

Techniques for Large, Slow Bearing Fault Detection

Eric Bechhoefer¹, Rune Schlanbusch² and Tor Inge Waag³

¹*GPMS Inc, Cornwall, VT, USA
eric@gpms-vt.com*

²*Teknova AS, Kristiansand, Norway
rune.schlanbusch.2008@ieee.org*

³*MHWirth AS, Kristiansand, Norway
tor.waag@mhirth.com*

ABSTRACT

Large, slow turning bearings remain difficult to analyze for diagnostics and prognostics. This poses a critical problem for high value assets, such as drilling equipment top drives, mining equipment, wind turbine main rotors, and helicopter swash plates. An undetected bearing fault can disrupt service, and cause delays, lost productivity, or accidents. This paper examines a strategy for analysis of large slow bearings to improve the fault detection of condition monitoring systems. This helps reduce operations and maintenance cost associated with these bearing faults. This analysis is primarily concerned with vibration, and is compared to temperature and grease analysis. Data was available from three wind turbines, where one of the turbine was suspected of having a faulted main bearing.

1. INTRODUCTION

This paper demonstrates a number of techniques in the diagnostics of large, slow turning bearings. Specifically, the use of: vibration (from 3 wind turbines), with temperature and grease analysis those three wind turbine and six other machines for comparison. The data is from the wind turbine main bearing. This analysis was initiated because one machine is suspected of having a main bearing fault. That initial fault assessment was based on temperature and grease analysis. The primary focus of the analysis is vibration-based diagnostics.

1.1. Bearing Fault Phenomenology

Large, slow rotating bearings are based on a rolling element bearing design. They are common in all large rotating machines and one of the most frequent reasons for machine failures. In general, because rolling element bearings are so fundamental to rotating machines, their vibration signals have been widely studied.

When a rolling element strikes a local fault (or the rolling element fault hits the inner or outer race), an impact is produced on the structure. The resonance is a result of the natural frequency of the bearing/support structure. The modulation rate is associated with the bearing pass frequency, such as:

- Cage Pass Frequency (CPF):
$$f_r \frac{1}{2} (1 - BD/PD \cos(\beta)) \quad \text{Eq 1.}$$

- Ball Fault Frequency (BFO):
$$f_r \frac{PD}{BD} \left[1 - (BD/PD \cos(\beta))^2 \right] \quad \text{Eq 2.}$$

- Ball Pass Frequency Inner Race (BPFI):
$$f_r \frac{n}{2} (1 + BD/PD \cos(\beta)) \quad \text{Eq 3.}$$

- Ball Pass Frequency Outer Race (BPFO):
$$f_r \frac{n}{2} (1 - BD/PD \cos(\beta)) \quad \text{Eq 4.}$$

Where:

n = number of balls or rolling elements

f_r = relative rate between inner and outer races (e.g. usually the shaft rate for a fixed bearing)

PD = pitch diameter: (inner + outer race) / 2

BD = ball diameter, and

β = contact angle.

Consider the example inner race fault used in Bechhoefer (2013). The bearing rolling element diameter is 0.235 inches, pitch diameter of 1.245, with eight rolling elements and zero contact angle. The BPFI is 4.7550×25 (with a shaft rate of 25 Hz), or 118.2 Hz. One revolution of the shaft takes 0.04 seconds, in which the inner race impact will be seen approximately four times, or in two revolutions (0.08 seconds), nine times (Figure 1). The impacts are seen to be 1/BPFI in time, or spaced 0.0084 seconds. Note that the magnitude of the impacts varies with 1/Rev, as the inner race fault passes under the load zone of the bearing. In the frequency domain, the 1/Rev modulation results in side

Eric Bechhoefer et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

bands around the fault frequency. This is an indicator of an inner race fault.

Zooming into one impact (time 0.0135 to 0.017), the resonance of the bearing structure is seen. Usually, there are a number of resonant modes. The period is approximately 0.0027 seconds, or a first resonant mode of 380 Hz.

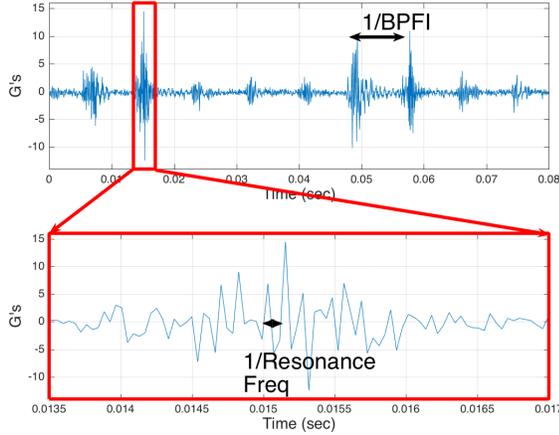


Figure 1 Inner Race Fault, Time Domain

Because this is an inner race fault, the impacts are modulated in time over the shaft rate. This results in 1 per revolution side bands around the ball pass frequency inner race rate (Eq 3). For a ball pass frequency outer race (BPFO) fault, the amplitude of the impacts would not be modulated by shaft rate (Figure 2).

2. LARGE SLOW BEARING ANALYSIS USING VIBRATION

Large slow bearings present an analysis challenge due to four considerations: resolution, small acceleration feature, inexperience (e.g. lack of test data) and nonstationary signals. These issues in one way relate to the nature of the operating environment. That said, these techniques used here are derived from those techniques developed for bearing analysis in general.

2.1. Cyclostationarity Signal Analysis

This class of signals is characterized by signals whose statistical properties change periodically with time (Antoni, 2007). While not periodic and random in nature, these signals are the result of a periodic phenomenon. In the time domain, these signals exhibit periodic variation of statistical descriptors, such as the instantaneous power, or auto-correlation.

By defining a cyclostationary signal (CS) $X[n]$ with n as the temporal index, the signal has a joint probability density function which is quasi periodic function of time. This implies that the ensemble statistics are stationary and

ergodic. For the purposes of bearing analysis, one can model the CS as periodically modulated white noise:

$$X[n] = p[n] \cdot W[n] \quad \text{Eq 5.}$$

Where $p[n] = p[n + N]$ is a N periodic function and $W[n]$ is strictly Gaussian noise. This model accounts for random process including slippage of the bearing elements (e.g. no Hertzian contact of the rolling element with the inner/outer race, see Boskoski 2013).

The signal $X[n]$ from an outer race fault (Figure 2) exhibits cyclic amplitude modulation of the rolling element over the outer race spall. This examples shows that the CS does not in general have a periodic waveform, but a quasi-random behavior that is caused by a periodic mechanism. The period, N , e.g. the cycle of the signal, is the ball pass frequency outer race fault frequency.

