1. Report No. RailTEAM VT-5	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Application of Doppler Lidar Sensors for Assessing Track Gauge Widening in Curves and Locations with High-Lateral Forces		5. Report Date October 2023 6. Performing Organization Code:
7. Author(s) Mehdi Ahmadian https://orcid.org/0000-0003-1171-4896 and Ian A. Larson		8. Performing Organization Report No. VT-5
9. Performing Organization Name and Address Virginia Tech Department of Mechanical Engineering 445 Goodwin Hall, 635 Prices Fork Road - MC 0238 Blacksburg, VA 24061		10. Work Unit No. 11. Contract or Grant No. 69A3551747132
12. Sponsoring Agency Name and Address Office of Research, Development and Technology (RD&T) US Department of Transportation 1200 New Jersey Avenue, SE Washington, DC 20590		13. Type of Report and Period 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

The primary purpose of this study is to evaluate the use of Doppler Lidar sensors for assessing track weakening that would indicate early stages of track instability. A series of tests were performed at the Transportation Technology Center's High Tonnage Loop, where two sections of track were "doctored" to have weaker lateral strength, one on a tangent and another one in a curve. Multiple tests were performed at speeds ranging from 10 – 40 mph during which the lateral and vertical deflections of the rail were measured under the weight of the passing wheels of a heavily loaded gondola. The track weakness is created by removing the rail spikes from eight consecutive ties. The measurements from the soft sections were compared with a track section on a tangent length of track that is determined to have nominally sufficient ("good") stiffness.

The measurement system consists of four Doppler Lidar units, two oriented toward the rail gage face to measure lateral rail movement and two directed to the top of the rail to measure vertical rail movement. The combination of the vertical and lateral measurements is used as an indicator of lack of rail stability if larger-than-normal movement of the rail is detected in either direction.

The data collected is analyzed through various methods designed to differentiate sections of track including Gaussian Mixture Model sorting algorithms, inspection via Short Time Fourier Transforms, Discrete Wavelet Transforms, and manual inspection. None of the methods can be done automatically. They each require a different amount of set up and pre-processing before the raw data can be made suitable for the analysis offered by each. The pre-processing can account for dropped data and can be used to identify some false positives such as switches or lubricators.

The test results indicate that the system provides a distinctly different measurement in the sections that are doctored to have less track stability than the section with nominally sufficient stiffness. The detection of the loose track in the tangent sections, however, proved to be less reliable. For those, a mostly ad hoc approach was necessary to match the measured data with video images to pinpoint the exact location of the measurements. It is not clear to what extent such approaches would be feasible in practice

17. Key Words Doppler Lidar sensors, rack weakening, Gaussian Mixture Model sorting algorithms, Short Time Fourier Transforms, Discrete Wavelet Transforms,	18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161. http://www.ntis.gov

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 146	22. Price

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized.**

USDOT Tier 1
University Transportation Center
on Improving Rail Transportation
Infrastructure Sustainability and Durability

Final Report VT-5

APPLICATION OF DOPPLER LIDAR SENSORS FOR ASSESSING TRACK GAUGE WIDENING IN CURVES AND LOCATIONS WITH HIGH-LATERAL FORCES

Mehdi Ahmadian, J. Bernard Jones Chair and Director

and

Ian A. Larson, Graduate Research Assistant

Center for Vehicle Systems and Safety
Railway Technologies Laboratory
Virginia Tech
3103 Commerce Street
Blacksburg, VA 24060

October 2023

Grant Number: 69A3551747132

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

ABSTRACT

The primary purpose of this study is to evaluate the use of Doppler Lidar sensors for assessing track weakening that would indicate early stages of track instability. Such track weakening could lead to gage widening or track buckling due to rail thermal expansion. A series of tests were performed at the Transportation Technology Center's High Tonnage Loop, where two sections of track were "doctored" to have weaker lateral strength, one on a tangent and another one in a curve. Multiple tests were performed at speeds ranging from 10 – 40 mph during which the lateral and vertical deflections of the rail were measured under the weight of the passing wheels of a heavily loaded gondola. The track weakness is created by removing the rail spikes from eight consecutive ties. The measurements from the soft sections were compared with a track section on a tangent length of track that is determined to have nominally sufficient ("good") stiffness.

The measurement system consists of four Doppler Lidar units, two oriented toward the rail gage face to measure lateral rail movement and two directed to the top of the rail to measure vertical rail movement. The combination of the vertical and lateral measurements is used as an indicator of lack of rail stability if larger-than-normal movement of the rail is detected in either direction.

The data collected is analyzed through various methods designed to differentiate sections of track including Gaussian Mixture Model sorting algorithms, inspection via Short Time Fourier Transforms, Discrete Wavelet Transforms, and manual inspection. None of the methods can be done automatically. They each require a different amount of set up and pre-processing before the raw data can be made suitable for the analysis offered by each. The pre-processing can account for dropped data and can be used to identify some false positives such as switches or lubricators.

The test results indicate that the system provides a distinctly different measurement in the sections that are doctored to have less track stability than the section with nominally sufficient stiffness. The detection of the loose track in the tangent sections, however, proved to be less reliable. For those, a mostly ad hoc approach was necessary to match the measured data with video images to pinpoint the exact location of the measurements. It is not clear to what extent such approaches would be feasible in practice

Further evaluations of the test data may be used to shed more light on practical analysis methods—possibly wavelets—that are more automated and less ad hoc. They can also provide alternative system setups or design of experiments for future tests at the Transportation Technology Center (TTC) or on revenue service tracks

CONTENT

LIST OF FIGURES

LIST OF TABLES

Chapter 1 Introduction

The purpose of this study is to explore and validate the potential for a non-contact system to identify sections of laterally weak or damaged track in revenue service. For the past several years, the Center for Vehicle Systems and Safety (CVeSS) has used its Doppler Lidar system to investigate dynamic components of train and track conditions. One of the issues has been the lack of ground-truth validation of the capability of the Lidar system to differentiate loose or damaged track from track in good condition. Previous implementations of the Lidar systems implemented by CVeSS has yielded promising results, enabling assessing railroad track condition in a somewhat automated manner.

Although the revenue-service test results were promising, they did not offer for an independent assessment against ground truth data. Specifically, it has not been possible to independently crosscheck and validate the predictions against track measurements or even track engineers' observation. The main challenge remaining has been the time elapsed between when the measurements are made and when the analysis is completed, and any changes to the track made in the meantime. Another challenge has been the lack of exact spatial correlation between the data and track events. If the data indicates a particular track issue, it is often difficult to pinpoint the location from GPS data, mainly due to inaccuracies in the GPS measurements and their asynchronous relation to the Lidar measurements.

One of the main questions raised by previous tests has been what types of track conditions were being detected by the Doppler Lidar system. While we knew that sections such as switches, junctions, and bridges would contain data indicative of deviations from normal track measurements, we were not able to determine the exact cause or validity of these other locations with sufficient reliability. We knew of several cases that could cause a change in track response, such as poor tie fasteners, a disjointed or cracked rail, or faulty ballast. We did not, however, have confirmation that we knew exactly how these phenomena would present themselves in the Lidar data, and as such which forms of track issues we were identifying with the greatest consistency. If an anomaly was detected, then the question remained as to its cause, and so we did not know if it was a broken rail, poor tie connection, or some other track irregularity.

The tests described in this study are intended to better connect the Lidar data with track events, by possibly determining the various track events that appear in the data through ground truth measurements in a controlled environment.

1.1 Motivation

The motivation of this study is to evaluate the effectiveness of the Doppler Lidar sensing system to evaluate track conditions while connected to a moving train. In particular, this study focuses on track which has a significantly reduced connection to its ties due to loose or removed tie spikes. The reduced connection of the track to the ties would cause a change in the stiffness and damping between the two of them, and as such would produce a greater deviation from the normal movement of the track. This additional movement is dangerous as unintended changes in the track geometry can cause accidents or derailments.

Additionally, the types of disconnection from the tie can influence the ways that the track might respond. Loose vertical or lateral retaining spikes each influence the direction in which the track responds. Beyond the translational effect, these changes could be made even more notable via orbital changes in the track position.

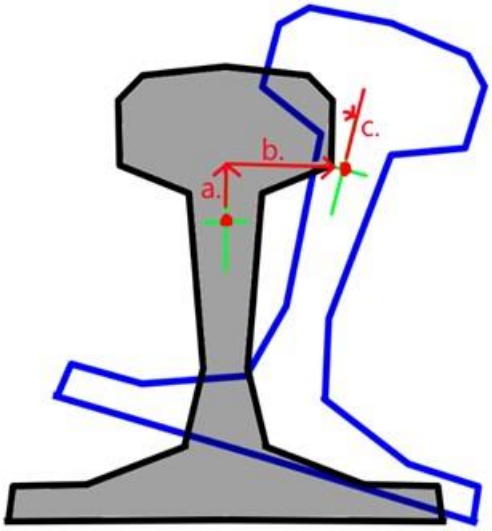


Figure 1-1: Rail motion as defined by: a. vertical translation b. lateral translation c. orbital rotation

As stated, the changes of these track dynamics are of great importance to the railroad and their operators because large deviations in track positioning could pose a risk for causing additional wear-and-tear or even for catastrophic failure in the case of derailment. If the track shifts far enough away from its expected position, the wheel could shove the rail out of the way enough to fall within the width of the track, and hence would result in a derailment. Alternatively, it could shift the orientation of the bogeys, causing the wheels to roll over the outside of the rail. While loose ties themselves may not be solely responsible for this result, they could combine with other track conditions to generate issues for the track.

As such, an evaluation of the Lidar system used to measure these phenomena is necessary to determine to what extent we can measure the track dynamics and prevent excess track motion. Evaluating the speeds and reliability with which the sensor system can measure these conditions at also allows for the appropriate implementation of them for future data collection.

1.2 Objectives

The objectives of this project are to:

1) Evaluate the capabilities of the Doppler Lidar sensors to measure changes in track conditions in both tangent and curved track,
2) Compare the effect of speed on the reliability of the track condition detection,
3) Assess the application of different analytical tools to determine if an automated detection process is feasible, and

4) Provide recommendations for further study into the measurement of track dynamics with Doppler Lidar systems.

1.3 Approach

On-track testing was conducted to evaluate the capabilities of the Doppler Lidar system. These tests were conducted on the 2.7-mile long Hight Tonnage Loop at the Transportation Technology Center. Tests were conducted from between 10 to 40 miles per hour, and recorded both vertical and lateral motion of both ties using the Lidar sensors. Two roughly 15-foot sections of track had their vertical ties removed, one in a tangent, the other in a curve. After the data was collected, it was analyzed through various processes to determine if the sections of loose rail could be identified or automatically detected. From these results, further suggestions were made for the continued evaluation of the Doppler Lidar system.

1.4 Contribution

This study furthers the validation of Doppler Lidar system developed by the Center for Vehicle Systems and Safety (CVeSS) in the use of characterizing railroad track conditions. In particular, this study provides information on the limits of detecting removed vertical tie spikes, but these processes could be applied to other conditions such as lateral track looseness or other damaged track conditions such as cracks. Expanding the range of applications for which Doppler Lidar can be used in allows for greater reliability when monitoring track conditions.

1.5 Report Organization

This document has seven chapters. Chapter 1 introduces the study. Chapter 2 provides background information on the current sensing practices of the railroad industry for at-risk rail conditions, as well as previous implementations of Lidar systems in the use of vehicle safety systems and measuring velocity. It also discusses the technical characteristics of our sensors and information on the track which the experiment was conducted on. Chapter 3 describes the process of developing the test, as well as the steps taken to prepare and conduct the data collection. Chapters 4 through 6 analyze the data collected; first through the use of an unsupervised sorting algorithm, then from a frequency domain analysis, and finally a manual review of the data. Chapter 7 provides the summarized results of the analysis as well as further recommendations for research.

Chapter 2 Background

This background provides information on the track conditions and their evaluations, along with a background in Lidar sensing systems. This chapter starts with a review of track irregularity classifications and the current monitoring methods. It then provides information on the function of Lidar sensors and the specifics of the Lidar system developed for this study. The additional sensors in the track evaluation system are also described. Finally, this chapter includes the reasoning for our choice of track used in this test.

2.1 Track Irregularities

When built, railroad tracks are built to within tolerances specified for their class of rail, but during operation, repeated loading and rocking of the train, along with other impact and frictional forces, can cause damage to accrue on the tracks. Failures from these conditions present themselves as growing rail defects or irregular track geometry. They can also cause changes to the track structure itself. The rail defects often grow from small cracks, holes, and fractures until rail failure occurs. Changes in track geometry conditions alter the orientation or structure of the track out of its prescribed alignment.

The rail defects can be categorized into external and internal defects. These defects are spots from which the damage grows and potentially threatens worse conditions for the rail. The external defects can be found via visual inspection of the rails, but the internal failures need other advanced inspection tooling to identify [1]. Rail defects can be created from many different causes, some of which include irregularities during manufacturing, impact forces, friction forces, and corrosion. All rail defects, if not prevented and allowed to grow, pose a threat to the continued safe function of the track and should be detected and fixed before that point.

The track geometry also has a significant role in the safety of a railroad track. As with rail defects, if the track geometry is allowed to deviate from the prescribed values, it poses a threat to safe use of the track. Some examples of important elements of track geometry are: the alignment of the track, the profile, gage, and cross level.

In this following section, these track alignment conditions are expanded upon.

2.1.1 Alignment

The track alignment is defined as the lateral deviation of the rail gage face from its designed position, or in other terms is a measurement of its uniformity/straightness. It can be measured manually by placing the ends of a string of a set length (in the US, typically around 62 feet) along the side of the rail and measuring the distance from the midpoint of the string to the face of the rail. In a tangent section of track, any distance from zero is a deviation in the alignment of the track.

2.1.2 Profile

Similar to alignment, track profile is defined as the vertical deviation of the rail from its designed position. It can also be measured through the same string method; however, deviations of the track profile indicates either hump or dip deviations for higher and lower deviations from the rail.

2.1.3 Gage

The gage of the track is the distance between the left and right rail as measured from 5/8" below the top of the railhead. Gage can be measured manually via a gage measuring stick. Gage can also refer to the change in distance from left and right rails compared to its prescribed distance, with positive being measured as an increase in distance and negative a decrease. US railroads have nominal gage distance of 56.5".

2.1.4 Cross Level

The cross level of the tracks is the change in the vertical distance between the track heights compared to their set height. A set height difference of greater than zero is referred to as the superelevation of the track. A positive cross-level is when the left rail is comparatively higher than the right rail, while negative is when the right rail is comparatively higher than the left. It is measured from a right angle from the top of a reference rail and is done under load.

While all of these track conditions are relevant to safe track operation, the Doppler Lidar system in this study does not measure these conditions directly. Rather, it is designed to inspect the overall motion of the rails, outliers in which could represent issues with alignment or with defects in the rails themselves. As such, the system might be able to detect problems caused by rail defects or irregular track geometry. Locating sections of the track with irregularities improves the monitoring and inspecting capabilities of railroad inspectors. The next section expands upon the current inspection methods used to detect track irregularities.

2.2 Track Inspection

The railroad tracks are inspected regularly in order to make sure that rail defects don't grow and that the track geometry is maintained within allowable constraints. The frequent and accurate inspection of the rails, ballast, wheels, and cars is important to prevent the building up of damage, which could lead to permanent damage to the structures, or even endanger passengers. As such, many methods are used in conjunction with one another to inspect a large variety of conditions related to railroad and train health.

