ABSTRACT

Defects in railway axle bearings can affect operational efficiency, or cause in-service failures, damaging the track and train. Healthy bearings produce a certain level of vibration and noise, but a bearing with a defect causes substantial changes in the vibration and noise levels. It is possible to detect bearing defects at an early stage of their development, allowing an operator to repair the damage before it becomes serious. When a vehicle is scheduled for maintenance, or due for overhaul, knowledge of bearing damage and severity is beneficial, resulting in fewer operational problems and optimising the fleet availability.

This paper is a review of the state-of-art in condition monitoring systems for rolling element bearings, especially axlebox bearings. This includes exploring the sensing technologies, summarising the main signal processing methods and condition monitoring techniques i.e. wayside and on-board. Examples of commercially available systems and outputs of current research work are presented. The effectiveness of the current monitoring technologies are assessed and the p-f curve is presented. It is concluded that the research and practical tests in axlebox bearing monitoring are limited compared to generic bearing applications.

1. INTRODUCTION

With the increasing popularity of the data driven railway and the use of measurement technologies to support decision making within the railway domain, there is increased focus on the practicalities and quality of such measurements. Traditionally, safety based decisions within the railway have been made using a combination of high precision inspection against set standards or, where inspection is impractical, periodic action. While safety is still the primary concern, there is now increased consideration being given to business focus in maintenance decisions. In this area, ongoing condition based monitoring is being used to inform decision making processes.

Condition monitoring systems are designed to identify the condition of an asset. They can increase the cost-effectiveness of maintenance regimes by reducing the incidence of urgent and costly unplanned interventions and lead to improvements in performance and safety. By their nature, condition monitoring systems are applied in an ongoing and, ideally, pervasive fashion. This presents a difficulty as measurement systems must either be “one-to-many”, i.e. single measurement systems observing a large area or number of units, or cost-effective enough to be produced and deployed in volume.

Investigation and research in bearing failures in the railway has been an area of interest for more than a century. In 1900 Robert Job from the Philadelphia & Reading Railway Company investigated the causes of hot axleboxes on steam locomotives and methods to prevent them [1].

Hot axleboxes started to become of interest as the speed and weight of trains increased. Job stated that the causes were mainly due to the incorrect proportion of the metals in the bearings material, crystallization of the material caused by high temperatures and oxidation and large amounts of gas in the metal structure [1]. Interest in bearing life prediction was also seen in the early 20th century; a theory for bearing life prediction was founded by Palmgren (SKF) [2]. A report published by the Glen Research Centre, NASA, explains the progression, concept, and application of rolling element bearing life prediction throughout the 20th century [2].

Axle bearings in railways are vital, and the selection of a suitable bearing for each application is key to fault-free operation. However, operating circumstances may vary from the calculated specifications. This can be caused by wear and tear, or changes in the overall vehicle dynamics, which affects the bearing service life.
In general, bearing faults can be categorised into three types [3]:

- **Abrupt**: symptoms appear unexpectedly
- **Intermittent**: symptoms are not continuously present; they can appear for a period of time and then disappear, which makes them challenging to identify
- **Incipient**: a fault which develops gradually over a period of time, with worsening symptoms

In axlebox bearings, abrupt and intermittent faults, depending on their level of damage, can become incipient. An example of this is damage caused by electrification and grease contamination [4]. Abrupt faults can be reduced by the selection of an appropriate bearing, correct installation, and maintenance procedures. Intermittent faults may be found through special inspection techniques such as oil analysis. Incipient faults are when the fault is permanently developed and it progresses and worsens with time. These faults can generally be detected using current monitoring techniques.

Typical railway axlebox bearings consist of a double row of tapered bearings with roller elements, as shown in Figure 1. The bearing has four main components: rollers, a cage, and outer and inner races.

![Figure 1: Double row tapered roller bearings, two-row double-outer race](image)

A report by bearing manufacture SKF showed that causes of bearing damage can generally be divided into four main categories, as shown in Figure 2 [6].

![Figure 2: Causes of damaged/failed bearings](image)

A report, also by the SKF group, on bearings in the railway and investigation of their failures, summarises the design problems in a block diagram [7], shown in Figure 3.
Figure 3: Simplified design problems [7]

Defects in railway axlebox bearings can affect operational efficiency, or cause in-service failures, damaging the track and train. For train operators, life cycle cost is also important. This cost can be categorised as purchase, maintenance, and scrap [8]. These are the main part of the life cycle, although other costs such as track access, charges, accidents and related delays can also affect the life cycle cost.