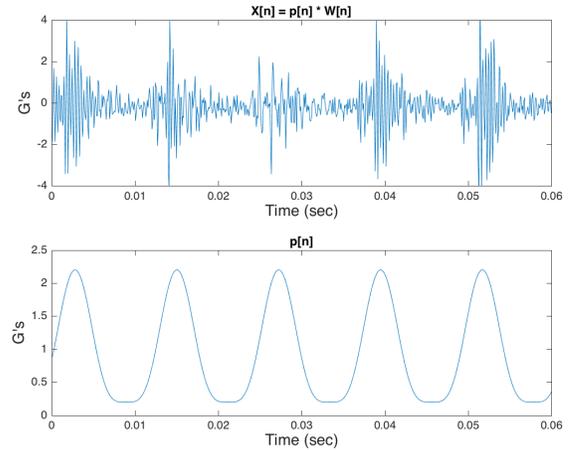


Figure 2 Outer Race Fault signal $X[n]$ and its instantaneous auto-correlation function.

2.1.1. Temporal and Frequency Description of CS

The second order statistics, which of the instantaneous auto correlation function (as per Antoni, 2013) is:

$$\mathcal{R}_{2X}[n, \tau] = E\{X[n + \beta\tau]X[n - \bar{\beta}\tau]^*\} \quad \text{Eq 6.}$$

where $\beta + \bar{\beta} = 1$, with $\beta = 1/2$ for the symmetric instantaneous auto-correlation function. The random signal from the bearing impact will then have a periodic instantaneous auto-correlation function of:

$$\mathcal{R}_{2X}[n, \tau] = \mathcal{R}_{2X}[n + N, \tau] \quad \text{Eq 7.}$$

This defines the signal as a second order CS.

By definition, the instantaneous auto-correlation function of CS has a Fourier representation

$$\mathcal{R}_{2X}[n, \tau] = \sum_{\alpha_i \in \mathcal{A}} \mathcal{R}_{2X}[\tau; \alpha_i] e^{j2\pi\alpha_i n \Delta} \quad \text{Eq 8.}$$

over the spectrum $\mathcal{A} = \{\alpha_i\}$ of cyclic frequencies α_i , where Δ denotes the sampling period.

The instantaneous autocorrelation function $\mathcal{R}_{2X}[n, \tau]$ is a function of the time variables n and τ . For analysis, the frequency domain representation allows the identification of α , the cyclic frequency (e.g. the bearing material response) and f , the bearing component frequency. The frequency domain is a 2D Fourier transform of the two frequency variables α and f :

$$\mathfrak{R}_{2X}(\alpha, f) = \Delta^2 \sum_{n=-\infty}^{\infty} \sum_{\tau=-\infty}^{\infty} \mathcal{R}_{2X}[\tau; \alpha_i] e^{-j2\pi\alpha n\Delta} e^{-j2\pi f\tau\Delta} \quad \text{Eq. 9.}$$

This is the spectral correlation of the power distribution of the signal with respects to the spectral frequency f (the bearing fault rate) and the cyclic frequency α (the cyclic evolution of the waveform as a result of the bearing material response to f). For a more detailed analysis, see Antoni 2009. Figure 3 is the spectral correlation of the main bearing for machine 2, which is suspected of a bearing fault. The spectral frequency has number harmonics of 4.3 Hz, which happens to be the BPFO frequency. The cyclic frequency indicates change in phase response of the bearing support material 0 and about 25 Hz.

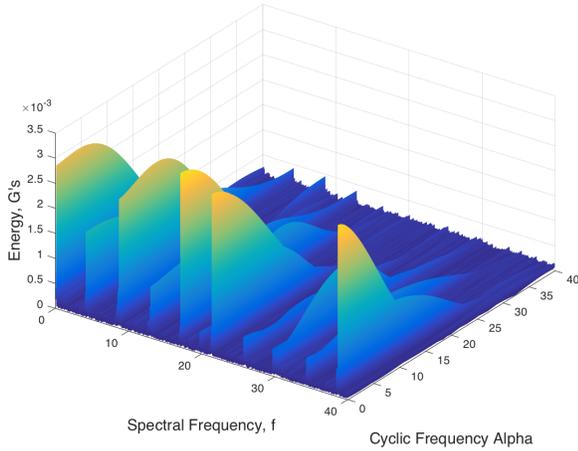


Figure 3 Spectral Correlation of Main Bearing, Machine 2

2.2. The Relationship between Cyclostationarity and Envelope Analysis

The spectral correlation is the 2D Fourier transform of spectral frequency f (the bearing fault rate) and the cyclic frequency α . Fixing α to a given frequency, the spectrum is the envelope analysis of the signal. Defining the envelope for a signal instance of α greatly reduces the computational burden and allows for more automated/embedded diagnostics. The implementation of the embedded diagnostic system then requires proper window selection (Ganeriawala, 2006). This is done by holding constant the cyclic frequency α , and selecting both a window bandwidth and spectrum length. Then for each fault frequency (Eq 1 – Eq4), the energy spectral density associated with the

frequency spectrum is the fault condition indicator. Eq 9 then reduces to:

$$\mathfrak{R}_{2X}(f) = \Delta^2 \sum_{\tau=-\infty}^{\infty} \mathcal{R}_{2X}[\tau] e^{-j2\pi\alpha n\Delta} e^{-j2\pi f\tau\Delta} \quad \text{Eq 10.}$$

The pseudo code for this is simply:

$$yh = y \times \exp(-i2\pi\alpha t)$$

lowpass filter filter yh to remove image

$$Y = \mathfrak{F}(yh)$$

$$\text{Spectrum}(f) = Y \times Y^* \quad \text{Eq 11.}$$

where y is the vector of vibration data, yc is the heterodyned (complex) vibration signal, and as per definition $t = n\Delta$. Note that in implementation, $\exp(-i2\pi\alpha t)$ is calculated using Euler's formula, where the $\text{real}(yh)$ is the product of $y[n]$ and the cosine of $2\pi\alpha t$, and the $\text{imag}(yh)$ is the product of $y[n]$ and sine of $2\pi\alpha t$. Because $p[n]$ is a periodic signal it is modeled as the cosine of fault frequency f . It then follows that the trigonometric identity applies:

$$\cos(f) \times \cos(\alpha) = 1/2[\cos(f - \alpha) + \cos(f + \alpha)] \quad \text{Eq. 12}$$

This requires low pass filtering to remove the image of yh (e.g. $\cos(f + \alpha)$). The bandwidth of the low pass filter must be greater than maximum fault frequency, and should encompass resonance bandwidth of the cyclic frequency. For example, Figure 3, the bandwidth should be greater than 25 Hz. It is usual practice to decimate the resultant by the ratio of Nyquist of the sample rate to the filter bandwidth. The product of the Fourier transform of yh with its conjugate is the envelope power spectrum of y .