The derailment of the trains is of great concern, and so care is taken to try to prevent these instances from occurring. Of the primary causes of train derailment, broken rails and/or welds is the largest contributor across all classes of rail. The conditions of improper track geometry, excluding wide gage, is the second most prominent for class 2 rail and above. The second most common cause of derailment for class 1 rail, and the third most common for higher class rail is gage widening [2]. In the case of hazardous material releases, the frequency of mainline accidents caused by rail conditions follows this same pattern of being most frequently caused by broken rails or welds, followed by bad track geometry and then wide gage [3].

To combat these issues, tracks are frequently inspected to prevent accidents. Of the current commonly used practices, manual inspection is the dominant method of inspecting the health of the track. The Federal Railroad Administration requires that all track of class 1 through 3 to be inspected between one to two times a week depending on the frequency and cargo of the train [4]. This is commonly done with the inspector walking the tracks or riding on a high-rail vehicle. This process must be conducted slowly enough (around 5 mph) that the inspector can sufficiently observe the rail and track, and is limited to the observation of only surface defects without the addition of tools to measure the track geometry conditions [1,4].

In addition to the manual visual inspection of the rails, track geometry cars are commonly implemented to inspect the conditions of the track geometry. These cars are capable of measuring the gage, horizontal and vertical alignments, cross-level variance, and twist of the track [5]. These geometry cars can also include inertial measurement and optical systems to measure the car accelerations caused by the track and to include additional visual inspection data [6].

While these methods have proven useful for the detection of issues in both rail geometry as well as track or weld damage, they are slow, and can interfere with the operations planning for many railroad companies, along with being labor-intensive and exposing inspectors to potentially hazardous conditions [1]. As such, new methods are often explored to try to improve the reliability of inspection, or to improve the speed at which inspections can be conducted.

One method of improving the reliability of maintenance and damage detection is through the use of historical data collected by these current methods and generating predictive models. From data collected repeatedly on a track which measures the changing conditions and correlating them to the emergence of damage, models can be generated to improve the planning and prediction of maintenance needs [7,8].

Beyond these inspection methods, several others are being explored or implemented on some rails. These new methods especially focus on automated inspection methods in order to increase the speed and improve maintenance and inspection decisions. These inspection methods include ultrasonic, induction, eddy current, Lidar, and inertial sensing, along with the implementation of computer vision. Some of these methods do not necessarily increase the speed at which the inspection can be conducted, but do allow for more reliable detection of faults within rails, and can detects issues not found via visual inspection or measurements of track geometry.

Ultrasonic inspection of the rails allows for the inspection of the structural conditions of the rails through the generation of an ultrasonic output into the rail head, where damage in the rail head produces a change in the outgoing ultrasonic frequency. This method does not need direct contact with the rail, although some implementations currently being explored induce the ultrasonics through direct contact with custom wheels [9]. The implementation of this method can also be done manually through the use of a sled or cart pushed along the rail by an inspector.

In addition to ultrasonic sensing, eddy current sensing is also able to detect changes in track health through a close-range non-contact sensing. Current eddy sensors generate a magnetic field by summing an alternating current through a coil, which when near the rail produces an opposite secondary field. These sensors measure the change in either the amplitude, phase, or both of these

fields, and through doing so can distinguish sections of track indicative of damage [10]. This process can only sense to a shallow depth into the track as the electric field falls off rapidly as the distance increases from the magnetic field. Ultrasonic sensing and eddy current sensing can be combined when inspecting rails, which works to increase their speed and effectiveness of inspection [11].

In a similar fashion to the eddy current sensor, induction testing works by creating an electromagnetic field across the top of the railroad track. Instead of generating a magnetic field from a coil, the field is generated by placing a pair of electrodes 120 cm from each other on the rail and providing a DC current between the two of them. An induction sensor above the rail then measures the magnetic field, where damage to the rail will present itself as a distortion. This method, however, is limited to around 22 mph [1].

As stated before, inertial sensors and visual inspection modules can be attached to track geometry cars. The inertial sensors are used to monitor the changes in acceleration of the train in sections of the track to analyze and eventually predict when repairs will be necessary. Computer vision also works to improve these sensing characteristics by monitoring the track for any changes in contraction, or by looking for missing or damaged track supports [11,12].

Lidar sensors have also been used in several different applications for railroad and track monitoring. This active sensing technology is able to measure accurately even while moving at high speeds, and to measure the distance and velocity of its target. For the application of our experiment, I will expand on the current implementation of Lidar technology for the purposes of sensing and instrumentation.

2.3 Doppler Lidar technology

Doppler Lidar systems work by generating a specific frequency laser pulse which is directed towards the target. The light which hits the target is scattered backwards and detected by the sensor. Based on the relative speeds of both the sensor and the target, the wavelength of the light emitted is shifted due to the Doppler effect when scattered back. The system measures the time of flight of the pulse as well as the Doppler shift in the returning frequency of the laser to determine the relative velocity of the target.

This shift occurs due to the change in apparent frequency caused by the differences in the velocity of the sensor and target. With respect to the target of the sensor, the apparent shift in the frequency of light emitted can be expressed as

$$f_1 = f_0\left(1 + \frac{\Delta v_1}{c}\right) \tag{2-1}$$

where the Δv_1 is the difference between the velocity of the target and the sensor as expressed by $\Delta v_1 = v_s - v_t$ with v_s being in-line component of the velocity of the sensor and v_t being the in-line component of the velocity of the target. c is defined as the speed of light in a vacuum. We also assume that for practical applications, the velocity of the sensor and the target are significantly smaller than c. The light is then scattered by hitting the target, and is then detected by the sensor.

This new apparent frequency at the sensor with respect to the scattered frequency can be expressed by

$$f_2 = f_1\left(1 - \frac{\Delta v_2}{c}\right) \tag{2-2}$$

With the change in velocities having the opposite sign since the direction of the light has been reversed and Δv_2 representing the negative of the difference between the two velocities. Together, these two equations create the function

$$f_2 = f_0\left(1 + \frac{\Delta v_1 - \Delta v_2}{c}\right) \tag{2-3}$$

And given that $\Delta v_1 = -\Delta v_2$, with the Doppler shift defined as $f_D = |f_2 - f_0|$, the representation of a Doppler shifted frequency becomes

$$f_D = 2f_0\frac{\Delta v_1}{c} \tag{2-4}$$

From (2-4) it is clear that the returning frequency is only a function of the emitted frequency and the relative velocity between the sensor and the target. The frequency of the emitted laser is known, as it can be directly calculated from the wavelength of the laser. As such, the relative velocity can be calculated by measuring the shifted frequency of the returning laser.

Since the relative velocity is limited to the vector direction between the two bodies, the only speed measured is that in the radial direction from the Lidar sensor. For example, if a stationary sensor was looking at a moving object at angle θ, then the velocity component in the Doppler shift equation would be $\Delta v_1 = V\cos\theta$ where V is the magnitude of the velocity of the target. The effect of the angle between the movement of the object and the sensor needs to be included in the calculation if the sensor cannot be fully aligned with the direction of the motion of interest.

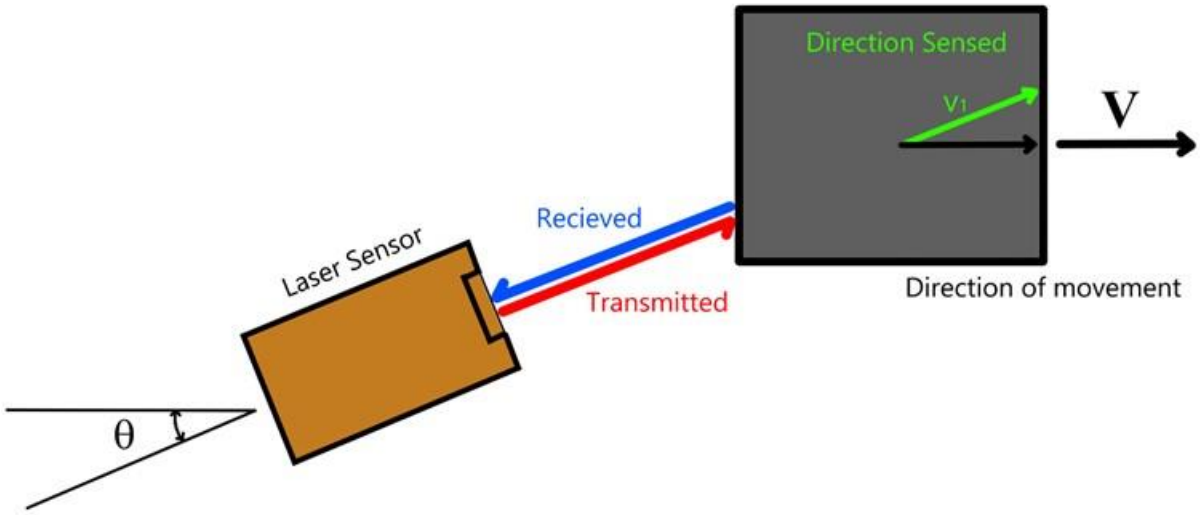


Figure 2-1: When an object is not moving directly in line with the path of the laser, the component of the velocity measured is a function of the angle of difference.

The same effect is true if the sensor itself was moving rather than the target, since the velocities of interest are relative in nature. In our application, the sensor is mounted to the train which acts as a moving platform but the equation holds true whether the target is moving or the sensor is moving.

2.4 Virginia Tech's Lidar System

The Lidar sensor system developed by CVeSS uses these same Doppler Lidar principles when measuring the velocity of the rails. This allows for the fast and accurate non-contact sensing of the relative velocity between the rail and the train. The sensors attach to the underside of the train and have two pairs of laser beams which are aimed at the rails. One pair of lasers are aimed at the inside gage face of the left and right rails while the other pair is aimed at the top of each rail. As seen in Figure 2-2, the solid red arrows represent the outgoing beam and the dashed blue arrows show the reflected beam which is used to measure the Doppler shift.

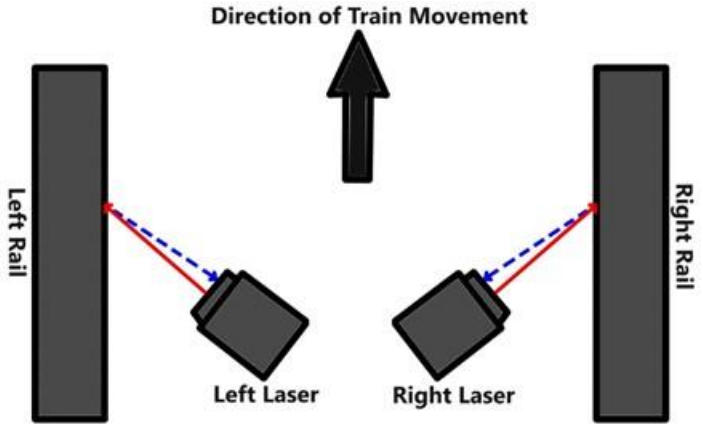


Figure 2-2: Virginia Tech's Lidar system design for measuring the relative speed between the rails and the sensors

The Lidar system transmits the beam through a pair of telecommunication optical fibers. These fibers guide the laser to the lenses, which both emit the laser at the track head, and also receive the returning scattered laser. The returning light data is then passed through a High-Fidelity RF device and then processed by a National Instruments PXI central processing unit. This unit generates a Fast Fourier Transform (FFT) of the data, which is used to determine the velocity of the object based on the Doppler shift in the frequency of the data. From the transformed data, the PXI calculates the velocities of each rail and the data is displayed on a monitor and saved to the hard drive. The saved data is later post-processed for further investigation of the motions of the rails to determine locations of laterally weak rail.

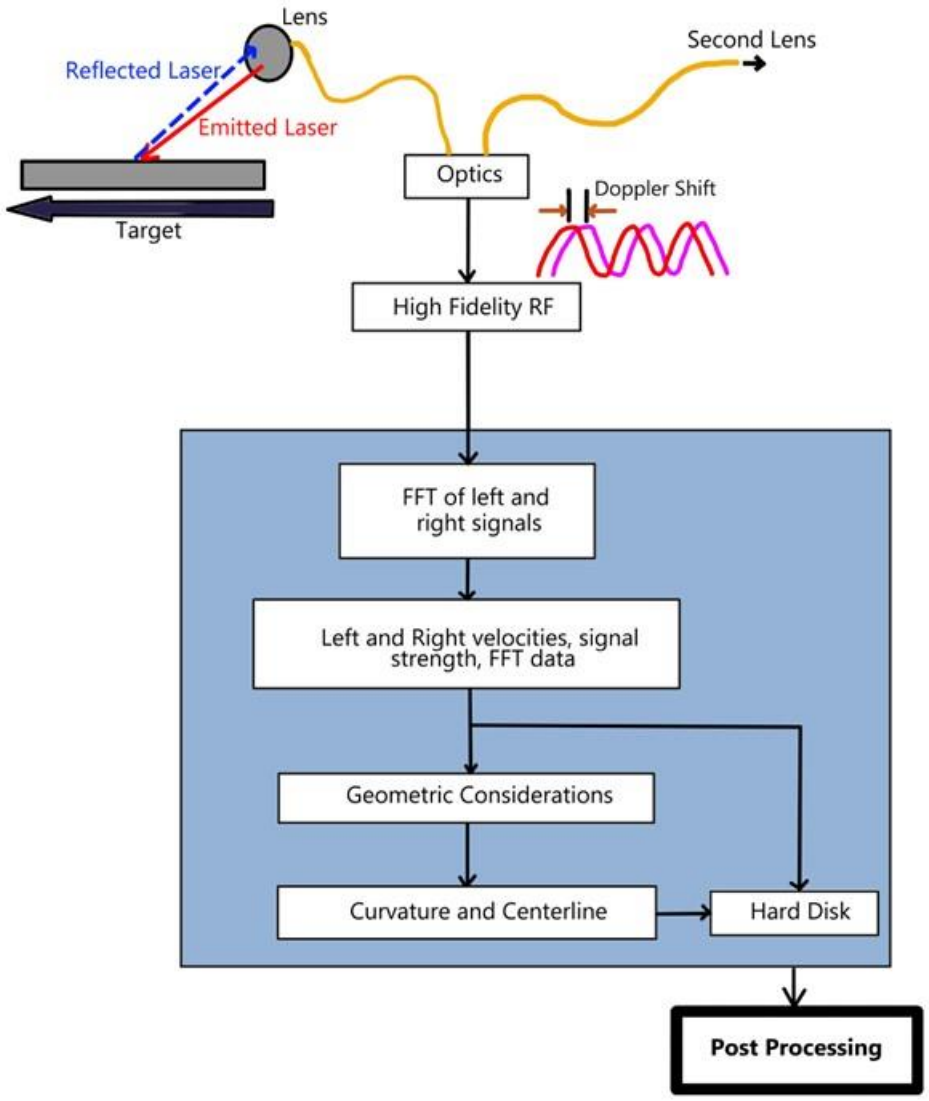


Figure 2-3: Block diagram of the Lidar system processing and storage for non-contact sensing

The laser is a homodyne instrument, which will return only the absolute value of the speed it measures. As such, if the changes in speed measured of an object is about zero, the results would not differentiate between positive and negative velocity. As such, an offset is needed, and in this case, the forward speed of the train allows for this addition. A nominal forward angle of 10 degrees is applied to the sensors aimed at the rails, which incorporates a slow changing longitudinal velocity offset to the data. With a known angle, the lateral and vertical velocity components can be calculated, and a high pass filter is applied to remove the longitudinal speed of the train.

