Gears and bearings defaults: from classification to diagnosis using machine learning for SURVISHNO Conference 2019

Sylvain BARCET[1], Valentin BARON[2], Alexandre CARBONELLI[1]

[1]VIBRATEC, F-69130 Ecully, France
[2]MICRODB, F-69130 Ecully, France

sylvain.barcet@vibratec.fr
valentin.baron@micro-db.fr
alexandre.carbonelli@vibratec.fr

Abstract

Gears and bearings are more and more used in every industrial area mainly due to their strong reliability. Nevertheless as every mechanical transmission system, failures appear during time life. It induces critical damage, time cost for maintenance services to repair the fault potentially on duty. A wide part of work in the scientific community already provides a large quantity of features to follow health status of these systems (e.g., RMS, kurtosis, crest factor, FM0) in order to detect the fault as soon as possible.

Since few years, methods developed in signal post-processing are coupled with Machine Learning (ML). ML allows ability to detect novelty or fault based on a trained algorithm. According to the literature [1], to identify the type of damage, a supervised algorithm is needed. Consequently an accurate diagnosis implies labelled data which are often difficult to obtain practically.

The aim of this paper is to provide keys, based on our knowledge about features in Structural Health Monitoring (SHM), to get higher information level in classification by adding a qualitative analysis (type of damage) without label or information about the type of fault.

Work carries on a measurement database. The assumption is made about two classes “healthy” / “faulty” using a supervised algorithm. The contribution of our work brings a new step in the default analysis by adding a probability for a defect case to be identified. Indeed, by combining some sensitive features selected for their relevance to describe a type of fault, a probability to have this particular default can be given. This classification is tested against three fault classes: bearing, gear generalized, gear localized.

Results show that a probability for having bearing fault can be identified using this method contrary to the gear generalized and localized fault which are more complex to characterize. This new step enables to help maintenance services to focus more efficiently on the incriminated faulty part of the system, inducing a reduction of time to repair for maintenance services, a shorter out of order time leading to a significant productivity gain.

[redacted]Introduction

Gears are used in a huge quantity of mechanical systems. As a consequence, monitoring their possible faults the most accurately possible is a major issue in the field of Structural Health Monitoring (SHM) as they can provoke critical damages.

With the increasing use of ML techniques, different algorithms have emerged to deal with this problem. Support Vector Machine, Neural Network, Random Forest are examples of ML methods currently used to classify faulty and healthy sample.

Looking at industrial maintenance services requirements, the needs in terms of monitoring may be resumed as:

1- Find efficient condition indicators (CI) to monitor their systems,
2- Use ML algorithms to allow a continuous monitoring and an high efficiency of faulty detection,
3- Have a minimal cost and time to repair the faulty equipment.

The first and the second point are already addressed in literature. This paper proposes to industrials a method to complete their process by the third part: a qualitative analysis of fault, leading to a reduction of cost and time for maintenance services.

[redacted]Presentation of the study case

PHM Society proposes a challenge for monitoring and fault detection. They provide a measurement database (measured on a test bench). Students, researchers and companies can participate. Each one proposes their own method to classify the given database. This work is based on the database provided for the 2009 challenge. Figure 1 presents the test bench used to build this measurement database.

