Wheel Bearing Fault Detection for Automobiles using Wheel Speed Sensor

Samba Drame[1], Graeme Garner[2], Xinyu Du[3], and Hossein Sadjadi[4]

[1, 4]*General Motors Company, Canadian Technical Centre, Markham, Ontario, L3R 4H8, Canada*
samba.drame@gm.com
hossein.sadjadi@gm.com

[3]*General Motors Global R&D, Warren, MI, 48092, USA*
xinyu.du@gm.com

[2]*Hexagon's Autonomy & Positioning division, Calgary, Alberta, T3K 2L5, Canada*

garnergraeme@gmail.com

ABSTRACT

Wheel bearing fault diagnosis and monitoring for automobiles are getting more attention as the market shifts towards electric vehicles (EV) and autonomous vehicles (AV). Electric drivetrains are expected to have longer lifespan than the reliable life of low-cost wheel bearing designs. The probability of having a failed wheel bearing in an EV thus increases. In an AV setting, most of the rides are taken by people who are unfamiliar to the vehicle, removing the human as a viable sensor to detect wheel bearing defects. Fleet managers will be required to pay a high price to inspect potential defects. The implementation of a wheel bearing fault detection system can help reduce these costs.

In our previous work, a method for injecting wheel bearing brinelling failure was developed. In this paper, we propose a wheel bearing brinelling fault detection algorithm using wheel speed sensor. The proposed algorithm detects wheel bearing faults by identifying peaks in the wheel speed spectrum at bearing critical frequencies. To implement this approach, we leverage phase domain transform technique to normalize against effect of varying speed. A wavelet filtering technique is employed to enhance the bearing fault signatures prior transforming the wheel speed signal into frequency domain. Post-processing techniques are developed to smooth the high-variance wheel speed spectra and improve the detectability of the wheel bearing faults. A regression model is then developed to calculate the wheel bearing health indicators. The results show that the proposed algorithm achieves high performance in detecting wheel bearing fault.

1 INTRODUCTION

Bearing fault diagnostics and condition monitoring is a well-established field that is mainly focused on industrial applications such as manufacturing, power generation and locomotive transportation. This is due to their high failure rate relative to other components. For example, bearings account for 70% of gearbox failures in wind turbines (Machado de Azevado, Araujo, & Bouchonneau, 2016).

Automotive applications, in contrast, has received less attention as wheel bearings are designed to last the vehicle life, have lower failure rate, and can be easily detected by the vehicle's owner in an event of a failure.

As the industry shifts towards AV and EV, the vehicle lifespan is expected to reach million miles (Motavalli, 2020). Designing a wheel bearing to last million miles becomes unrealistic due to the higher quality material quality required and associated cost. This increases the bearing failure rate within the vehicle life, making a solid case for a wheel bearing fault detection system for EV and AV. This paper demonstrates wheel bearing fault detection capabilities, using only signals that are widely available on modern vehicles.

1.1 Background

1.1.1 Wheel Bearing Failure Modes

Wheel bearings are critical components of a vehicle's chassis system. They are designed to translate the wheel rotational motion into linear vehicle with minimal friction (Lee, 2018) and without noticeable noise or vibration (Sutherlin, 2017).

The most common failure modes in automotive wheel bearings, are contamination ingress and Brinell damage (Sutherlin, 2017). Contamination ingress is caused by seal damage, which exposes the bearing to contaminant and water ingress. This will ultimately lead to the lubricant degradation and corrosion of the bearing.

Bearing Brinelling is caused by a heavy impact load during operating cycle, such as striking a pothole or curb. This leads to permanent indentations, known as Brinell marks, on the bearing outer raceway (fixed to the chassis), the inner raceway (rotates relative to the fixed outer raceway), or the rolling elements themselves. A Brinell mark caused by a curb strike will lead to an increase in vibration and audible noise (Fahrni, G. & Crichton, D., 1999) (Daggupati, Karedla, B., Chavan, C., & Risam, G., 2016).

1.1.2 Bearing Condition Monitoring

Bearing condition monitoring (BCM) is a widely studied field, and many algorithmic detection techniques rely on identifying impulses at the bearing critical frequencies, summarized in Table 1.