Poor operating conditions such as shaft misalignment, moist or contaminated areas, excessive load, overheating, lack of lubrication, electrical damage, wheel damage (flats, cavities) and manufacturing defects are the main causes of bearing faults. For example, excessive loads usually cause spalling on the race, this gradually increases and the bearing will eventually fail to operate [9]. Also removing and re-fitting the bearings during the overhaul procedures can cause damages to the asset. This is one of the reasons that fleet managers would prefer to avoid overhauls if they can be sure that the expected life cycle can be met. Bracciali et al have also introduced a new wheelset design that simplifies the maintenance cycle, reducing the possibility of damage being caused to the bearings during overhaul [8].

Healthy bearings produce a small level of vibration and noise, but a bearing with a defect causes substantial changes in the vibration and noise levels. It is possible to detect bearing defects at an early stage of their development, allowing an operator to repair the damage before it becomes serious [10]. When a vehicle is scheduled for maintenance, or due for overhaul, knowledge of bearing damage and severity will be beneficial, resulting in fewer operational problems and optimising the fleet availability.

Bearing wear has been categorised into four stages [11]:

- **Stage 1**: Microscopic level of damage but the bearing is still operational and the risk of catastrophic failure is low. Most condition monitoring techniques cannot detect this level of damage.
- **Stage 2**: The damage is visible and can be detected by some condition monitoring techniques. At this stage, repair at the next inspection is recommended. The risk of a failure is moderate.
- **Stage 3**: The damage is now spread around the rollers and races. Depending on mileage (operational duty) and load, a repair might be needed within a month. There is a significant risk of an in-service failure.
- **Stage 4**: The damage is audible (to a nearby human). Immediate attention is required.

This paper reviews the sensing technologies, signal processing methods and condition monitoring systems currently in use, or the subject of current research, on axlebox bearings.

2. CONDITION MONITORING SYSTEMS

Condition monitoring systems are data collection systems with intelligent algorithms that can detect faults prior to failures, identify faults, and assess the severity of a fault. A condition monitoring system that predicts or detects
critical faults can either be used to improve safety or to make the existing safety level more affordable. Another benefit is that the availability and operational reliability of the fleet can be enhanced by predictive monitoring [12].

Condition monitoring and maintenance of wheelsets, especially bearings, for in-service trains is an on-going priority for operators and infrastructure managers. Monitoring systems can be divided into two categories: on-board and wayside (lineside). The potential-to-functional failure curve (p-f curve) can be used to represent the performance of condition monitoring technologies. Functional failure is the point at which the bearing does not perform, such as a seized bearing, and will require replacement. The development of a fault over time can have various symptoms, which can be detected with different monitoring technologies [13, 14, 15].

Figure 4 is an example of a p-f curve that can illustrate the stages in bearing degradation. The general overview of a p-f curve is to quantify and allow an appropriate reaction to be defined over time [16]. The p-f curve demonstrates the time between the potential failure and the event of failure. Condition-based maintenance and condition monitoring allow strategic scheduling of corrective maintenance, and hence cost savings [17].

Some condition techniques are available for use on moving trains. They can assess an asset on a daily basis and provide information that can be used to guide maintenance routines and failure prevention procedures. Some of the techniques can be used to predict faults and in some cases can be used for prognostic purposes, which require high quality signatures being extracted from the measurements of a fault. To achieve these levels of extractions, complex algorithms and high quality data with an appropriate signal to noise ratio (SNR) are required [18, 19].

2.1 On-board
In this approach, sensors are physically coupled to each bearing. As the sensors are close to the bearing, the signal-to-noise ratio is high, which provides an efficient way to detect defects at an early stage. This approach is costly as it is necessary to install sensors and processing units on every axlebox. The resulting cost could be higher than the existing overhaul and maintenance regime.

Temperature and vibration analysis are well-known techniques for condition monitoring of rolling elements both for in-service operation and during maintenance. These techniques are generally implemented by fitting each bearing directly with condition monitoring equipment. In order to do this, a number of technical and commercial challenges must be overcome. Commercial systems, making use of energy harvesting and wireless technologies, can provide the technical capability, but for most vehicle operators there is still some concern over the cost of fitting such systems to every bearing in a fleet. Technology penetration studies have been carried out in order to establish the level of instrumentation required to provide statistically significant coverage, but ultimately a system that measures every bearing is preferable.

