An approach to investigating rolling element bearing damage under impact loading

Matti Savolainen*, Arto Lehtovaara
Group of Tribology and Machine Elements
Materials Science, Engineering and Natural Sciences, Tampere University
PO BOX 589, 33014 Tampere University, Finland

Abstract

This paper presents an approach to studying rolling element bearing damage under the interference of impact loading. In the experimental part, a series of bearing tests was performed by using the twin-disc test device with artificially damaged bearings. This was followed by analysis of the measured acceleration response data in impact-free condition as well as under the influence of the impact loading. The results showed successful detection of the bearing outer race damage by using typical bearing damage detection approaches regardless whether the impact loading was applied to the system or not. In turn, recognition of the bearing rolling element damage required specific signal processing.

Keywords: Rolling bearings, testing, impact load, envelope analysis, signal processing

1. Introduction

Rolling element bearings are commonly used components in mechanical engineering because of their good trade-off between size, load-carrying capacity, durability and cost. Since extensive research has been conducted over the years, their life expectancy can be calculated relatively accurately, especially when the loading and the working conditions are well known. In cases where these parameters vary heavily during use, the estimation of lifespan may become challenging. Moreover, unexpected bearing failures may lead to machine breakdowns and consequently financial losses or even health-damaging incidents. An approach to guaranteeing the reliable operation of machinery is to monitor the possible occurrence as well as progression and severity of damage. This enables precise decision-making and planning of maintenance.

One of the most widely used techniques to diagnose possible damage in roller element bearings is the envelope analysis, where the envelope of the measured acceleration response signal is frequency analysed instead of the signal itself in order to find bearing damage frequencies [1, 2]. Basically, envelope analysis uses amplitude demodulation of the signal to find bursts initiating from damage [3]. One of the challenges in the envelope analysis has been the selection of a suitable frequency band for the demodulation. Probably the most widespread band selection tool is the fast kurtogram [4] because of its high efficiency [5, 6].

Measured signals often include interference that makes detection of relatively weak signals, originating from bearing damage, challenging. Several techniques have been developed over the years to address the problem. In this paper, a couple of them have been applied. Wiggins [7] originally presented the minimum entropy deconvolution (MED) technique, and since then it has been successfully used to deconvolve the impulsive excitation from a mixture of responses to find indications on rotating machine faults in several papers [8-12]. Furthermore, in many studies the autoregressive (AR) approach has given good results also when combined with MED [8, 13, 14], and quite recently it was effectively utilised to predict the extent of damage on rolling element bearing components [15]. The development of self-adaptive noise cancellation (SANC) was based on the idea of separating the periodic signal from the broadband signal by means of only

one sensor [16]. This is done by using a delayed version of the primary signal as a reference. Over the years, SANC has been successfully used to separate planetary gear meshing and bearing signals [17-19].

Rolling element bearing testing can be roughly divided into bench-type testing and full-scale bearing testing. In the first case, the fatigue strength of a limited number of rolling elements is tested with a specific machine imitating the real operating conditions of a bearing. The obtained results have been used to indicate trends and rank materials as well as lubricants [20], or the research has concentrated on the development of bearing failure modes, such as micropitting, smearing or white etching cracks [21-23]. Over the years, various full-scale bearing testing arrangements have been used to find the durability of roller element bearings. These test arrangements often include the possibility of testing multiple bearings at the same time, the introduction of specific operation conditions as well as load sequences that vary in the axial and in the radial directions in an application-specific manner [24, 25].

The basic idea of the twin-disc test device is to reproduce pressure distribution between the discs, which is comparable with the distribution seen between gear teeth, for instance, in spiral bevel gears. This is enabled through the precise definition of the geometries of the discs [26, 27]. Research with the twin-disc test device has focused on investigating phenomena mainly occurring in contact of the discs. Over the years, the device has been successfully utilised in studying formation of micropitting [28, 29] as well as the influence of surface roughness, pitting, wear, and behaviour of friction in the contact under different load conditions [30–33]. Furthermore, in other papers by the authors the occurrence of scuffing under varying load conditions and subsurface-initiated fatigue has been explored [34-36].

In many comprehensive papers dealing with damage detection in bearing components through analysis of measured data, the experimental validation arrangement in laboratory environment has been typically rather simple including an electric motor rotating the observed bearing with artificial damage [8, 11, 13, 37]. Another popular approach has been to analyse artificially generated signal that includes impulsive events imitating response from a bearing defect [5, 38, 39]. Relatively recently there has been also development in understanding the vibration content originating from local defects in the rolling element bearings by utilizing numerical simulation models [40, 41]. However, studies in which a controllable realistic interference in the form of external loads is affecting the observed device and the measured signals, are much more challenging to find.

Furthermore, it is generally recommended that the detection of a rolling element bearing damage in real machinery is performed when the observed machine is operating in well-known load condition, often idling, and is separated from sources of disturbing signals [42, 43]. In such cases, the bearing check may require costly pausing of operation of the entire production line, which this machine is part of, increasing the demand for an approach that is able to reliably conduct bearing damage detection also during variating operation conditions and under the interference of external signals.

Although the twin-disc test device has been mainly developed to study lubricated contacts, the modification possibilities of the apparatus can further extend its range of use. This paper introduces an approach to investigating rolling element bearing damage under realistic impact loading in laboratory environment. First, the twin-disc test device and the method of making bearing loads are presented. After conducting a series of tests with damaged and damage-free rolling bearings the obtained acceleration content is analysed. The analysis results show that the envelope analysis by using the traditional Hilbert transformation approach can indicate bearing damage at the outer race regardless whether the impact load is applied or not. However, the weaker signal originating from damaged rolling elements seems to be slightly challenging to detect. Therefore, in the latter part of this paper means to process the measured signals to reveal indication about the damage are applied successfully.

2. Test arrangement

The tests were carried out with a twin-disc test device that was originally developed in-house. This device is equipped with dynamically adjustable radial force as well as axial displacement loading possibilities as it is described in detail in Refs. [35-37]. Although the main purpose of the device has been to research the phenomena occurring in contact of the discs, its usability can be widened to study also rolling bearings on shafts. Fig. 1 illustrates the principle of bearing testing with the twin-disc system. The load frame of the device is presented in Fig. 2.

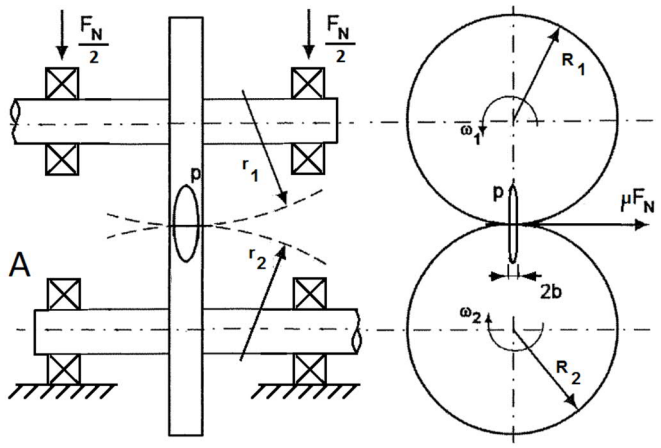

Fig. 1. The principle of the twin-disc test device. The test bearing in this study is marked with “A”.

The structure of the twin-disc arrangement is such that one of the shafts (upper in Fig. 1) can be moved in the direction of the radial force F_N. A hydraulic cylinder is used to produce this force that moves the shaft and pushes the discs together resulting in the formation of pressure distribution p in the contact of the discs. The cylinder is driven by a control valve that adjusts the loading by using measured signal from a force transducer situated behind the cylinder. The rotation speed and the direction of rotation of the shafts (illustrated as ω_1 and ω_2 in Fig. 1) are generated by electric motors that are controlled with frequency converters. Flexible couplings are used to connect the motors and shafts allowing slight misalignment between the axles.

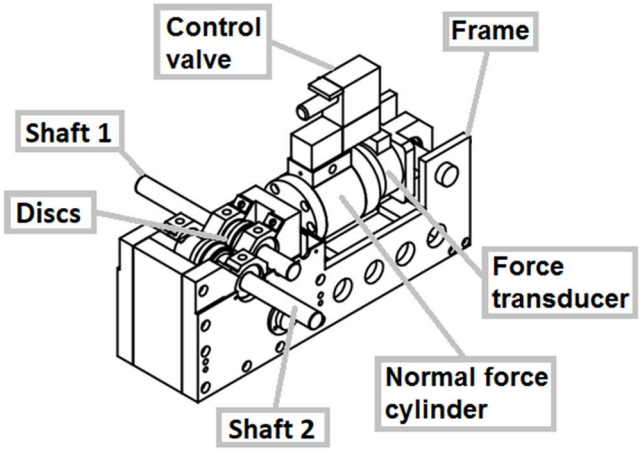

Fig. 2. Load frame.