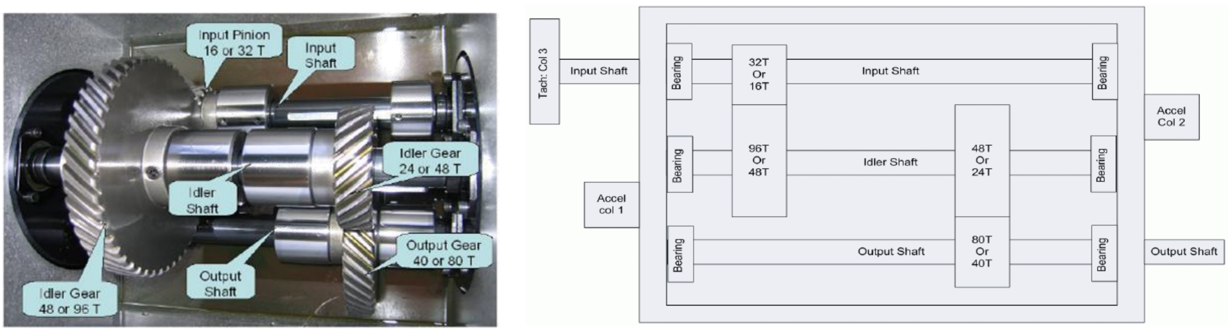


Figure 1: Presentation of the test bench used for measurements coming from PHM Society 2009 Challenge

This test bench is built with two gear stages mounted on three shafts with six bearings. Two gear geometries are used: one using spur gears, the other one using helical gears.

Table 1 presents the gear parameters for both configurations.

Shaft	Gear	Spur gear	Helical gear
Input shaft	input gear	32	16
Idler shaft	1st idler gear	96	48
Idler shaft	2nd idler gear	48	24
Output shaft	output gear	80	40

Table 1 : Gear description for the two geometries: spur and helical

Whatever the gear geometry, the gear ratio between each gear stage is the same, leading to keep the same global gear ratio. Thus, from input to output the gear reduction ratio is 5 to 1 reduction:

$$gear\ ratio = \frac{16}{48} \times \frac{24}{40} = \frac{1}{5}\,. \quad (1)$$

The instrumentation is composed of a limited number of sensors with two accelerometers mounted on the housing and one tachometer on the input shaft (see Figure 1). The tachometer delivers 10 tops by rotation. The sampling frequency is the same for the three sensors and fixed at 66666.67 Hz. Each sample of the database is composed of three raw data columns of one second length.

Several configurations are listed such as:

- 14 gear cases:
    - 8 configurations on spur gear [7 faulty and 1 healthy],
    - 6 configurations on helical gear [5 faulty and 1 healthy],
- 5 rotational speeds [30 Hz, 35 Hz, 40 Hz, 45 Hz, 50 Hz],
- 2 load cases [high, low].

This leads to 140 different configurations. Each configuration is repeated four times to give at least 560 measurement inputs.

[redacted] Method

The methodology developed and exposed here is composed of three main steps:

- CI computing,
- Classification using ML,
- Qualitative analysis using relevant selected features for each chosen default.

[redacted] Condition indicators computing

Figure 2 and Figure 3 present the post-processing used to build a matrix with all indicators.