2.2 Wayside
Wayside systems are installed at the side of the track either in the four foot, or at a distance from the track. For operators with a sufficiently large number of assets, a wayside measurement system capable of capturing data from every passing vehicle could be more cost effective than fitting equipment to each bearing. The most common system for checking the state of axle bearings from the wayside is the hot axlebox detector. However, these are only able to detect a faulty bearing when the fault is at an advanced stage. A wayside condition monitoring
technique that can be used to detect early stage (minor) bearing faults is highly desirable, for both train operating companies and infrastructure managers.

3. SIGNAL PROCESSING METHODS

Bearing damage causes interaction between the damaged area and components such as the races and rollers within the rolling element that generates additional noise and vibration.

A number of signal processing methods can be applied to the measurements to detect and/or identify the type and severity of axle bearing defects. Signals can be analysed in time, frequency or time-frequency domains.

The geometry of the bearing can lead to certain frequencies being important, dependent on the type of fault. The following equations define the characteristic frequencies of a bearing:

FTF, Fundamental Train Frequency, fault on the cage or mechanical looseness

\[ f = \frac{f_r}{2} \left( 1 - \frac{RD}{PD} \cos \beta \right) \]  

(1)

RPFO, Roller Passing Frequency Outer Race, local fault on outer race

\[ f = n \frac{f_r}{2} \left( 1 - \frac{RD}{PD} \cos \beta \right) \]  

(2)

RPFI, Roller Passing Frequency Inner Race, local fault on inner race

\[ f = n \frac{f_r}{2} \left( 1 + \frac{RD}{PD} \cos \beta \right) \]  

(3)

RFF, Roller Fault Frequency = 2 * RSF, Roller Spin Frequency, local fault on rolling element

\[ f = \frac{PD}{RD} \frac{f_r}{2} \left( 1 - \left( \frac{RD}{PD} \cos \beta \right)^2 \right) \]  

(4)

where \( f \) is the frequency of the item [Hz]; \( f_r \) is the shaft rotation rate [Hz]; \( RD \) is the roller diameter [m]; \( PD \) is the mean roller race diameter [m]; \( \beta \) is the contact angle [rad]; and \( n \) the number of rollers [20, 21].

These equations also show the importance of knowing the rotational speed and the geometry of the bearing for fault diagnosis.

Commonly used statistical analysis techniques for monitoring rotating machines are based on the assessment of peak-to-peak values, RMS, kurtosis, crest factor and power spectrum [22, 23]. However, there are other analysis techniques such as using wavelet transform and B-spline that have been assessed, which are currently less frequently seen in industry [24, 25].

3.1 RMS

The root mean square of a signal is the simplest way to assist in identifying faults within the bearing elements. This method can be used to indicate abnormality around the area that the monitoring system is focused on. However, it has been shown that a rise in RMS value can indicate certain types of damage, but this does not necessarily mean that it can indicate all bearing faults. A combination of band-pass filtering followed by RMS computation in the frequency bands that tend to have high kurtosis values can be more effective [26].

3.2 Crest Factor

Crest factor (CF), equation (5), is the maximum peak divided by the RMS value of the signal, which provides a value for the impact caused by the strikes in a damaged bearing [11].

\[ CF = \frac{\text{max peak value}}{\text{RMS}} \]  

(5)

Depending on the severity of the fault, the CF usually increases and the fault could appear in the amplitude spectrum along with a few related harmonics.
3.3 Skewness and Kurtosis

The density and shape of a signal distribution can be visualised by moments about the means [28]. To calculate these the mean of a set of data, which is the estimated of the mean of the whole data, is first computed using:

\[ \bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \]  

(6)

where the \( x_i \) are the recorded data samples and \( N \) is the number of samples. The \( n \)th central moment (about the mean) is defined as

\[ \mu_n = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^n \]  

(7)

where \( \bar{x} \) is the true mean. When there is a large number of samples then \( \bar{x} \) can be used instead of the \( \bar{x} \) for engineering purposes. The ‘central’ part of the ‘central moment’ means that the moment is computed about the mean, or centre of the signal. The second central moment is the ‘variance’. The variance is assumed to be a property of some underlying distribution [29] and the equation above gives an estimate of the variance. The square root of variance is the standard deviation, \( \sigma \), the second defining parameter of a normal distribution after the mean.

Skewness is a measure of the degree of asymmetry of a distribution, derived from the third central moment by [30]

\[ S = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{\mu_2^{3/2}} \]  

(8)

Data distributed symmetrically about the mean has zero skewness. As the mean has been eliminated, skewness does not show that there are more positive than negative data, but it means that the positive going peaks are in some sense more significant than the negative going peaks, or vice versa. In principle, the skewness indicates the stability of the collected data and can be used to detect sensor faults.