Pressurised lubrication circulation is provided by a separate hydraulic unit. The same lubricant is directed to the disc contact as well as to the rolling bearings on the shafts. The overall flow of the lubricant is controlled by means of frequency-converter-driven pump/motor and the division of the flow between the bearings and

the disc contact can be adjusted with manual valves. The temperature of the lubricant is automatically adjusted by a microcontroller and can be set anywhere between 25 °C and 120 °C, while the total flow can be altered between 0.5 l/min and 20 l/min.

The lubricant used was a mineral oil –based commercial gear oil equipped with an EP-additive system. The oil is specifically designed for heavy-duty industrial gears and the specifications are shown in Table 1.

Table 1. Specification of the used lubricant.

		Value
Kin. viscosity @40ºC	[mm^2/s]	150
Kin. viscosity @100ºC	[mm^2/s]	15
Density @15ºC	[kg/m^3]	897
VG class	[-]	150
Flash point (COC)	[°C]	240
Pour point	[°C]	-24

2.1. Discs

The discs were made of 18CrNiMo7-6 (EN 10084). After turning, they were case-hardened reaching a surface hardness of 59-61 HRC. In the next manufacturing phase, the contact surfaces were ground as discs assembled on the shafts in order to minimise the risk of any eccentricity error between the bearings and the contact surface. The outer diameter of the discs at the contact surface is 70 mm, and the width of the ground contact area is 10 mm. The discs on the shaft 1 in Fig. 2 were ground to a radius of 100 mm in the axial direction (illustrated as r_1 in Fig. 1), while the discs on the shaft 2 were ground to be flat (infinite radius r_2 in Fig. 1). This geometry results in an elliptic pressure distribution (p in Fig. 1) in the contact.

In order to generate impact-like loading, three grooves were manually ground onto the contact area of the disc on shaft 2. The grooves were positioned at 90 degrees from each other along the circumference of the disc. The width of a groove is approximately 1.5 mm in the rolling direction and the depth about 1 mm as illustrated in Fig.3. The groove in Fig. 3 is not fully reaching the edge of the ground area on the left side, but this is acceptable since the midpoint of contact is in the middle of the ground area and the sides are not in contact at all because of the low radial load and the curvature of one of the discs.

Fig. 3. The disc with grooves.

2.2. Bearings

The shafts of the test device contain a total of four bearings; two per shaft on each side of the disc. All the bearings are two-row spherical roller bearings. The outer diameter is 62 mm, the bore diameter is 30 mm, and there is a total of 30 roller elements in a bearing divided into two rows. The basic dynamic load rating according to the manufacturer is 66.1 kN, and the basic static load rating is 60 kN.

Two artificial fault cases were generated. In the first one, three of the roller elements were damaged by manually grinding a groove across the width of three rollers. In the second case, three grooves were ground into the inner side of the outer race by using an electric multi-tool. The bearing damages are presented in Fig. 4.

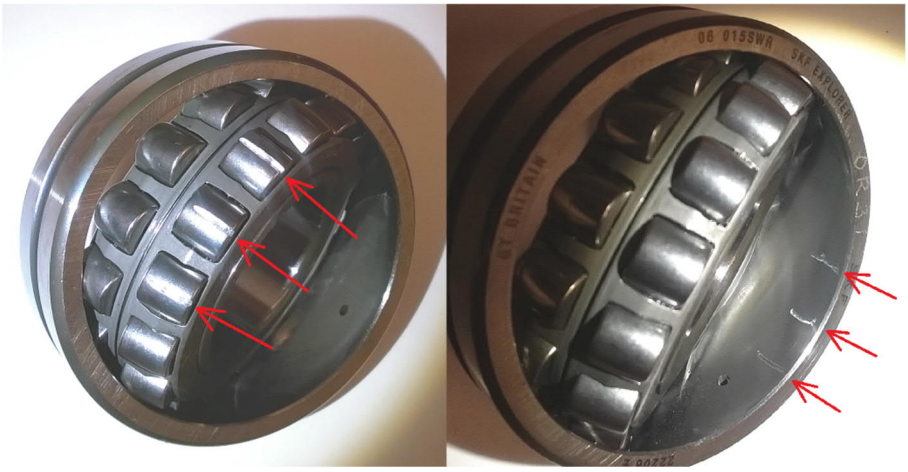

Fig. 4. The bearing damage on the roller elements on the left and the damage on the outer race on the right.

The bearing's characteristic frequencies at the rotation speed of 899.4 RPM (rotation speed of the shafts in the tests) are as follows [44]:

- Ball pass frequency outer, BPFO = 92.6 Hz
- Ball pass frequency inner, BPFI = 131.3 Hz
- Ball spin frequency, BSF = 80.9 Hz

Continuous pressurized oil flow was directed through the bearings throughout the tests providing sufficient lubrication in the tests (viscosity ratio κ being roughly 2), where the dynamic load rating of the bearing is significantly higher than the applied load.

2.3. Measurements

Two data acquisition systems were used in the experiments. The first one collected information about the lubricant temperature, the normal force and the rotation speeds of the shafts. In addition, a separate system recorded the acceleration response from the device frame aiming to catch also bursts originating from the damage in the bearing.

2.3.1. Acceleration measurement

The acceleration response was measured in the frame near the observed bearing by using a tri-axial accelerometer with a measurement range up to 50 g. The sampling frequency during the data recording was 20 kHz. The sensor was mounted by a thread to a pad that was glued to the frame, as illustrated in Fig. 5.

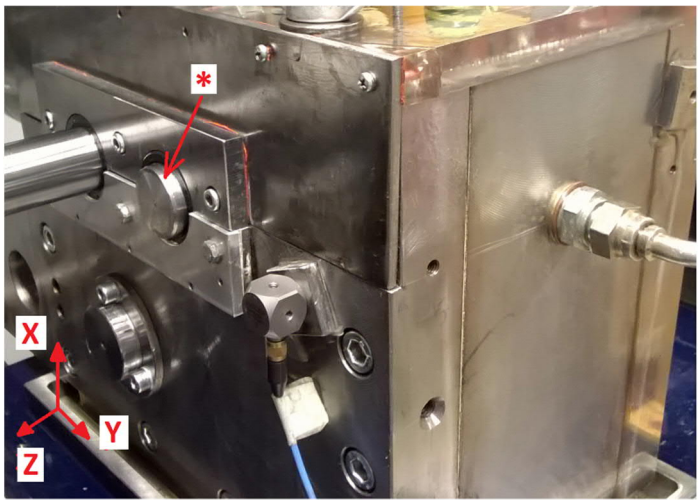

Fig. 5. The accelerometer and the measurement directions. The observed bearing is on the sensor side of the frame assembled on the shaft marked *.

3. Experimentation and analysis

Each test began by rotating both shafts at the same speed of 600 RPM for approximately an hour without any radial loading, but with lubrication circulation. During this period the temperatures of the bearings, shafts and the lubricant were stabilised. The lubricant temperature was set at 60 °C, which is a typical lubricant temperature in powertrain transmissions [45, 46].

The rotation speed of the shafts was 900 RPM in all the tests. The radial load applied to the bearings through the disc contact was 3.0 kN, yielding a radial load of 1.5 kN on the observed bearing. The axial load on the bearing was estimated to be insignificant; since it originated mainly from deformations of the system, which are negligible at the used load level.

The acceleration response recorded during the tests was used in the analysis. The following cases were analysed:

- **A.** Reference: No damage in the bearing
- **B.** Three grooves on the bearing outer race
- **C.** One groove on three rolling elements

All the tests were repeated with discs having no grooves on the contact surface (indicated with “a”, e.g. Aa) and with grooves on one of the test disc surface (indicated with “b”, e.g. Ab).

3.1. Analysis approach

Firstly, a 100 s period of the data was chosen from the recorded acceleration signal. Secondly, band-pass filtering from 5 Hz to 10 kHz was applied to this data. The subsequent examination included the calculation of kurtosis as well as envelope analysis with specific pre-enhancements to improve the possibility of detecting bearing damage.

Kurtosis is a statistical parameter observing the deviations in the measured data. Kurtosis K_t can be regarded as comparing the tails of the probability density function with that of the normal distribution [47]:

$$K_t = \frac{M_4}{\sigma^4} \quad (1)$$

where σ^2 is variance and M_4 is the fourth central moment, which can be written as expectation $E(l)$ with the time varying signal $x(t)$ and estimator of the mean $\bar{x}$ [47]:

$$M_4 = E(l) = E[(x - \bar{x})^4] \quad (2)$$

An undamaged bearing is assumed to produce a normally distributed vibration content and yield a kurtosis value of 3. If the kurtosis value is higher, the distribution has “larger tails” and vice versa in the case that the kurtosis is lower than 3. The presence of larger tails indicates high peaks in the observed signal (impulsivity), which may be initiated by damage in the bearing parts.