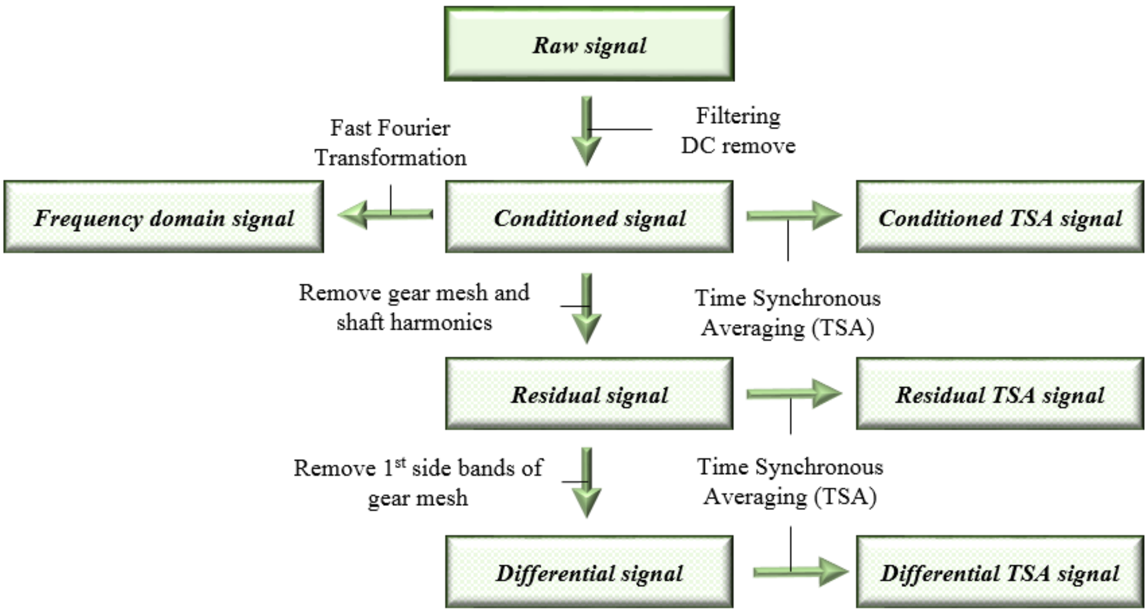


Figure 2 : Signal post-processing used to access to the different needed types of signals

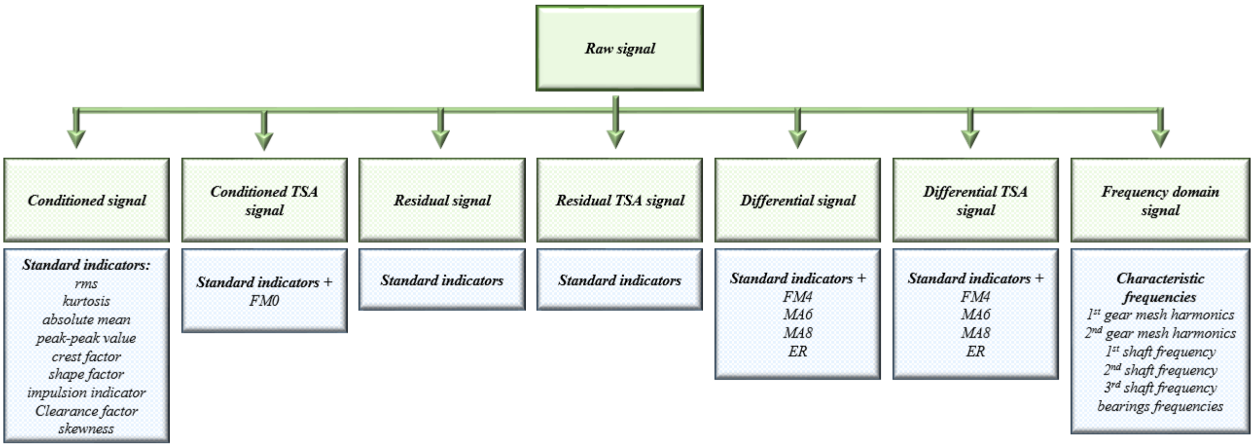


Figure 3 : List of indicators calculated on each signal

From raw signals, based on literature on bearings and gearboxes monitoring, some CI are extracted. Taking into account the difficulty to add a lot of different sensors, most of usual CI used in industry are directly computed from accelerometers raw signals (conditioned signal) such as RMS, kurtosis, crest factor. Based on [3], some new indicators, using residual or differential signals, requiring a tachometer information allow to increase strongly the efficiency of fault detection.

In order to complete this list of CI, some other indicators are computed from the frequency domain, such as gear mesh harmonics, shaft rotational frequencies and specifics bearings frequencies.

All these frequencies are calculated and added in the panel of features using the following equations:

$$Ball\ Pass\ Frequency\ of\ Inner\ ring: BPFI = \frac{N_b}{2} f \left(1 + \frac{D_b}{D_p} \cos(\alpha)\right), \qquad (2)$$

$$Ball\ Pass\ Frequency\ of\ Outer\ ring: BPFO = \frac{N_b}{2} f \left(1 - \frac{D_b}{D_p} \cos(\alpha)\right), \qquad (3)$$

$$Ball\ Spin\ Frequency: BSF = \frac{D_p}{2\,D_b} f \left(1 - \left(\frac{D_b}{D_p}\right)^2 (\cos(\alpha))^2\right), \qquad (4)$$

with f the number of revolutions per second, D_b the ball diameter, N_b the ball number, D_p the pitch diameter and α the contact angle.

$$Gear\ mesh\ frequency = f_{sh(i)} \times n_{teeth(i)}, \qquad (5)$$

with $f_{sh(i)}$ the rotational speed frequency of the shaft "i" and $n_{teeth(i)}$ the number of teeth on the gear mounted on the shaft "i".