Kurtosis is a non-dimensional quantity that measures the relative peakiness of a signal compared to a Gaussian distribution. It is particularly high for signals containing impulsive features. It is defined as the fourth central moment (\( \mu_4 \)) divided by the square of the variance (\( \sigma^4 \) or \( \mu_2^2 \)) [31]:

\[ K = \frac{\mu_4}{\mu_2^2} \]  

(9)

Kurtosis is independent of the mean and the variance. For normally distributed data (e.g. a healthy component), the kurtosis is 3. For a sinusoidal signal, the kurtosis is around 1.5 and for signals with repetitive impulsive type responses, the kurtosis can become very high. Therefore, high values of kurtosis indicate that the data contain spikes.

In [32], laboratory tests showed that the use of kurtosis on unprocessed vibration signals is a practical way to find abnormal behaviour in ideal conditions. But the kurtosis values of raw acoustic signals do not behave in a similar way to vibration signals and show small changes for faulty conditions. On a simple test rig where the sensor is directly coupled to the bearing, kurtosis analysis can be effective. Within acoustic signals, the noise from the drive train or other components in the system affects the sound coming from the bearing therefore kurtosis analysis responds to other signals not related to the bearing. One way to improve the use of kurtosis is to compute the kurtosis for various frequency ranges rather than for the original signal.

3.4 Spectral Kurtosis

The Spectral Kurtosis (SK) [33] of a signal is obtained by splitting the original signal into frequency bands and finding the kurtosis of each frequency band. This shows how the peakiness of a signal varies with frequency, and can be used to indicate the frequency ranges within which faults are present. SK results are somewhat sensitive to the frequency bands chosen, the choice being application-dependent. An overview of the SK procedure is shown in Figure 5.
Figure 5: Overview of the SK method

Figure 6 illustrates the application of spectral kurtosis to identify a frequency band within which a signal has the most impulsiveness behaviour. Further analysis on the selected band can be used for fault diagnosis.

Figure 6: Spectral kurtosis of a simulated healthy and faulty bearing

Examples of the use of spectral kurtosis in bearing fault detection in railways can be found in [34, 35, 36].

3.5 High Frequency Resonance Technique

The high frequency resonance technique (HFRT), also known as Envelope Analysis, extracts fault-related high frequency features of a signal, which are compared to known bearing fault characteristic frequencies [37]. The procedure for this nonlinear signal processing method is as follows:

1. High-pass or band-pass filtering of the original signal to emphasise the frequency contents associated with the impulsive signals generated by a fault. The filter is centred around the resonant frequency of the bearing structure.

2. Envelope analysis using a Hilbert Transform [38] or another high frequency envelope detection and modulation technique to identify individual impact events within that signal. Further peak detection and interpolation algorithms can be applied for a smoother envelope.

3. Conversion of the signal envelope obtained in the previous step into the frequency domain to identify the characteristic frequencies associated with a fault.

The resolution of the frequency domain of the final envelope depends on the length of the data. To achieve a sufficient level of accuracy at high rotational speeds, a short period of data collection can be satisfactory, however, at low speeds a long length of data is required.

Envelope analysis can be applied to signals within frequency bands showing high kurtosis in the SK, forming a hybrid technique between HFRT and SK which has been shown to be effective for detecting and diagnosing bearing faults [34, 36, 21]. Crest factor can also be used to make a decision using the HFRT spectrum; the largest
peaks can be identified based on the overall spectrum RMS and thus identifying the featured frequencies in the spectrum [27]. Figure 7 is an overview of the HFRT using the SK to select the band-pass filter.

![HFRT and SK procedure](image)

**Figure 7: HFRT and SK procedure**

### 3.6 Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is a self-adaptive signal processing technique, also known as a data driven technique, originally proposed by Huang et al. [39]. Promising results of the application of EMD to various signals obtained from mechanical systems have been published in [40, 41, 42]. EMD focuses on extracting the stationary points of a signal [43]. In EMD, the signal is decomposed into sub-signals, called intrinsic mode functions (IMF). These sub-signals are used for later signal processing, these results are similar to frequency or wavelet analysis but the main difference is that the procedure is based in the time domain.

To illustrate EMD on a signal, the process was applied to white noise and the frequency spectrum of the results generated. The spectrum of the first five IMFs along with the spectrum of the original signal are shown in Figure 8.