3.1.1. Blind band selection

The idea of the spectral kurtosis (SK) is to identify frequency bands containing a signal of maximum impulsivity [48]. In machinery, these impulses may originate for instance from a failure in the rolling bearing. As the outcome, the SK indicates how the possible impulsiveness (non-Gaussian or non-stationary behaviour) of a signal, is distributed in the frequency domain. To improve the efficiency of the original SK analysis to cope with a wider class of non-stationary signals and to decrease computational cost, the fast kurtogram (FK) was developed. In the FK analysis the signal is subject to multiple parallel band-pass filtering operations where the SK is calculated in each band to pinpoint the centre frequency and the bandwidth of high non-stationary content [4].

The FK is probably the most commonly used band selection method and was also utilized in this paper by using the 1/3-binary-tree-structure for the division of the bands (maximum level was set at seven). The frequency band with the highest SK value was used in the subsequent analysis.

3.1.2. Self-adaptive noise cancellation

Self-adaptive noise cancellation (SANC) can be used to separate discrete frequency components (deterministic part) from random components in the signal. The aim is to remove masking components which reside in the same frequency range as the bearing damage components and therefore cannot be removed by using band-bass filtering. The working principle of SANC is to use a delayed version of the primary signal as a reference signal to make the separation. If the delay is longer than the correlation length of the random signal, the adaptive filter will not recognize the relationship, but it can then find the transfer function between the delayed version of itself and the deterministic part. The recursive algorithm can be expressed as follows [49]:

$$W_{k+1} = W_k - \mu\nabla_k \quad (3)$$

where the gradient vector can be written

$$\nabla_k = \frac{\partial E[\varepsilon_k^2]}{\partial W_k} \quad (4)$$

and μ is a convergence factor. The algorithm was used in this study by using a conservative approximation presented by Ho [50] for W_{k+1}:

$$W_{k+1} = W_k + \frac{2\mu_n \varepsilon_k X_k}{(L+1)\sigma_k^2} \quad (5)$$

where W_k is the vector of weight coefficients of the adaptive filter at the kth iteration, μ_n is the normalized convergence factor, μ is the convergence factor ($\mu = \mu_n/((L+1)\sigma_k^2)$), ε_k is the output error at kth iteration, X_k is vector of the input values at the kth iteration, L is the order of the adaptive filter, $(L+1)$ is the number of filter coefficients, and σ_k^2 is the exponential-averaged estimate of the input signal power at the kth iteration.

3.1.3. Spectral kurtosis and autoregressive model

By using 'autoregressive' (AR) linear prediction models, the deterministic part of the signal can be obtained. Once the deterministic parts have been removed from the data, impulses initiating from bearing damage may appear in a clearer way. Spectral kurtosis can be used to further accent them [8].

In AR filter design for linear prediction, a value of the signal x at time k is normally based upon a linear combination of prior values as follows [51]:

$$x_k = -\sum_{i=1}^{p} a_i x_{k-i} + \mathcal{E}_k \quad (6)$$

where a_i is AR coefficient and ε_k is the residual error that is the difference between the actual signal and the AR prediction. Basically, the AR filter predicts the deterministic pattern of the signal, but not the sudden impulses initiating from localized damage, which are left in the residual ε_k. Spectral kurtosis can be applied to the residual signal, while high values may indicate about the presence of impulses from faults and their clearer extraction can be obtained. The AR based approach was implemented according to [8] in this paper by using the Yule-Walker method to compute the parameters for an AR model of order 100.

3.1.4. Minimum entropy deconvolution

Externally measured vibration signals are often significantly distorted by the transmission paths from the source to the transducer. This is typically the case for signals that result from internal sharp impacts initiating, for instance, from local damage in bearings. The minimum entropy deconvolution (MED) method is designed to reduce the spread of impulse response functions (IRF), aiming to obtain signals that are closer to the original impulses that gave increase to them. The idea is to find an inverse filter that counteracts the effect of the transmission path. The assumption is that the original excitation is impulsive and therefore has a high kurtosis. Basically, in the MED method the optimum filter coefficients of the inverse filter are determined by maximization of the kurtosis of the inverse filter output y. The filter f is modelled as a FIR filter with L coefficients and v is the input signal including the damage impulses [49]:

$$y[n] = \sum_{l=1}^{L} f[n] v[n-l] \quad (7)$$

where the f[i] must invert the system IRF h[i] such that:

$$f[i] * h[i] = \delta(i - lm) \quad (8)$$

The delay l_m allows the inverse filter to be causal. The MED filter was implemented in this study as presented in Ref. [8] by first removing the periodic part with the AR method and then by utilizing the objective function method (OFM) [52] in maximizing the kurtosis of the MED output.

3.1.5. Envelope analysis

Failure detection through envelope analysis is based on the recognition of acceleration bursts originating from localised damage in the various components of a rolling bearing by utilising an amplitude-demodulation

procedure. In the frequency spectrum of the envelope signal, the presence of clear peaks at specific bearing frequencies, their multiples and side bands on both sides of the specific frequency can be considered as indicators of damage in the bearing elements.

Two approaches were used to generate envelope signals. Firstly, the envelope curve was produced in a traditional way by utilising a FIR Hilbert transform filter of length 100. In this approach the Hilbert transform of a signal $x(t)$ is used to find an approximation of the imaginary part $\tilde{x}(t)$ (by taking into account also delay) of the so-called analytical signal $z(t)$ (complex variable) of the signal $x(t)$ [47]:

$$z(t) = x(t) + j\tilde{x}(t) \quad (9)$$

The envelope $E_{Hil}(t)$ of the signal $x(t)$ is defined as

$$E_{Hil}(t) = |z(t)| = \sqrt{x^2(t) + \tilde{x}^2(t)} \quad (10)$$

The second method of generating the upper envelope signal was a log-based approach, which has been showing promising results in displaying cyclostationary bearing-fault symptoms [5, 53]. In this study the log-based envelope curve $E_{log}(t)$ of the signal $x(t)$ was produced as follows:

$$E_{log}(t) = log(|x(t)|^2) \quad (11)$$

It should be noted that in this study the logarithm is calculated from the square of the signal.

4. Results and discussion

During the tests, the lubricant temperature and the radial load varied slightly according to the measurements. Typically, the lubricant temperature remained between 60 ± 1.5 °C. In the radial force signal peaks deviating roughly ± 10 % from the setting value were noted. It can be assumed that the scatter in these parameters had only a neglectable effect on the detection of the bearing damages from the captured acceleration response.

The upper graphs (a, b) in Fig. 6 show the shape of a typical measured acceleration signal in the time domain. In graph b, the three impulses resulting from the grooves on the disc during one shaft rotation are clearly visible. The effect of the impacts can be seen also in the frequency domain (Fig. 6 d) as several heavily amplified resonance peaks, for instance, at approximately 1 kHz and 2.5 kHz, which spectral amplitude is much higher than in Fig. 6 c (no impact load). Additionally, graphs e and f illustrate the deviation of the acceleration peaks from the mean level. In this analysis, a peak is the maximum/minimum value during a (local) signal period, while the signal values are above/below the mean level. Furthermore, the normal distribution was fitted to the data and scaled to the maximum of the histogram as shown in Fig. 6. Although the peaks in test Aa seem to be divided into two areas, the correlation is relatively high at 0.92. However, the data recorded in test Ab is heavily non-gaussian and the correlation is 0.48. This is caused by the tails at both ends of the histogram originating from the impact loading.

According to the frequency spectrum, the rotation speed was 899.4 RPM, which equals to 14.99 Hz. Correspondingly, the BPFO was then 92.6 Hz, the BPFI 131.3 Hz and the BSF 80.9 Hz.