As stated before, two pairs of Lidars were implemented as part of the system, with one pair sensing the lateral motion of the track and aimed at the inside gage face of the rails, and the other sensing the vertical motion of the track and aimed at the top of the rails. The vertical motion inspection system was nicknamed the Corsair system, and so information recorded by it is referred to as

Corsair data. The lateral system's nickname was Blackbeard, and as such its data is Blackbeard data. For both of these systems, a channel 0 and a channel 1 were designated. The channel 0 signal refers to the left rail and will typically be colored in green, while the channel 1 signal refers to the right rail and will typically be colored in red. Hence, we had four Lidar lenses implemented to record the conditions of the track: Blackbeard Channel 0 and Channel 1, and Corsair Channel 0 and Channel 1.

2.5 Lidar in industry

Several industries have already implemented Lidar technologies in widespread use. Autonomous vehicles have been taking advantage of the capabilities for Lidars to accurately measure distances rapidly for their application in sensing surrounding objects. Rather than just aiming at a single target, the Lidar, or battery of Lidars, scan the surrounding area and return the distances of objects around the car [13,14]. These methods use a time-of-flight implementation of the Lidar sensors, however other applications in autonomous driving use Doppler Lidar to sense for pedestrians through a combination of both time of flight and distance sensing [15].

Doppler Lidars in specific find widespread use in the measurement of atmospheric conditions. The Doppler Lidar is able measure small particles to indicate information regarding the growing of rain droplets [16] or wind and dust effects [17]. The application of this technology has proven useful in these fields.

Even within the field of railroad track inspection, Lidars are currently being explored for their usefulness in detecting faults or helping to predict and manage maintenance. Time of flight sensors, which implement a point cloud system similar to that found in autonomous vehicles can be used to extract information from aerial vehicles for the purpose of identifying track [18].

In addition to this point cloud implementation used for track identification, the Doppler Lidar system of measuring the velocity of the rail has been tested for use in track analysis. The application of Doppler Lidar has been used for the measurement of track curvature from the differences in velocities between the left and the right rail, as well as the distinguishing track irregularities [19].

This last type of Lidar application is the one which this experiment aims to expand on. The analysis of rail conditions through Doppler Lidar has been promising, however the extent and full capabilities of the technology have not been completely explored. In this study, Doppler Lidar sensors are implemented in an attempt to monitor for conditions that could be indicative of laterally weak rail, which could pose issues for gage widening or other track safety conditions.

2.6 Additional Sensors

To support the CVeSS Lidar system, a surrounding series of other sensors was included, including an accelerometer and a pair of GPS receivers. We also included a manual signal switch attached to the data acquisition system to allow for the manual identification of locations of note throughout the track by an operator. To review for any issues found during testing, a pair of GoPro™ Hero 8s were mounted to the underside of the train and filming the location of the rails which were being

measured by the Lidars. On the side of the track, fish scales were implemented on the ties to measure the maximum lateral displacement of the track head during testing.

The accelerometer was constructed with an ADXL327 triaxial accelerometer. The accelerometer has a full range of +/- 2 g and was configured to sense at 50 Hz. The GPS units were Garmin RS 232 marine sensors. The receiver was a GPS 19x HVS, which had up to a 10 Hz update rate. They also could track GLONASS, GALILEO, and QZSS. The receivers were housed in IPX7 waterproof housing pole mounts. All of these sensors were connected through a Measurement Computing USB-1608FS which is a 16-bit 8 channel A/D converter. The A/D converter was controlled through a Lenovo ThinkPad PC and was set to collect at a rate to match the Lidars with a sample rate of 33 Hz.

While the GPS receivers had up to a 10 Hz update rate, the actual rate was a function of the computation necessary to determine its location. As such, it did not have a consistent sample rate, nor was the location measured consistent. Through testing, there also was clearly a delay in the update which caused the GPS location to lag behind the rest of the data.

Figure 2-4: The pair of Garmin RS 232 Marine Sensors connected to the top of the freight car

The fish scales measure the maximum field side lateral deflection of the head of the rail during a test. The sensors are rods positioned laterally at the same height as the rail head, with the end of the rod placed against the side of the track. After the train has passed, the change of the rod's displacement indicates the maximum field side deflection the rail experienced as the train passed that location. This can be measured and reset between each trial.

Figure 2-5: A fish scale connected to field side gage face of the rail.

2.7 Track Classes

In addition to the sensors needed to validate the Doppler Lidar system, the track is important to the test. Railroad track can be classified into different classes of rail, which determines the speed which the train can operate at as well as the limits for the required parameters of track conditions. The classes as defined by the Track Safety Standards are presented as class 1 through class 5 track, with maximum allowable speeds set for both freight trains and passenger trains.

Table 2-1: Maximum Allowable Train Speeds based on Track Classification

Track Classification	Maximum operating speed for freight trains (MPH)	Maximum operating speed for passenger trains (MPH)
Excepted	10	NA
Class 1	10	15
Class 2	25	30
Class 3	40	60
Class 4	60	80
Class 5	80	90

In addition to the maximum allowable speeds, the allowable deviated conditions of the track are also based on the class of rail. For example, the alignment of class 1 tangent track is allowed to be up to five inches offset from the mid-point of a 62-foot line, while class 5 track is allowed no more than a ¾" deviation from the mid-point of a string of the same length. If a track fails to meet one or more of its requirements, it is considered the next lowest class of track.

The different classes also tend to be used for different purposes. Class 1 rail is common for yard and short line travel, and class 2 is common for branch lines and secondary main lines. The most common class of track for passenger and long-haul freight is class 4 rail.

2.8 Track Selection

The Transport Technology Center (TTC) was selected as the location for our testing due to both our already established contact with Transportation Technology Center, Inc. (TTCI), now known as MxV Rail, as well as their extensive capabilities for trackside testing. The center has 10 tracks on ground, including the Railroad Test Track (RTT), Wheel/Rail Mechanism Loop (WRM), and the High Tonnage Loop (HTL). These tracks all form continuous loops, which allows us to easily conduct several tests in succession without needing to reverse directions.

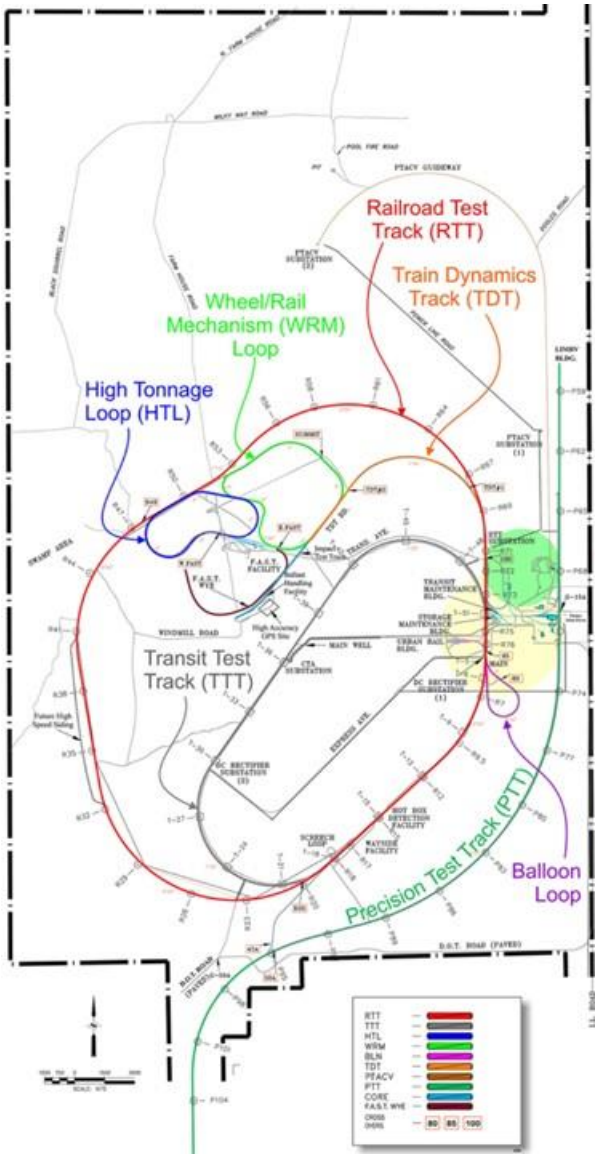


Figure 2-6: The as-built plans for the Transportation Technology Center testing tracks

The TTC also has a large selection of locomotives, passenger cars, and freight cars. Because the site is designed for testing track conditions already, they match our needs exceptionally well, and can provide us with the resources to test. Specifically, they could provide us with a closed off track, the cars necessary to create a consists containing a locomotive, passenger car, and freight car, as well as the ability to change some of the conditions of the track. They also had the necessary

staff both to coordinate and to conduct the experiments. The additional people on the track allows for the addition of track-side measurements to validate the deflection of the rail at our points of interest, and to compare those results with the information gathered by our Doppler Lidar system. Working with the staff, we were able to set up a date for our experiment to be conducted over a week-long period during the summer.

The selected track on which we were to conduct our experiment was the High Tonnage Loop. The HTL is a short track, with a length of about 2.7 miles, including two tangent sections, three 5-degree curves, and one 6-degree curve. This track is Class 4 rail with an operating speed of up to 40 mph. The loop is designed for heavy and overloaded trains, which allowed us to deflect the track by the greatest amount due to the weight of the train.

The track was also selected since it would potentially have additional "weak" sections that could be used for further Doppler Lidar characterization given its high tonnage use. The track accrues 1 million gross tons a day for standardized testing and 140 million gross tons a year of freight testing wear. The track also has other common elements such as a bridge, concrete ties, lubricators, switches, etc. which would provide an environment typical of natural track. These elements are useful for seeing if the Doppler Lidar system can characterize or distinguish between them and irregular track.

Figure 2-7: Additional elements found on the HTL such as switches, bridges, concrete ties, and a lubricator

The track also contains both tangent and curved track, which allows us to test the capabilities of the Doppler Lidar system in detecting laterally weak rail in both tangent and curved track. In one of the tangent sections, section 40, a set of eight ties was selected as the control section as the on-site workers determined them to be typical of the condition of the track. In this same tangent section, a different set of eight ties had their vertical tie spikes completely removed to emulate an

extremely loose vertical connection to the ties. In one of the curved sections, section 25, we also selected eight ties to be decoupled vertically from the track. Section 25 was a 6-degree curve with about 3 degrees of superelevation

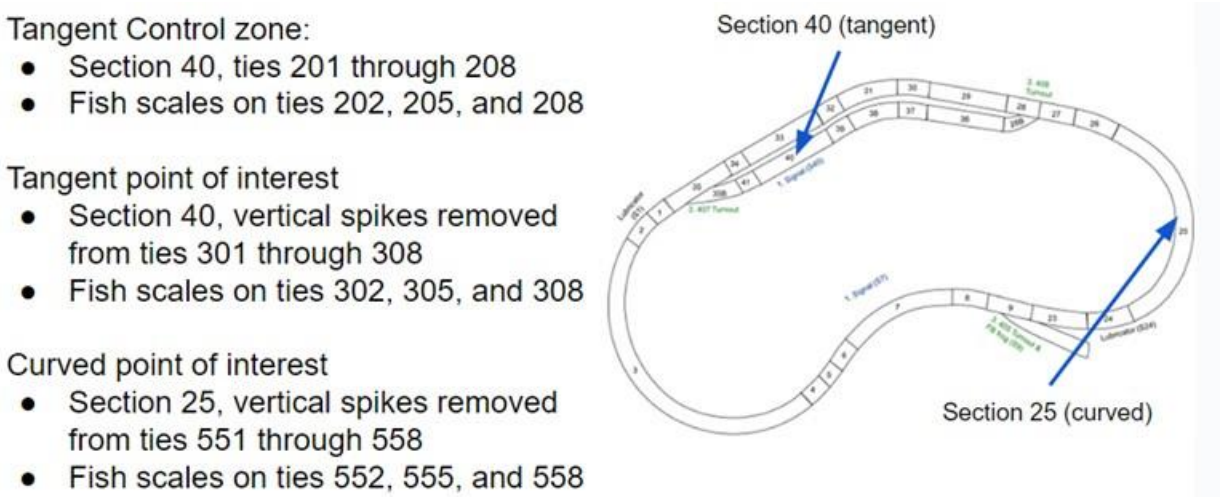


Figure 2-8: Locations of control and altered track on the HTL

While we were not able to know the models of locomotive, passenger car, or freight cars ahead of time, we knew that we would have a consist containing one of each. We planned to have the freight car loaded as heavily as reasonably possible in order to excite the rail to the greatest extent and hence to produce the strongest signal. The freight cars used on the HTL regularly have a weight of 315,000 pounds. With this large weight causing a significant deflection, we also planned to connect our Lidar sensors, GPS units, and accelerometer to the freight car for testing.

Chapter 3 Test Preparation

This chapter reviews the process of preparing for and collecting the data to validate the Doppler Lidar system. For the data collection session in Colorado, we knew we were going to run our tests on the HTL with a consist containing a locomotive, passenger car, and freight car. While preparing for our tests, however, we did not know what models of each we would have access to. Due to the scheduling issues related to MxV moving to a new location, we would be shown what was available to use on the first day of our testing. To prepare for this, in addition to our Lidar testing practice and preparation, we designed several possible mounting options to account for a variety of different mounting conditions that could be presented. To start with, we practiced using the Lidar sensors to understand their capabilities and limitations.

3.1 Lidar Practice

Our first key step in the pre-test preparation process was getting familiar with the mechanics of the Lidar sensors. After being provided a description of their operation, we first mounted the lasers to a desk in the lab and aimed them at a rotating disk which provided a known and consistent rotational velocity for us to measure. We practiced the steps of turning on, collecting, and exporting the data collected by the Lidars. From there, using the same setup, we moved the lenses closer and farther from the disk to observe what the data would look like if the target was too close or too far from our focal distance. In either case, the result was that the signal would drop to zero if the target was outside of the lens's focal range, and that there was inconsistency in the signal near the edges. In the same experiment, we also observed the changes in the data as the disk was rotated from completely perpendicular to completely parallel to the lens. At either extreme, no signal was detected. Once the disk was angled between either of these extremes, the signal returned. When the disk was close to perpendicular, the signal indicated a slow speed, while at an extreme angle, the velocity was measured to be much greater. This reflects the Doppler Lidars measure the velocity component directed in line with the target itself.