Frequency	Formula
Ball Pass Frequency Outer	$\frac{N}{2}\left(1-\frac{B_d}{P_d}\cos(\phi)\right)$
Ball Pass Frequency Inner	$\frac{N}{2}\left(1+\frac{B_d}{P_d}\cos(\phi)\right)$
Ball Spin Frequency	$\frac{P_d}{2B_d}\left(1-\left(\frac{B_d}{P_d}\right)^2\cos(\phi)^2\right)$
Fundamental Train Frequency	$\left(1-\frac{B_d}{P_d}\cos(\phi)\right)$

Table 1. Bearing Critical Frequencies (Randall & Antoni, 2011)

Note that the frequencies are presented in units of samples per rotation of the inner race, not samples per second (Hz), so bearing rotational speed is not a factor. These frequencies are a function of the number of rolling elements, N, and bearing geometric parameters. More details on the geometric variables needed to calculate these critical frequencies can be found in our previous work (Garner, Drame, Du, & Sadjadi, 2021).

Various techniques have been studied to identify bearings faults. Most of the published fault detection methods use accelerometer to record the bearing vibration (McFadden & Smith, 1984),) (Sawalhi, Randall, & Endo, 2007), (Abboud, Elbadaoui, Smith, & Randall, 2018).

The goal of this research work is to adapt techniques from the existing literature to develop a bearing condition monitoring algorithm using wheel speed sensor signals. The remainder of the paper is organized as follows: Section 2 outlines the materials and methods used for fault injection, data collection, and the signal processing algorithm used to detect bearing faults. Section 3 then shares results and discussion from over 1400 vehicle-level tests, including discussion of detection performance.

2 MATERIALS AND METHODS

2.1 Experimental Setup and Fault Injection

The proposed wheel bearing fault detection algorithm was developed and validated using vehicle data with healthy and faulty components. Experimental bearing samples with varying levels of degradation were created by applying the Brinell fault injection procedure described in our previous work (Garner, Drame, Du, & Sadjadi, 2021).

The experimentally fault injected bearings with various fault levels are summarized in Table 2.

Health State	Class	ID	Vibration [X]	Duration [h]	Distance [km]
Healthy	Dev	R1	0.6	50	720
		R2	1.3	34	488
	Val	R3	1.0	29	374
		R4	1.3	29	369
Mild Fault	Dev	R5	3.2	10	144
		R6	3.5	10	139
		R7	4.2	15	213
	Val	R8	3.7	2	32
		R9	4.5	2	29
Severe Fault	Dev	R10	5.1	14	185
	Val	R11	5.9	4	59

Table 2. Experimentally fault injected bearings for algorithm development and validation

The injected bearings are divided into two sets: development and validation, labeled as ‘Dev’ and ‘Val’ respectively. Development bearings were used in algorithm development, whereas validation bearings data were not look at until the algorithm performance assessment phase. The vibration column refers to the bearing ground-truth defined as a relative “vibration multiple”. Given that wheel bearing vibration is function of speed, the speed-average vibration is measured at four speeds: 400, 600, 820, 950 RPM. The area under the vibration vs. speed is then used to define the speed-average vibration. Finally, the vibration ground-truth is provided by the ratio between an individual bearing’s vibration and the average vibration of a healthy bearing, as summarized in Figure 1. More details are provided in our previous work (Garner, Drame, Du, & Sadjadi, 2021). The duration and distance columns on the other hand,

summarize the total amount of test data collected for each bearing.

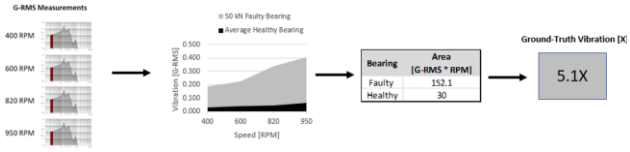


Figure 1. Bearing ground-truth calculation

2.2 Data Collection

In total, 1400 vehicle-level tests, with a variety of healthy and faulty wheel bearings, were conducted to collect data on a MY17 Bolt EV. The main signals of interest included Wheel Speed (WS), Steering Wheel Angle (SWA), Longitudinal Acceleration (AX), Brake Torque (BT), and Axle Torque (AT). These signals were collected with a sampling frequency of 100Hz under various noise factors as shown below:

- Maneuver (lane change, parking lot, cruise)
- Passenger count (0, 3 passengers)
- Road surface (rough, smooth)
- Tire pressure (22, 30, 38, 47psi)
- Tire type (summer, worn, winter)

2.3 Fault Detection Algorithm

In this section the components of the wheel bearing fault detection algorithm is discussed in more detail, including motivating evidence, mathematical formulations, and design considerations. Algorithm performance and results will be discussed in section 3. The end-to-end fault detection algorithm is shown in Figure 2.