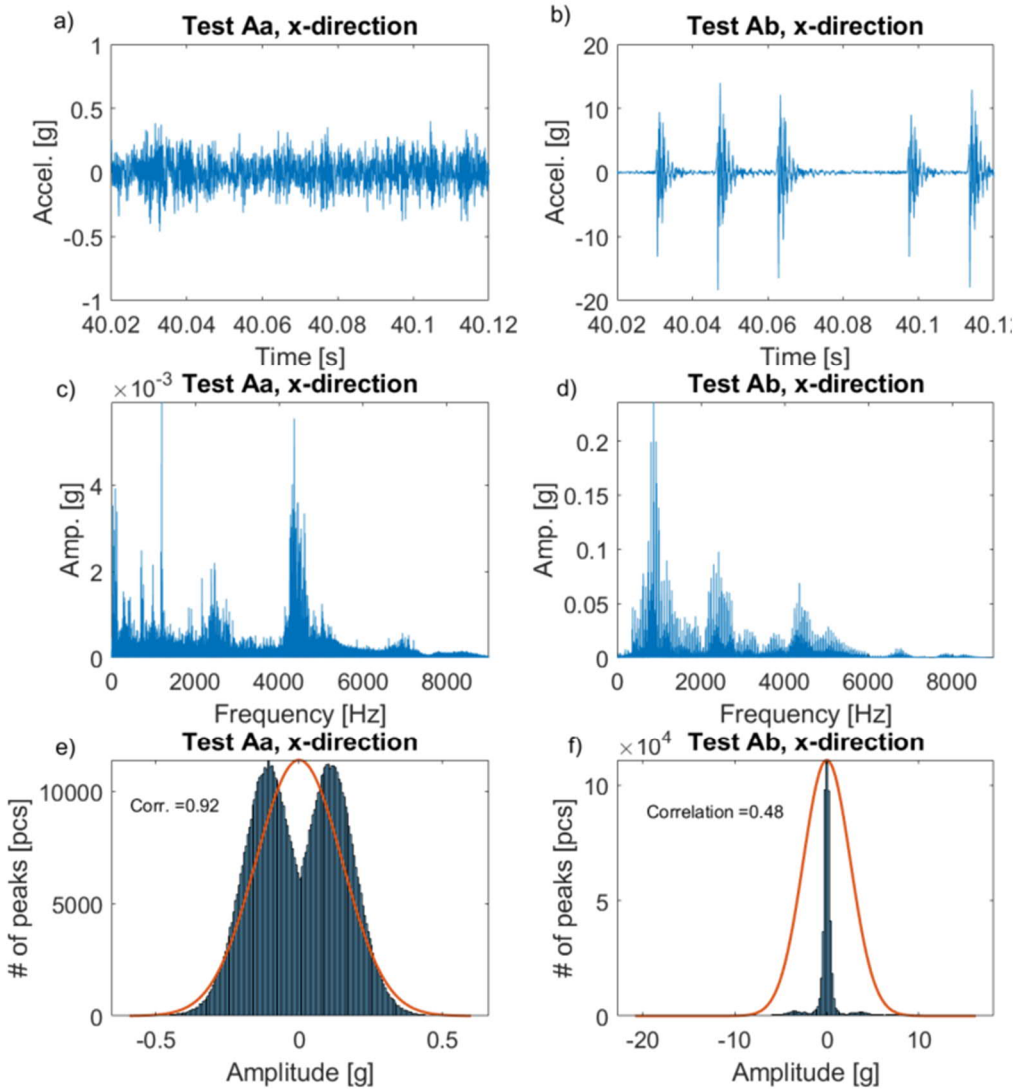

Fig. 6. Acceleration signals without and with impact load in time (a, b) and frequency domains (c, d). Distribution of peaks and fitting of the normal distribution (e, f).

4.1. Presence of damage

The signal analysis started by generating the envelope spectrums directly based on the measured data. To improve the visibility of the interesting peaks in the spectrum, the bearing characteristic frequencies were noted with the help of two adjacent vertical lines, where the particular frequency peak should appear in between, if it is visible in the spectrum. The abbreviation of the frequency is marked above the graph and the sidebands are presented as single lines. In Fig. 7, the envelope spectrums for all the test cases without the impact loading are presented by utilizing the Hilbert transform approach.

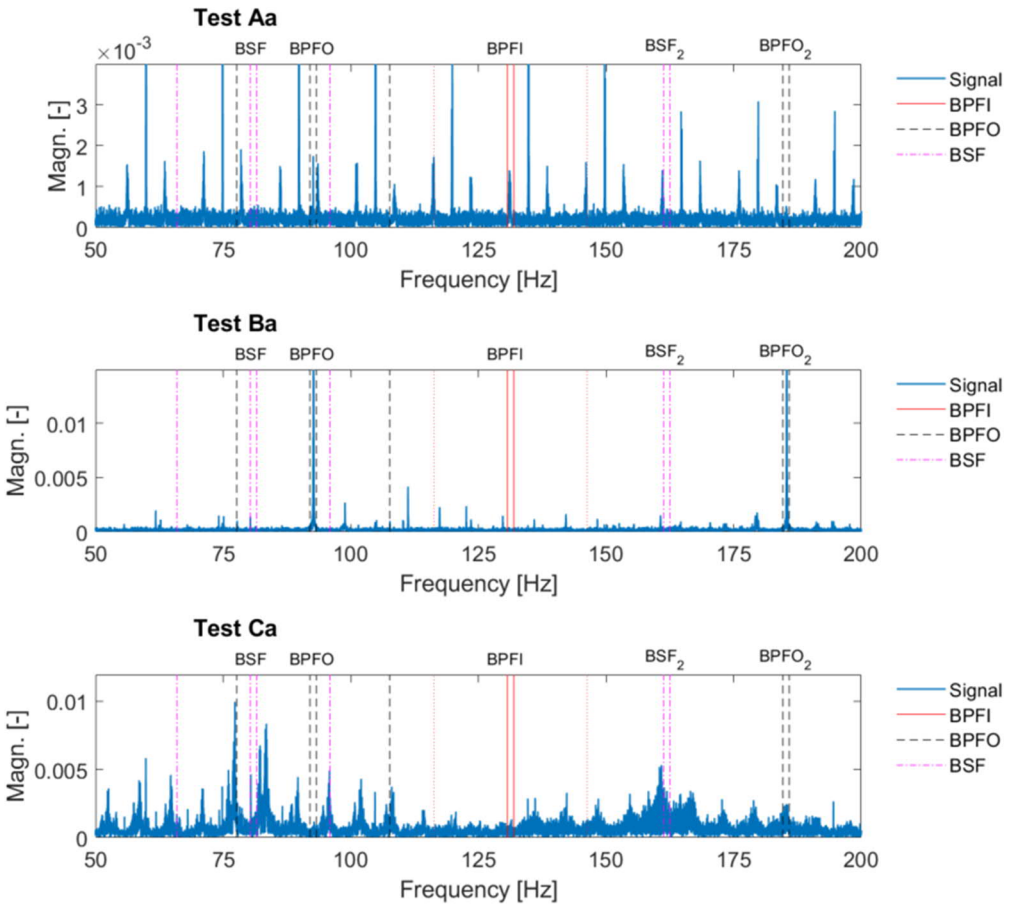

Fig. 7. Envelope spectrums with Hilbert transform-based approach (E_{Hil}) for tests Aa, Ba and Ca.

Several clear peaks are visible in all the graphs at multiples of the rotation frequency (14.99 Hz). For test Aa, where there isn't any damage present in the bearing, the pass frequencies are not seen except as a minor peak in the outer race pass frequency (BPFO). On the other hand, in test Ba, where outer race damage is present, peaks for the pass frequency (BPFO) and its second multiplier clearly separate from the frequency content.

The rolling element spin frequency (BSF) in the test Ca is visible along with a sideband on the right side. In addition, a rise in the spectrum can be seen also at the second multiple of the BSF. These peaks together indicate about damage although not as clearly as in case of the outer race damage (Ba).

Fig. 8 illustrates the envelope spectrums by using the log-based approach to create the envelope curve for tests Aa, Ba and Ca. The content of these graphs regarding the bearing pass frequencies is quite like the results in Fig. 7.