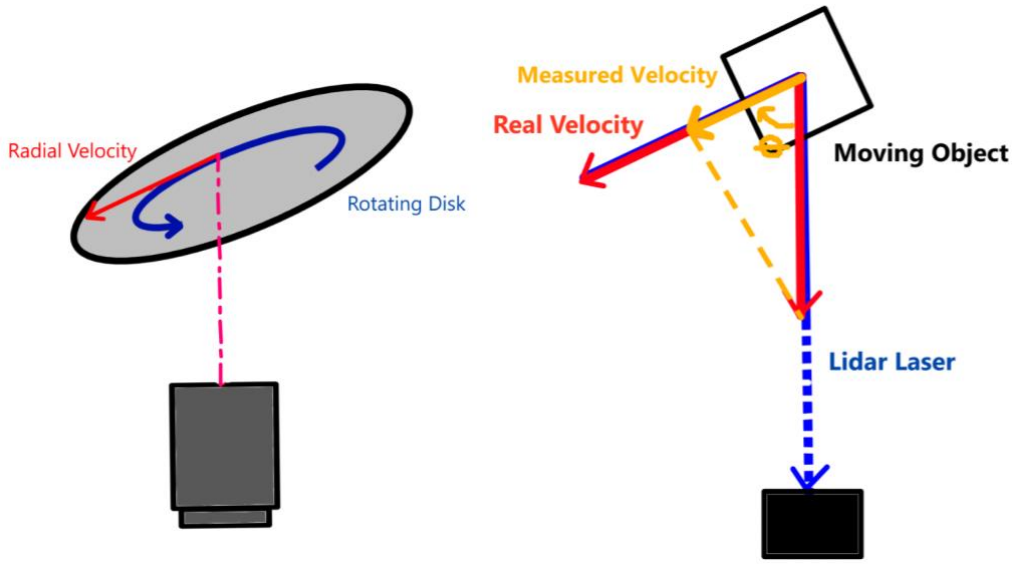
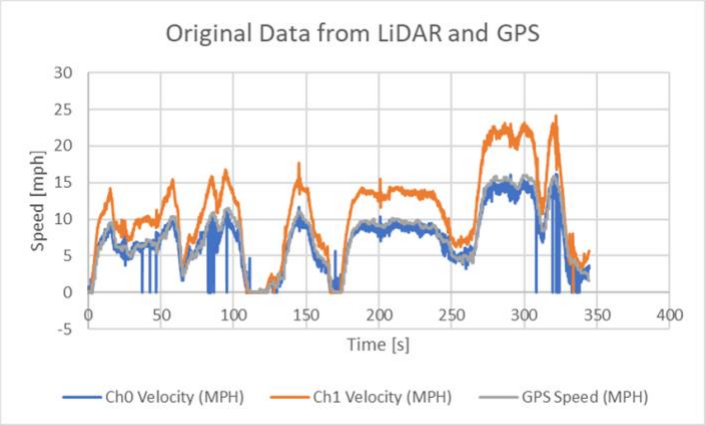
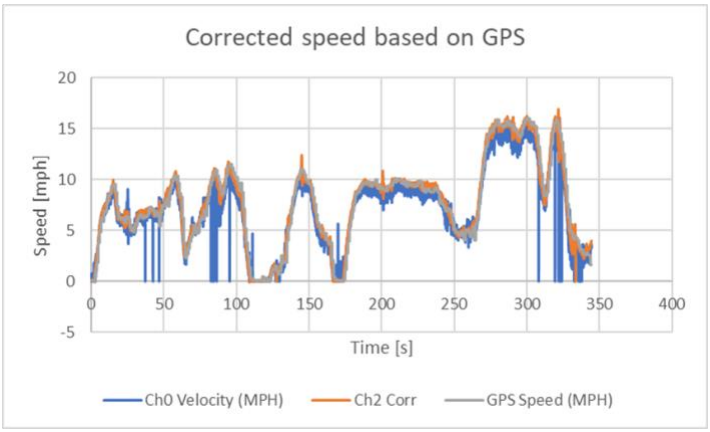


Figure 3-1: Lidar measuring the velocity of a rotating disk (left); Lidar measuring the radial component of velocity relative to the lens (right)

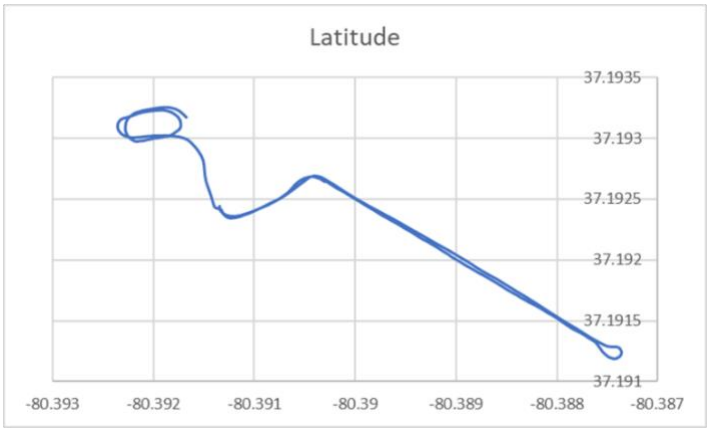
From there, to familiarize ourselves with the data collection process while attached to a vehicle and as an initial practice with mounting, we connected the Lidar sensors to the lab truck and drove the truck in the parking lot of the laboratory, recording the velocity of the ground beneath the truck. This improved our practice with identifying curves in the data as well as observing the impact of vibrations or slow speeds on the data. From driving in the lot around the lab, we were able to correctly determine the velocity of the truck through the Lidar data. We also found that at speeds at around 2 mph or less, the returned signal lost clarity and had an increase in outlying speed measurements. We then mounted the lenses to the truck a second and third time, with a laser angle which was too shallow and then too deep. During these tests, we also mounted the GPS receivers to the roof of the truck, which we used to normalize the observed speeds of the test to compare between each other. From these results, we were able to see the extent to which the angles of the sensors affected the precision of the results. Since the Lidar software accounts for an assumed 10-degree angle to the surface, the shallower and steeper angles each needed to be scaled larger or smaller to represent the correct speeds. At an angle close to perpendicular with the ground surface, the results were precise but were less able to differentiate between upwards and downwards movement of the road due to the homodyne laser. The near parallel mounting to the surface of the road required greater scaling to account for the angle and had less precision in the speed. Figure 3-2 shows an example set of data using the GPS to normalize Doppler Lidar data. The channel 1 lens was mounted with a shallower angle to the surface of the road.



a.



b.



c.

Figure 3-2: Results of initial truck mounted Lidar testing with the (a) raw results, the (b) corrected results, and (c) the GPS data for basic road testing.

We also found that the GPS signal lagged behind the location of the car and that there were some sections not representative of the motions actually taken during our tests. The cause of this issue, we found, was an insufficient connection of the GPS receivers to the satellites. During the trials with worse representations of the path of travel, we found that there were only around two to three satellites connected to the system. When we drove into a more open location, and were connected to six to nine satellites, the representation of the path of travel was significantly more accurate. With a better understanding of the Lidars and their capabilities, we needed to prepare a method to mount them to the train for testing.

3.2 Sensor Mounting

Having practiced using the Lidars and GPS receivers, the next step of preparation was to design a mounting system for the train in the Colorado test. We knew that we needed to determine the location of the connection for the lenses. We reviewed the images of previous mounting

configurations and tests conducted with this Lidar sensor system. Since we still did not know exactly what kind of locomotives or cars we had available to us, we considered both locomotive and freight car mounting configurations. As such, we considered each of the challenges that either style would present to us.

For mounting the Lidar sensors to a locomotive, the undercarriage of the train would provide a challenge, especially with regards to the lateral laser setup due to the large amount of fuel and machinery underneath. The undersides of locomotives have little space beneath them, and tend to not have easy access to locations to fix lenses to. Previous tests that mounted to a locomotive solved this by finding a section beneath the stairs at the front of the train which allowed for enough of an angle to spot the inside of the opposite rail. These types of mounting also tended to connect via the stairs, either in the front of the locomotive or off of the rear. Of course, with all configurations, the lasers needed to be positioned near the wheels, and so the exact model of locomotive would impact our choices greatly. In addition to the physical capability of mounting to the train, the issue of the lasers falling off the top of the track while rounding corners presented itself as a consideration. The laser's distance from the wheels and the length of the locomotive were the two main factors in determining the degree of curvature which would cause the laser to fall off the track. Hence, even though we were not sure whether the locomotive would be a viable mounting platform, we purchased mounting options such as c-clamps and rail clamps to allow us to mount to the train staircases if the opportunity presented itself.

Figure 3-3: Staircase mounting locations on a locomotive

In the case of needing to mount to a freight car, we also reviewed our previous mounting configurations from prior years' tests. In contrast to some of the previous locomotive mounting setups, which used clamps and optical posts, the previous freight car Lidar setup used a purpose-built frame which clamped to either side of the car. This frame, while a great and stable platform to mount the Lidars onto, was unfeasible as a choice for our purposes, as it would require specific knowledge of the dimensions of whichever freight car we would have access to on the day of the test. With a purpose-built frame out of the question, we considered what other mounting options we could produce. We considered if the prior tools such as the C-clamps that were used to mount to the locomotive would have issues finding an appropriate surface to clamp onto.

Figure 3-4: Purpose built frame from the 2021 testing attached to the NS 38 track geometry car

There were several different types of freight cars that could have been provided to us, and each of them have their own potential issues when trying to mount onto them. For example, different freight cars have different distances between their wheels and the front/rear of the car, and some of those cars will have stairs and others will not. As such, the option to mount to stairs on freight cars, like with locomotives, was not guaranteed. In addition, if the freight car contains hoppers, they could make mounting to the undercarriage a much more difficult prospect, as the hoppers would not provide anything protruding to clamp onto and cause the underside of the car to have less space since they are lower down. The height of the underside of the freight car could also provide its own challenges. High undercarriages could make lowering the lenses to an appropriate height while still not inducing interfering vibrations due to long rods much more difficult.

Figure 3-5: Hopper freight car (top) with lack of easy connection points and a boxcar (bottom) with open space beneath

We considered and tested several other mounting options, one of which was the option to use heavy duty magnets to connect to the train cars. While we found that most of the body of the train cars would be made out of aluminum, the frames of the cars tended to be made of steel, which would provide a good surface for mounting the magnets to. We also already had a couple of Magswitch 400's and Magswitch 600's in the lab, which had a hold force of 400 and 600 lbs. respectively. We considered whether one magnet would be appropriate or not, and ultimately created modular brackets that could combine two Magswitch 600 magnets. These would also allow for the connection of an 80/20 structure or optical posts to either a single magnet or a pair of magnets.

We also practiced mounting these magnets, but since we did not have a freight car frame to practice with, we instead used the metal frameworks of our workbenches as practice. From this, we also found that any optical posts which were greater than a foot in length would contain too much vibration for a consistent and accurate signal. The 80/20 beam mounting provided a much more stable platform to mount onto for longer reaching setups. We then returned to the truck and mounted the lenses via the magnets for additional mounting practice with both the Lidars and the magnetic platforms.

Figure 3-6: Magnetic Lidar test mounting on the lab car frame

From this, we had two main mounting platforms that we had practiced with for the attachment of the Lidar sensors; the clamp-style attachments and the magnet mounts. With these, we planned to be able to modify our configuration to match whatever model of either locomotive or freight car we were provided for testing. A large number of optical posts and 80/20 supplies were set aside for mounting, including swiveling optical post connectors which would allow for compound angles without chaining several small posts together with 90-degree connectors. These posts were used in previous tests to create the angles and positions of the lenses necessary to align with the track, and were ½" in diameter.

3.3 Track Setup

In order to verify the capability of our Lidar system to detect sections of laterally weak rail while in service, we decided to gather data on a track where we had sections purposefully altered to have reduced connections to their ties. To do this, the experiment was set up at the Transportation Technology Center on their High Tonnage Loop. The loop allowed us to conduct our tests at 10, 20, 30, and 40 mph with a heavily loaded freight car.

With the help of the MxV employees who worked on the HTL, we selected three parts of the track to be our points of interest. Two of these sections had their vertical tie spikes removed to reduce their stability, and the third was selected as a reference to compare to. One of the sections with removed ties was placed in a tangent section of track in section 40, with the vertical tie spikes

removed from ties 301 to 308. The other loose section was placed in the curved section 25. The vertical spikes were removed from ties 551 to 558. The reference section was also on the tangent track in section 40, specifically the ties 201 to 208. This would allow us to observe the impacts of reduced track stability both in tangent track and curved track, which are the two types of track which we are most interested in confirming our efficacy on. The reference section was selected by the MxV employees for its representative stiffness of the track. This selection should hopefully provide a more distinct difference between the control and emulated loose track than if the two sections were chosen from different locations on the track.

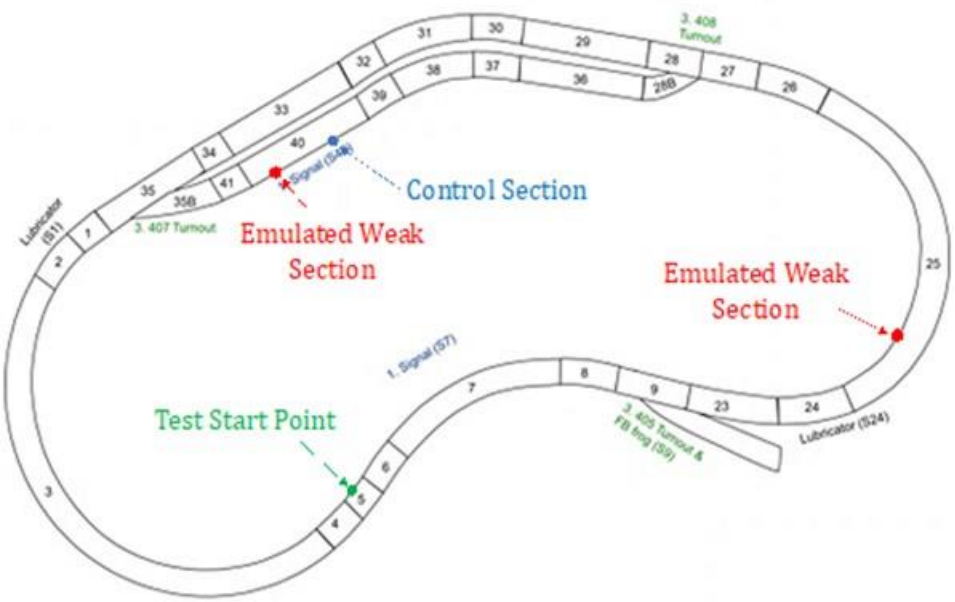


Figure 3-7: Labeled HTL track sections

The ties are spaced longitudinally apart from one another by about 20-21", which makes each section of interest roughly 14 to 15 feet long. Additionally, in order to measure the maximum lateral displacement of the head of the rail at the points of interest, fish scales were mounted on the field side of the rails. Fish scales measure the maximum lateral displacement of the head of the rail as the train passes that section of track. The locations of these removed ties and fish scales are as follows.

Table 3-1: Locations of interest and their ties

Zone	Range of Ties	Removed Tie Spikes	Fish scales
Control in Section 40	Ties 201 to 208	None	Ties 202, 205, 208
Tangent Test in Section 40	Ties 301 to 308	All vertical spikes	Ties 302, 305, 308
Curve Test in Section 25	Ties 551 to 558	All vertical spikes	Ties 552, 555, 558

The consecutive sets of "loose" ties should provide enough instability to be detectable in our data. With too short of a section, the train might not even excite the rail in a way that is different from its fully supported condition. Too long of a section, and the structure of the track itself might be at risk, and so for safety we limited the length of the removed spikes to within 15 feet. Once both the Lidar mounting options and track conditions were prepared, we went to the HTL and began preparing for our data collection.

3.4 Lidar Configuration

For our testing at the Transportation Technology Center, we needed to prepare all of the sensors, equipment, and mounting methods to allow us to collect the data we needed. This included the two paired sets of Lidars, the pair of GPS sensors, the accelerometer, as well as the fish scales track-side. Once we were on site, we began to prepare the specifics for the experimental setup.

Given that the Lidar sensors measure the radial velocity component with respect to the sensor, we implemented both a vertical and lateral Lidar sensor to measure the vertical and lateral motions of the rails. Each sensor was mounted behind the front wheel set of the freight car on the train consist. The Lidars were mounted to the front of the loaded freight car for two main reasons. The first was that the higher weight of the freight car would cause the rail to experience a large loading force and hopefully deflect the rail. The second was that the Lidar sensor needed to be close to the point of contact so that it can measure the deflection before the track settles back into its original position. The front of the freight car was also the closest to the passenger measurement car, and therefore offered the easiest wiring between the two. They were mounted in the wheel well of the freight car since the frame of the car had two flanges that allowed for the magnets to mount to them on each side, as well as supports that were positioned almost above the rails. These offered easy mounting locations for the clamps holding the vertical Lidars to measure the vertical component of the rails. In Figure 3-8, the mounting configuration of the Lidars can be seen.