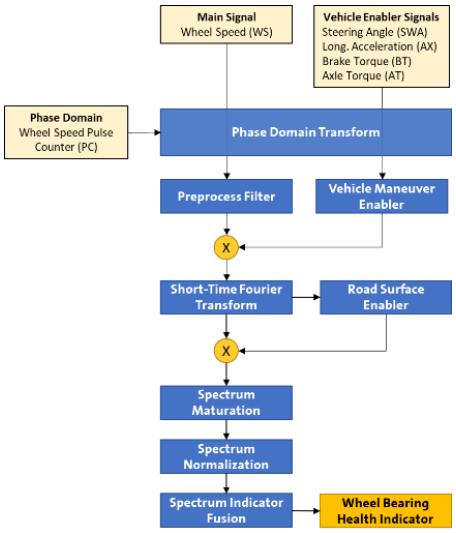
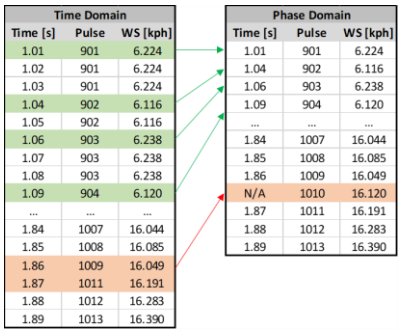


Figure 2. Wheel bearing fault detection algorithm block diagram

2.3.1 Phase Domain Transform

The first step of the wheel speed spectral analysis is normalization for rotational speed by transforming the wheel speed signal to the phase domain. In this domain, each data sample is evenly spaced by rotational phase of the wheel, not by time. Figure 3 shows an example of the phase domain transform above and below the quantization speed, which is 15 kph for the MY17 Bolt EV.



Time Domain		
Time [s]	Pulse	WS [kph]
1.01	901	6.224
1.02	901	6.224
1.03	901	6.224
1.04	902	6.116
1.05	902	6.116
1.06	903	6.238
1.07	903	6.238
1.08	903	6.238
1.09	904	6.120
...	...	...
1.84	1007	16.044
1.85	1008	16.085
1.86	1009	16.049
1.87	1011	16.191
1.88	1012	16.283
1.89	1013	16.390

Phase Domain		
Time [s]	Pulse	WS [kph]
1.01	901	6.224
1.04	902	6.116
1.06	903	6.238
1.09	904	6.120
...	...	...
1.84	1007	16.044
1.85	1008	16.085
1.86	1009	16.049
N/A	1010	16.120
1.87	1011	16.191
1.88	1012	16.283
1.89	1013	16.390

Figure 3. Illustration of the phase domain transformation above and below the WSS quantization speed

When the vehicle is travelling below the WSS quantization speed, the pulse count signal is updated at a slower rate than data is sampled. Therefore, the phase domain transformation requires down sampling to only have one data point per phase.

Above the quantization speed, the pulse count signal is updated at a faster rate than data is sample. In these cases, the calculated wheel speed is interpolated to estimate a value for the skipped pulse.

2.3.2 Preprocess Filter

Transient impulses in the phase-domain wheel speed signal are good indicators of a bearing defect. To amplify these transient impulses, a wavelet filter is applied to the wheel speed signal. This technique has been used in numerous bearing fault detection methods (Kankar, Sharma, & Harsha, 2011) (Guo, Liu, Li, & Wang, 2020).

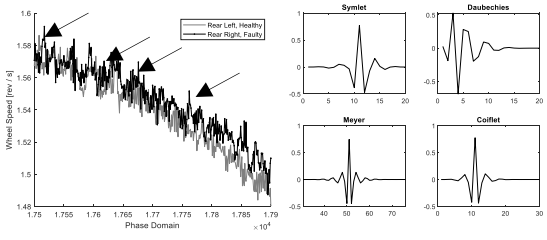


Figure 4. Comparison of phase-domain wheel speed data and wavelet shapes

Figure 4 shows phase-domain wheel speeds from a healthy bearing and a faulty bearing (ID R10). In this example, local peaks in the wheel speed are observed on the faulty bearing. These local peaks occur when a rolling element passes through a damage portion of one of the bearing raceways. More than 30 different wavelet types were investigated. The results show that any wavelet similar in shape to those shown in Figure 4 yields good performance.