Window selection (e.g. α , and bandwidth) greatly affects the performance of analysis. Figure 4, with a bandwidth of 100 Hz, shows that for $\alpha = 1$ Hz, the CPF is modulating the

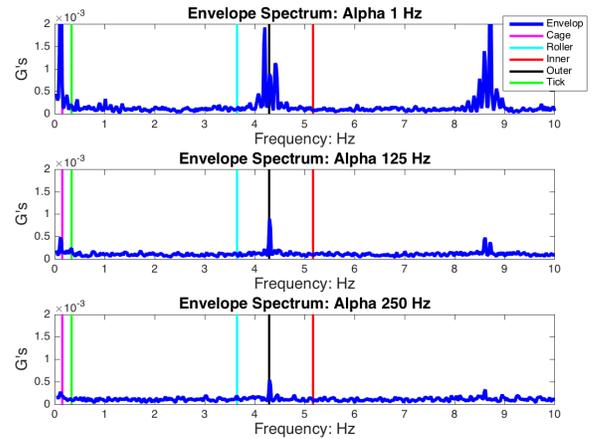


Figure 4 Comparison of Cyclic Frequency of Machine 2

BPFO, causing side bands around the fault frequency of 0.12 Hz. An automated system would select the energy associated with the side band vs. the BPFO fault. When $\alpha = 125$ Hz, the CPF sidebands are eliminated, the peak is associated with the BPFO. Further the energy associated with the BPFO is the same for $\alpha = 1$ Hz and $\alpha = 125$ Hz.

For $\alpha = 250$ Hz, the energy associated with the BFPO is greatly reduced (α is now outside of the resonance of the structure). For this study, $\alpha = 125$ Hz was selected. Note that for $\alpha = 0$ Hz, equation 10 reduces to:

$$\mathfrak{R}_x(f) = \Delta^2 \sum_{\tau=-\infty}^{\infty} \mathfrak{R}_X[\tau] e^{-j2\pi f\tau\Delta} \quad \text{Eq 13.}$$

which is just the Fourier transform. No bearing fault features were visible in the raw spectrum, which is why the envelope analysis is used (see Figure 7).

2.3. Resolution

Resolution is the ability to discern one feature from another. Consider analysis of the main bearing of a commercial, 1.25 MW wind turbine. The main rotor shaft speed is approximately 20 RPM, or 0.34 Hz. The main bearing fault rates are given as:

- Cage fault frequency is 0.45 Hz,
- Ball/Rolling element fault frequency is 10.8,
- Inner race fault frequency is 15.3 and
- Outer race fault frequency is 12.7.

Assume that one needs at least 10, preferable 30 bins between the frequencies of interest. To distinguish between DC and the cage CPF, the frequency resolution must be between 0.015 and 0.005 Hz.

For a sample rate of 3,000 samples per second (Nyquist of 1,500 Hz), this is 300,000 data points. Since the Fast Fourier Transform (FFT) typically uses radix 2, i.e. $2^{\lceil \log_2(300,000) \rceil}$, or 524,288 data points, or 175 seconds of data!

Of course, in most applications, the torque/load is not constant over time. In this wind turbine application, due to tower shadow and wind shear, there is a three per revolution change in speed (Figure 5).

In this two minute acquisition, the mean rotor speed was 20.39 RPM with a standard deviation of 0.2 RPM, or 0.003 Hz. Numerically, for inner race fault of 15.3 Hz, 68% of the time (given the 0.005 Hz/bin resolution), the spectral content is smeared across +/-3 bins in the frequency domain. In order to correct for variation and shaft rate and improve the envelope analysis, a time synchronous resampling algorithm was implemented (Bechhoefer et al., 2013).

2.3.1. Resampling vs. Time Synchronous Averaging

The time synchronous average, (TSA, McFadden, 1987) has long been used for gear and shaft analysis to control

variation in shaft speed. However, this technique is not appropriate for bearing analysis. The TSA is only appropriate in instances where the features are both an integer of the shaft rate and synchronous with the shaft. This is not the case for bearing analysis, where the fault features are not an integer of the shaft, and are only quasi-synchronous (due to slippage of the bearing).

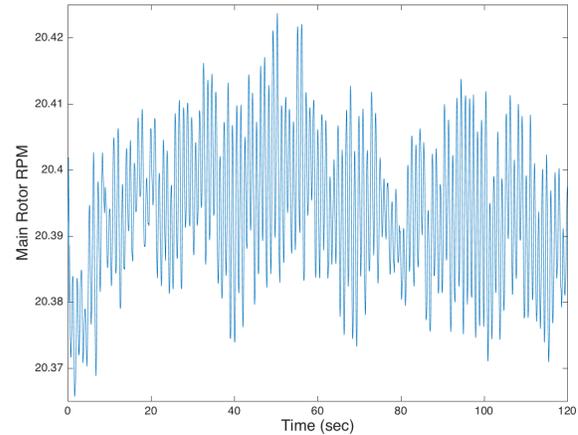


Figure 5 Main Rotor RPM over 120 Seconds

While the resampling algorithm is similar to the TSA, it does not average the data points over each revolution. Instead, it builds a new array of data, which is locked synchronous with shaft position by using a tachometer key phasor (Figure 6).

For example, say the sample rate was 1000 samples per second, and the lowest shaft rate was 10 Hz, for a .5 second acquisition. The resample length, m , is 128. The number of data points between each key phasor was: 87, 92, 100, 95, 89, 37. For each shaft revolution, the data is resampled to length l : Rev 1: 87->128, Rev 2: 92->128, Rev 3: 100->128, Rev 4: 95->128, Rev 5: 89->128. Note that for half of a second of data, there is 640 data points – the remaining 37 data points were in the next, incomplete revolution, so the last 37 data points are dropped. The resample length is taken as the next largest power of 2 over the maximum length of all revolutions, again, assuming a radix 2 DFT.

Because of interpolation, the sample rate for each revolution is now changed. To accurately determine the frequency associated with a DFT bin, an apparent sample rate is calculated. This is the original sample rate \times length of the resampled data / length of the original data: $1000 \times 640 / (463)$, or 1382.

For the main shaft, given the approximately 0.34 Hz shaft rate, it takes approximately 2.9 seconds for a revolution. For a sample rate of 3,000 sps, there are, on average, 8,800 data points. In the resampling algorithm, for each revolution, the data is up-sampled to 16,384 data points. This effectively removes any variation in shaft speed.

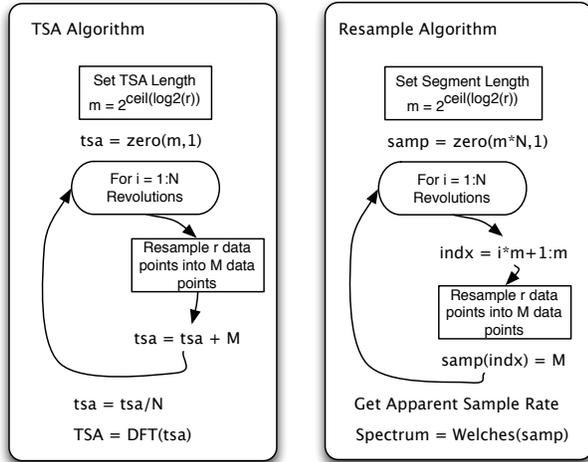


Figure 6 TSA vs. Resampling Algorithm

2.4. Energy Associated with the Bearing Fault Features

Faults associated with slowly rotating equipment are difficult to measure because the energy is small, as a result of measuring acceleration, and not displacement. Consider a 30 Hz output shaft, with a 0.001" (1 mil) displacement fault on the outer race (e.g. BPFO rate of 5). Because acceleration is the second derivative of displacement, the estimated acceleration of this fault would be $0.0511 \times (30 \times 5)^2 \times 0.001 = 1.15$ Gs. This is not an unreasonable value, as we can see in Figure 1, an RMS of 1.76 Gs on a 25 Hz shaft.