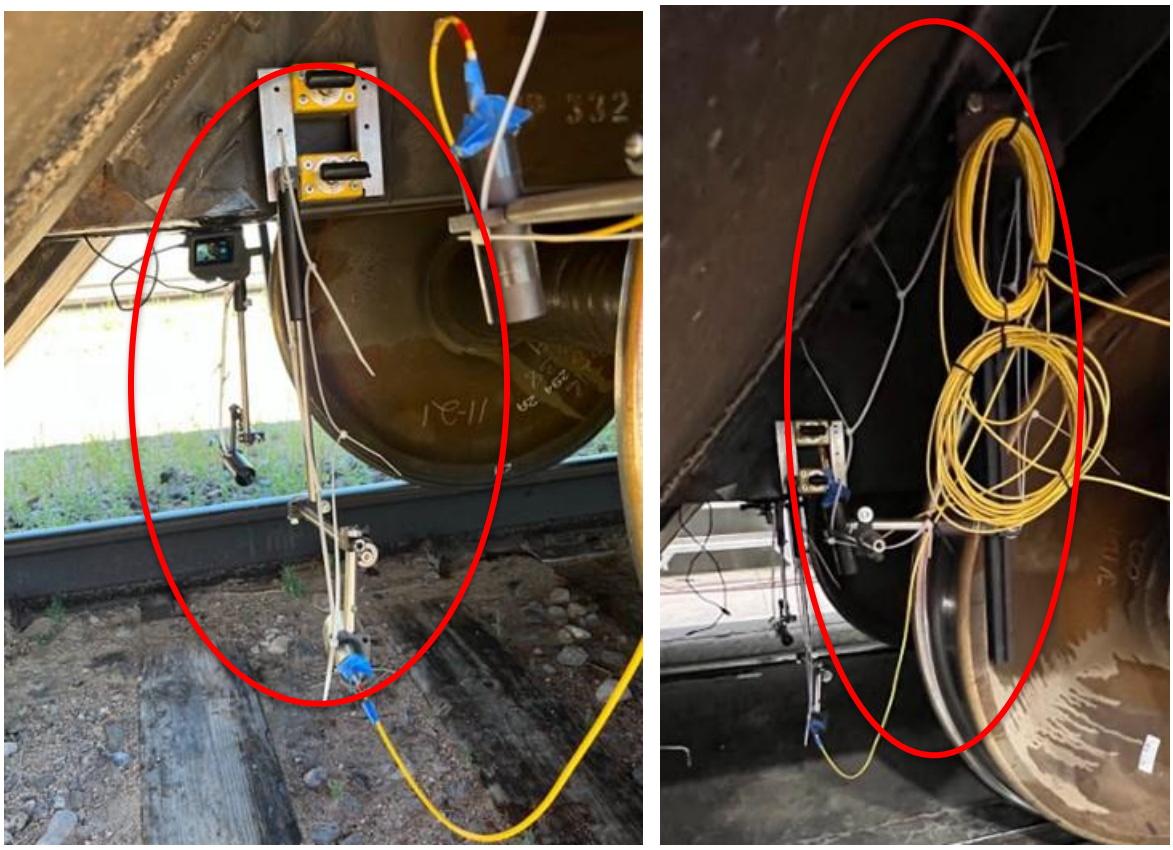


Figure 3-8: Lateral Blackbeard mounting (left) and vertical Corsair mounting (right)

The lateral Lidars are affixed to the center steel frame through the pair of the magnet mounts, with optical posts bolted to their frame. Each optical post then extends downwards where additional posts are extended outwards to accommodate for the focal range of the sensors; the end of which holds the posts that directly mount to the lenses.

The vertical Lidars were initially going to use the 7-inch focal length lenses, but the distance which the optical posts would have needed to be lowered to be in range caused vibrations which seen to the naked eye during the mounting process, so we switched to the 22-inch lenses. To mount these vertical lenses, the clamps were bolted to the flange connecting the wheel well to the bottom of the frame, where a foot and a half long, one-inch beam of 80/20 extended straight downwards. Along that beam, one of the optical posts extended straight backwards, where the lens and its mounting post were connected.

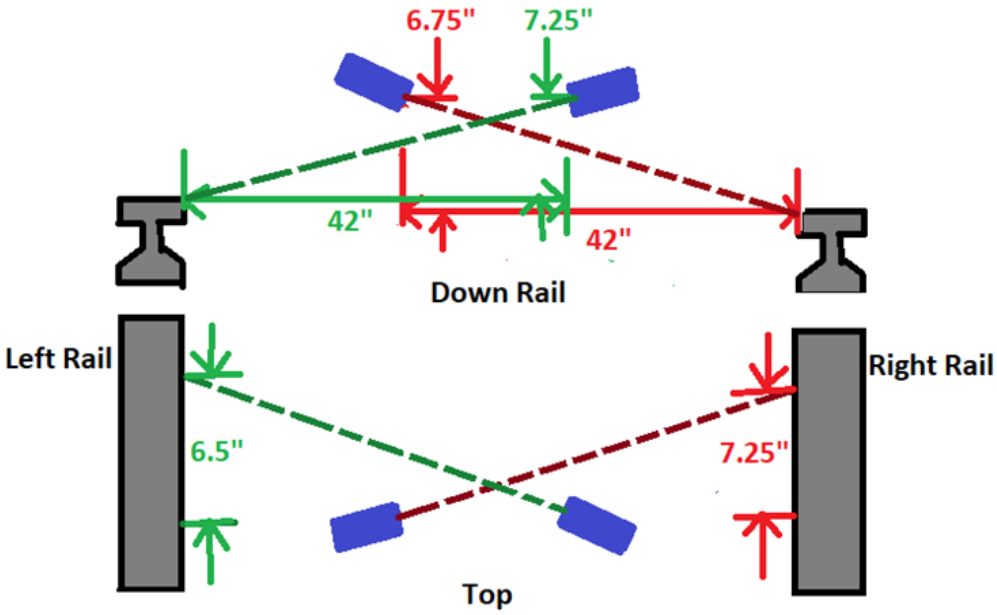


Figure 3-9: Lidar mounting locations for the lateral lasers

All four of the Lidars' angles were calculated by connecting a visible red laser to the lenses. Once they were connected, the lenses were aligned vertically to the track, marking the location of the lasers, and then using similar triangles to calculate how far forward the laser needed to be moved to be as close to the 10-degree angle as possible. Using a ruler along the rail, we measured the target distance and moved the laser to that point. To start, we aligned the lasers to the center of the tops and sides of the rail head. The calculated distances, as well as the final distances measured at the end of testing and adjustments are shown in

Table 3-2.

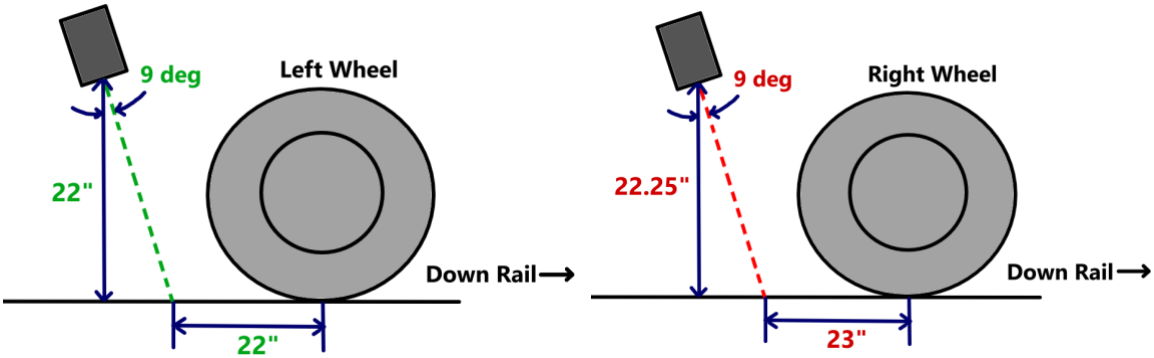


Figure 3-10: Lidar mounting locations for the vertical lasers

Table 3-2: Target focal ranges and mounting distances for Lidar lenses

Laser	Target Focal Dist. (inch)	Longitudinal Dist. (inch)	Lateral Dist. (inch)	Vertical Dist. (inch)	Actual Dist. (inch)
Blackbeard, Left	43	6.5	42	7.25	43.1
Blackbeard, Right	43	7.25	42	6.75	43.1
Corsair, Left	22	3.5	0	22	22.3
Corsair, Right	22	3.5	0	22.25	22.5

The fiber optic cables that connect the lenses and the Lidar box were strain relieved through the use of loosely affixed zip ties and tape, taking care to prevent small loops from occurring in the cable. We also made sure that if a sudden pull on the cables were to happen, the cables would not be able to yank directly on the computers themselves. Finally, in the case of extreme failure on the parts of the mounting methods, the Lidars, magnets, and clamps were zip tied in several places to the frame of the train so that they could not fall on the tracks or get caught in the wheels in the case of a disconnection.

Figure 3-11: Fiber optic cable connected from the rear of the passenger car to the frame of the freight car and down into the wheel well

3.5 Additional Sensors and Onboard Computers

Both the accelerometer and the GPS receivers were attached to the loaded freight car as well. Unlike the Lidars, and due to the much shorter cable connections, the accelerometer was attached to the front of the freight car, and the GPS receivers were attached to the top front of it. This placed both the accelerometer and GPS sensors about 9 feet in front of the lasers. The GPS receivers were placed at the top of the freight car to maximize their exposure to the sky. The greater the exposure of the GPS receivers to the sky, the more reliable their results should be as a greater number of satellites can connect and corroborate their location.

All of the sensor cables were connected back into the passenger car where the Lidar boxes, desktops, junction box, and monitors were contained. All of the on-board electronics were powered by a generator on the passenger car. The equipment was either strapped or taped down to prevent it from sliding or hitting things during operation. The test operators also were given chairs to sit and observe the data during the tests and to communicate with the engineer about the speeds and number of loops that were intended to be conducted.

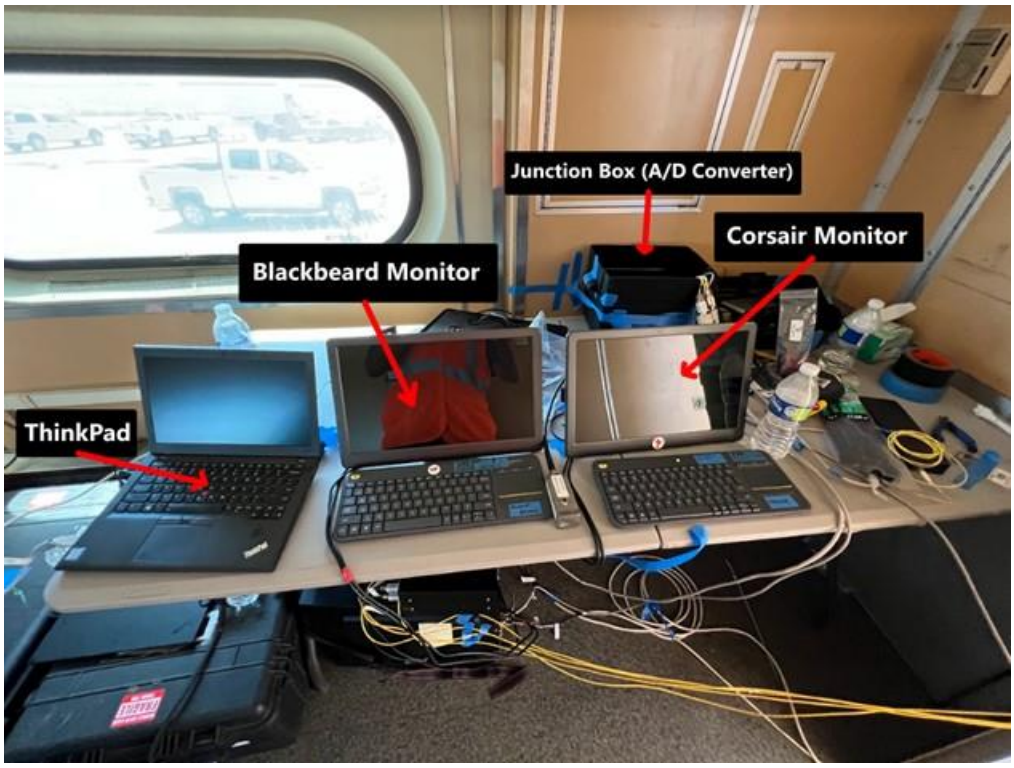


Figure 3-12: Onboard setup of sensor interfaces

The computers themselves were placed on a table with the chairs in front of them. The location of the table and the chairs at the rear of the passenger car allowed for the short GPS and accelerometer cables to access the freight car, and also for the test operators to observe the movement of the freight car itself.

3.6 On-location Test Preparation

This section describes the steps taken at the HTL to mount and calibrate our system to the consist provided to us.

3.6.1 Track Setup

Working with the staff at the TTC, they prepared the HTL for us before we arrived. As stated earlier, the vertical railroad spikes were removed from ties 201 to 208 and 551 to 558. Additionally, they placed the fish scales on some of these ties to measure the maximum lateral displacement of the rail relative to the ties. In section 40, the fish scales were placed on ties 202, 205, and 208 as well as on ties 302, 305, and 308. In section 25, the fish scales were placed on ties 552, 555, and 558. We had also printed forms to be filled out with the corresponding displacements for each of the fish scales by the people on the sides of the track who were monitoring them.

The track was cleared for use, and the days for all the tests were clear and hot. There was no noticeable wind, and the sun was out. The weather on all of the days was consistent. In addition,

there was no other traffic on the HTL during our testing, and so the train consist was the only system which couple have produced a change in the track condition aside from slightly varying track temperatures during the day.

3.6.2 Mounting Process

When we first arrived at the HTL, we were given a consist which had a locomotive and a passenger testing car, but not a freight car. In order to prepare the Lidars, we had to find a location for the computers as well as a place for the Lidars to be mounted. The best location for the sensors with regard to signal strength was just in front of or just behind a heavily loaded wheel since it would best show the displacement of the rail caused by heavy loading. While we had planned to mount to a freight car, they could not provide one to use that morning, and as such we first began to mount to the underside of the passenger car. This was a suboptimal location since the passenger car weighed significantly less than a loaded freight car and therefore would cause significantly smaller displacements.

Throughout this process, we were discussing what other options for cars would be available to us for use during our testing. After a few hours, we had managed a preliminary setup of the Lidar sensors on the passenger car, however we were offered a fully loaded freight car as another choice to mount our Lidars to. We decided the extra weight of the car and more convenient mounting locations was worth the loss of time unmounting from the passenger car and remounting the lenses to the loaded freight car. Thankfully, the car had an exposed frame support that the magnets could attach to, as well as a few other flanges that made the use of our clamp mounts convenient, especially since they were located just a few inches out from the center of the top of the rail. This freight car had an unloaded weight of 24 tons, and was fully loaded with a weight of 133 tons, for a total of 157 tons.

We also had to determine the best location for the Lidar computers within the passenger car. The on-board power requirements for our setup meant we needed to be within range of one of the outlets with the use of an extension cord. Additionally, with a 25 ft cable for the accelerometer, the table with the junction box needed to be placed near the rear of the consist. A folding table was set up, where it was pinned between our Pelican cases used to transport the sensors and the wall. The computer towers were placed under the table and ratchet straps were used to connect them to the table legs so they were less likely to fall over in the event of the train rocking. On top of the table was the pair of monitors, junction box, and laptop so that all of the data could be observed during collection. This allowed us to look for any discrepancies between the physical and digital events and to monitor the health of the data. Finally, the pair of chairs were placed by the table so that the observers would be able to watch both data health and to mark with the digital flag the locations of interest as we passed them.