2.3.3 Vehicle Maneuver Enabler

To ensure the bearing fault detection algorithm is enabled during normal driving cycle, a set of enabling conditions is defined below:

A data sample i shall be enabled if all the following are true:

$$MinSpeed < WS[i] < MaxSpeed \quad (1)$$

$$BrakeTorque[i] < MaxBrakeTorque \quad (2)$$

$$AxleTorque[i] < MaxAxleTorque \quad (3)$$

$$|SWA[i]| < MaxSteeringAngle \quad (4)$$

At high speeds, wheel speed samples are interpolated to populate the under-sampled phase domain. The maximum speed is set based on Nyquist interval to avoid aliasing. At low speed any speed fluctuations related to a bearing fault are not as detectable. Thus, a minimum speed limit shall be set. During acceleration or braking events, external forces are being applied to the wheel hub assembly. These external forces may influence the fault detection performance, hence the need to set a limit on the maximum brake torque and maximum axle torque. Additionally, artifacts are introduced to the wheel speed spectrum by lateral wheel slip during high steering events. It is then required to set a limit on the maximum steering wheel angle to minimize slippage between the tires and the road surface. The parameters for this enabler are calibratable and tuned using healthy and faulty bearings data, with the goal of identifying the optimal driving maneuvers that result in the best wheel bearing fault detection performance. Table 3 shows the default enabling parameters for MY17 Bolt EV.

Parameter	Default Value
MaxSpeed	15 kph
MinSpeed	8 kph
MaxBrakeTorque	700 Nm
MaxAxleTorque	500 Nm
MaxSteeringAngle	150 degrees

Table 3. MY17 Bolt EV Enabling Parameters

2.3.4 Short-Time Fourier Transform

As summarized in previous section the goal of the proposed algorithm is to identify fault signatures at the bearing critical frequencies. For Brinell dent failures, damage on both the inner and the outer race are expected.

A Short Time Fourier Transform (STFT) with partial enablement is used to implement a real-time fault detection algorithm. This technique requires both a minimum and maximum window length, and computes spectra only for windows of data that are continuously enabled. The Fast Fourier Transform (FFT) will be calculated if the minimum length criteria are met, even if the enablement is interrupted before the buffer is full. The need for full buffers is therefore eliminated, allowing the algorithm to run passively in unconstrained maneuvers and yield a sufficient detections rate.

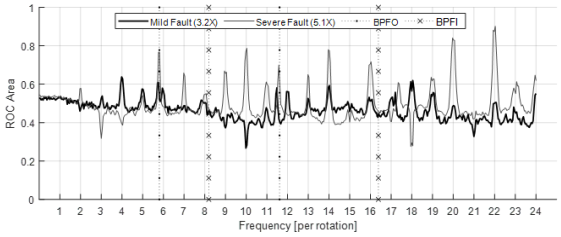


Figure 5. STFT Frequency Detection Performance

A Receiver-Operating Characteristics (ROC) curve is used to assess which frequencies in the wheel speed spectrum yield promising fault detection results. The area under the ROC curve at each frequency in the STFT of the two faulty bearings is shown in Figure 5.At both harmonics of the BPFO, a large spike is observed in the ROC detection area of the severely faulted bearing, and a small peak in ROC detection of the mildly faulted bearing. However, no significant peaks are observed at BPFI. Peaks related to anomalies such as wheel misbalances or shaft misalignments are also observed at integer orders (shown by grey gridlines). These peaks are not related to the bearing state-of-health and are therefore ignored in this analysis. This analysis confirms that BPFO and its harmonics are promising frequencies to consider for bearing failure detection. However, the BPFO ROC area of 0.75 is not sufficient as a classifier and requires further signal post-processing to improve the fault detection performance.

2.3.5 Road Surface Enabler

The interaction of the wheels with irregularities on the road surface causes more random input to the wheel speed signal on rough roads. This variation in road surface will impact the fault detection robustness,

hence the necessity to disable the algorithm when the vehicle is on a rough road. This can be done using a road roughness indicator.

$$RoadRoughness = \sum_{i=i_{low}}^{i_{upper}} DFTWS[i]^2 \quad (5)$$

where $DFTWS$ is the amplitude spectrum of the wheel speed signal, i_{lower} is the index of a lower bandpass frequency in the wheel speed spectrum, and i_{upper} is the index of an upper bandpass frequency in the wheel speed spectrum.