For the wind turbine example, which is shaft rate of 0.34 Hz and a BPFO frequency of 4.3 Hz, the same damage, 0.001", would generate an acceleration of $0.0511 \times 4.3^2 \times 0.001 = 0.00094$ Gs. This is a very small value to measure in the presence of noise.

2.5. Few Data Sets

While not a technical hurdle, programmatically there is a lack of experience in working with large bearings. This is due in part to little, if any, documentation of the structural resonance of these bearings. This is, as noted, essential for envelope analysis. Further, because of the size, weight and cost of these types of bearings, there are few seeded fault tests conducted to learn about this type of bearing fault phenomenology.

3. LARGE SLOW BEARING EXAMPLE

This example is based on the availability of three commercial wind turbines (1.25 MW), where one of the turbines had a suspected main bearing fault. The machines were equipped with a bused condition monitoring system, which collected vibration data on two accelerometers mounted on the main bearing, and a tachometer signal. The accelerometers were based on MEMS technology, where

- Accelerometer 1 has a bandwidth of 0 to 32 KHz, and a noise density of $4\text{mg}/\sqrt{\text{Hz}}$, +/-70Gs
- Accelerometer 2 has a bandwidth of 0 to 2.5 KHz, and a noise density of $110 \mu\text{g}/\sqrt{\text{Hz}}$, +/-18Gs

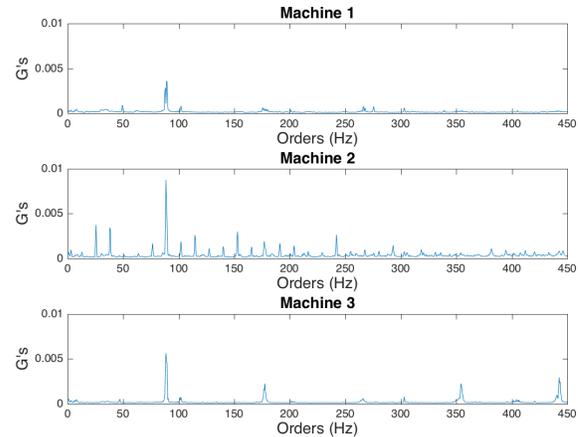


Figure 7 Resampled Spectrum in Order Domain

Each sensor was configured to sample at 3,052 sps for 120 seconds using a 24-bit ADC. The resampled spectrum, in order domain (which corrects for the variance in the main rotor RPM across the three machines) is seen in Figure 7. The order domain is the frequency divided by shaft rate. This is done to normalize each plot by the machine RPM. Each machine has a large spectral tone at order 89 (this is equivalent to 30.52 Hz), which is associated with the planetary gearbox mesh frequency. This was also clearly visible in the time domain.

Machine 2 shows multiple harmonics order 76 and 140 (26 and 140 Hz). This is likely due to resonance. As noted the envelope window $\alpha = 125$ Hz, with a bandwidth of 100 Hz. This means that for operation on Eq. 7, the vibration signal y was low pass filtered with a normalized frequency of: $100 / (5,580 / 2) = 0.03585$, then decimated by 27 prior to the Fourier transform being taken. This allows for improved spectral resolution for a given FFT length. Further, the accelerometer's spectral noise, which is a function of $1/\sqrt{\text{bandwidth}}$, improving SNR by 7.2 dB.

In Figure 8 (zoomed to show 0 to 40 Hz), Machine 2 indicates a BPFO (outer race fault), with multiple higher order harmonics associated with the outer race fault frequency. These fault features are missing from machine 1 and 3. The fault energy in Machine 2 is approximately 10 times the noise density (e.g. floor energy) of Machine 1 and 3.

Note that after resampling the apparent sample rate is 5,580 sps, with a Nyquist of 2,760 Hz. The bandwidth of the window, being 100 Hz, allows a large decimation in time, by 27, decreasing in the accelerometer spectral noise by 7.1

dB. Even with a modest 16,384 point FFT (about the largest that can be done without numerical errors with 32 bit floating point resolution) the bin width of 0.0125 Hz. This gives enough resolution to resolve the cage fault from DC.

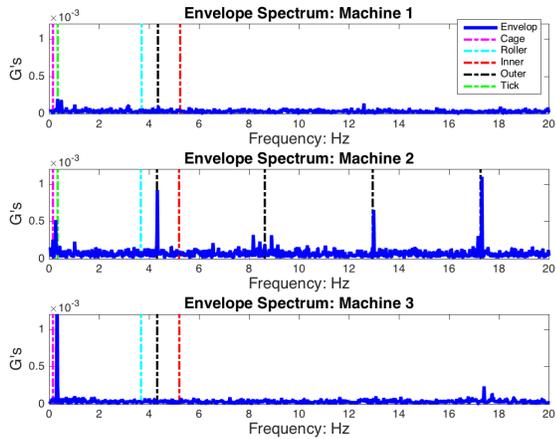


Figure 8 Envelope Spectrum Clearly Showing an Outer Rate Fault (first 4 Harmonics) on Machine 2

3.1. Cepstrum Analysis

The real cepstrum analysis was defined by Oppenheim (1965) and is a way to de-convolute homomorphic systems. The cepstrum is defined as:

$$cep = |\mathcal{F}^{-1}\{\log(|\mathcal{F}\{f(t)\}|)\}| \quad \text{Eq 14.}$$

This is the real part of the inverse FFT of the log of the FFT. The concept is that the product of two signals is the sum of the logs of the signals. Since a bearing fault is the modulation of the fault frequency on the resonance frequency (e.g. multiplication of two signals), the cepstrum should be sensitive to this fault (Figure 9). Because the inverse Fourier analysis has been called, the cepstrum is a time domain analysis.

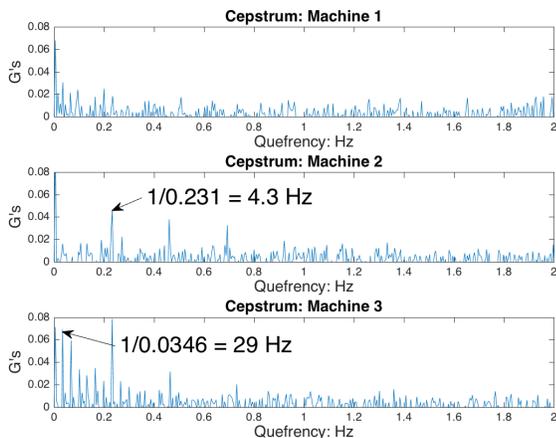


Figure 9 Cepstrum of the Main Bearings

The x-axis of the analysis is called the queffrequency (in seconds), and is the reciprocal of the frequency spacing.

