Analysis of the Rolling Element Bearing data set of the Center for Intelligent Maintenance Systems of the University of Cincinnati

William Gousseau, Jérôme Antoni, François Girardin, Julien Griffaton

Analysis of the Rolling Element Bearing data set of the Center for Intelligent Maintenance Systems of the University of Cincinnati

William Gousseau, Jérôme Antoni, François Girardin
Univ Lyon, INSA-Lyon, Laboratoire Vibrations Acoustique, F-69621 Villeurbanne, France
william.gousseau@insa-lyon.fr, jerome.antoni@insa-lyon.fr,
francois.girardin@insa-lyon.fr

and Julien Griffaton
julien.griffaton@snecma.fr
Safran Aircraft Engines, Rond point René Ravaud, 77550 Réau, France

Abstract

Due to its importance in the industry, vibration-based diagnosis and prognosis of rolling element bearings (REB) attract more and more attention from the research community. Moreover, it is now possible to test newly developed methods on test rig data that are open-access. The NASA provides a list of data bases concerning prognosis and some REB data-sets, such as the one from the Center for Intelligent Maintenance Systems(1) (IMS), University of Cincinnati. This data-set has so far been used by many authors. However, even if the state of health of the tested bearings is clearly provided and sometimes supported by photographs taken at the end of endurance, the diagnosis from the provided vibration signals is not so obvious. This paper tries to establish what can be clearly concluded from the analysis of the signals. In order to fulfill this objective, several diagnostic techniques are used: time analysis, spectral analysis, blind deconvolution, spectral coherence, envelope spectrum. A diagnosis is therefore returned whenever possible, together with recommendations for prognosis in those cases which have been identified as difficult.

1. Introduction

1.1 Motivations

Todays, vibration-based diagnosis and prognostics of rolling element bearings attracts more and more attentation in the scientific community. Therefore, providing open-access databases is of prime importance. There are so far few available datasets. The Case Western Reserve University data(2) are one example dedicated to diagnostics which has been widely used in numerous works. The IMS dataset is another example which constitutes a precious database for prognostics. Indeed, endurance tests with natural degradation are rare. In most cases, damage is artificially initiationed in order to accelerate the damage growth. Whereas for diagnosis studies seeded faults may be acceptable, prognostis requires an evolution of the damage starting ideally from a natural initiation. The IMS dataset has therefore interesting potential both for diagnosis and (mostly) for prognostics.

As opposed to several studies based on the direct application of some *ad hoc*trending features (such as statistical moments), this paper follows an explicative

approach where the vibration signals of the IMS dataset are thoroughly analysed by dedicated signal processing tools.

1.2 IMS dataset in the literature

This IMS database has been so far used in different research works published in the specialized literature. Many papers used it for illustrating remaining useful life (RUL) estimation based on statistical indicator[3,4,5,6] (Mahamad & all 2010) (Soualhi & all 2013) (Ben Ali & all 2015) (Ben Ali & all 2014). Other papers used it to test new signal denoising techniques[7] (Guo & all 2015) often with the goal of acheiving an early detection of the damage[8,9] (Qiu & all 2006) (Mortada & all 2011). Paper of Qiu and all[8] is the reference paper concerning this dataset.

2. Dataset description

2.1 The test rig

The IMS bearing dataset has been collected on an endurance test rig of the University of Cincinnati and relased in 2014[1]. The test rig (shown in Fig.1) has the following characteristics:

- 4 double row bearings of type Rexnord ZA-2115,
- 2000 rpm stationary speed,
- 6000 lbs load applied onto the shaft and bearing by a spring mechanism,
- PCB 253B33 High sensitivity Quart ICP® accelerometers.