This indicator serves as an additional enabler. If the road input energy for a given wheel speed STFT buffer is too high, that buffer will be rejected from further analysis. This way, results can be limited to only those calculated on smooth road surfaces.

2.3.6 Spectrum Maturation

In development of the wheel bearing fault detection algorithm, significant wheel speed variance impacting the fault detection performance is observed as the vehicle drives. To reduce the effect of wheel speed variance, a low-pass filter is applied to the spectra.

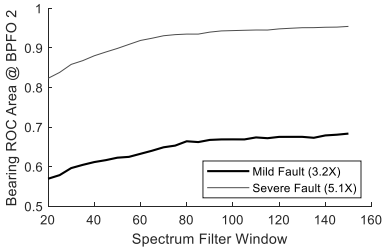


Figure 6. Spectrum Maturation Performance

Figure 6 shows the effect of using a moving mean filter to mature the wheel speed spectra. An improvement in ROC detection area at the 2nd harmonic of the BPFO can be observed. Note that the two bearings shown here are the same as those summarized in Figure 5. An unmatured ROC area of 0.82 is enhanced to 0.94 for the severe fault. For the mild fault, some improvement can be observed, but a maximum ROC area of 0.65 is still very poor, and these mild faults are unlikely to be detected.

2.3.7 Spectrum Normalization

The detectability of wheel bearing faults in the wheel speed spectra is further improved by normalizing the amplitude at the BPFO to the noise floor at surrounding frequencies. This results in an enhancement of the BPFO peaks. The following peak height normalization metric is used.

$$PeakHeight(k) = \frac{max\left(\{DFTWS(m)\}_{m=k-w}^{k+w}\right)}{median\left(\{DFTWS(n)\}_{n=k-W}^{k+W}\right)} \quad (6)$$

This normalization formula above calculates the ratio of the maximum value in a small window (w) versus the median value in a larger window ($W > w$) about some center frequency index k.

2.3.8 Spectrum Indicator Fusion

After calculating the wheel bearing health indicators (peak height and amplitude), the final stage of the wheel bearing fault detection algorithm is to fuse together these health indicators to an estimate of the overall level of vibration. A variety of classification and regression methods were explored, and in order to avoid the risk of overfitting to the exact injected faults, it was decided to select the simplest method with the least number of parameters. The raw and normalized spectra result at different harmonics of the BPFO are fused together using a linear regression model:

$$HI_{final} = C + a_1 * DFTWS(1 * BPFO) + b_1 * PeakHeight(1 * BPFO) + \cdots + a_3 DFTWS(3 * BPFO) + b_3 * PeakHeight(3 * BPFO) \quad (7)$$

Here, C is the regression constant, a is the coefficient for amplitude health indicator at BPFO harmonics, and b is the coefficient for peak height health indicator at BPFO harmonics. The regression coefficients are tuned on vehicle data using classic least-squares regression on experimental data from a variety of bearing health stages, ranging from brand new healthy to severely faulty. The regression coefficients are calibratable parameters that may vary for the front and rear wheels, given their differences in geometry.

3 RESULTS AND DISCUSSION

The results of applying the wheel bearing fault detection algorithm shown in Figure 2 to the full data set are shown below. All development bearings are shown in black, and all validation bearings are shown in lighter grey.

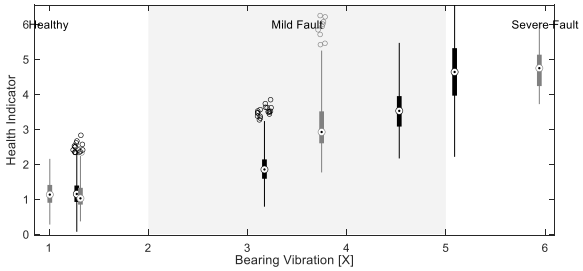


Figure 7. Bearing results

The boxplots in Figure 7 show the outputs of the bearing fault detection algorithm health indicators.

Each box shows the 25% to 75% percentile range of the health indicators calculated for each wheel bearing.

Overall, the algorithm performance on the rear wheels is exceptional. The validation bearings (in lighter grey) show the algorithm adapts well to bearings not used in development. With these outputs, it is possible to choose a threshold such that there are no false positives, 97% true positives for the two severe faults, and an average of 54% true positives for mildly faulted bearings. Figure 8 shows the positive detection rate for each bearing with a threshold of 2.7.