![IMFs frequency ranges](image)

**Figure 8: IMFs frequency ranges**

In bearing fault analysis EMD can be used instead of the filtering in the HFRT process, avoiding artefacts caused by filtering. This method showed successful results in vibration signal analysis [23]. A weakness of this method
is the processing time, which is considerably higher than using simple band-pass filters, and choosing the correct IMF is complicated.

3.7 Moving train tracking techniques

In wayside monitoring systems, it is necessary to locate the position of the axlebox bearings. For example, in wayside acoustic (sound) systems, sound emitted from an axlebox of a moving train must be tracked by the microphones, and to provide this, wheel detecting systems are required. To run a diagnostic algorithm, it is also necessary to know the speed of the shaft (train speed). This can be determined by using a number of wheel detectors along the track with known distance.

The other issue with tracking sound from a moving train is the Doppler effect. Knowing the train speed helps to eliminate the Doppler effect. Also, due to the environmental noise and the distance to the track, the SNR is low. There are research articles that have addressed these issues and which have presented promising outcomes [44, 45, 46].

To track the location of the axlebox bearings and calculate the train speed, Infra-Red light gates or high speed cameras together with specific vision processing algorithms have been used [47, 48].

4. Sensing Technologies

Various sensing techniques can be used to monitor the condition of axlebox bearings [49, 50]:

1) Acoustic emissions / ultrasound
2) Vibration / acceleration
3) Microphone / sound measurement
4) Thermal

The first three technologies are used to detect early stage defects. When bearing defects become severe, they generate heat that can be detected by thermal sensors. Lubricant analysis is another method for detecting bearing defects, although this is only used during routine inspection rather than as part of a condition monitoring system.

4.1 Acoustic emission

Acoustic emission (AE) is usually defined as the elastic waves generated by a sudden redistribution of molecules inside or on the surface of a material, for example during crack development [51]. AE can be used as a technique for structural health monitoring. External stimuli such as temperature or load on the material can cause a built up of strain that, if released rapidly for example as a crack propagates, produces energy in the form of stress waves that can be recorded by a piezoelectric AE sensor. There are other applications of AE such as monitoring bridges, railways and aerospace structures manufactured from different materials including alloys, plastics and composites [52]. Recently, AE has been used to monitor rotating and reciprocating machinery, in most cases as a supplementary technique to vibration analysis.

AE was originally developed and used for non-destructive testing of static structures, however, its uses have been extended to the monitoring of rotating machines. AE techniques, compared to traditional alternatives can achieve earlier fault detections due to the high operating frequency characteristics, which is highly desirable for condition monitoring systems. AE sensors have demonstrated better performance in fault detection in bearings compared to other techniques [49, 22].

The upper frequency range of AE sensors is typically between 20 kHz and 1 MHz, but in most applications the upper frequency limit is 100-500 kHz [52]. Excitation signals caused by defects in bearings are a broadband phenomenon that produces stress waves in a wide frequency range, particularly above 100 kHz. Thus, AE signals often have a higher SNR compared to more common vibration measurements. This allows the presence of a bearing defect to be detected noticeably earlier by an AE system than other techniques [53].

There are some disadvantages associated with AE as a diagnostic tool. The high frequency AE signals are subject to significant attenuation with distance so it is necessary to install the AE sensors as close as possible to the asset [49]. Another drawback of AE sensing systems is the need for a significantly high sampling rate, which makes collecting, managing and processing the data more challenging and time consuming.

The formation of subsurface cracks due to the stress induced by the rolling action of bearing elements in contact with the inner or outer race generates AE activity. Surface-breaking defects impact on the opposing surface of the
bearing (e.g. races or rollers), generating high energy AE signals. Measurement of these signals can lead to early fault detection [54, 55].

Applications of using the AE technique for monitoring the lubricant condition in rolling element bearings have been demonstrated in [56, 57, 58]. A lack of lubricant and induced contamination conditions were tested. They both showed a significant change in the mean of the measured signal. The RMS value in AE signals has also been used for lubricant monitoring [59].

Tandon et al [22] published a review of bearing condition monitoring using vibration and acoustic emission and it was claimed that AE sensors are able to find damage earlier than vibration sensors. Price et al and Al-Ghamd et al [60, 61] also carried out tests to compare the techniques and reached the same conclusions.

The use of AE for axle bearing monitoring was also carried out by Papaelias et al and Amini et al [10, 62, 34, 49, 63]. They have demonstrated the use of acoustic emission on test trains and in-service trains. The rise in RMS values and kurtosis analysis have clearly demonstrated the existence of defects. This system is also able to identify the condition of the wheel without any additional hardware [49]. Acoustic emission sensing technology was also demonstrated as part of a European Commission funded project, MAXBE (Interoperable Monitoring, Diagnosis and Maintenance Strategies for Axle Bearings). This work has been presented in [64, 10].