For machine 2, the peak at queffrequency of 0.231 presents a frequency of 4.3 Hz. This coincides with the outer race fault. For machine 3, the queffrequency peak at 0.0346 represents approximately 30 Hz in the frequency domain, which was seen in the envelope spectrum (Figure 8). This is likely due to gear mesh and is not a feature of a bearing fault.

As the cepstrum is a time domain analysis, a number of condition indicators can be generated from it for automated analysis. In this example the cepstrum kurtosis or cepstrum RMS was tested as fault indicators. RMS was selected as it is a standard measure of noise. The cepstrum of a signal with multiple harmonics will have greater noise than the cepstrum with no harmonics present. Kurtosis is used to measure the “tailedness” of the PDF. Similarly, harmonics as a result of impact will have a more heavily tailed PDF in the cepstrum.

As seen in Table 1, machine 3 generated largest cepstrum RMS and kurtosis (see Table 1). This was generated by the 0.0346 queffrequency features that corresponded to gear mesh. These condition indicators values were based on the average of 10 acquisitions. This information would not be actionable for the bearing fault, but may indicate a potential planetary gear ring fault.

Table 1 Cepstrum Condition Indicators

Machine	Cepstrum RMS	Cepstrum kurtosis
Machine 1	0.0038	63.33
Machine 2	0.0036	48.28
Machine 3	0.0045	98.42

It is felt that additional work needs to be performed on automated cepstrum analysis. It is likely that other cepstrum based condition indicators could work better than RMS and kurtosis.

For automated fault detection, a method is needed to both quantify the bearing damage and to recommend a maintenance action when appropriate. The quantification of damage is done using a condition indicator (CI), which is a descriptive statistics of the component under analysis. Condition indicators for bearings include the energies associated with the bearing’s pass frequencies: CPF, BPO, BPFI, BPFO. Other possible CIs are the average spectral energy, the spectral kurtosis (although more commonly used for window selection, see Randall 2011), and as noted, the cepstrum RMS and kurtosis.

4. BEARING CONDITION INDICATORS AND HEALTH

For this experiment, 16 acquisitions were available from Machine 1, nine acquisitions from Machine 2, and 15 from Machine 3. This allowed a population of CIs to be developed for the cage, ball, and inner and outer race energies.

4.1. Threshold Setting and Component Health

Because the physics of failure is poorly understood (e.g. from a given CI value, the absolute level of damage cannot be calculated), threshold setting is typically a data driven process: maintenance is performed when a statistically set threshold is exceeded. The concept of thresholding was explored by Byington et al. (2004) where for a given, single CI, a probability density function (PDF) for the Rician/Rice statistical distribution was used to set a threshold based on a probability of false alarm (PFA). This is contrasted with Dempsey (2008), where the relationship between CI threshold and PFA was explored to describe the receiver operating characteristics (ROC) of the CI for a given fault. Additionally, Dempsey used the ROC to evaluate the performance of the CI for a fault type. These methods support a data driven approach for condition monitoring by formalizing a method for threshold setting.

Because no single CI has been identified that works with all fault modes, the concept of fusing n number of CIs into a bearing health indicator (HI) was presented in Bechhoefer (2012). Computationally, the use of an HI is attractive. The HI provides a decision-making tool for the end user on the status of the system health. The HI consists of the integration of several CIs into one value that provides the health status of the component to the end user.

This concept is similar to that proposed by Djurdjanovic et al. (2003) of a Watchdog Agent. The Watchdog Agent fuses multiple inputs and extracted features relevant to the components performance. This performance is then evaluated with a performance confidence value (CV) ranging between 0 and 1. In this paradigm, the higher CV signifies performance closer to nominal.

Highlighted in (Bechhoefer, 2007) are a number of advantages of the HI over CIs, such as: controlling false alarm rate, improved detection, and simplification of user display. This approach allows the use of well established statistical methods. Further, it is a generalized process for threshold setting, where the HI is a function of distribution of CIs, regardless of the correlation between the CIs.

Prior to detailing the mathematical methods used to develop the HI, a nomenclature for component health is needed. To simplify presentation and knowledge creation for a user, a uniform meaning across all components in the monitored machine should be developed. The measured CI statistics (e.g. PDFs) will be unique for each component type (due to different rates, materials, loads, etc.). This means that the

critical values (thresholds) will be different for each monitored component. By using the HI paradigm, one can normalize the CIs, such that the HI is independent of the component.

The HI can be designed such that there are two alert levels: warning and alarm. This paradigm also provides a common nomenclature for the HI, such that:

- The HI ranges from 0 to 1, where the probability of exceeding an HI of 0.5 for a nominal component (e.g. no damage) is the PFA.
- A warning alert is generated when the HI is greater than or equal to 0.75.
- An alarm alert is generated when the HI is greater than or equal to 1.0. Continued operations could cause collateral damage.

This nomenclature does not define a probability of failure for the component, or that the component fails when the HI is 1.0. Rather, it suggests a change in operator behavior to a proactive maintenance policy: perform maintenance prior to the generations of cascading faults. For example, by performing maintenance on a bearing prior to the bearing shedding extensive material, costly gearbox replacement can be avoided.

4.1.1. Controlling for the Correlation Between CIs:

All CIs have a probability distribution (PDF). Any operation on the CI to form a health index (HI) is then a function of distributions. For this study, the HI function is taken as the norm of n CIs (energy).

In general, the correlation between CIs is non-zero. This correlation implies that for a given function of distributions to have a threshold that operationally meets the design PFA, the CIs must be whitened (e.g. de-correlated). Fukunaga (1990) presented a whitening transformation using the Eigenvector matrix multiplied by the square root of the Eigenvalues (diagonal matrix) of the covariance of the CIs: $\mathbf{A} = \mathbf{\Lambda}^{1/2} \mathbf{\Phi}^T$, where $\mathbf{\Phi}^T$ is the transpose of the Eigenvector matrix and $\mathbf{\Lambda}$ is the Eigenvalue matrix. The transformation is not orthonormal; the Euclidean distances are not preserved in the transformation. While ideal for maximizing the distance (separation) between classes (such as in a Bayesian classifier), the distribution of the original CI is not preserved. This property of the transformation makes it inappropriate for threshold setting.

If the CIs represented a metric, such as bearing energy, then an HI can be constructed, which is the square of the normalized power (e.g. square root of the squared acceleration). A generalized whitening solution can be found using Cholesky decomposition (see Bechhoefer and Dempsey 2011). The Cholesky decomposition of a Hermitian, positive definite matrix results in $\mathbf{A} = \mathbf{LL}^*$,

where L is a lower triangular, and L^* is its conjugate transpose. By definition, the inverse covariance is a positive definite Hermitian. It then follows that if

$$LL^* = \Sigma^{-1} \quad \text{Eq 15.}$$

then

$$Y = L \times CI^T \quad \text{Eq 16.}$$

The vector CI is the correlated CIs used for the HI calculation, and Y is 1 to n independent CIs with unit variance (one CI representing the trivial case). The Cholesky decomposition, in effect, creates the square root of the inverse covariance. This in turn is analogous to dividing the CIs by their standard deviations (the trivial case of one CI). This creates the necessary independent and identical distributions required to calculate the critical values for a function of distributions.