An AC motor, coupled by a rub belt, keeps the rotation speed constant. The four bearings are in the same shaft and are forced lubricated by a circulation system that regulates the flow and the temperature. It is announced on the provided "Readme Document for IMS Bearing Data" in the downloaded file, that the test was stopped

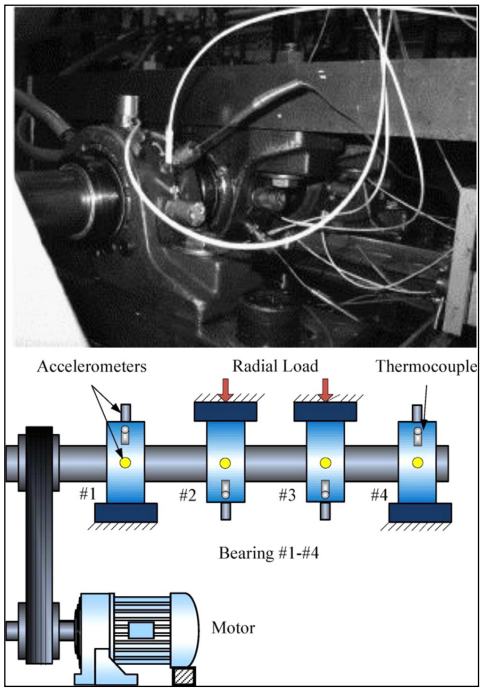

Figure 1. IMS test rig layout

when the accumulation of debris on a magnetic plug exceeded a certain level indicating the possibility of an impending failure.

The four bearings are all of the same type. There are double range pillow blocks rolling elements bearing. Their characteristics are given in table 1.

Table 1. Bearing characteristics

Rexnord ZA-2115 Characteristics		
Pitch diameter	2.815 inch	71.5mm
Rolling element diameter	0.331 inch	8.4mm
Number of rolling element per row	16	16
Contact angle	15.17°	15.17°
Static load	6000 lbs	26690 N

2.2 Data description

Three datasets are provided on the downloaded file, composed of numerous files of one second each. Each file is made of 20,480 samples. Althougth it is mentioned that the sampling frequency is 20 kHz, it is thus believed that it was actually 20.48 kHz(and the spectral analyses presented in this paper all seem to support this assumption, since more consistent results are then obtained with respect to expected fault frequencies). A one second acquisition has been made every ten minutes. There is an exception for the first dataset for which the first forty-three files have been acquired every five minutes.

The timeline of the signals files gives important information about the recording. In particular time intervals are identified in the first dataset where recording was switched off.

The description of the datasets is given the table 2 Concerning dataset 1, there are two accelerometers on each bearing (x and y positions). The two other datasets only have one accelerometer on each bearing.

Table 2. Datasets description

	Number of files	Number of channels	Endurance duration	Duration of recorded signal	Announced damages at the end of the endurance
Dataset 1	2156	8	49680 min 34 days 12h	36 min	Bearing 3: inner race Bearing 4: rolling element
Dataset 2	984	4	9840 min 6 days 20h	16 min	Bearing 1: outer race
Dataset 3	4448	4	44480 min 31 days 10h	74 min	Bearing 3: outer race

The fault frequencies required for the diagnosis of the rolling element bearing have been calculated from the bearing characteristics. They are reported in table 3.

Table 3. Characteristic frequencies of the test rig

Characteristic frequencies	
Shaft frequency	33.3 Hz
Ball Pass Frequency Outer race (BPFO)	236 Hz
Ball Pass Frequency Inner race (BPFI)	297 Hz
Ball Spin Frequency (BSF)	278Hz (2x139 Hz)
Fundamental Train Frequency (FTF)	15 Hz

2.3 Time analysis

Before proceeding to sophisticated signal processing tools, a simple analysis in the time domain can already provide interesting information.

Dataset 1 has been subjected to a series of interruption during the recording, which makes the time history not continuous. This is illustrated in Fig 2 for the time signal of accelerometer 1 displayed as a function of the real time (and not recording time which makes the signal continuous).

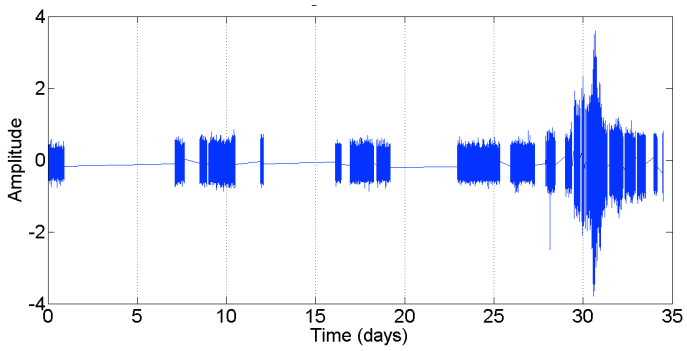

Figure 2. Raw signal and its real time history scale for dataset 1

Another point concerns the announced faults. There is no fault on bearing 1 and 2. Bearing 1 and 2 are considered as healthy and their analyses are not shown hereafter.