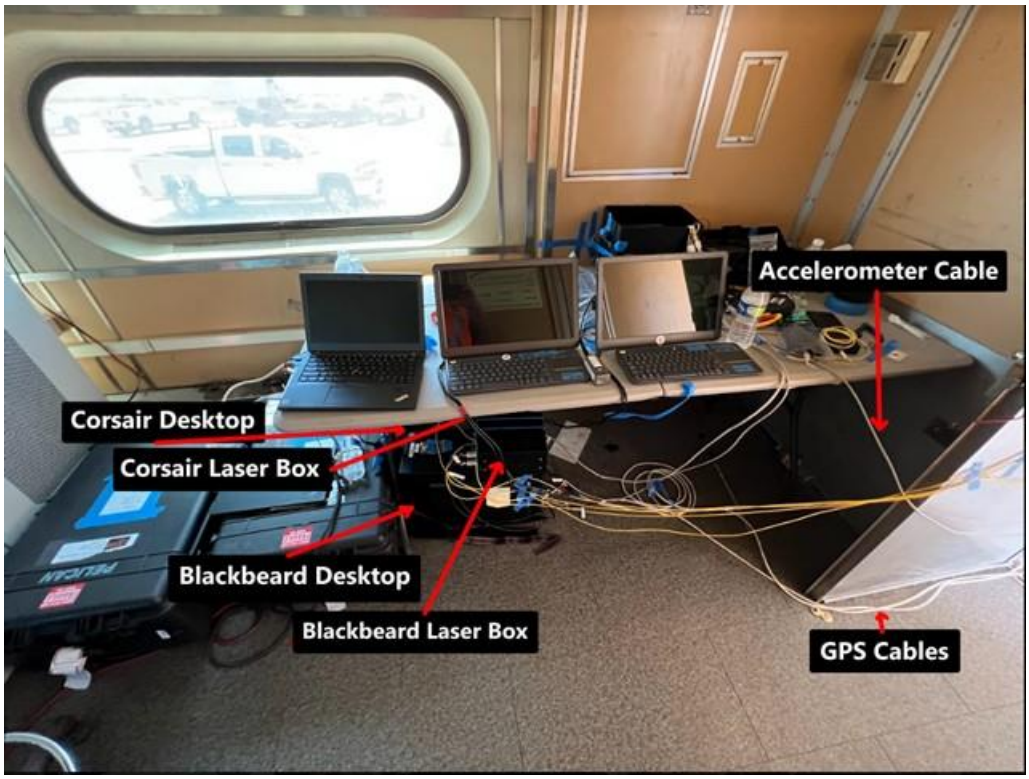


Figure 3-13: Onboard setup of sensor electronics and cables

With the freight car chosen, we began to mount the lenses to its frame. The process of mounting these lenses started with identifying the locations which would work best for mounting the lenses with the hardware available. A tape measure was placed across the rails and aligned perpendicular to them. Using the distances previously calculated for the target angles, we were able to identify which sections on the underside of the train would work best for mounting. Once the magnet mounts were attached to the steel frame, the optical posts were adjusted to get as close as possible to the correct distance from the rail. Several methods were used to verify this distance and the starting angle of the lasers. Using a combination of the red laser attached to the lenses, and verifying the alignment with the rulers and tape, the angles were derived by similar triangles. Afterwards, once the red laser was replaced with the operation-ready Lidar infrared laser, an infrared reactive card was placed to verify the final location of the laser point. This step was challenging to complete in the starting building because the building's floor between the rails was raised to the height of the top of the rails. This elevated floor interfered with the laser's direct line of sight to the rail and as such the final position on the side of the rail had to be estimated.

Figure 3-14: Elevated floor within the staging building

A similar process was followed for the vertical Lidar lenses. The initial setup involved a pair of 14-inch optical posts, but during the mounting process, clear vibrations were observed at a light touch. To solve this, one of the optical posts was replaced with an 80/20 beam to reduce the amount of vibration. In addition, a shorter secondary post was used. This substitution meant changing from the original 7-inch focal length lens to a 22-inch focal length lens. The same process of using the red lasers and measuring tape followed by confirming the distances with the in-use infrared laser and IR reactive card was used for confirming the location of the vertical lenses.

After the lenses were mounted to the frame and aligned as best as possible while within the staging area, we connected the fiber optic cables to the computers in the passenger car. The cables were run out the rear through the passenger car. Using a combination of zip-ties and tape to prevent the loose cable from falling into the gap between cars and getting caught or damaged in the linkage, the cables were connected to the lenses.

The GPS cables were also connected to the freight car through the rear opening on the passenger car. The GPS sensors themselves were attached to the top of the ladders on the front of the freight car using zip ties, as seen in the figure below

Figure 3-15: GPS receivers and accelerometer mounted to the freight car

The sensors were placed at the top of the car so that the frames of the surrounding cars would not interfere with their connection to the satellites and to get a more accurate and consistent reading of our location. In addition to the GPS mounting, the accelerometer was also attached to the front of the rail car. Instead of being placed at the top, however, it was mounted to the front of the freight car. It was oriented in a vertical position, so that the X-reading indicated lateral motion, the Y-reading indicated vertical motion, and the Z reading indicated longitudinal motion.

3.6.3 Calibration

Once all the sensors were connected, we did some preliminary testing before moving the consist to run on the track itself. While the accelerometer orientation was confirmed during installation, the GPS system would not provide a signal while inside of the building. Additionally, we started preliminary validation of the initial mounting of the Lidars. We did this through the use of the same spinning disk as discussed in the previous chapter, as well as the use of a sheet of paper sliding back and forth to simulate moving rail. The disk provided a consistent speed to compare the sensitivity of each of the lenses with, however it presented a challenge due to the offset from the rail caused by the thickness of the disk. As such, the sheet of paper was moved back and forth over the rail with another person looking at the live data feed. This would allow for testing of the lenses on a target which had a negligible offset from the rail surface. Once more, the issues of the raised floor meant that we were forced to measure the lateral laser positions offset from the rail and to look for a much weaker signal when validating the data feed.

After this initial check, the train was brought out of the staging building and we were able to finish our initial calibrations. Here, we were able to see that the GPS receivers were connected to their satellites, and we were receiving GPS data that seemed reasonable for our location. They were connected to at least seven satellites consistently throughout testing. We also were no longer hampered by the raised floors from inside the building and could finish the sensor calibrations using both the IR reflective card and the disk and paper methods to check the signal strength itself, this time positioned correctly on the inside face of the head of the rail.

After the sensors were confirmed to be in working condition and we were ready to check the quality of the data, we started collecting data and watched the signal while we went from the front of the staging building to our starting position on the HTL. During this travel, we noticed that the signals would drop out for significant portions of the time for both the vertical and horizontal Lidars on both sides. As such, before we went around our first track run, we continued to adjust the sensors to correct what we thought was a combination of general misalignment and the lasers falling off the track due to the effects of curvature.

Our first loops around the track were also used for further calibration of the sensors, and while we succeeded in getting the majority of the data to show measured track velocity, there was still a clear presence of data dropping. These effects can be seen in Figure 3-16, where the vertical lines indicate that the sensor is not receiving a sufficient or accurate signal.

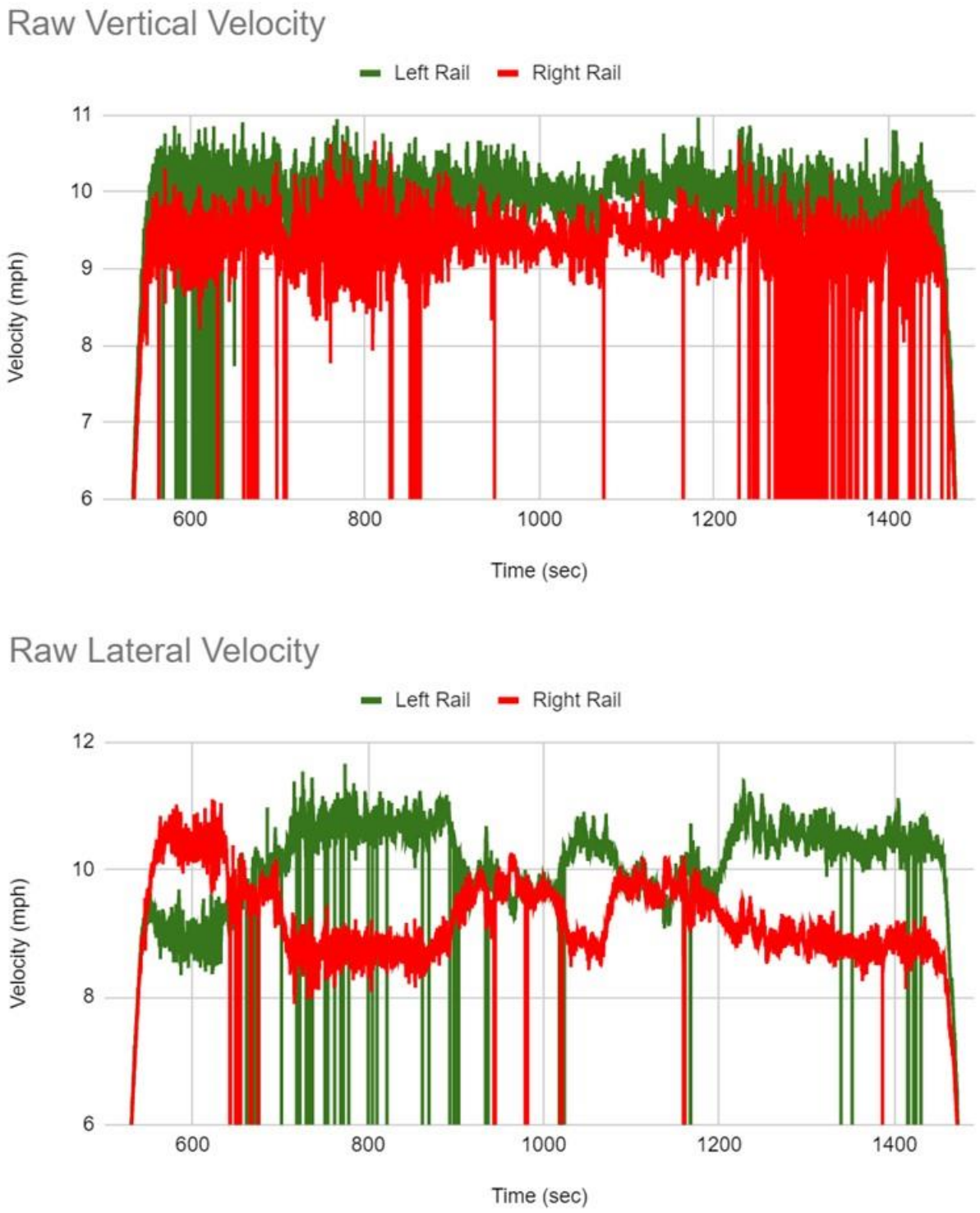


Figure 3-16: Raw vertical and lateral velocity data for the second 10 mph trial

These first calibration runs were recorded; however, they were not intended to be included in the data processing because the lenses were adjusted after each run. They were useful, however,

because we were able to see the issues occurring during the entire loop around the HTL. During these runs, we found that the lasers would fall off the rail for a large proportion of the loop. Additionally, we had to tighten the mounting connections for all of the sensors, which helped a significant amount with reducing the dropouts. While we were unable to maintain consistent data without dropouts for the majority of the run, we got the best sections to mostly overlap with our sections of interest, so these were the orientations we settled on for the final data collection.

After the first calibration runs of each of the run speeds, and the general location of the sensors was set, we attached our GoPro cameras to the wheel well. The cameras were mounted in the center of the undercarriage, one for each side, and were positioned to see the lenses, track, and wheels on each side. The cameras were mounted through the use of high strength double sided tape on their base plates and additional duct tape for security. This type of mounting provided no issue with regards to camera vibration when reviewing the footage.

Figure 3-17: GoPro cameras mounted to the center frame of the freight car

3.7 Data Collection

Each loop in the data collection process followed a similar procedure but the order and combinations of speeds were varied throughout. In general, for each pass around the HTL the Lidar, accelerometer, GoPros, and GPS sensors recorded during the entire loop. In addition, several people were stationed at the point of interest in section 25 and the two points of interest on the tangent section in section 40. After the train had passed these points of interest, and it was signaled safe to do so, the people on the side of the track would record the maximum displacement of the track for each loop and reset the fish scales.

Figure 4-1: Fish scales mounted on the tie to measure the lateral deflection of the rail

The method of data collection throughout the experiment varied based on which trial was being conducted. The first two runs of each speed (the first of which being the test runs for sensor adjustment) were conducted with the train at the start position in section 5, following the loop, and then stopping at the same position. The data collection was then stopped, named, and saved, and the next data recording session set up for the next trial. The time spent between these runs was used to validate that there was enough time to safely check and reset the fish scales before the train would have finished a loop, as well as to check on the sensor system to make sure it has not loosened up or fallen out of alignment.

After each of these beginning trials, we stopped the train and got out to check the sensor configurations. The cameras were confirmed to still be recording, the lasers were checked for their alignment, and the optical posts were tightened. We also verified that the GPS system was connected to enough satellites. The posts did not need to be tightened in the large majority of trials and as such it was assumed that more infrequent checks could be used for future tests.

The third run for each of the speeds was then done continuously. As such, the 10, 20, 30, and 40 mph loops were run back-to-back in successive order without stopping either the train or the data collection. Finally, for the remaining runs, they were conducted back-to-back within their own speeds. For 10 and 20 mph, there were three continuous loops, and for the 30 and 40 mph runs, two continuous loops were conducted. This generated a total of six 10 and 20 mph trials, and five 30 and 40 mph trials.

Initially, we had planned to conduct the experiment with 6 runs for each speed, however we removed the final 6th run for each of the 30 and 40 mph speeds due to both running out of time for that day of testing as well as the lower quality of data for speeds greater than 20 mph.

3.8 Issues Encountered

Several issues presented themselves during the data collection process. The most apparent of these problems was the health of the data collected. We believe we have identified two main reasons for this issue, as well as a few thoughts on what might have caused additional issues regarding the health of the data. First, it seems that the track was not within the focal range for sections of each run, which could be caused either from focal range issues or errors with aligning to the track face. Another source of the data health problem was the vibration and rocking of the freight car.

Regarding the laser's ability to stay within focus and on the face of the rail, many issues presented themselves. For the focal range itself, while we spent much of the previous day and that morning on measuring and aligning the lenses, we did not find a fully clear position which produced as strong of a signal strength as seen in the lab. This could certainly be indicative of issues in alignment angle or distance, but our final measurements show that the angles and distances we had were still within the operating ranges for the lenses. Another possible issue we identified, which was much harder to address, was the effects of the combination of the curve of the track and the superelevation within those curves. The curve of the track moves the vertical laser's position away from the center of the top of the rail's head. For example, a clockwise curve of the track would push the right vertical laser further to inside of the rail. The lateral lasers would also be subject to this change of focus position due to curvature. These lateral lasers would need to travel a longer distance to reach the inside face of the track, and as such could be pushed to the edges or beyond of the laser's focal range.