There are commercial products available to check the bearing condition using acoustic emission but they have not been used on in-service trains for continuous monitoring. However, these products can be a useful tool during overhaul and maintenance operations. An example of these commercial products uses acoustic emission sensing technology and reports a distress level value which can indicate bearing degradation and inadequate lubrication [65].

4.2 Vibration

Vibration is the oscillation or repetitive motion of an entity around an equilibrium position [66]. Vibration signals in bearings can provide information on the cause of the vibration. A bearing is generally a source of noise and vibration, and an increase in the level of the vibration is caused by a change in the system dynamic. This leads to detection of abnormal behaviours and therefore faulty condition, which tend to significantly increase the vibration level [22].

Vibration analysis has been a condition monitoring tool for bearing fault detection and diagnosis since the early twentieth century [67]. The earliest review on the vibration analysis of bearings was carried out by Houser et al [68]. Houser et al provided an inclusive research on bearings, their rotational frequencies, failure modes, resonance frequencies and several vibration analysis techniques.

Vibration levels change with the speed, mass, and alignment of the bearing. In order to accurately monitor the bearings it is necessary to identify parameters such as the speed of the rotation and the geometry using the vibration signals [69].

There are commercial systems available for bearing monitoring. Some can be used during inspection and maintenance, and some are on-board for in-service monitoring. The Fluke 810 Vibration Tester is an example of a tool for mechanical analysis of a system. In this system, connections of the components of a mechanical system are given to the tester and the vibration sensor can be installed in different components, where it analyses each component using a tri-axial accelerometer [70]. This system provides scores for the severity of the damage, and based on these scores can provide some diagnostic information. However, this is not designed to be used on in-service axle bearings.

Common analytical methods used to analyse vibration data are RMS, crest factor, kurtosis, high resonance frequency technique and power spectral density analysis. A combination of time- and frequency- domain, and some statistical tools have also been applied [22, 49].

In 1998, Sneed et al presented their research on monitoring the bearing using an on-board vibration sensor. They concluded that this technique could be the basis of a new generation of fault detection and failure prevention for the future [71].

The U.S. Federal Railroad Administration funded a programme called the Safety IDEA project [72]. Section 16 of this project was to investigate the use of accelerometers on the rail to assess the transmissibility of bearing signals through wheel-rail contact. The results were compared with on-board accelerometers; the on-board
accelerometer system could clearly detect the bearing fault but the wayside system could not. It was concluded that the amplitude of the transmitted signals was reduced by the wheel/rail path. This may be due to the mechanical components and their damping effects. It was also established that background noise interferes with the weakened bearing signals, meaning the ability to detect faults is further reduced [73].

Corni et al have presented the use of an energy harvesting system (EHS) that measures the axlebox bearing vibrations to detect bearing defects [74]. This work showed that the vibration in axlebox bearings are complex and that it is difficult to distinguish a faulty bearing. However, analysis of the data collected using this EHS has shown good results in detecting abnormality within the axlebox, wheel and track. Such a product is currently available and started to be used in rail applications [75]. A similar approach is also under development by the SKF group and the preliminary results of this work have been presented [76].

4.3 Shock pulse method
The shock pulse method (SPM) was introduced by E. Schoel in 1966 [69]. This method relies on the mechanical shock waves generated by contact between the roller or race and the damaged area or debris in the bearing. The SPM requires parameters such as the shaft speed and the bearing geometry. In this method, a quantitative result is provided that can indicate the state of the bearing and help with scheduling maintenance. The waves generated by the impact cause oscillation at the transducer and the peak amplitude of the received wave is used to determine the bearing condition [69, 77].

In SPM the transducers are usually focused around 32 kHz, which eliminates low frequency components and therefore cannot identify problems such as an unbalanced shaft or misalignment [78]. However, this method can be applied at rotational speeds as low as 10 RPM. This method is also able to estimate the thickness of the lubricant [79]. SPM can detect contamination but it does not perform as well as AE [80].

Shock pulse method systems are simple to use as they do not require highly skilled operators [22]. Even though companies such as SKF have investigated and worked on this method [69] the authors could not find any evidence of the use of the shock pulse method for in-service train bearing monitoring. This could be because of the sensitivity of the SPM to background noise and other sources of shocks in a complex system (e.g. trains) [81].