Considering the two accelerometers, a total of 298 indicators are extracted.

[redacted] Classification using Machine Learning

Based on all our indicators, the second step of the methodology consists in the classification of the 560 signals in two classes: healthy or faulty. Although the labels are not given in the database from the PHM Society, in [2], authors give the label of all healthy samples they have classified. Based on this, an approach using supervised algorithms is possible for the classification.

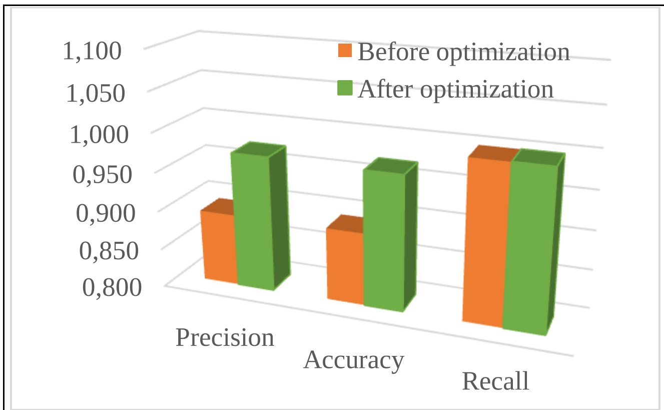
A wide range of supervised algorithms exists in the literature. Among them, a classification is proposed using four of them implemented within the Scikit-learn Python module [5]:

- Nearest Neighbors classifier (KNN) [6],
- Random Forest classifier (RFC) [7],
- Support Vector classifier (SVC) [8],
- Multi Layer Perceptron Classifier (MLPC) [9].

Each algorithm works with a specific method to classify. SVC determines a boundary between the two classes using only the data of each class which are close one to the other, the so-called support vectors. KNN looks at the same class nearest samples of a particular observation to build the boundary between classes. RFC builds decision trees and it combines them together to give its final classification. Finally MLPC relies on a trained neural network to decide that class the tested observation belongs to.

The considered input signals for this step are the 560 one second duration signals provided by the PHM database. So an observation for the following of the contribution refers to a vector of dimension 298 gathering all the features computed from one of these 560 signals. These 560 signals are splitted in two categories: a train set and a test set. 80 % of the database is used for the train set (448 signals) and 20 % for the test set (112 signals).

The efficiency of these algorithms without any optimization is around 90% of good classification. An efficient solution to increase the performance of a ML algorithm is to optimize some hyper parameters. Table 2 presents results before and after optimization on SVC algorithm.



SVC algorithm	Before optimization	After optimization
Accuracy	0.893	0.973
Precision	0.893	0.971
Recall	1.000	1.000

Table 2 : Results of Support Vector Classifier before and after hyper parameters optimization

The optimization phase enables to increase significantly the accuracy of the considered algorithm from 89.3 % to 97.3 %. After an optimization of the hyper parameters on every algorithm, each individual algorithm gives interesting results with a percentage of good classification between 90 % and 97.1 %.

Finally a method to increase strongly the performances relies on combining different algorithms with different approaches. This method called “ensemble learning” reduces individual weaknesses of each algorithm. In the present case, the precision reaches 100 % using ensemble learning on the three best algorithms, namely RFC, SVC and MLPC.

[redacted] Qualitative analysis

The classification in healthy or faulty cases enables to alarm maintenance service immediately after or ideally a short time before the failure. However, no information is given about the type of fault.

The methodology proposed in this work consists in adding a qualitative information in the classification.

The work presented here proposes a method to estimate the probability for a faulty observation to correspond to one of these three types of faults:

- Bearing faults,
- Gear generalized faults,
- Gear localized faults.