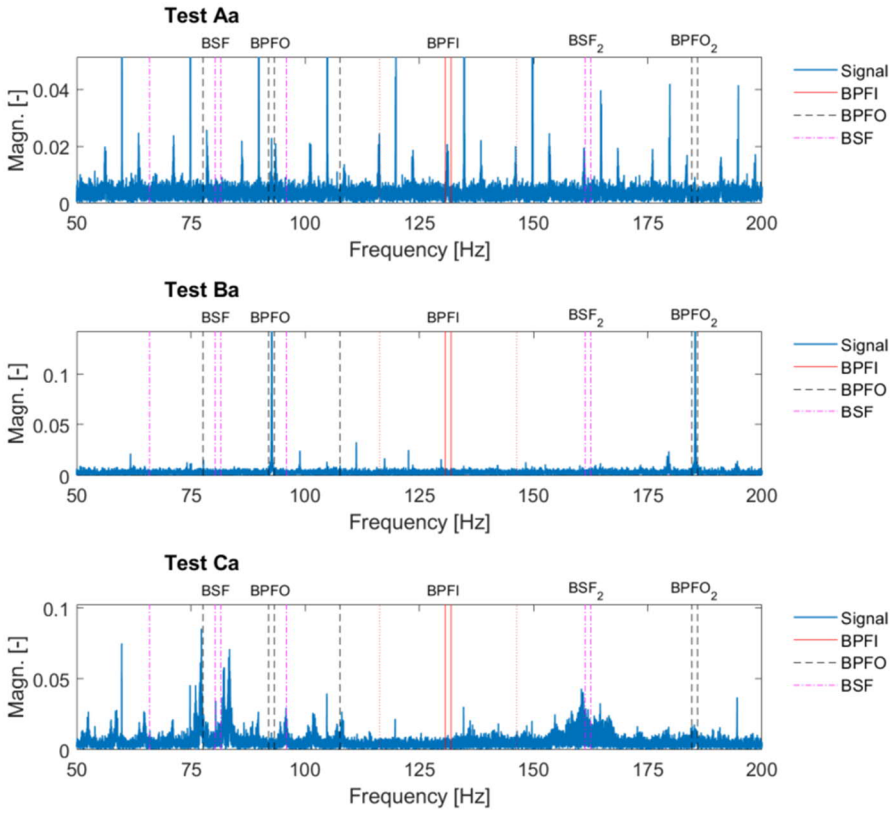

Fig. 8. Envelope spectrums with log-based approach (E_{log}) for tests Aa, Ba and Ca.

Kurtosis was calculated for tests Aa, Ba and Ca resulting in 3.1, 6.7 and 22.8. The results for case Aa indicate damage-free components as it should (the value is roughly 3). In cases Ba and Ca damage is present according to the kurtosis values. In addition, in test Ca some other source than the bearing error might have had an influence on the high kurtosis value, which is also implied by the general high magnitude of the frequency peaks for test Ca in comparison to the graphs of the other tests.

4.2. Effect of impact loading

Multiples of the impact (shaft rotation) frequency can be seen as high peaks in the frequency spectrum by using both of the methods in the generation of the envelope signal as shown in Figs. 9 and 10. The outer race pass frequency seems to be visible in all the tests and with both envelope methods, but its second multiple is visible only in test Bb, as it should. In addition, the peaks for the sidebands around the BPFO frequency are more clearly visible in Fig. 10 for the test Bb than in the corresponding spectrum in Fig. 9. This suggests that the use of the log-based approach can give stronger indication about the bearing condition in comparison to the Hilbert transform-based approach.

In test Cb, with the damaged rolling elements, the corresponding pass frequency (BSF) does not stand out clearly from the spectrums in Fig. 9 or 10 even though one sideband peak on the left side can be noticed in both figures. The peak of its second multiple (BSF_2) seems to be elevated when using the Hilbert transform-based analysis approach although it is clearly visible also in test Ab and Bb in Fig. 9 where it probably should not be, because of the absence of the damage in the rolling elements in these tests.

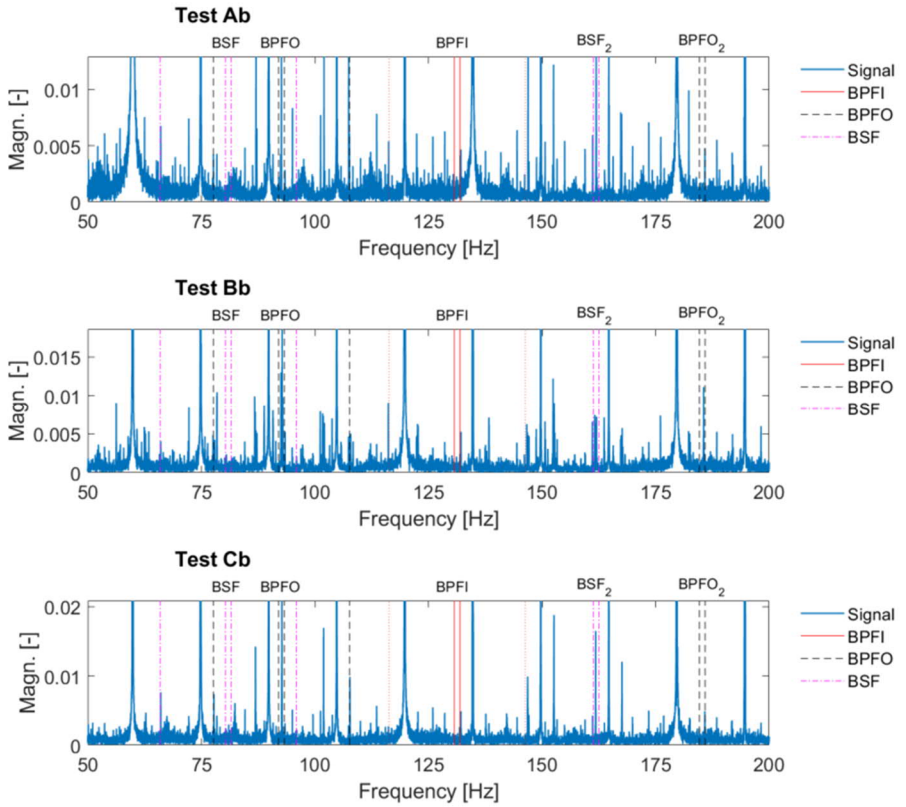

Fig. 9. Envelope spectrums with Hilbert transform-based approach (E_{Hil}) for tests Ab, Bb and Cb.

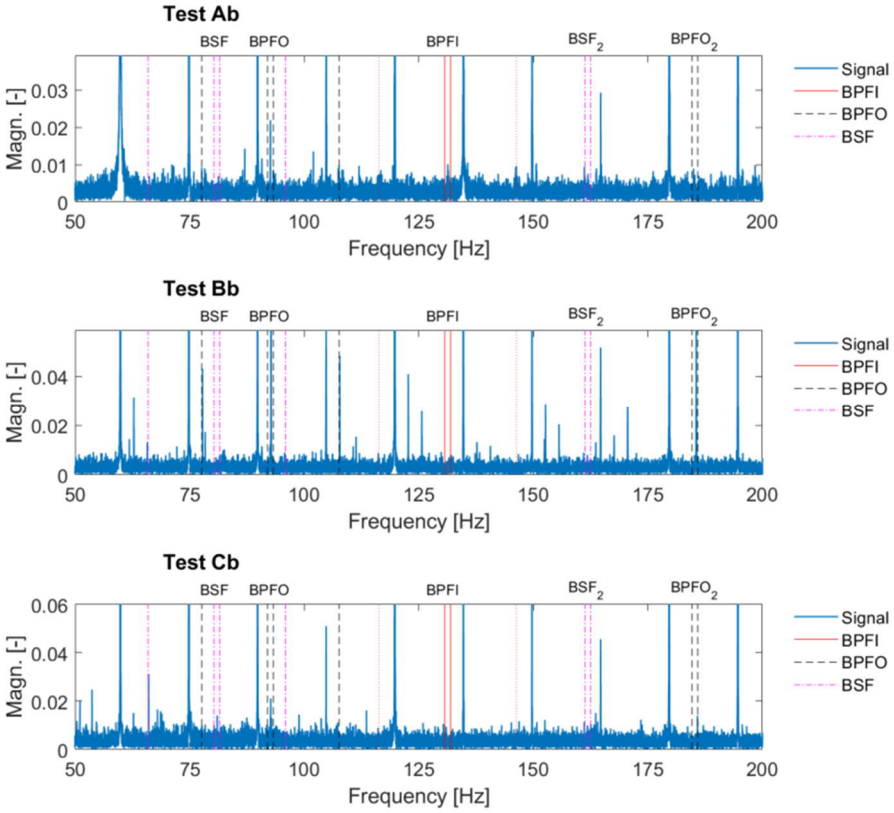

Fig. 10. Envelope spectrums with log-based approach (E_{log}) for tests Ab, Bb and Cb.

The kurtosis results for the tests Ab, Bb and Cb are 24.4, 42.7 and 52.4, respectively. These relatively high values are clearly influenced by the impact load, producing unreliable indications about bearing damage.

Generally, the results suggest that damage in the bearing outer race seems to be detectable even when the frequency content is significantly affected by impact loading. However, for the bearing rolling element damage the case is the opposite. One of the reasons for this may be that the response from the outer race is stronger because all the rolling elements meet the damage on the outer race during bearing operation, while the three damaged rolling elements initiate vibration response only when the damaged location is directly facing the outer or inner race of the bearing.