Dataset 2 and 3 have been recorded without interruption. Records of 1s – hereafter denoted as frames -have been captured every 10 min all along the endurance test.

3. Methodology

This section introduces the signal processing methods which are used to extract the diagnostic information contained in the three endurance datasets. Two preprocessing methods are first presented which have been used to enhance the signal.

3.1 Preprocessing

In order to enhance the results, pre-whitening and blind deconvolution are used in a pre-processing stage.

The pre-whitening operation used here consists in normalizing the spectrogram by its averge value taken over fifty frames. This is supposed to remove the constant baseline endured all along the test, which includes structural resonances of the system as well as s of the machinery (belt, motor). This operation has been found important to better enhance the contrast between the beginning and the end of the test. For example, a spectral peak at 236 Hz appears in all frames all along the endurance, which in all likelihood corresponds to the ball passage frequency seen by the sensors which are placed very close to the bearing outer race. This frequency is therefore not to be confused with that of an outer race fault, the reason why it is important to remove from the beginning. Applying pre-whitening will thus avoid confusion in the diagnosis.

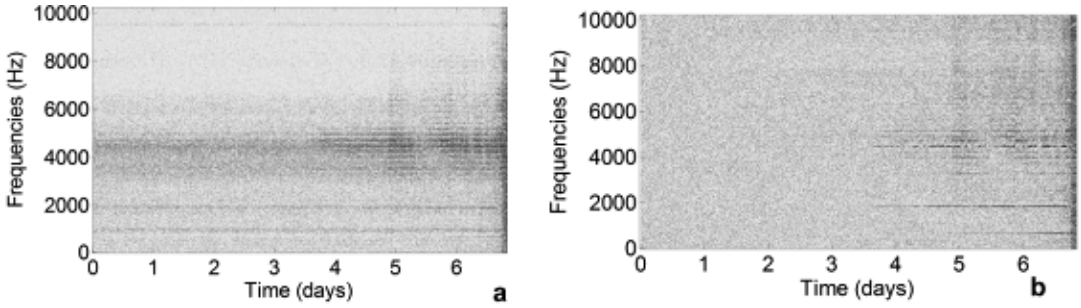

Figure 3. a) raw spectrogram, b) pre-whitened spectrogram

Blind deconvolution is applied to the time signal in order to enhance the emergence of impulses in the signal due to the initiation of the fault. The blind deconvolution algorithm consists in maximizing the Kurtosis[10].

3.2 Signal processing

This section introduces the signal processing tools used to detect the possible presence of damage as early as possible in the endurance test. They pertain to the current state-of-the-art in the domain[11].

Time-frequency analysis is based on the Short Time Fourier Transform on each bloc of one second, with a 1 Hz frequency resolution without overlap (because of time discontinuity between each frame). Results are represented by means of the spectrogram on a logarithmic scale. This allows easy tracking of the frequencies along the tests.

The Squared Envelop Spectrum (SES) is used as indicated in Ref. (11) after filtering the signal in an optimal band that maximized the kurtosis as found by application of the Kurtogram[12].

As demonstrated in the literature, bearing defaults produce cyclostationary signals. The spectral coherence[13] has been found as anefficient tool to detect the presence of cyclostationarity and to identify the bearing fault frequency.

4. Results

This section presents diagnostic results based on the signal processing tools introduced above. Results are presented by means of tables and figures for each dataset. The goal is to confirm that diagnosis of the bearing faults is feasible and, in certain cases, to explain some unexpected results.