The wide quantity of indicators to monitor bearings and gearboxes given in the first section does not describe the same input signals' features. Consequently they are not equally sensitive to the different types of faults [4].

Where RMS represents the energy of the signal leading to an accurate indication about the general state of the complete system, the peak-peak value is in opposition, very sensitive to any localized phenomenon on the signal, enabling to discriminate a localized fault such as crack on tooth or more critical case like a missing tooth. A solution based on the physics described by each indicator is investigated to predict the type of fault.

3.3.1 Feature selection

The selection proposed is composed of:

- 70 indicators for generalized faults:
    - Rms on TSA signal:
        - Rms on conditioned TSA signal on the 1^{st} shaft, 2^{nd} shaft and 3^{rd} shaft,
        - Rms on residual TSA signal on the 1^{st} shaft, 2^{nd} shaft and 3^{rd} shaft,
        - Rms on differential TSA signal on the 1^{st} shaft, 2^{nd} shaft and 3^{rd} shaft.
    - Absolute mean:
        - Absolute mean on conditioned signal,
        - Absolute mean on conditioned TSA signal on the 1^{st} shaft, 2^{nd} shaft and 3^{rd} shaft,
        - Absolute mean on residual signal,
        - Absolute mean on residual TSA signal on the 1^{st} shaft, 2^{nd} shaft and 3^{rd} shaft,
        - Absolute mean on differential signal,
        - Absolute mean on differential TSA signal on the 1^{st} shaft, 2^{nd} shaft and 3^{rd} shaft.
    - MA6:
        - MA6 on differential signal,
        - MA6 on differential TSA signal on the 1^{st} shaft, 2^{nd} shaft and 3^{rd} shaft.
    - Gear mesh harmonics:
        - 1^{st} gear mesh frequency, harmonics 1 to 5,
        - 2^{nd} gear mesh frequency, harmonics 1 to 5.
- 14 indicators for localized faults:
    - FM0:
        - FM0 on conditioned TSA signal on the 1^{st} shaft, 2^{nd} shaft and 3^{rd} shaft.
    - Peak-peak value on residual signal:
        - Peak-peak value on residual signal,
        - Peak-peak value on residual TSA signal on the 1^{st} shaft, 2^{nd} shaft and 3^{rd} shaft.
- 18 indicators for bearing faults:
    - BPFO:
        - BPFO on 1^{st} shaft, 2^{nd} shaft and 3^{rd} shaft.
    - BPFI:
        - BPFI on 1^{st} shaft, 2^{nd} shaft and 3^{rd} shaft.
    - BSF:

- BSF on 1st shaft, 2nd shaft and 3rd shaft.

From the 298 indicators, a total of 102 indicators are extracted for this qualitative analysis, they are summarized in Table 3.

Type of fault	Type of CI	Number of indicators (for 2 accelerometers)
Generalized	RMS on residual signals	18
	MA6	8
	Absolute mean	24
	Gear mesh harmonics	20
Localized	FM0	6
	peak-peak value on residual signal	8
Bearing	BPFO	6
	BPFI	6
	BSF	6

Table 3: List of CI used for the qualitative analysis

3.3.2 Characterization method

Once feature selection is performed according to their relevance for each of the investigated fault, the method to highlight the cases associated with a particular fault relies on the Principal Component Analysis (PCA) method [10]. The goal of PCA is to find, in a point cloud, the direction on which the projection of its point cloud has a maximum variance. This direction is called the first principal component and the next ones are the orthogonal directions of this first one that again explain the maximum variance. In this contribution each feature selection for each fault gives a point cloud of high dimension:

- Dimension 70 for the gear generalized faults,
- Dimension 14 for the gear localized faults,
- Dimension 18 for the bearing faults.

Consequently the PCA is used in this case to perform a dimensionality reduction in order to be able to represent in a two-dimensional space data which came from these high-dimensional spaces. From this representation the observations which represent a particular fault are expected to be significantly far away from the healthy and the other default cases.