4.2.1. Rolling element damage

As noted earlier the damage on the rolling elements appears to be slightly challenging to be detected especially when distorting components are present. In addition, the use of the traditional method of creating the envelope analysis based on the Hilbert transform seems to rise a peak at the second multiple of the rolling element spin frequency (BSF_2) in the envelope spectrum even though any failure isn't present in the rolling elements. Therefore, in this section the focus is on further processing of the signal before the envelope analysis to highlight the ball spin frequencies in the spectrum generated by only using the log-based approach.

Three signal processing enhancement approaches were applied on the recorded data from tests Ca and Cb before calculating the envelope spectrum. Firstly, band selection alone was done by using FK approach. The second method was inspired by the work in refs. [8, 9], where the AR manipulation was followed by the MED calculation. In the third attempt, first the FK selected the interesting frequency band and then the SANC analysis was used to elevate the ball spin frequencies.

Fig. 11 illustrates the effects of the FK, the AR and MED and the SANC enhancements in the test Ca. The efficiency of the approaches was investigated by the trial-and-error method. As the outcome the MED filter size of 10 as well as 35 for SANC (with convergence factors 0.01 and 7.1) were found to give elevated performance in terms of highlighting ball spin frequencies (BSF, BSF_2) in the spectrum.

The graph in the Fig. 11 shows that the use of the FK doesn't significantly increase the BSF or BSF_2 peaks in the spectrum in comparison to the spectrum without any additional processing, although a clear frequency band was highlighted as illustrated in Fig. 12. The main influence of the FK calculation can be seen as an increase in the height of the sideband peak on the right side of the BSF frequency. However, the AR+MED calculation seems to mildly elevate the second ball spin frequency, but at the same time the first one is slightly decreased. Slight decrease can be noted also in the peak of the BSF sideband. The SANC method applied after FK band selection results in an increase in the peak near the BSF as well as the right-side sideband peak raising the probability of detecting the failure.

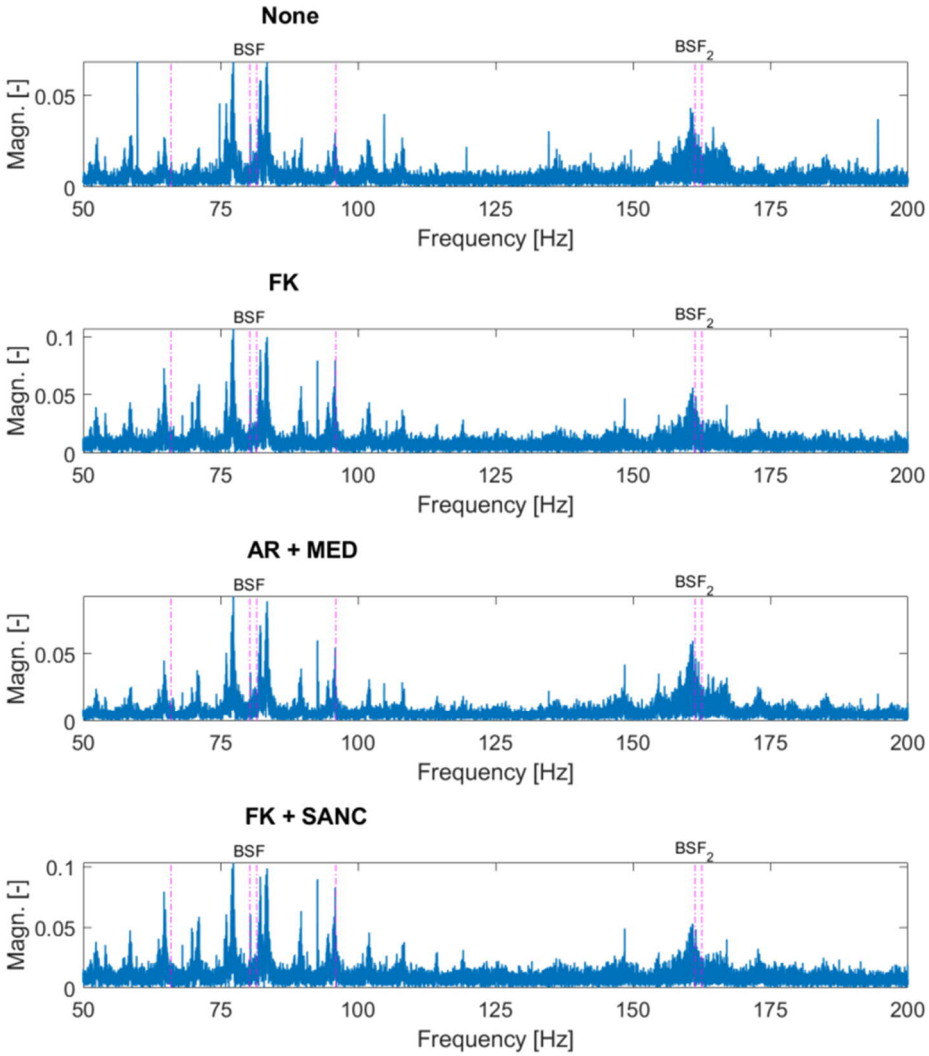

Fig. 11. Envelope spectrums in case Ca. The uppermost graph shows the original spectrum, while in the lower graphs the signal was processed by using FK, AR and MED and FK combined with SANC.

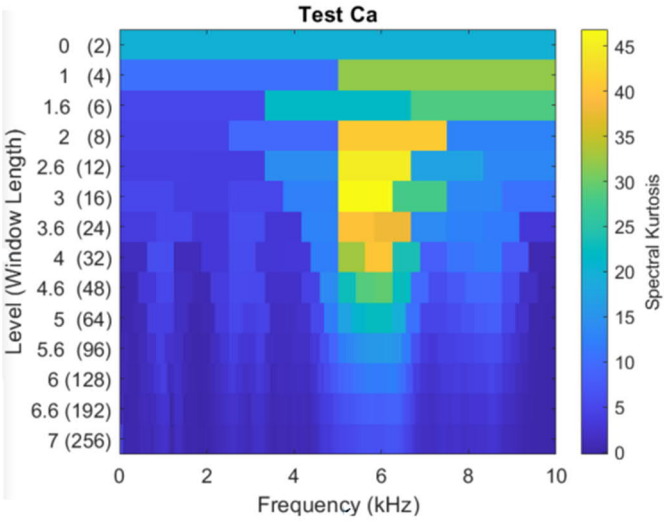

Fig. 12. Kurtogram in test Ca. The frequency band on level 3 with the centre frequency of 5.625 kHz and bandwidth 1.25 kHz was used in the analysis.

The effects of the FK, the AR and MED and the SANC enhancements in the test Cb are presented in Fig. 13. The kurtogram related to case Cb is illustrated in Fig. 14. For elevated efficiency the filter sizes for MED and SANC were 98 and 174, while the convergence factors were 0.01 and 2.1 respectively.