The following categorization of results is used:

Table 4. Categorisation of diagnosis

Category	Explanation
TP	« True Positive », when a fault has been announced and is confirmed by the diagnosis
TP2	« True positive 2 », confirmation of what has been announced needing some more explanation.
FN	« False Negative » when a fault has been announced but is not confirmed by the diagnosis
TN	« True Negative » Nothing has been announced and the diagnosis confirms it
FP	« False Positive » Nothing has been announced but the diagnosis detects some signatures

4.1 Dataset 1, inner-race damage on bearing 3, ball damage on bearing 4

This dataset contains 2 accelerometers per bearing. Since analyses carried out on sensors placed on the same bearing are found quite similar, only results obtained on one of the two accelerometers are shown. This observation could indeed explain why there is only one accelerometer available per bearing for the other datasets.

Weak harmonics are visible at twice and three times the BPFI (table 3) with modulation sidebands spaced by the shaft speed(33.3 Hz). This signature is typical of an inner race fault. In the reference paper [8] is also shown a BPFO fault at the end of endurance on bearing 4. It is also visible on the analyses but not shown in this paper.

In the spectrogram (Fig. 4) the inner race fault signature is visible from 33,8 days, as evidenced by a spectral comb. Twice and three times the BPFI are the dominant frequencies. The shaft modulations, typical of an inner race fault, are also easily identified. This is so considered as a TP.

For test 1, the SES has been computed in the entire frequency bandwidth – see Fig. 5. Results are close to those of the spectrogram, with dominant frequencies at twice

and three times the BPFI. However, the fault signature is visible slightly sooner, from 32 days. The sideband due to shaft modulations are also visible at 34 days.

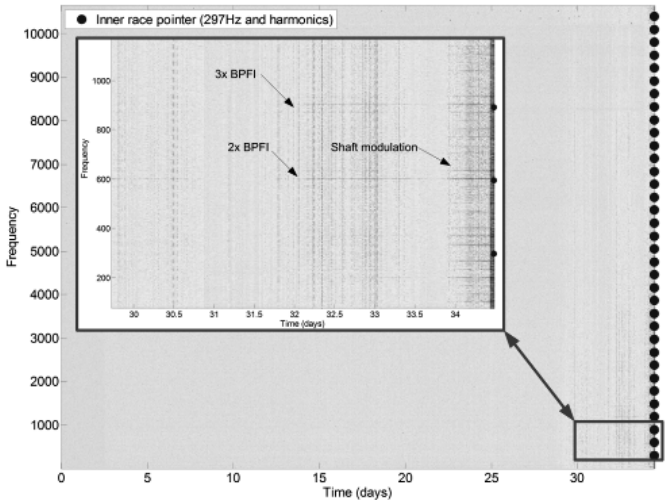

Figure 4. Spectrogram, dataset 1 accelerometer 3

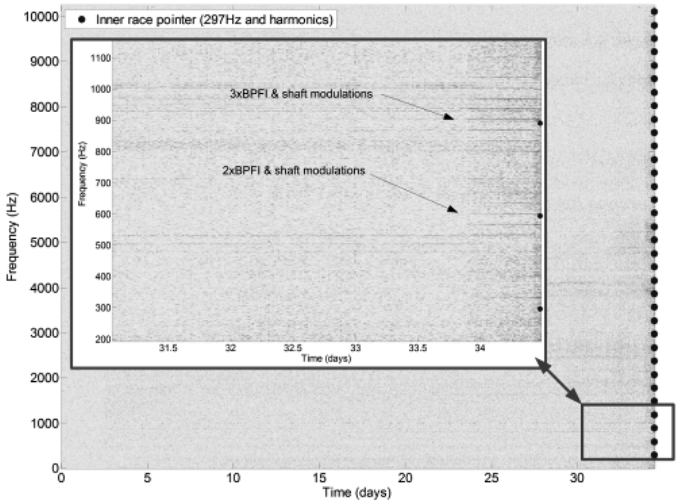

Figure 5. Squared Envelop Spectrum, dataset 1 accelerometer 3

The spectral coherence probably returns the most interesting results. Although the detected frequencies are the same (twice and three times the BPFI), the signature of the inner race fault is visible from 29.2 days (Fig. 6). The shaft modulation sideband is also visible later, around 32 days.