Therefore once the PCA is performed on each case, the probability density function (PDF) derived from the healthy cases is estimated. According to isolines corresponding to specific probability values three categories are differentiated:

- A category with a low probability to have the studied fault type whose data are close to the healthy ones after PCA,
- A category with a medium probability to have this fault which are a little bit further from the healthy data,
- A category with a high probability to have this fault which are far away from the healthy data.

[redacted]Results and analysis

4.1.1 Bearing defaults

Once the 18 features which should represent well the bearing faults are selected, the PCA is performed and Figure 4 is obtained.

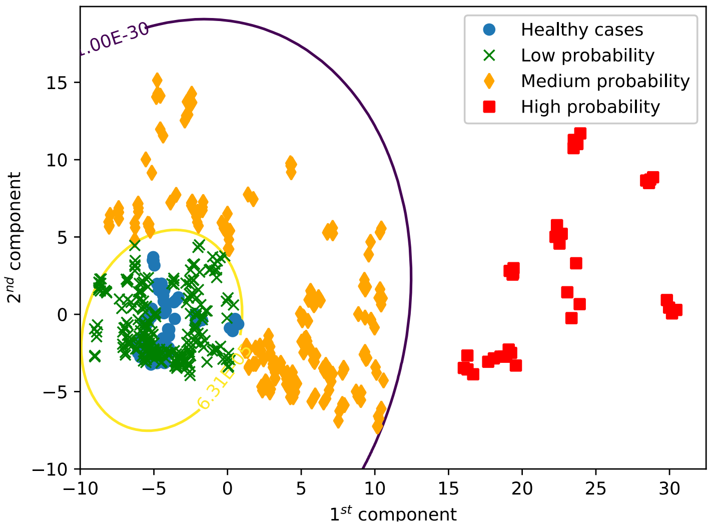


Figure 4: PCA result with the bearing faults feature selection. Yellow and purple lines represent respectively the $10^{-4.2}$ and 10^{-30} isolines of the probability density function derived from the healthy data. Healthy cases are given by blue dots, cases with a low / medium / high probability to have a bearing faults are given by green crosses / orange diamonds / red squares.

It clearly shows a distinct group further than the 10^{-30} probability isoline that gives candidates for having bearing fault. Then in between the two probability isolines the observations are classified as possible to have bearing fault. Finally data which are lying among the healthy cases are not likely to have bearing fault at all.

Consequently the feature selection made beforehand has allowed a qualitative analysis of the data. Indeed, thanks to their position compared to the healthy cases some observations can be classified within a bearing fault category with a given probability.

4.1.2 Localized and generalized defaults

Concerning gear localized and generalized faults, the results are more difficult to interpret as it is shown in Figure 5 with the PCA result of the localized fault case and in Figure 6 with the generalized fault case.

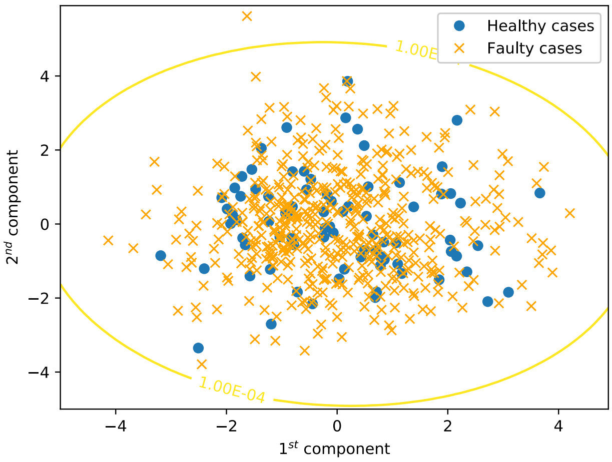


Figure 5: PCA result with the gear localized fault feature selection. The yellow line represents the 10^{-4} isoline of the probability density function derived from the healthy data. Healthy cases are given by blue dots and faulty cases are given by orange crosses.