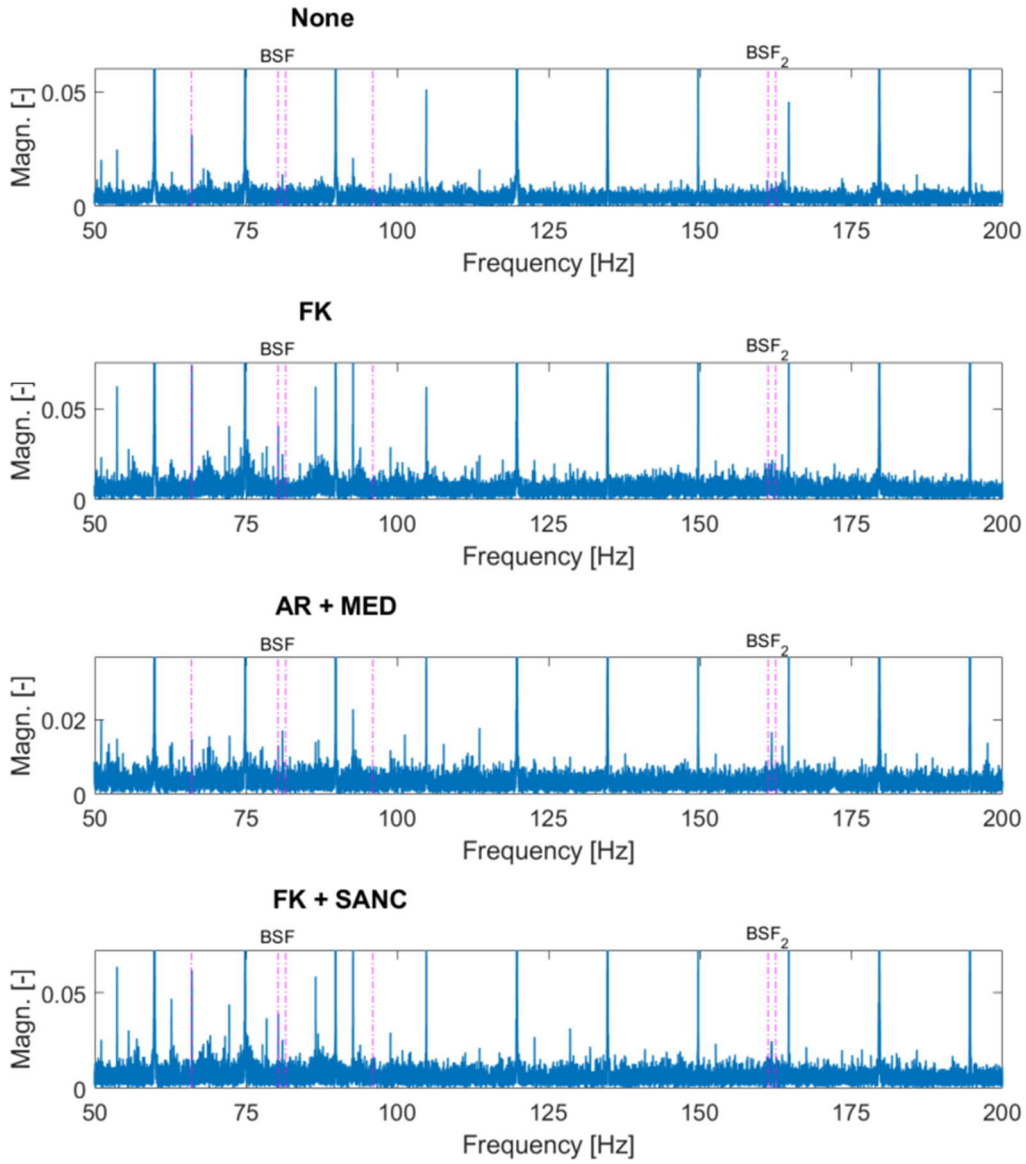

Fig. 13. Envelope spectrums in case Cb. The uppermost graph shows the original spectrum, while in the lower graphs the signal was processed by using FK, AR and MED and FK combined with SANC.

According to the graphs in Fig. 13 the BSF and BSF_2 are not highlighted in the upmost spectrum indicating failure-free bearing when any enhancements aren't applied. Also, the use of FK band selection alone doesn't change the conclusion about the bearing condition although the BSF peak and a sideband on the left of the BSF are elevated because of using the discovered frequency band illustrated in Fig. 14. However, processing with the AR+MED method the peaks at the BSF and BSF_2 are elevated in the spectrum indicating not only the damage, but also the capability of the algorithms in these methods to pick and enhance the weak signal originating from the rolling element damage from the full measured acceleration content disturbed with

external signals. The disadvantage of the use of this approach seems to be that the sideband peak on the left of the BSF is slightly decreased although it still remains visible.

The FK+SANC approach seems to deliver similar results, as the AR+MED, excluding the downgrade of the left sideband peak. However, this approach tends to also elevate other peaks in the spectrum more than the AR+MED method as it can be seen when comparing the spectrum contents in Fig. 13 for instance between 50 Hz and 100 Hz. This may impede the damage detection process in some cases.

Despite of the promising results it should be noted that a careful selection of the parameters for these approaches are needed to clearly reveal the BSF and the BSF_2 peaks from the spectrum.

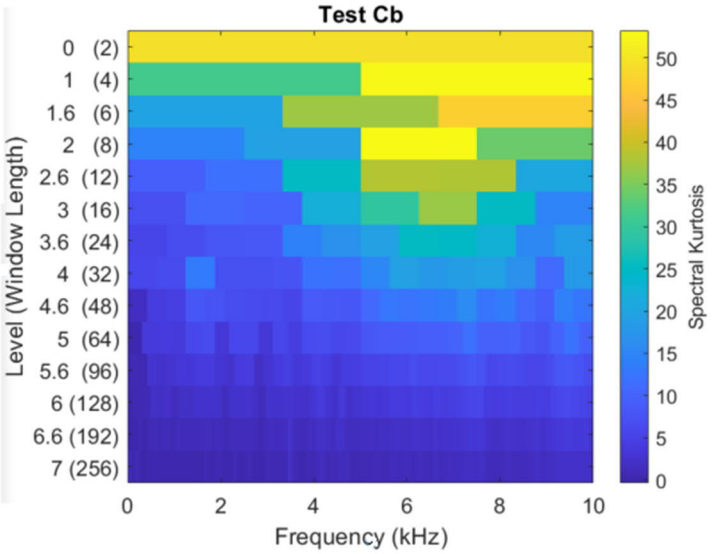

Fig. 14. Kurtogram in test Cb. The frequency band on level 2 with the centre frequency of 6.25 kHz and bandwidth 2.5 kHz was used in the analysis.

5. Conclusions

In this study, a unique experimental approach to investigate rolling element bearing damage, both as impact-free and under the influence of impact loading, was presented with analysis results of the related measurement data. The main conclusions are as follows:

- The twin-disc test device suits well to study different forms of rolling element bearing damage and their detection methods in varying load conditions.
- Detection of damage on the bearing outer race seemed to be quite straight-forward by utilizing the traditional approaches like calculating kurtosis or performing envelope analysis using the Hilbert transform.
- Detection of damage on the rolling elements, which vibration response is relatively weak, can be challenging especially under the influence of impact loading:
    - The choice of the method to generate the envelope signal affects the success of the detection process. The log-based approach showed promising results.
    - Processing of the obtained signal by using the AR-combined-with-MED -approach significantly enhanced recognition of the failure by highlighting the ball spin frequency and its second multiple (BSF and BSF2) in the envelope spectrum of the impact-load-disturbed signal. Alternatively, the FK-SANC -approach showed similar performance. Nevertheless, it should

be noted that to achieve good result, a careful choice of the settings for these methods is required.

The used well-known signal processing approaches showed relatively good performance in detecting damage in situations when the system is heavily influenced by impact loading. This suggests that they could be utilized successfully also when evaluating the condition of bearings in real machinery working in harsh environment. However, several challenges remain for further studies. Firstly, in the tests the impact loading was applied to the system on a regular basis (perfectly periodic), while in many real applications, it may be more random. Secondly, the damage in the bearings was artificially made, meaning that it was clearly noticeable on the component surface and very local. In some bearing failure modes, small damage marks may occur in large areas of the bearing instead, which may influence the likelihood of detection. Nevertheless, the flexibility of the twin-disc test device in terms of adjustability and structural changes makes it possible for the future research to focus on specific types of failure modes, loads, lubrication conditions as well as analysis approaches.

6. Acknowledgements

The research work has been carried out under the auspices of the project OPMO - Operation monitoring of mineral crushing machinery. This activity has received funding from the European Institute of Innovation and Technology (EIT RawMaterials project number 18253). This body of the European Union receives support from the European Union's Horizon 2020 research and innovation programme. The funder has had no influence on the content of this paper.

References

1. McFadden PD, Smith JD. Model for the vibration produced by a single point defect in a rolling element bearing, Journal of Sound and Vibration, 1984;96 (1);69–82.

2. Darlow MS, Badgley RH, Hogg GW. Application of high frequency resonance techniques for bearing diagnostics in helicopter gearboxes, Technical Report, US Army Air Mobility Research and Development Laboratory, 1974. Pp. 74–77.

3. Randall RB, Antoni J. Review. Rolling element bearing diagnostics—A tutorial. Mechanical Systems and Signal Processing, 2011;25;485–520.

4. Antoni, J, Randall RB. Fast Computation of the Kurtogram for the Detection of Transient Faults. Mech Syst Sig Proc, 2007;20 (1);108–24.

5. Smith WA, Borghesani P, Ni Q, Wang K, Peng Z. Optimal demodulation-band selection for envelope-based diagnostics: A comparative study of traditional and novel tools. Mech Syst Sig Process, 2019;134;1–24.

6. Zhang Y, Randall RB. Rolling element bearing fault diagnosis based on the combination of genetic algorithms and fast kurtogram, Mech. Syst. Signal Process, 2009;23;1509–1517.

7. Wiggins RA. Minimum entropy deconvolution. Geoexploration 1978;16;21–35.

8. Sawalhi N, Randall RB, Endo H. The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis. Mech. Systems and Signal Processing 2007;21;2616–2633.

9. Endo H, Randall RB. Application of a minimum entropy deconvolution filter to enhance autoregressive model based gear tooth fault detection technique. Mechanical Systems and Signal Processing Journal 2007;21 (2);906–19.

10. Boumahdi M, Lacoume J. Blind identification using the kurtosis: results of field data processing. IEEE Transactions on Signal Processing, 1995;1960–83.