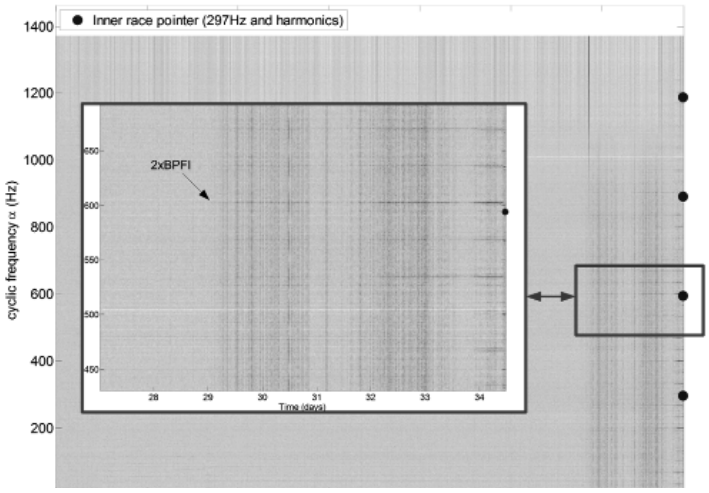

Figure 6. Spectral Coherence, dataset 1 accelerometer 3

The BSF (table 3) and harmonics are difficult to find and appear late in the endurance test.

The signature of the fault in the spectrogram is only visible at the very end of the test, around 34 days, with clear modulation sidebands ate the shaft speed (Fig. 7). Only the sixth BSF harmonics can be identified from 18 days, with FTF modulations. This is enough the diagnose the ball fault, which is thus categorized as a TP.

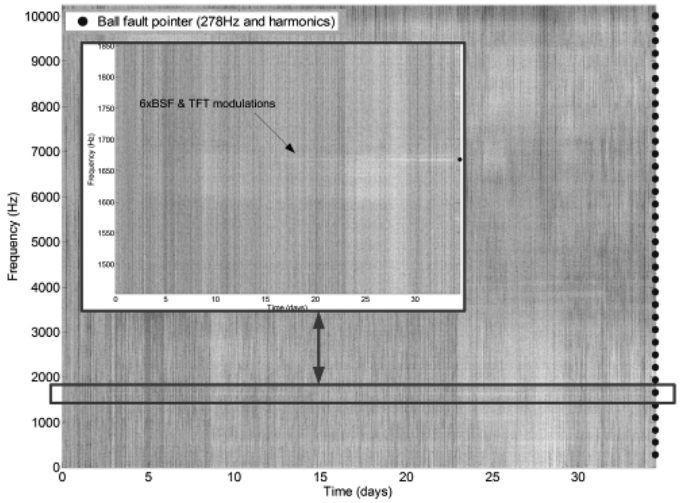

Figure 7. Spectrogram, dataset 1 accelerometer 4

The SES has been computed in the entire frequency bandwidth. Although the signature is weaker than in the spectrogram, its is more consistant with a ball fault: a peak is visible at twice the BSF beginning around 25 days. Some modulations at shaft speed are also visible at the very end, from 34 days.

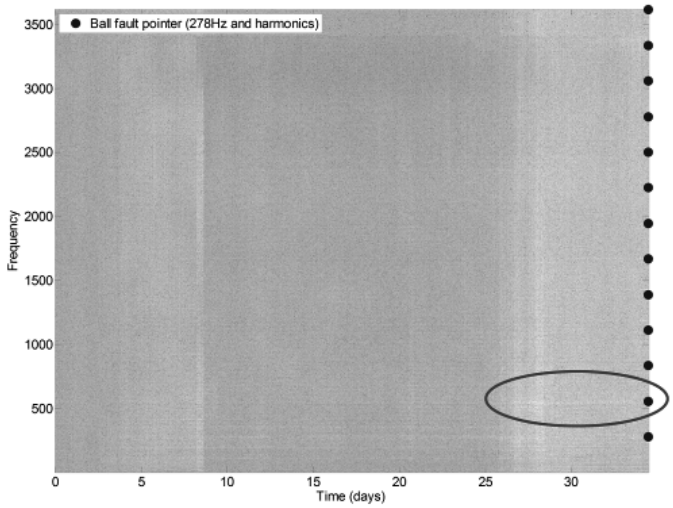

Figure 8. Squared Envelop Spectrum, dataset 1 accelerometer 4

The spectral coherence returns similar results as the SES, with more pronounced FTF modulations. In addition, there is a higher increase of the the peak magnitudes at the BSF and its harmonics from 23 days. This case is also considered as a TP.