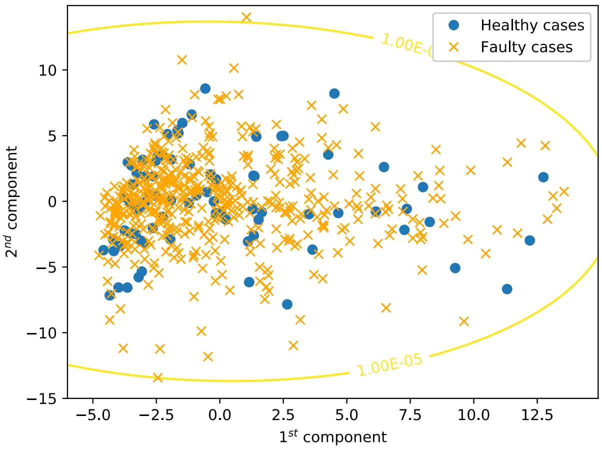


Figure 6: PCA result with the gear generalized fault feature selection. The yellow line represents the 10^{-5} isoline of the probability density function derived from the healthy data. Healthy cases are given by blue dots and faulty cases are given by orange crosses.

In these cases the feature selection has not allowed to distinguish different categories within the fault cases. It means that the selected features are not a relevant set enough to predict both gear localized or generalized faults. Some work is in progress to find better sets of features in order to be able to reproduce the results obtained with the bearing fault identification.

[redacted]Conclusion

This work presents the full method to diagnose and follow the healthy condition of a rotating equipment: from the conditions indicators to the classification. From 298 relevant conditions indicators obtained from accelerometer signals, a first supervised classification step enables to decide whether or not a considered observation is healthy or faulty using ensemble learning with three combined ML algorithms. To add a qualitative analysis of the faulty cases, a new step is performed using principal component analysis on a feature subset of the 298 ones. Three different faults are studied, bearing faults and gear localized and generalized faults, bringing three different subset of features of dimension 18, 14 and 70, respectively. The dimensionality reduction obtained using the principal component analysis allows to represent in a two-dimensional space, corresponding to the two first principal components, both healthy and faulty cases. From this representation a probability density function of the healthy cases is determined and the faulty cases can be marked as having a low or high probability to indeed, having this specific fault from their position compared to this function.

Results show that bearing faults can be identified using this methodology with three different categories highlighted: low, medium and high probability for having this fault type. However the feature selection made for the gear localized and generalized faults have not allowed to distinguish clearly between healthy and faulty cases and consequently there is no specific observation which can be identified with these fault types.

Work is in progress on the feature selection to be able to reproduce the bearing fault results and other methods are investigated to perform a better dimensionality reduction such as manifold learning techniques. These two ideas could bring the missing part to be able to give a complete qualitative assessment of the fault types.

References

[1] K. Worden and al., *The fundamental axioms of structural health monitoring*, 2007

[2] H. Al-Atat, and al., *A systematic methodology for gearbox health assessment and fault classification*, Journal of Prognostics and Health Management, 2011

[3] A. S. Sait, Y. I. Sharaf-Eldeen, *A Review of Gearbox Condition Monitoring Based on vibration Analysis Techniques Diagnostics and Prognostics*, 2011

[4] V. Sharmaa, A. Pareya*, *A review of gear fault diagnosis using various condition indicators*, *12th International Conference on Vibration Problems*, ICOVP 2015

[5] F. Pedregosa et al., *Scikit-learn : Machine Learning in Python*, JMLR 12, pp. 2825-2830, 2011.

[6] R. Duda et al., *Pattern Classification*, 2nd Edition, John Wiley & Sons, 2012.

[7] L. Breiman, *Random Forests, Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001 January 01.

[8] C. Cortes and V. Vapnik, *Support-vector networks, Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995 September 01.

[9] D. Rumelhart et al., *Learning representations by back-propagating errors*, Nature, vol. 323, no. 6088, pp. 533-536, 1986 October.

[10] I. Jolliffe, *Principal Component Analysis*, 2nd Edition, Springer Series in Statistics, 2002.