4.4 Thermal
High temperature makes the lubricant (usually grease) in the bearing less viscous so it can escape from the housing, followed by bearing failure. A derailment could be the result of an unidentified overheating axle bearing. Therefore, the temperature of the axlebox bearings and the surroundings (metal coupled) such as wheels and brakes are monitored on railways to ensure safety [82].

The monitoring of bearing temperature can be carried out on-board or wayside. On-board systems are available that continuously monitor the bearing temperature [83]. These systems can be retrofitted or integrated into new products. Wayside systems use contactless temperature measurement devices such as infra-red (IR) sensors to measure the heat emitted by the asset [84]. There are single, dual and multi beam systems which can provide different scanning points based on the application [82].

In railways, hot axlebox detectors (HABD) and hot wheel detectors (HWD) are used to monitor the temperature of the axleboxes and brakes. This is a safety requirement for the UK’s railway network [85]. Figure 9 shows a HABD installed on a high speed line in the UK. The measurement system is fast enough to monitor temperature while the target object is moving past at up to 500 kmh\(^{-1}\) [82]. The HABD temperature range is usually around 0-150 °C and HWD can go up to 600 °C. According to the European Standard EN 15437-1 the alarm in HABD must be generated when the temperature is over 95 °C or the differential temperature is above 56 °C [85].
Early stage damage in bearings does not generate significant amounts of heat. Williams et al did not find any correlation between the rise in temperature and a damaged bearing [26]. It has also been reported that bearings with half a race damaged, as shown in Figure 10, on an in-service UK high-speed train did not trigger any hot axle box detector.

4.5 Sound measurements (Air-borne acoustic)

When rolling elements strike a local fault on the inner or outer race, or a fault on a rolling element strikes the inner or outer race, an impact is generated. The impact causes vibration and emits sound. Depending on the size and characteristics of the damage, sound and vibration levels vary [22]. SKF has published a list of faults where the bearing is noisy [86]. The use of microphones to detect bearing faults has been of interest as this technique does not require a physically coupled sensor and in the railways it can be carried out wayside [36, 87]. The earliest article on wayside systems was published in the 1990’s. Cline et al [88] demonstrated the use of acoustic data to detect bearing faults on freight trains. In the same decade Tandon et al [22] compared a specific frequency band of the sound pressure signals to other sensing technologies on bearings with microscopic defects.

As this technique listens to sound carried through the air, it is susceptible to environmental noise. The early stage faults typically generate little volume, so the SNR is low and it can be difficult to detect the damage. To address this issue, approaches such as the use of directional microphones [89] or microphone arrays and beamforming algorithms have been used [36, 87, 90].

There are only a few wayside acoustic monitoring systems commercially available such as Railsam and TADS [91, 92]. The authors could not find any documents or publications on the performance of the TADS system. However, the early performance of Railsam was published in [87] and showed good potential for a wayside system. Entezami et al have also carried out research on monitoring axlebox bearings, especially on high-speed rail, and the preliminary results have confirmed that the technique has potential [47, 36]. These wayside systems also use a wheel detection system to identify the speed of the passing trains, to count wheelsets, and to compensate for the Doppler effect.
The analysis techniques used on acoustic signals are essentially the same as those used for analysing vibration signals [32].

4.6 Sensing technologies comparison

Williams et al carried out experiments that involved running a new bearing through to failure [26]. This work demonstrated the use of AE and vibration sensors and their effectiveness in early-stage bearing fault detection. It was also shown that the temperature does not rise until the damage becomes significant.

Tandon et al compared vibration with AE sensing technologies on a small bearing defect (depth of 150 µm and a 500 µm in diameter). It was determined that the acoustic emission can detect the defects at lower speeds than vibration analysis (down to 100 RPM) and it was also concluded that the AE performs better overall [22, 93, 94]. The work also tested the shock pulse method and recorded the sound pressure. The results of the experiments show that the AE performs the best while sound pressure is the weakest technique.