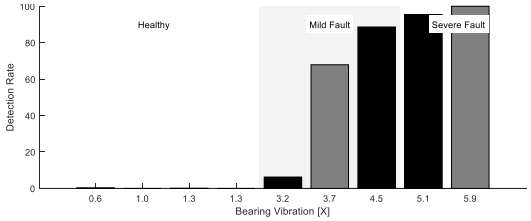


Figure 8. Detection rates with a threshold at HI > 2.7

These positive detection rates give the probability that any one sample from these bearings will be detected as faulty. Note that this rate in monotonically increasing with bearing fault severity, evidence that the algorithm generalizes well and that the health indicator is correlated to the defined ground-truth of the wheel bearing. Therefore, in the cases where a false negative is issued, there is still a chance the next sample will issue a true positive. Overall, the wheel bearing fault detection algorithm meets the performance expectations, as it can detect severe faults with a high true positive rate, and mild faults with a moderate true positive rate.

4 CONCLUSION

This paper presented a novel health indicator calculation and fault detection algorithm for automotive wheel bearings that consumed only signals that are widely available in modern vehicles. It was shown that this algorithm achieves high performance on the test vehicle's wheels.

The performance results of the wheel bearing fault detection algorithm was assessed on a single vehicle by comparing a health indicator to a threshold. In a fleet-management setting, there is added depth by allowing for fleet-wide comparisons of health indicators. Ideally, this algorithm should be implemented such that all vehicles in the fleet can share their calculated health indicators with a central repository. This repository could take advantage of advanced techniques, such as anomaly detection, to improve upon the presented fault detection performance.

ACKNOWLEDGEMENT

The authors would like to thank the following people for their contributions to this project: Zachary Kroeze, Mohammad Hosseini, Milad Jalali, Tooba Sheikh, Alaeddin Bani Milhim, Hamed Kazemi, and Negin Lashkari for their support with test data collection; Xinyu Du and Yilu Zhang for their initial work towards the idea of automated bearing fault detection; and Regan Dixon for her feedback and support for this research work.

5 REFERENCES

Abboud, D., Elbadaoui, M., Smith, W., & Randall, R. (2018). Advanced bearing diagnostics: A comparative study of two. *Elsevier*.

Daggupati, G., Karedla, B., Chavan, C., & Risam, G. (2016). FE Based Steering Bearing Design Optimization for Angular Contact Ball Bearings. *SAE International*.

Fahrni, G., & Crichton, D. (1999). Benefits of Profiling Tapered Roller Bearings - Matching Driveline Component Performance to the Demands of Higher Horsepower Engines of Heavy Duty Trucks. *SAE International*.

Garner, G., Drame, S., Du, X., & Sadjadi, H. (2021). Brinell Fault Injection to Enable Development of a Low-Cost Wheel Bearing Fault Monitoring System for Automobiles. *PHM*.

Guo, J., Liu, X., Li, S., & Wang, Z. (2020). Bearing Intelligent Fault Diagnosis Based on Wavelet Transform and Convolutional Neural Network. *Shock and Vibration*.

Kankar, P. K., Sharma, S. C., & Harsha, S. P. (2011). Rolling element bearing fault diagnosis using wavelet transform. *Neurocomputing, 74*(10), 1638-1645.

Lee, S. (2018). Bearing Life Evaluation for Automotive Wheel. *SAE International*.

Machado de Azevado, H. D., Araujo, A. M., & Bouchonneau, N. (2016). A review of wind turbine bearing condition monitoring: State of the art and challenges. *Renewable and Sustainable Energy Reviews*, 368-379.

McFadden, P. D., & Smith, J. D. (1984). Model for the vibration produced by a single point defect in a rolling element bearing.

Journal of Sound and Vibration, 96, 69-82.

Motavalli, J. (2020, 11 12). *Million-Mile Batteries? They're Coming*. Retrieved from https://www.autoweek.com/news/a34620676/million-mile-batteries-theyre-coming/

Randall, R. B., & Antoni, J. (2011). Rolling element bearing diagnostics - A tutorial. *Mechanical Systems and Signal Processing, 25*, 485-520.

Sawalhi, N., Randall, R. B., & Endo, H. (2007). The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis. *Mechanical Systems and Signal Processing, 21*(6), 2616-2633.

Sutherlin, R. G. (2017). Wheel Bearing Brinelling and a Vehicle Curb Impact DOE to Understand Factors Affecting Bearing Loads. *SAE International*.