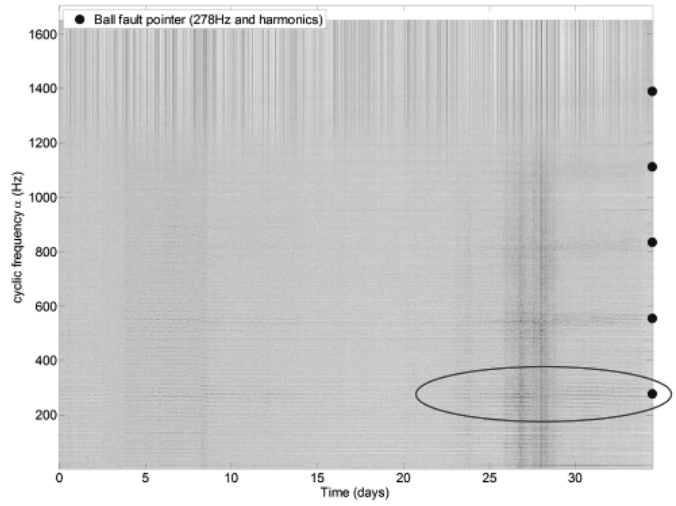

Figure 9. Spectral Coherence, dataset 1 accelerometer 4

Table 5 summarizes the results concerning dataset 1. It reports the categorization of each analysis and the time from which the fault diagnosis seems possible.

Table 5. Summary of dataset 1 results

	Inner race fault				Ball fault			
accelerometer	acc3x	acc3y	acc4x	acc4y	acc3x	acc3y	acc4x	acc4y
T-F analysis	TP (33.8 days)		TN		TN		TP (18 days)	
Envelop	TP		TN		TN		TP	

	(32 days)			(25 days)
Spectral Coherence	TP (29.2 days)	TN	TN	TP (23 days)

4.2 Dataset 2, outer race damage on bearing 1

This case is easy to diagnose, and all methods give relevant results.

The spectrogram (see Fig.10) provides enough information to make a first diagnosis. The fundamental frequency is not the most relevant ones to monitor here, in particular because if corresponds to the ball pass frequency seen by the sensor placed close to the outer-race, even in the absence of a fault. The three-time BPFO harmonic appears from around 3,5