Table 1: Results from the bearing tests [22]

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Measurement parameter</th>
<th>No Defect</th>
<th>Inner race</th>
<th>Outer race</th>
<th>Roller</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vibration (ms(^2))</td>
<td>Overall acceleration</td>
<td>0.67</td>
<td>1.79</td>
<td>12.8</td>
<td>1.07</td>
</tr>
<tr>
<td>Sound pressure (dB)</td>
<td>Overall pressure</td>
<td>65.3</td>
<td>66.6</td>
<td>83.8</td>
<td>65.8</td>
</tr>
<tr>
<td>Shock pulse (dBm)</td>
<td>Max value</td>
<td>14.5</td>
<td>22</td>
<td>38</td>
<td>21</td>
</tr>
<tr>
<td>AE (dB)</td>
<td>Peak amplitude</td>
<td>60.7</td>
<td>70</td>
<td>78</td>
<td>71</td>
</tr>
<tr>
<td>AE</td>
<td>Number of Peaks over threshold</td>
<td>5450</td>
<td>175000</td>
<td>40000</td>
<td>2800</td>
</tr>
</tbody>
</table>

A comparison between acoustic (air-borne) and vibration sensors was also carried out by Entezami et al [32] and it was concluded that vibration sensors can detect faults sooner than acoustic sensors (free-field microphones).

The application of integrated high-frequency acoustic emission with vibration analysis has been extensively studied for both on-board and wayside measurements [10, 34, 49, 62, 95, 96, 97, 98]. This research in the UK was issued a certificate of acceptance, which allows trackside instrumentation [97]. Trains are monitored for both wheel and axle bearing defects [97, 98]. The detection of a suspected axlebox bearing defect has been reported in [98]. However, this is pending its final validation by the infrastructure manager.

Figure 11 a p-f curve concluded from the reviewed works that can potentially illustrate the relative fault detection positions for current state-of-the-art axle bearing condition monitoring technologies.

Figure 11: P-F curve in bearing and sensing technologies

A review paper on prognostics for rolling element bearings by Jammu et al [99] compared vibration, oil analysis, temperature, and acoustic emission technologies. A summary of the advantages and disadvantages of these techniques is presented in Table 2.
Table 2: Pros and cons of some bearing condition monitoring technologies [99]

<table>
<thead>
<tr>
<th>Sensing technologies</th>
<th>Vibration</th>
<th>Thermal</th>
<th>Acoustic Emission</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advantages</strong></td>
<td>Reliable and well-established</td>
<td>Can detect abnormality in an interest area</td>
<td>Performs fault detection and diagnostics well</td>
</tr>
<tr>
<td></td>
<td>Sensitive to small changes</td>
<td>Used in safety critical conditions</td>
<td>Sensitive to microscopic changes</td>
</tr>
<tr>
<td></td>
<td>Ability to perform fault diagnostics</td>
<td></td>
<td>High SNR</td>
</tr>
<tr>
<td><strong>Disadvantages</strong></td>
<td>Expensive</td>
<td>Too late for predictive maintenance when rise in temperature occurs</td>
<td>Very expensive</td>
</tr>
<tr>
<td></td>
<td>Requires physical contact</td>
<td>Does not perform fault diagnostics</td>
<td>Requires high Sampling frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Requires physical contact</td>
</tr>
</tbody>
</table>

The comparison does not include air-borne acoustic methods.

5. CONCLUSIONS

A large amount of research has been carried out into bearing condition monitoring and a variety of signal processing methods have been developed to detect and identify faults. This paper started with an exploration of condition monitoring techniques and their effectiveness. The paper aimed to review monitoring systems for axlebox bearings and it has come to the authors’ attention that the number of research publications and industrial work on axlebox bearing monitoring systems are limited compared to the generic bearing monitoring applications. However, it is shown that the techniques can be, to some extent, adapted from the extensive works in other industrial areas. The application of these methods and technologies to railway trains is limited.

In this paper, condition monitoring systems for axlebox bearings were divided into two main categories: on-board and wayside. The bearing faults were categorised into four stages, which can be used in decision making for maintenance routines. The stages were also presented in a p-f curve, which demonstrates the severity of each stage and how it fits with the predictive and preventive maintenance schedule.

A number of signal processing techniques were presented and compared. It was claimed that AE can detect defects, including lubrication contamination, earlier than the other sensing technologies. However, its application in the railway industry is in its infancy and requires extensive research and assessments due to the nature of AE and other mechanical noise emitted by the vehicle. It was also explained that hot axlebox detectors are used as a safety component in the railways, rather than being used for predictive maintenance.

To demonstrate the capability of the current technologies in prediction and prevention maintenance, a p-f curve of axlebox bearing condition monitoring was presented.

Currently, only a limited number of on-board or wayside monitoring techniques are commercially available within the railway industry. Apart from the hot axlebox detectors, which identify faulty bearings too late to be useful for maintenance scheduling, there are vibration and acoustic (sound) techniques that can be used to detect early stage defects. However, these systems are still developing and their processing techniques and performance still require improvements to achieve more sophisticated bearing fault detection and diagnostics systems.

References