11. Barszcz T, Sawalhi N. Fault detection enhancement in rolling element bearings using the minimum entropy deconvolution. Arch Acoust 2012;37(2):131-141.

12. Endo H, Randall RB. Enhancement of autoregressive model based gear tooth fault detection technique by the use of minimum entropy deconvolution filter. Mechanical Systems and Signal Processing, 2007;21(2);906-19. https://doi.org/10.1016/j.ymssp.2006.02.005.

13. Al-Bugharbee H, Trendafilova I. A fault diagnosis methodology for rolling element bearings based on advanced signal pretreatment and autoregressive modelling. Journal of Sound and Vibration, 2016;369;246-265.

14. Guo J, Zhang H, Zhen D, Shi Z, Gu F, Ball AD. An enhanced modulation signal bispectrum analysis for bearing fault detection based on non-Gaussian noise suppression. Measurement, 2020;151(107240). https://doi.org/10.1016/j.measurement.2019.107240.

15. Bastami AR, Vahid S. Estimating the size of naturally generated defects in the outer ring and roller of a tapered roller bearing based on autoregressive model combined with envelope analysis and discrete wavelet transform. Measurement, 2020;159(107767). https://doi.org/10.1016/j.measurement.2020.107767

16. Antoni J, Randall RB. Unsupervised noise cancellation for vibration signals: part I—evaluation of adaptive algorithms. Mechanical Systems and Signal Processing, 2004;18(1);89-101. https://doi.org/10.1016/S0888-3270(03)00012-8

17. Randall RB, Li L. Diagnostics of planetary gear bearings in the presence of gear vibrations. Second International Conference on Gearbox Vibration and Diagnostics, Imeche, London, 1995, pp. 73–80.

18. Zhou L, Duan F, Mba D, Wang W, Ojolo S. Using frequency domain analysis techniques for diagnosis of planetary bearing defect in a CH-46E helicopter aft gearbox. Engineering Failure Analysis, 2018;92;71-83. https://doi.org/10.1016/j.engfailanal.2018.04.051

19. Zhou L, Duan F, Corsar M, Elasha F, Mba D. A study on helicopter main gearbox planetary bearing fault diagnosis. Applied Acoustics 2019;147;4-14; https://doi.org/10.1016/j.apacoust.2017.12.004

20. Vlcek BL, Zaretsky EV. Rolling-element fatigue testing and data analysis – A tutorial. Trib. Transactions, 2011;54;523-41.

21. Brizmer, V, Gabelli A, Vieillard C, Morales-espejel GE. An experimental and theoretical study of hybrid bearing micropitting performance under reduced lubrication. Trib. Transactions 2015;58;829-35.

22. Fowell M, Ioannides S, Kadiric A. An Experimental Investigation into the Onset of Smearing Damage in Nonconformal Contacts with Application to Roller Bearings. Trib. Transactions 2014;57;472-88.

23. Pape F, Terwey JT, Wiesker S, Averbeck S, Muhmann C, Lipinsky D. Tribological research on the development of white etching cracks (WECs). Forsch Ingenieurwes 2018;82;341–352. https://doi.org/10.1007/s10010-018-0289-6

24. Galbato, AT. Methods of Testing for Rolling Contact Fatigue of Bearing Steels. Rolling Contact Fatigue Testing of Bearing Steels, Hoo, J. J. C. (Ed.), American Society of Testing Materials: Philadelphia, 1982, pp 169-189.

25. Harris TA., Kotzalas MN. Advanced Concepts of Bearing Technology: Rolling Bearing Analysis, Fifth Edition, CRC Press LLC, Taylor & Francis Group, 2007.

26. Kleemola J, Lehtovaara A. An approach for determination of lubricant properties at elliptical elastohydrodynamic contacts using a twin-disc test device and a numerical traction model. Proc Inst Mech Eng 2008;222:797–806.

27. Simon V. Load distribution in spiral bevel gears. ASME. J Mech Des 2006;129:201–9 http://dx.doi.org/10.1115/1.2406090.

28. Ahlroos T, Ronkainen H, Helle A, Parikka R, Virta J, Varjus S. Twin disc micropitting tests. Tribol Int 2009;42;1460-6.

29. Oila A, Bull SJ. Assessment of the factors influencing micropitting in rolling/sliding contacts. Wear 2005;258;1510-24.

30. Kleemola J, Lehtovaara A. Experimental evaluation of friction between contacting discs for the simulation of gear contact. TriboTest 2007;13(1):13–20.

31. Johansson J. On the influence of gear oil properties on pitting life. Luleå Sweden: Dissertation, Luleå University of Technology; 2015.

32. Tassini N, Quost X, Lewis R, Dwyer-Joyce R, Ariaudo C, Kuka N. A numerical model of twin disc test arrangement for the evaluation of railway wheel wear prediction methods. Wear 2010;268:660–7.

33. Höhn B-R, Michaelis K, Kreil O. Influence of surface roughness on pressure distribution and film thickness in EHL-contacts. Tribol Int 2006;39:1719–25.

34. Savolainen M, Lehtovaara A. An experimental approach for investigating scuffing initiation due to overload cycles with a twin-disc test device. Tribol Int 2017;109:311–8.

35. Savolainen M, Lehtovaara A. An experimental investigation of scuffing initiation due to axial displacement in a rolling/sliding contact. Tribol Int 2018;119:688–97.

36. Savolainen M, Lehtovaara A. An approach to investigating subsurface fatigue in a rolling/sliding contact. Int J Fatigue 2018; 117: 180-188. https://doi.org/10.1016/j.ijfatigue.2018.08.014

37. Junchao Guo a, Hao Zhang a, Dong Zhen a,⇑, Zhanqun Shi a, Fengshou Gu b, Andrew. D. Ball. An enhanced modulation signal bispectrum analysis for bearing fault detection based on non-Gaussian noise suppression. Measurement, 2020;151(107240);

38. Borghesani P, Pennacchi P, Chatterton S. The relationship between kurtosis- and envelope-based indexes for the diagnostic of rolling element bearings. Mech Syst Sig Process, 2014; 43(1–2); 25-43. https://doi.org/10.1016/j.ymssp.2013.10.007

39. Schmidt S, Heyns PS, Gryllias KC. An informative frequency band identification framework for gearbox fault diagnosis under time-varying operating conditions. Mech Syst Sig Process, 2021;158; https://doi.org/10.1016/j.ymssp.2021.107771

40. Niu L, Cao H, He Z, Li Y. A systematic study of ball passing frequencies based on dynamic modeling of rolling ball bearings with localized surface defects. Journal of Sound and Vibration, 2015;357; 207-32. https://doi.org/10.1016/j.jsv.2015.08.002

41. Liu J, Shao Y. Dynamic modeling for rigid rotor bearing systems with a localized defect considering additional deformations at the sharp edges. Journal of Sound and Vibration, 2017; 398; 84-102. https://doi.org/10.1016/j.jsv.2017.03.007

42. ISO 13373-1:2002. Condition monitoring and diagnostics of machines -- Vibration condition monitoring - Part 1: General procedures.

43. ISO 20816-1:2016. Mechanical vibration -- Measurement and evaluation of machine vibration -- Part 1: General guidelines.

44. https://www.skfbearingselect.com/#/size-lubrication/single-bearing [Accessed 13.8.20]

45. AGMA 925-A03. Effect of lubrication on gear surface distress 2003.

46. Hammami M, Martins R, Abbesb MS, Haddarb M, Seabra J. Axle gear oils: Tribological characterization under full film lubrication. Tribol Int 2017;106;109–22.

47. Brandt A. Noise and Vibration Analysis: Signal Analysis and Experimental Procedures. Wiley, 2010. 438 p.

48. Dwyer DF. Detection of non-Gaussian signals by frequency domain kurtosis estimation, in: International Conference on Acoustic, Speech, and Signal Processing, Boston 1983, pp. 607–610.

49. Randall RB. Vibration-based condition monitoring. Industrial, aerospace and automotive applications. Wiley, 2011. 289 p.

50. Ho D. Bearing diagnostics and self-adaptive noise cancellation. PhD dissertation (2000), University of New South Wales.

51. Wang W, Wong AK, Autoregressive model-based gear fault diagnosis, Transaction of ASME, Journal of Vibration and Acoustics, 2002;124;172–9.

52. Lee JY, Nandi AK. Extraction of impacting signals using blind deconvolution. Journal of Sound and Vibration, 1999; 232 (5);945–62.

53. Borghesani P, Antoni J. CS2 analysis in presence of non-Gaussian background noise–Effect on traditional estimators and resilience of log-envelope indicators. Mech Syst Sig Process, 2017;90;378–398.