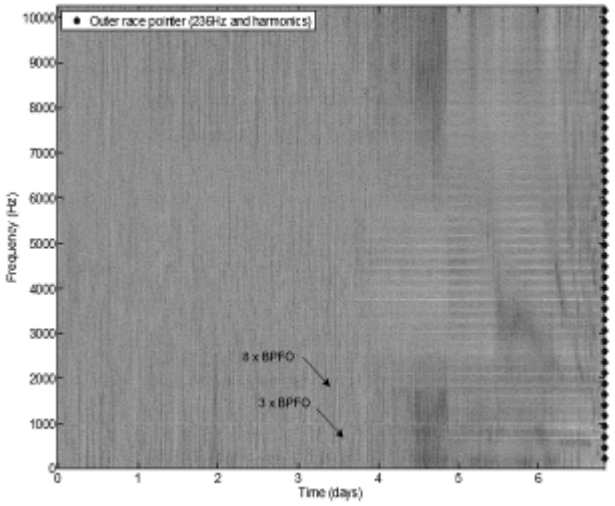

Figure 10. Spectrogram, dataset 2 accelerometer 1

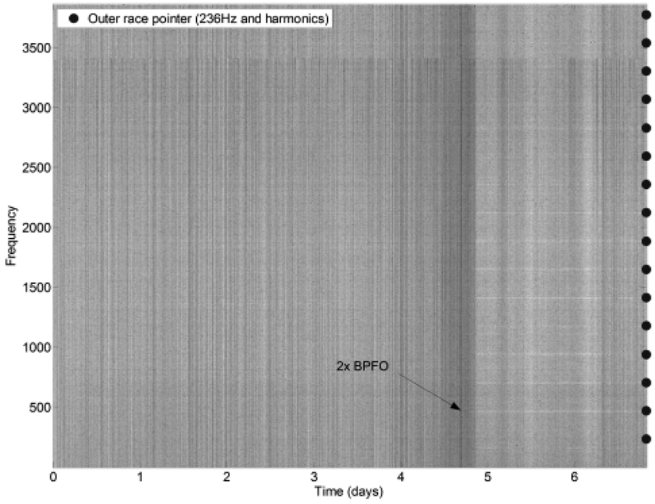

Figure 11. Squared Envelop Spectrum, dataset 2 accelerometer 1

The Squared Envelop Spectrum (see Fig. 11) has been computed after bandpass filtering in the optimal band between 6830 Hz and 10240 Hz as returned by the Kurtogram. Even if the BPFO harmonics are visible they appear later -- from 4,8 days -- than in the spectrogram. Once again,because of its dominance, the twice BPFO harmonic seems the most efficient to follow.

The spectral coherence returns the best diagnostics results. Again, the BPFO frequency appears to not be the best one to track; although it is clearly visible, it appears from the beginning of the test and increases only after 3,5 days. Twice and three times BPFO seem again more relevant because of their dominance

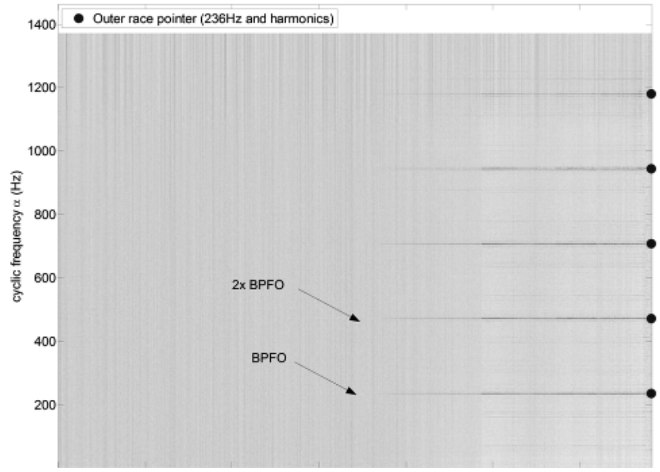

Figure 12. Spectral Coherence, dataset 2 accelerometer 1

This case is clearly considered as a TP.

The three other accelerometers of dataset 2 show similar results which confirm the diagnosis made from accelerometer 1. Although they are located on other bearings (see Fig. 1), they are well able to sense the BPFO fault in bearing 1.. This is categorized here as a TP2. Having the possibility to monitor bearing from distant location can be considered as very useful.

Analysis of the signals reveals that the fault signature appears at the same time whatever the distance to the faulty bearing. The spectrogram, the SES and the spectral coherence of accelerometer 4 are nearly identical to those of accelerometer 1 shown in Fig.10, 11, and 12.

Table (6) summarizes the results concerning dataset 2. It reportes the categorization of each analysis and the time from which the fault diagnosis seems possible.

Table 6. Summary of dataset 2 results

	OUTER RACE FAULT			
	acc1	acc2	acc3	acc4
T-F analysis	TP (3.5 days)	TP2 (3.5 days)	TP2 (3.5 days)	TP2 (3.5 days)
Envelop	TP (3.5 days)	TP2 (3.5 days)	TP2 (3.5 days)	TP2 (3.5 days)

Spectral Coherence	TP (3.5 days)	TP2 (3.5 days)	TP2 (3.5 days)	TP2 (3.5 days)

4.3 Dataset 3, Outer-race damage on bearing 3

The description of dataset 3 is similar to dataset 2 except for two differences. First the outer race on announced inbearing 3 instead of bearing 1 at the end of the endurance test and second the signals are longer (about 30 days of 1s records every 10 minutes resulting in 72 min of signals). Although similar results as with dataset 2 are expected, they eventually turn out very different. None of thethe accelerometers are able to detect the fault.

All other accelerometers return identical results. These results probably explain why dataset 3 has never been used in the literature. It so considers as a FN.