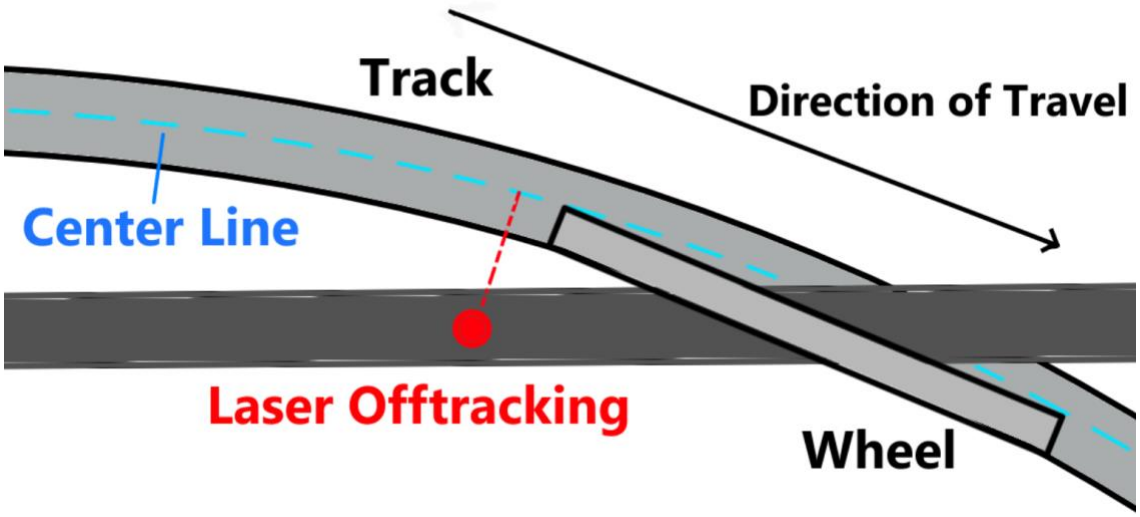


Figure 4-2: Vertical laser offtracking during a curve

In addition to the curvature of the track, the HTL curves contained superelevation of about 3 degrees [20]. This would further shift the focal ranges and focal points of the lasers on the track. With the freight car loaded on an angle, its center of gravity was shifted off the centerline of the track, which the suspension responds to by compressing further on the side of heavier loading. This angled body position would lift the outside lasers and lower the inside lasers. It would also have an even more distinct effect on the lateral lasers, where they could be pushed below the head of the rail or fully over the top of it. This explanation seems to hold some validity due to the observed issues in our first few runs since the right rail lateral laser on counterclockwise curves would not just have inconsistent information but would sometimes fully lose signal in its entirety, which could be demonstrative of the shift in center of gravity pushing the outward facing laser over the top of the rail. For the vertically aligned lasers, the effect was even more prominent, with the right laser falling off the track around left curves and the left laser getting more drop outs during right curves.

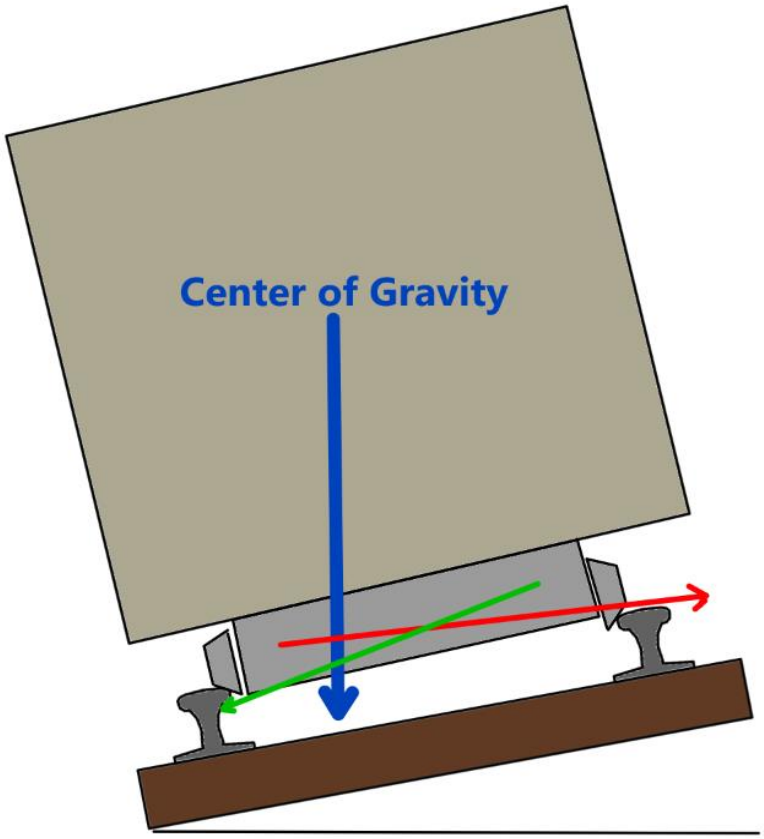


Figure 4-3: Effect of the superelevation on the lateral Lidar lasers

Another issue found during the data collection was the rocking of the freight car; it would oscillate by rolling back and forth. Its bogeys would also shift back and forth laterally, inducing an oscillating yawing motion. This could also contribute to the lasers falling off the track as mentioned above. The rocking continued throughout the entire loop, but was worse going into or coming out of the curved sections of the track. This rocking, however, provided us information about the track-train interaction dynamics, but it forces us to process the data through a different method than initially intended to observe it. The interplay between the rocking train and the track conditions influences how they both are observed, as the rocking train could excite the rail motion to a greater extent, while changing track stiffness can shift the speed and extent to which the freight car rocks. Rather than filtering out the motion of the train, some use could be found by keeping the data coupled between the two of them, as will be discussed in chapter 6.

Finally, one of the issues we found with our Lidar system was the broadened bandwidth of the absorbed laser frequency. This expanded bandwidth reduces the detectability of the exact velocity as the Doppler shift becomes less clear, and if the bandwidth is too broad, will produce no output. This could also be one of the main contributing factors to the dropout rate in our data. While the laser being on the edge of its focal range could produce this result, it is also possible that environmental factors were the perpetrator. For example, the reflectivity of the target impacts the returning signal strength, so if the track was heavily lubricated, or there was a sufficient amount

of dust on the lens, these could produce this broadened bandwidth as the returned laser was diffused.

The conditions of the fibers themselves could also add to the broadened bandwidth. This could be produced a couple different ways during testing. The first of these being caused by kinks, bends, or pressure on the fiber optics, as this would change the path of the light, or even damage the connection itself. Another way the cable connections could have interfered with the signal strength is the connection to the lens itself. After our first couple runs around the HTL, we disconnected and reseated the fiber optic cable to the Lidar sensor lens and tightened it further. This did produce a noticeable difference in the power returned during testing, somewhere in the range of an order of magnitude greater. One concern with this process, however, was that of the fragility of the fiber optic cables and the danger that overtightening could have on the condition of the cable itself. As such, we did not attempt to tighten it further than a strong hand tightening.

While all of these issues could contribute to the weaker signal strength and the inconsistent signal, we took steps to solve or mitigate them. Many of these have been discussed in the sections describing the test setup and procedure such as verifying the lens alignment. We also inspected the lenses after returning to the lab and found that they were not dusty or otherwise interfered with the signal strength themselves. We also worked to address these issues in the data quality through the processing of the data to condition it for analysis. These include the implementation of data cleaning procedures to remove the drop-ins and drop-outs found within the sets of data. The accelerometer results when aligned with the GPS signals allowed for the correction of small variations in the angles of the Lidar sensors to the track. Overall, while the data health was not to the level we were hoping for, the results, combined with our other sensor data, allowed us to make use of the data for the evaluation of the efficacy of non-contact Doppler Lidars in detecting laterally loose rail. We then approached this collected data with our previously used data analysis methods combined with several new approaches to fully understand the capabilities of the Doppler Lidar system. The next three chapters approach the problem of data analysis from different lenses, first starting with data cleaning and implementing an unsupervised sorting algorithm.

Chapter 4 Gaussian Mixture Model Method

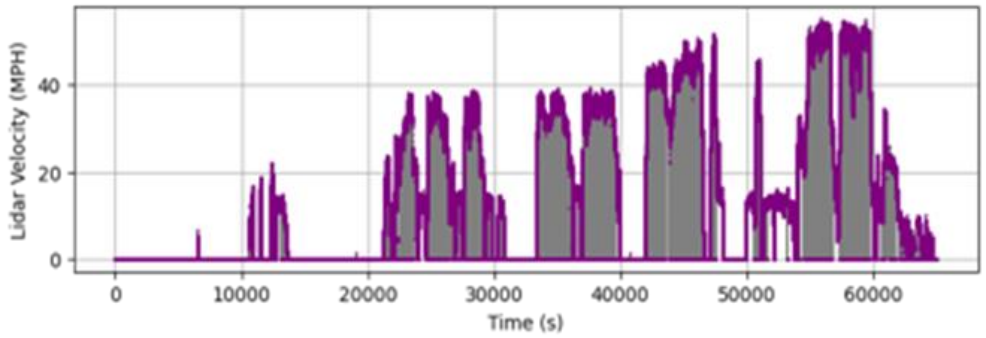
Once the data was collected, the goal was to identify sections of the track that indicated weak rail, or rail with increased vibrations. Sections of loose rail should produce several observable phenomena including a greater deflection distance and a less damped vibrational system. These excess motions of the track negatively impact the safety of the rail, as the train is more likely to derail or cause additional damage to the track. Overall, the purpose of these tests and data analyses were to corroborate or identify the locations along the track which were purposefully designed to be "loose" through the removal of vertical tie-plate spikes.

The first method used to inspect an evaluate the data collected from these trials is the implementation of a Gaussian Mixture Model, which sorts points in a data set into groups, which the differentiation of can indicate rail of interest. Before that method, or any other analysis method, could be conducted, the data needed to be cleaned and prepared first.

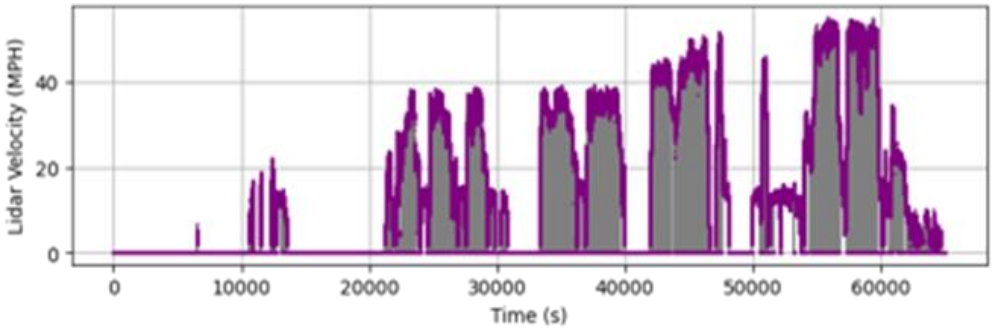
The processes in this chapter have been applied multiple times in previous examinations of the Doppler Lidar data collected from the CVeSS system. As such, much of the work presented in this chapter is applied from previous works by this lab. Even during this field evaluation, other members of CVeSS developed and executed the following procedures to analyze the data for characteristics of weak track as measured on the HTL. [21]

4.1 Data Cleaning and Preparation

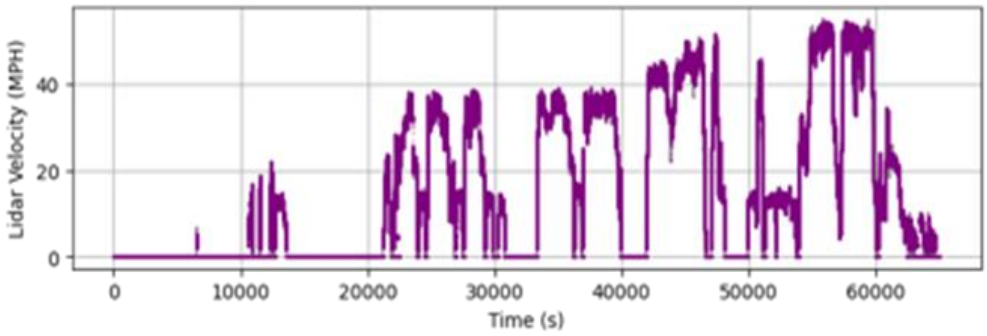
In order for any of the data processing methods to work effectively, the raw data must be prepared and processed into a form that most represents the measured movement. The first step in this process was to mark the start and the stop of the run. This included any sections which contained time where the train was stopped, as the Lidar may produce unreliable data due to noise peaks in the RF noise floor at sufficiently low speeds. As such, a threshold based on the GPS speed was applied to the data to remove sections which contained speeds lower than 2 mph. Additionally, cleaning techniques were implemented to remedy the issue of drop-ins which occur due to the reduced signal-to-noise ratio and therefore have a much higher or lower indicated speed than what it should have been measured. Through the comparison with the GPS velocity, any Lidar data which was two times greater or smaller than the velocity was removed. An example of the impact of these first two steps can be seen on this example from the 2021 Roanoke to Norfolk data set in image b. of Figure 4-1.



a.



b.



c.

Figure 4-1: 2021 Roanoke to Norfolk: (a) raw data, (b) low speeds removed, and (c) the dropouts removed [21]

The next step in the data cleaning and preparation process was to remove sections indicating rapid changes in forward velocity. The momentum and inertia of the train prevents rapid changes in speed from being possible, and as such, large changes in measured velocity indicate additional outliers. The method implemented for cleaning these outliers was to once more remove data which falls outside of a reasonable threshold, which in this case was a change of more than 20% the average of the prior second of velocity.

The removed data from these previous steps needed to be replaced. This was implemented through a linear interpolation of the data surrounding the removed data, which created a vector of velocity without discontinuities. Here is an example of the data collected from the Roanoke to Norfolk test after completing this interpolation process.

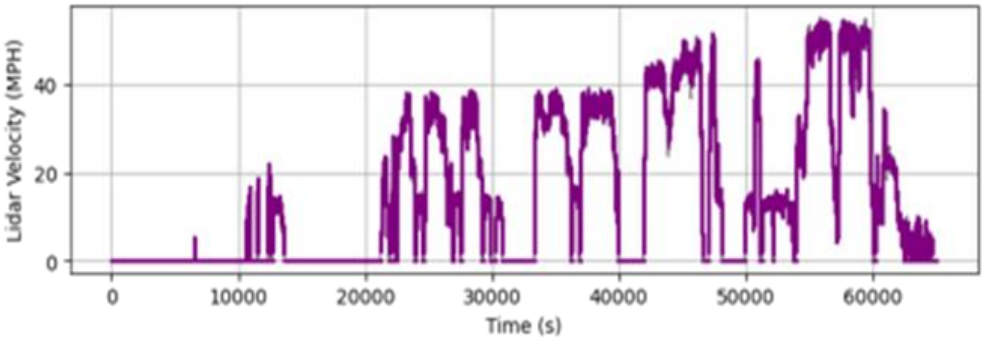


Figure 4-2: 2021 Roanoke to Norfolk – removed data replaced with interpolated data [21]

Due to the lateral and vertical Lidar data being recorded on different computers, the data was not recorded at the exact same time. The next step in the data preparation and processing was to synchronize the two sets of data. This was done by aligning the start of both of their runs as well as the start of their recorded acceleration data. This allowed for the comparison of both the lateral and vertical data simultaneously, and therefore to be able to observe coupled lateral and vertical motion.

The delays in the peak seeking algorithm of the Lidar produced a non-uniform data acquisition rate and needed to be corrected. In order to generate a uniform sample rate, a Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) was implemented. This allowed for some of the future data processing methods such as the linear time invariant filters to be applied on the now prepared data. The resulting data from this PCHIP implementation was a time vector with uniform spacing, rather than the initially varying time vector resulting from delays, along with the interpolated velocity data.

