diagnosis method based on second generation wavelet de-noising and local mean decomposition for rotating machinery. *ISA Transactions*, 61, 211–220.
- Loutridis, S. (2008). Self-similarity in vibration time series: Application to gear fault diagnostics. *Journal of Vibration and Acoustics*, 130(3), 031004.
- Luo, J., Huang, J., & Li, H. (2020). A case study of conditional deep convolutional generative adversarial networks in machine fault diagnosis. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-020-01579-w>.
- Luts, J., Ojeda, F., Van de Plas, R., De Moor, B., Van Huffel, S., & Suykens, J. A. (2010). A tutorial on support vector machine-based methods for classification problems in chemometrics. *Analytica Chimica Acta*, 665(2), 129–145.
- Medina, R., Alvarez, X., Jadán, D., Cerrada, M., Sánchez, R.V., & Macancela, J.C. (2017). Poincaré plot features from vibration signal for gearbox fault diagnosis. In *2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM)* (pp. 1–6). <https://doi.org/10.1109/ETCM.2017.8247500>.
- Medina, R., Macancela, J. C., Lucero, P., Cabrera, D., Cerrada, M., Sánchez, R. V., et al. (2019). Vibration signal analysis using symbolic dynamics for gearbox fault diagnosis. *The International Journal of Advanced Manufacturing Technology*, 104(5–8), 2195–2214.
- Mevel, B., & Guyader, J. L. (2008). Experiments on routes to chaos in ball bearings. *Journal of Sound and Vibration*, 318(3), 549–564.
- Motahar, H., Samani, F. S., & Molaie, M. (2016). Nonlinear vibration of the bevel gear with teeth profile modification. *Nonlinear Dynamics*, 83(4), 1875–1884.
- Moumene, I., & Ouelaa, N. (2012). Gears and bearings combined faults detection using hilbert transform and wavelet multiresolution analysis. In T. Fakhfakh, W. Bartelmus, F. Chaari, R. Zimroz,

- & M. Haddar (Eds.), *Condition monitoring of machinery in non-stationary operations* (pp. 319–328). Berlin: Springer.
- Pacheco, F., de Oliveira, J. V., Sánchez, R. V., Cerrada, M., Cabrera, D., Li, C., et al. (2016). A statistical comparison of neuroclassifiers and feature selection methods for gearbox fault diagnosis under realistic conditions. *Neurocomputing*, *194*, 192–206.
- Patil, M., Mathew, J., & RajendraKumar, P. (2008). Bearing signature analysis as a medium for fault detection: A review. *Journal of Tribology*, *130*(1), 014001.
- Randall, R. B., & Antoni, J. (2011). Rolling element bearing diagnostics—A tutorial. *Mechanical Systems and Signal Processing*, *25*(2), 485–520.
- Repo, J. (2010). Condition monitoring of machine tools and machining processes using internal sensor signals. Ph.D thesis, School of Industrial Engineering and Management, KTH Royal Institute of Technology, Sweden.
- Rocha, R., Andruccioli, G. L., & Medrano-T, R. O. (2010). Experimental characterization of nonlinear systems: A real-time evaluation of the analogous chua's circuit behavior. *Nonlinear Dynamics*, *62*(1–2), 237–251.
- Sait, A. S., & Sharaf-Eldeen, Y. I. (2011). A review of gearbox condition monitoring based on vibration analysis techniques diagnostics and prognostics. *Rotating Machinery, Structural Health Monitoring, Shock and Vibration*, *5*, 307–324.
- Sánchez, R. V., Lucero, P., Macancela, J. C., Cerrada, M., Vásquez, R. E., & Pacheco, F. (2017). Multi-fault diagnosis of rotating machinery by using feature ranking methods and svm-based classifiers. In *2017 International conference on sensing, diagnostics, prognostics, and control (SDPC)* (pp. 105–110). IEEE.