Table 7.Summary of dataset 3 results

	OUTER RACE FAULT			
	acc1	acc2	acc3	acc4
T-F analysis	FN	FN	FN	FN
Envelop	FN	FN	FN	FN
Spectral Coherence	FN	FN	FN	FN

5. Conclusion and discussion

This IMS database is frequently used in the specialized literature to test and validate the proposal of new diagnostic and prognositc algorithms. The goal of this paper was to establish in a comprehensive way which diagnostic information can be extracted from it. This should turn out most useful for its future use by other researchers.

The IMS dataset constitutes an interesting database for several reasons. First, signals are provided from a long endurance test with a natural degradation of the damage, without artificial initiation. Second, it includes three different types of fault with different expected signatures. Third, it contains signals simultaneous recordered by at least four accelerometers placed at different locations. Some of the main findings of the paper are the following.

1) It is not advisable to detect and track the fundamental frequency of the BPFO for diagnostic and prognostic purposes in Dataset 1 2 (outer-race fault); this frequency is present from the beginning of the test and corresponds to the natural passage of the balls as seen by accelerometer 1 mounted close to the outer race.
2) The spectral signature of the inner-race fault in dataset 1 is perhaps unusual as it mainly contains the second and thirs harmonics of the BPFI.
3) It is possible to detect the inner-race in bearing 3 fault and the roller in bearing 4 fault from accelerometers 1 and 2 mounted on the other bearings even if ot shown in this paper (even the outer-race in bearing 4).
4) It is possible to detect the outer-race in bearing 1 fault from accelerometers 2, 3 and 4 mounted on the other bearings.

5) Contrary to what is announced in Ref.(1), it seems that Dataset 3 does not contain any signature of an outer-race fault in bearing 3 (or any other bearing).
6) It seems that the sampling frequency announced in Ref. (1) should be fixed to 20.48 kHz (instead of 20kHz).

References

1. IMS bearings dataset (2014) http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/
2. Smith, W. and Randall R.B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study.Mechanical Systems and Signal Processing, vol. 64, p. 100-131, 2015.
3. Mahamad A.K., Saon S. and Hiyama T. "Predicting remaining useful life of rotating machinery based artificial neural network." Computers & Mathematics with Applications, vol. 60, no 4, p. 1078-1087, 2010.
4. Soualhi A. Razik H. and Clerc G., *et al.* "Prognosis of bearing failures using hidden Markov models and the adaptive neuro-fuzzy inference system." IEEE Transactions on Industrial Electronics, vol. 61, no 6, p. 2864-2874, 2014.
5. Ali, J.B., Chebel-Morello B. and Saidi L. and al. "Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network. Mechanical Systems and Signal Processing", vol. 56, p. 150-172, 2015.
6. Ali J.B., Saidi L. Chebel-Morello B. and al. "A new enhanced feature extraction strategy for bearing Remaining Useful Life estimation". In : Sciences and Techniques of Automatic Control and Computer Engineering (STA), 15th International Conference on. IEEE, p. 365-370, 2014..
7. Guo L. Gao H. Li J. and al. "Machinery vibration signal denoising based on learned dictionary and sparse representation". In : Journal of Physics: Conference Series. IOP Publishing, p. 012124, 2015.
8. Qiu H. Lee J. Lin J. and al. "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics".Journal of sound and vibration, vol. 289, no 4, p. 1066-1090, 2006,.
9. Mortada, M-A, Yacout S. and Lakis A. "Diagnosis of rotor bearings using logical analysis of data." Journal of Quality in Maintenance Engineering, vol. 17, no 4, p. 371-397, 2011.
10. Endo, H. Randall, R. "Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter". Mechanical systems and signal processing, vol. 21, no2, p.906-919,2007.
11. Randall RB, Antoni J. "Rolling element bearing diagnostics -a tutorial". Mechanical Systems and Signal Processing, vol.25, no2, p.485-520, 2011
12. . Antoni J. "Fast computation of the kurtogram for the detection of transient faults." Mechanical Systems and Signal Processing, vol.21, no1, p.108-124,2007
13. Randall R.B. Antoni J et Chobsaard S. "The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals". Mechanical systems and signal processing, vol. 15, no 5, p. 945-962, 2001.