4.1.2. HI Based on Rayleigh PDFs

The CIs for bearing used the spectral energy associated with the fault frequency of the cage, ball, inner and outer race. The energy was the square root of the power spectrum. Energy was used because is relatively linear with damage size and, assuming nominal bearing condition, has Rayleigh like PDFs (e.g. heavily tailed). Further, for a white noise signal, it can be shown that the Fourier transform for a given bin is Rayleigh distribution is Consequently, the HI function was designed using the Rayleigh distribution. The PDF for the Rayleigh distribution uses a single parameter, β , defining the mean

$$\mu = \beta \times (\pi/2)^{0.5} \quad \text{Eq 17.}$$

and variance

$$\sigma^2 = (2 - \pi/2) \times \beta^2 \quad \text{Eq 18.}$$

The PDF of the Rayleigh is:

$$x/\beta^2 \exp(x/2\beta^2) \quad \text{Eq 19.}$$

Note that when applying these equations to the whitening process, the value for β for each CI will then be: $\sigma^2 = 1$, and

$$\beta = \sigma^2 / (2 - \pi/2)^{0.5} = 1.5264 \quad \text{Eq 20.}$$

The HI function using the norm of n CIs can be shown to define a Nakagami PDF (Bechhoefer and Bernhard 2007). The statistics for the Nakagami are $\eta = n$, and

$$\omega = 1/(2-\pi/2) \times 2 \times n, \quad \text{Eq 21.}$$

where n is the number CIs used in the calculation of Y . The critical value for the HI, based on four CIs (cage, ball, inner and outer race energy), i.e. $\eta = 4$ and $\omega = 18.64$. For a PFA of 10^{-6} , the threshold 9.97, with the HI function calculated as:

$$HI = \frac{0.5}{9.97} \sqrt{YY^T} \quad \text{Eq 22.}$$

The 0.5 value normalized the HI, such that the probability of an HI being greater than 0.5 for a nominal bearing is 10^{-6} . The HI for the machines are given as in Figure 10.

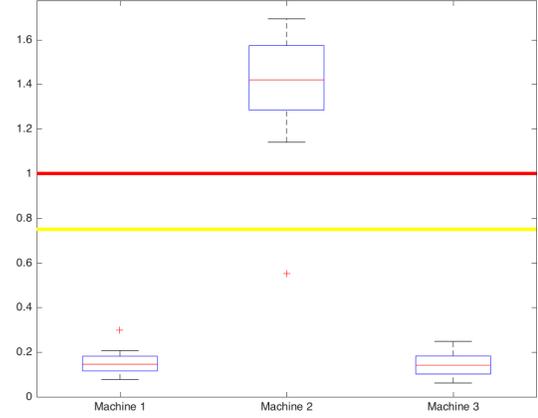


Figure 10 Main Bearing Health

5. EFFECT OF ACCELEROMETER NOISE DENSITY

The prior analysis was conducted with accelerometer 2, with a noise density $110 \mu\text{g}/\sqrt{\text{Hz}}$. This is because there was no fault detection with accelerometer 1, with a noise density of $4\text{mg}/\sqrt{\text{Hz}}$, i.e. accelerometer 1 had a $36\times$, or a 15.6 dB increase in the noise floor of accelerometer 2. By increasing the acquisition time by $4\times$, the signal to noise can be improved by 3 dB ($2\times$). This is because the Welch's spectrum averages the power, reducing random noise by $1/\sqrt{n}$, where n is the number of averages.

This was verified by increasing the data length to 480 seconds, and 960 seconds ($4\times$ or 3dB and $8\times$ or 4.5 dB , Figure 11).

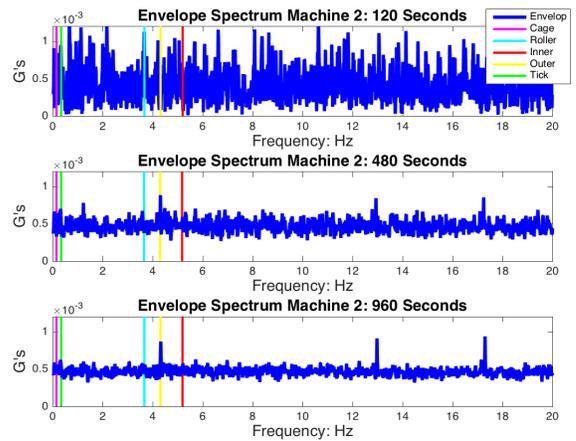


Figure 11 Reducing Noise by Increasing Sample Length

Increasing the Signal to Noise by 3 dB allows the outer race fault to be detectable. The acquisition is eight minutes and

1,464,960 data points. This may exceed the memory of some embedded systems. Additionally, because of the large variation in speed seen over eight minutes, it necessitates the use of a time synchronous resampling algorithm (Bechhoefer et al 2013).

6. OTHER CONDITION ANALYSIS TECHNIQUES

Both temperature (from the supervisory control and data acquisition, SCADA) and grease analysis was available for these machines. While it may be assumed that temperature analysis is a simple method, this is not the case. Temperature measured from the main bearing is related to ambient temperature (which was not available) and turbine power output. If the wind speed is low, or the turbine has been ordered to curtail production, the main bearing approaches the ambient temperature. If the machine is in production, the temperature is related to load (e.g. power production), ambient temperature and damage.

Further complicating this is that the wind at one area of the wind farm is likely to be different from another part of the wind farm, such that the power production from one machine varies in time with other machines (Figure 12).

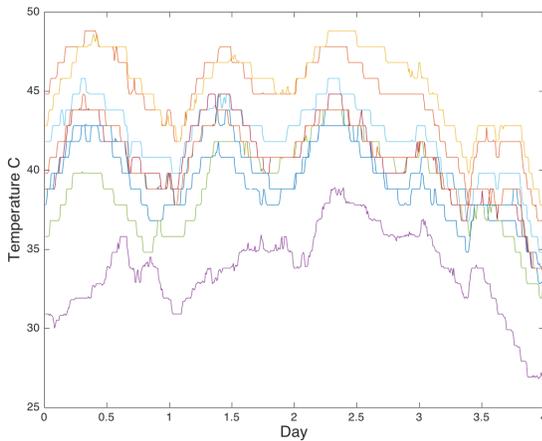


Figure 12 Temperatures from 9 Wind Turbines

The periodic daily variation in temperature is based on the increased wind speed at night. Increased wind speed increases the load on the main bearing, which in turn results in elevated temperatures. Initially, exploratory analysis was performed to measure the correlation between temperature and power production (11,000 data points):

M1	M2	M3	M4	M5	M6	M7	M8	M9
.59	.63	.59	.57	.41	.47	.61	.62	.51

This shows that approximately 50% of the variance is associated with power production. Note that machine 2 had the largest correlation between power and temperature.

A boxplot of the machines (1 through 9) shows that there is a large difference between the mean temperatures of each

machine (Figure 13). Machine 2 vibration and grease analysis (Table 2) suggest that it is the most damaged machine. Yet the median temperature of machine 2 is less than that of machine 3.