- Sánchez, R. V., Lucero, P., Vásquez, R. E., Cerrada, M., Macancela, J. C., & Cabrera, D. (2018). Feature ranking for multi-fault diagnosis of rotating machinery by using random forest and knn. *Journal of Intelligent and Fuzzy Systems*, *34*(6), 3463–3473.
- Sharma, V., & Parey, A. (2016). A review of gear fault diagnosis using various condition indicators. *Procedia Engineering*, *144*, 253–263.
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, *45*(4), 427–437.
- Soleimani, A., & Khadem, S. (2015). Early fault detection of rotating machinery through chaotic vibration feature extraction of experimental data sets. *Chaos, Solitons and Fractals*, *78*, 61–75.
- Sun, C., Wang, P., Yan, R., & Gao, R. X. (2016). A sparse approach to fault severity classification for gearbox monitoring. In *2016 19th international conference on information fusion (FUSION)* (pp. 2303–2308).
- Sun, Y. (2012). Fault detection in dynamic systems using the largest lyapunov exponent. Ph.D. thesis, Texas A & M University, USA.
- Szuminski, P., & Kapitaniak, T. (2012). Nonlinear control and chaotic vibrations of perturbed trajectories of manipulators. Copyright (c) 2012 Annual Review of Chaos Theory, Bifurcations and Dynamical Systems (ARCTBDS) ISSN 2253-0371 All Rights Reserved www.arctbds.com (p. 32).
- Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980* (pp. 366–381). Berlin: Springer.
- Tharwat, A. (2018). Classification assessment methods. *Applied Computing and Informatics*. <https://doi.org/10.1016/j.aci.2018.08.003>.
- Trendafilova, I., & Manoach, E. (2012). Vibration-based methods for structural and machinery fault diagnosis based on nonlinear dynamics tools. In G. Rigatos (Ed.), *Fault diagnosis in robotic and industrial systems*. New York: IConcept Press LTD.
- Tucker, W. (2002). Computing accurate poincaré maps. *Physica D: Nonlinear Phenomena*, *171*(3), 127–137.
- Von Oertzen, T., & Boker, S. M. (2010). Time delay embedding increases estimation precision of models of intraindividual variability. *Psychometrika*, *75*(1), 158–175.
- Wang, C., Gan, M., et al. (2017). Intelligent fault diagnosis of rolling element bearings using sparse wavelet energy based on overcomplete dwt and basis pursuit. *Journal of Intelligent Manufacturing*, *28*(6), 1377–1391.
- Wang, J., Li, R., & Peng, X. (2003). Survey of nonlinear vibration of gear transmission systems. *Applied Mechanics Reviews*, *56*(3), 309–329.
- Wang, Z., & Xue, X. (2014). Multi-class support vector machine. In *Support vector machines applications* (pp. 23–48). Berlin: Springer.
- Wang, Z., Han, Z., Gu, F., Gu, J. X., & Ning, S. (2015). A novel procedure for diagnosing multiple faults in rotating machinery. *ISA Transactions*, *55*, 208–218.
- Wu, S. T., Campos, S. P., & de Aguiar, M. A. (1998). Scientific visualization of poincaré maps. *Computers and Graphics*, *22*(2–3), 209–216.
- Xia, Z., Xia, S., Wan, L., & Cai, S. (2012). Spectral regression based fault feature extraction for bearing accelerometer sensor signals. *Sensors*, *12*(10), 13694–13719.
- Yan, R., & Gao, R. X. (2004). Complexity as a measure for machine health evaluation. *IEEE Transactions on Instrumentation and Measurement*, *53*(4), 1327–1334.
- Yin, Z., & Hou, J. (2016). Recent advances on svm based fault diagnosis and process monitoring in complicated industrial processes. *Neurocomputing*, *174*, 643–650.
- Zhou, Q., Yan, P., Liu, H., & Xin, Y. (2019). A hybrid fault diagnosis method for mechanical components based on ontology and signal analysis. *Journal of Intelligent Manufacturing*, *30*(4), 1693–1715.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.