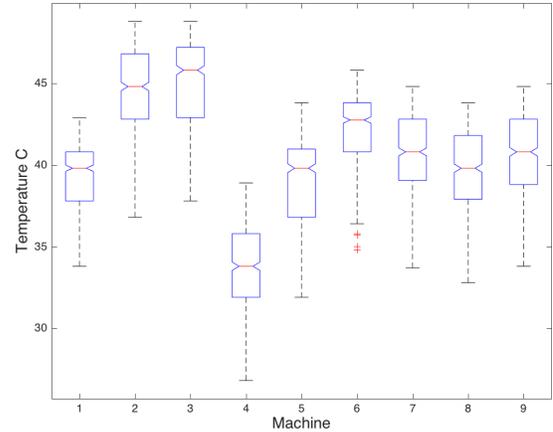


Figure 13 Boxplot of Temperature vs. Machine

Because machine 2 had the highest correlation between power production and temperature, a residual model was built that remove the mean temperature (Figure 14).

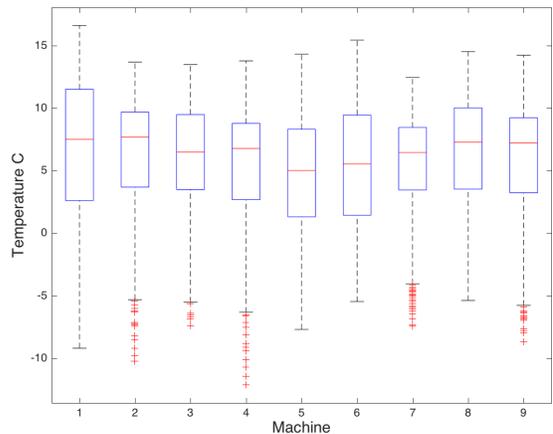


Figure 14 Boxplot of Residual Temperature vs. Machine

There is no significance between machine 2 and the other machines.

When ANOVA (analysis of variance) is conducted on the temperature, the F statistic with 8 degrees of freedom is 1083: the hypothesis is rejected that the temperature are the same. Because it is assumed (via grease analysis) that machines 2 and 3 are biased high, they can be removed from the ANOVA (along with machine 4, which is running cool). This should allow one to test the hypothesis that the temperatures of the nominal machines are the same. The F statistic for this test is 121.8 with 5 degrees of freedom. The probability of this exceeding F is 0, hence rejecting the null

hypothesis that the nominal machines have similar temperatures.

It is clear that for temperature analysis to be effective, the data must be normalized by ambient temperature (unfortunately unavailable from the SCADA system) and turbine power output. Without reducing the between machine variance, temperature appear to be a poor analysis tool for early stage, automated fault detection.

6.1. Grease Analysis

Oil analysis has had a long history of acceptance and is recognized as an effective condition monitoring technique. This is not the case with grease. Grease analysis presents a challenge in obtaining representative samples. One most important functions of grease is to stay in place – it is precisely this property that makes reliable sampling difficult. Grease does not readily circulate. This means that valid samples must be taken in an area close to the target bearings, know as the “live zone”.

Based on a number of studies (Electrical Power Research Institute, (2010), Møller, H., et. al., (2014)), the American Society of Testing and Materials (ASTM) Grease subcommittee established a working group. Over a two-year period, the committee worked on delivering a peer reviewed document to obtain industry consensus on grease sampling practices. This was balloted by the ASTM Main Committee, resulting in the release of ASTM D7718, “Standard Practice for Obtaining In-Service Samples of Lubricating Grease.”

The standard gives guidance on how to take representative grease samples. Much evidence suggest that grease samples, take in accordance with the standard, allows for trending of wear levels in bearings and gearboxes.

Grease analysis was available for 8 of the 9 machines on which temperature SCADA data was available. It was the case that machines 1, 2 and 3 had marginally higher levels of wear materials than the other machines. Since vibration data was available for machines 1, 2 and 3, only these are compared.

The grease analysis was performed by an ISO 9001:2008 and ISO/IEC 17025:2005 certified laboratory (Table 2).

Table 2 Main Bearing Grease Analysis (ppm)

Aqueous Metals	Machine 1	Machine 2	Machine 3
Silver, Aq. ppm	<9.37	<10.98	<10.05
Aluminum, Aq. pp	<10.8	**268.5	**175.6
Boron, Aq. ppm	<7.9	<9.3	<8.5
Barium, Aq. ppm	<5.8	<6.8	<6.2
Calcium, Aq. ppm	*34.1	*43.1	<17.8
Cadmium, Aq. ppm	<5.8	<6.8	<6.2
Chromium, Aq	<10.8	**136.4	**91.3

ppm			
Copper, Aq. ppm	**71.5	**16483.0	**11679.7
Iron, Aq. ppm	121.8	**27499.3	**16084.7
Potassium, Aq. p	<14.4	<16.9	<15.5
Magnesium, Aq. p	<7.2	<8.4	<7.7
Manganese, Aq. p	<5.8	*171.3	*77.2
Molybdenum, Aq.	<13.0	39.8	<13.9
Sodium, Aq. ppm	<15.9	<18.6	<17.0
Nickel, Aq. ppm	<7.2	<8.4	<7.7
Phosphorus, Aq.	*<50.5	*991.2	*669.1
Lead, Aq. ppm	<31.7	**236.3	**132.0
Antimony, Aq. pp	<16.6	<19.4	<17.8
Silicon, Aq. ppm	20.7	**241.5	**147.2
Tin, Aq. ppm	<20.9	<24.5	<22.4
Titanium, Aq. pp	<3.6	<4.2	<3.9
Vanadium, Aq. pp	<3.6	<4.2	<3.9
Zinc, Aq. ppm	*134.6	*10629.7	*7173.2

The table is coded by the laboratory such that (*) yellow indicates a caution limit and (**) red is an alarm limit. Machine 2 had the highest level of aqueous metals in the grease. Both machine 2 and machine 3 grease analysis indicates a fault, yet, machine 3 vibration health is nominal, i.e. not actually higher than machine 1. Further, both machine 2 and 3 tend to have higher temperatures than machine 1. Hence, there is some collaboration that machine 2 is faulted (high vibration, high grease particulates, and higher temperature). There is ambiguity in the condition on machine 3, as the vibration analysis appears nominal, but grease and temperature are elevated.

7. CONCLUSION

The availability of both vibration, temperature and grease analysis from three wind turbines, where one turbine is suspected of having damaged main bearing, allows for an opportunity to develop a strategy for fault detection of large, slow bearings. Approximately 10 acquisitions from two accelerometers (high bandwidth, high noise, and low bandwidth, low noise) and a tachometer were taken. Temperature was available from the SCADA system every 10 minutes. Grease analysis from an ISO certified laboratory was also available from these machines.

In order to detect faults using vibration data, long acquisition times are needed (120+ seconds). Because of the long acquisitions, there were large changes in shaft speed. This is detrimental to the analysis because it smears the spectral content in the Fourier analysis. A time synchronous resampling algorithm used tachometer data to correct for

changes in shaft speed, improving the analysis (The Fourier analysis assumes a stationary signal).

The envelope analysis was conducted using resampled vibration data (120 seconds at 3,052 sps) and a 50 to 150 Hz window with a 16,384-point spectrum. This facilitated the detection of an outer race fault on machine 2. Machine 1 and 3 were found to be normal. The fault was detected with the low noise density (110 $\mu\text{g}/\sqrt{\text{Hz}}$) sensor. The high noise density accelerometer could not detect the fault until the acquisition was increased to 480 seconds (3 dB improvement in SNR).

The cepstrum analysis was performed on the vibration data, but the results were not actionable. The cepstrum detected ring mesh gear on machine 1 and 3, which were assumed nominal. The condition indicators (cepstrum RMS, kurtosis) of the cepstrum analysis did not indicate if machine 2 had a faulted bearing. Further condition indicator development for cepstrum is recommended.

The health of the main bearing was calculated using a data driven method. Using the Nakagami distribution and a probability of false alarm of 10^{-6} , the bearing condition indicator data was mapped to a health indicator, where a health indicator of less than 0.5 is nominal, greater than 0.75 and less than 1 is warning, and greater than 1 is alarm. The health indicator for the main bearing of machine 2 was 1.45.

ANOVA analysis of the bearing temperatures rejected the hypothesis that the temperature of nominal machines was similar. In general, while machine 2 ran someone hotter than machine 4 through 9, the difference was not statistically significant. In fact, machine 3 ran hotter than machine 2.

The grease analysis indicated that both machine 2 and 3 were faulted, as many aqueous metals were in alarm. Grease analysis indicated that machine 2 was most damaged.

Typically, vibration based detection on large slow bearings is difficult. However, using long acquisition times to improve resolution and proper window selection, the vibration data were easier to interpret, and gave more actionable information, than temperature and grease analysis. With online condition monitoring, using the trend of the health indicator, it is likely that remaining useful life of the bearing could be calculated. The vibration data, as presented, can give operations and maintenance organizations valuable information to enhance logistic support of the machine.

ACKNOWLEDGEMENT

The research presented in this paper has received funding from the Norwegian Research Council, SFI Offshore Mechatronics, project number 237896.

REFERENCES

- Antoni, J., (2007) Cyclic spectral analysis in practice, *Mechanical Systems and Signal Processing* 21, 597-630.
- Antoni, J., (2009) Cyclostationarity by examples, *Mechanical Systems and Signal Processing*, 23(2009) 987-1036.
- Bechhoefer, E. (2013), *Condition Based Maintenance Fault Database for Testing Diagnostics and Prognostic Algorithms*. Retrieved from <http://www.mfpt.org/FaultData/FaultData.htm>
- Bechhoefer, E., & Bernhard, A. (2007). *A Generalized Process for Optimal Threshold Setting in HUMS*. IEEE Aerospace Conference, Big Sky.
- Bechhoefer, E., Duke, A., & Mayhew, E. (2007). *A Case for Health Indicators vs. Condition Indicators in Mechanical Diagnostics*. American Helicopter Society Forum 63, Virginia Beach.
- Bechhoefer, E., & Fang, A., (2012) Algorithms for Embedded PHM, *Prognostics and Health Management (PHM)*, IEEE.
- Bechhoefer, E., He, D., & Dempsey, P. (2011). *Gear Threshold Setting Based On a Probability of False Alarm*. Annual Conference of the Prognostics and Health Management Society.
- Bechhoefer, E., & He, D., *A Process for Data Driven Prognostics*, (2012), MFPT 2012: Dayton, Ohio.
- Bechhoefer, E., Van Hecke, B., & He, D., *Processing for Improved Spectral Analysis*, PHM Conference, 2013.
- Boskoski, P., & Juricic, D. (2013), Modeling localized bearing faults using inverse Gaussian mixtures, Annual Conference of the Prognostics and Health Management Society 2013.
- Byington, C., Safa-Bakhsh, R., Watson, M., & Kalgren, P. (2003). *Metrics Evaluation and Tool Development for Health and Usage Monitoring System Technology*. HUMS 2003 Conference, DSTO-GD-0348
- Darlow, M.S., Badgley, R. H. & Hogg, G.W., (1974). Application of high frequency resonance techniques for bearing diagnostics in helicopter gearboxes. US Army Air Mobility Research and Development Laboratory, Technical Report pp 74-76.
- Dempsey, P., & Keller, J. (2008). Signal Detection Theory Applied to Helicopter Transmissions Diagnostics Thresholds. NASA Technical Memorandum 2008-215262
- Djurdjanovic, D., Jay L., & Jun N., "Watchdog Agent—an infotronics-based prognostics approach for product performance degradation assessment and prediction." *Advanced Engineering Informatics* 17.3 (2003): 109-125.
- Electrical Power Research Institute, (2010), "Nuclear Maintenance Applications Center: Effective Gear Practices", Technical Report #1020247.
- Fukunaga, K., (1990), *Introduction to Statistical Pattern Recognition*, Academic Press, London, 1990, page 75.

- Ganeriawala, S., (2006) Some Observations of the Detection of Rolling Bearing Outer Race Faults, SpectraQuest, www.spectraquest.com
- Ho, D. & Randall, R. B. (2000) Optimization of bearing diagnostic techniques using simulated and actual bearing fault signals. *Mechanical Systems and Signal Processing*. 14 (5), 763-788.
- McFadden, P., (1987) "A revised model for the extraction of periodic waveforms by time domain averaging", *Mechanical Systems and Signal Processing* 1 (1) 1987, pages 83-95
- Møller, H., et. al., (2014) "Analysis of Grease in Wind Turbine bearing – a tool for condition monitoring. Part 2", LUBMAT 2014 Proceedings, Manchester, UK.
- Oppenheim, A.V. (1965) "Superposition in a class of nonlinear systems" (Ph.D. dissertation), Res. Lab. Electronics, Massachusetts Institute of Technology, Cambridge, MA.
- Randall, R. (2011) *Vibration-based Condition Monitoring: Industrial, Aerospace & Automotive Application*, John Wiley, New York, 2011.

BIOGRAPHIES

Eric Bechhoefer is the president of GPMS, Inc., a company focused on the development of low-cost condition monitoring systems. Dr. Bechhoefer is the author of over 100+ juried papers on condition monitoring and prognostics health management, and holds 23 patents in the field of CBM.

Rune Schlanbusch received his MSc in Space Technology from Narvik University College (NUC), Norway in 2007, and a PhD in Engineering Cybernetics from NTNU, Norway in 2012. He currently holds positions as Senior Researcher at Teknova, Norway and II Associate Professor at The Arctic University of Norway (UiT).

Tor Inge Waag works as Specialist Engineer at MHWirth AS and holds a part time position as Senior Researcher at Teknova AS in Norway. He has received both his MSc and PhD in Physics from NTNU, Norway. His primary work focus is at present on prognostics and health management, signal processing, and condition monitoring.