Finally, our interest in track dynamics lies in the lateral and vertical motion of the rail. In order to extract these components from these data sets, they must be decoupled from the longitudinal component they were combined with due to their forward angled mounting conditions. The prior scaling of the data matched the Lidar data velocities with the forward velocity of the train as presented by the GPS. To remove this forward component, the data was processed through a high-pass filter with a 3rd order Butterworth with a cutoff frequency of 0.65 Hz. This filter results in the extracted orthogonal rail motion, which further data processing methods will be conducted on. The first processing method used, and the focus of this chapter being the implementation of a Gaussian Mixture Model.

4.2 Sorting Via Gaussian Mixture Model

One of the methods implemented by the lab was the sorting of the data through an unsupervised sorting algorithm. Each run was divided into sections, dubbed micro-trips, which provided bounds

for the data to be analyzed. This algorithm would take the velocity data of each micro-trip and produce results which indicate points of interest for loose track conditions. In the case of our tests, since the track was sufficiently small, the micro-trips included the entire length of the track for one revolution.

These micro-trips were then subdivided further via segmentation. The resulting segments were used to sort the sections of the track via their velocity characteristics. These velocity characteristics include the average velocity as well as the variance in that velocity for that section. A greater variance would tend to indicate a greater amount of deviation of the rail, and as such is one of the criteria used in the sorting algorithm. That, in combination with the average velocity in the lateral and vertical direction of the track, act as the inputs for the sorting algorithm.

One of the issues with evaluating the stability of the track as a measure of deviance is that higher speeds will produce a greater deviation, even if the rail is not any weaker. This is due to a correlation of the lateral and vertical motions to the speed of the train. As such, rather than producing a threshold which may not be sufficient to consistently observe instability at slower speeds than greater ones, the data must be weighted to counteract this effect.

The weighting of the data during filtering aimed to normalize the variance to become consistent across the speeds. The first step in this process was to fit a linear regression to the velocity data, where the velocity data was the independent variable and the forward velocity was the dependent variable. The residuals of this regression were squared followed by a second-order regression line on the forward velocity. The squaring of the residuals acted to bring the values all in the positive domain, since the average tended around zero. The weights used were the inverse of the calculated regression line. Using these weights to normalize the data, the variances in velocity were decoupled from the speed of the train.

The instability of the track impacts the dynamics of the rail both before and after the particular points of instability. The division of the track into segments allows for the analysis of not just exact points in time, but the overall impact and result of the dynamics at the locations across the track. Each micro-trip was divided into 300 ft segments via spatial segmentation for this purpose.

The transient vibrations associated with other rail structures like switches, railroad crossings, and bridges can be observed in the data even if they are structurally sound. These vibrations would present themselves in the data as false positives for loose track conditions, and therefore should be differentiated in some way so as to enable observers to detect sections of track that are uncharacteristically unstable. To this end, through the use of GPS data overlaid with Google Earth, the track segments were categorized by the existing structures within them, which should allow for the distinguishing of “loose” rail conditions as a result of structures versus as a result of track behavior.

When dividing the segments based on their characteristic structures, they were divided into 5 different categories. These categories were switches, road crossings, bridges, tangents, and curved track. A significant portion of the track is characterized by either tangent or curved track, within which large variations in velocity could be indicative of potentially weak track segments. The remaining segments containing rail structures were excluded to improve the reliability of track

evaluation of weak segments and to reduce the false positives that would result from their inclusion.

The types of instability found within the track conditions can be found through slightly different means depending on how they present themselves. For example, if a segment contains a large lateral or vertical velocity component, its large motion could indicate the entire segment is subject to instability. If, for example, a segment contains relatively low variance in the velocity content, but a smaller portion of the data has a much greater variance, the measure of the mean absolute velocity would not be sufficient to recognize this section as possibly containing loose track sections. To address this issue, the comparison of the number of points that deviate from the variance of the segment could be used to indicate segments where smaller portions of the data contain potentially unstable conditions.

Hence for both the lateral and vertical Lidar data, the number of points which fall beyond two standard deviations from the mean absolute velocity along with the mean absolute velocity of the segments themselves were chosen to be the representative criteria for the segment. These parameters were scaled to a range of 0-1 to compare along similar scales. The HTL was divided into 48 segments of 300 ft each, and as seen below in the figure, the segment containing the emulated loose rail in section 25 emerges with the greatest average absolute lateral velocity of the track.

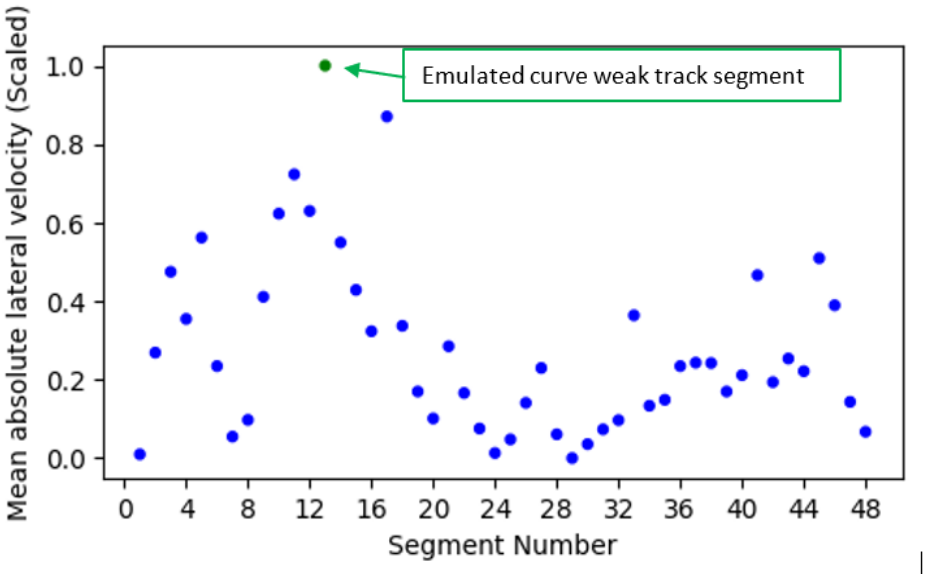


Figure 4-3: Mean absolute lateral velocity of each segment from the second 10 mph trial [21]

The sorting algorithm for these data points is a clustering method. The current algorithm uses a Gaussian Mixture Model to sort the segments. This style of data modeling has been used in other applications to evaluate the extent of damage a structure has suffered; however, they analyzed the same structure as a function of time to observe changing damage characteristics [22]. Rather than comparing a response to its earlier undamaged state, our implementation attempted to find sections

of the track that were sufficiently different from the known structurally stable track. A necessary step for the implementation of this model was the selection of the optimal number of clusters to sort the data into. To determine this optimal number, the Elbow method and the Bayesian Information Criterion (BIC) were applied for this current rail analysis. As the number of parameters increases, the likelihood to differentiate the data increases, but so does the likelihood of overfitting. The BIC uses a penalty function to prevent overfitting when selecting the number of parameters likely to represent the model. Typically, the lower the result the better, but for clarity, these results are presented as the inverse of their BIC values, and as such the higher the better. From the selected bounds, we can see that using 4 clusters was the optimal choice for the potential to detect the phenomena we were hoping to see. Fewer clusters would decrease the reliability of the results and a greater number would cause overfitting and overly complex algorithms.

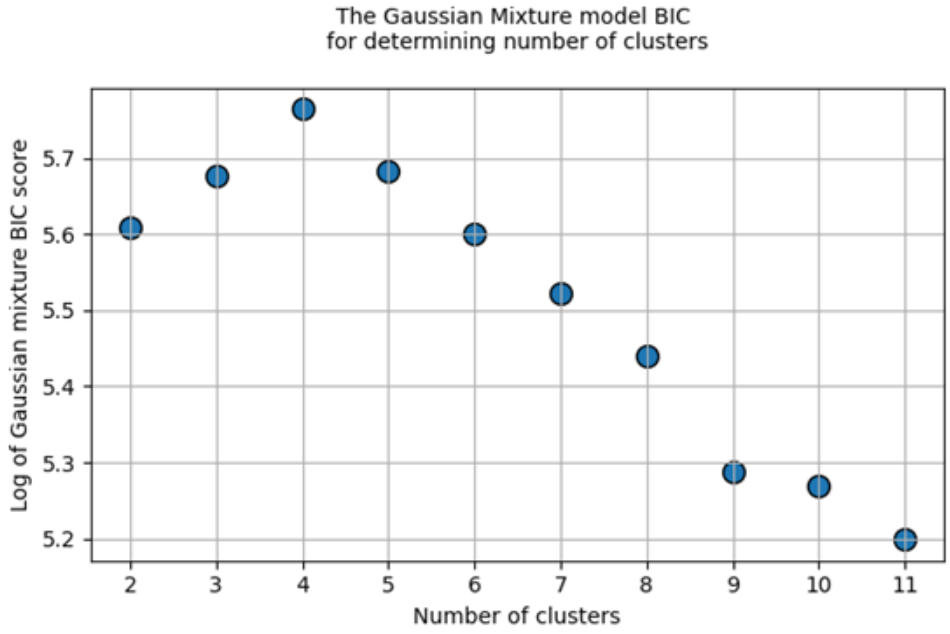


Figure 4-4: BIC to determine the optimal number of clusters for the Gaussian Mixture model [21]

The clustering analysis was performed on all runs from the TTC testing. To present and discuss these results, the four parameters were divided into four plots, each containing a relationship between three of the parameters. These four graphs together cover each combination of parameters which could be grouped together. The results of each parameter are scaled to a 0-1 range, and each graph shows the 48 segments as positioned based on their characteristics. The color of each data point represents which cluster it was sorted into. From this figure, we can see that cluster 3 stands out to the greatest extent when compared to the rest of the data and this segment also contains our point of interest in the curved section of track.

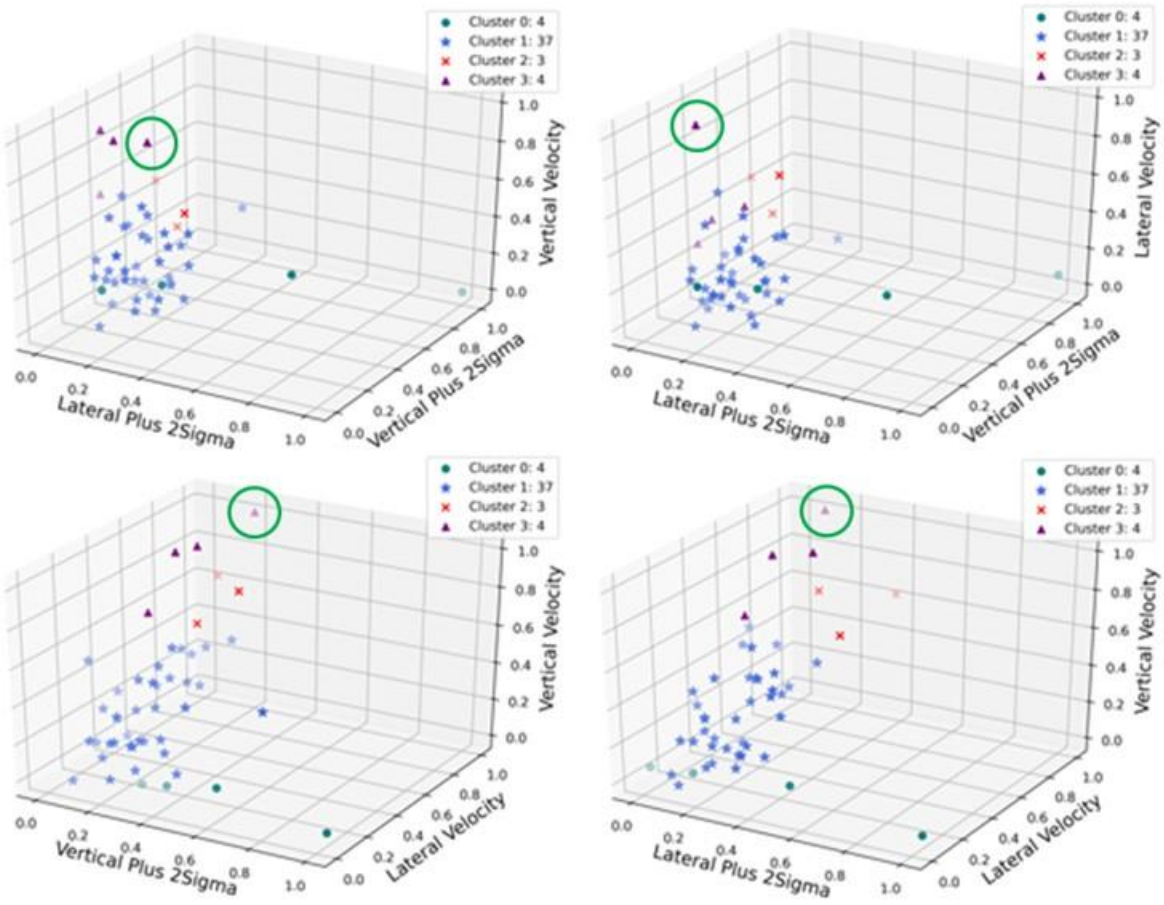


Figure 4-5: GMM clustering along all four parameters [21]

Cluster 3 also contains the points which had the greatest absolute values with regards to the vertical and lateral average velocities. In comparison, cluster 1 contains the points with relatively small absolute orthogonal velocities and low amounts of deviation. Clusters 0 and 2 present larger parameter variations. When looking at one of the parameters, in this case the mean absolute lateral velocity, the points in cluster 0 are associated with the starting and stopping motion of the train while cluster 2 presents itself surrounding the loose track sections.

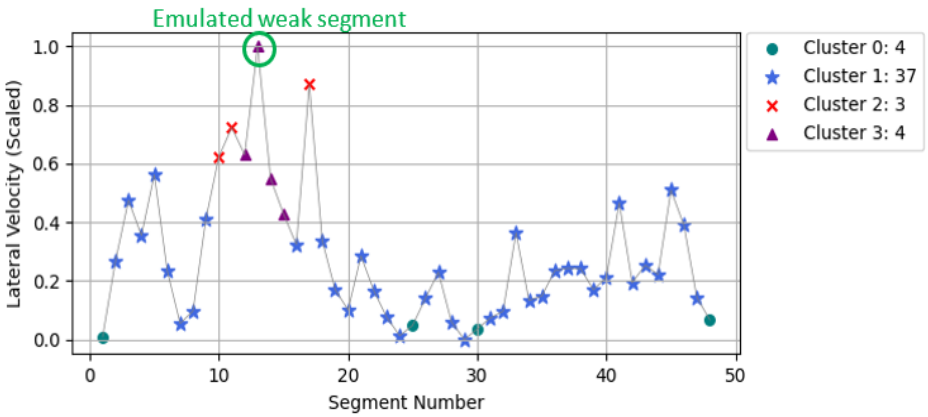


Figure 4-6: Sorted GMM results compared to the data [21]

From these results, cluster 3 was selected as the parameter in this case to indicate which section of the track was most likely to demonstrate loose characteristics. Out of the 16 loops performed at 20 MPH and below, 13 of the results demonstrated the emulated weak track being positioned within this cluster of interest. These results validate the possibility of detecting weak track conditions through the use of machine learning algorithms and the implementation of non-contact Lidar sensing. Above 20 mph, however, the data was not strong enough to use for reliable detection of the loose points of interest in the track. Additionally, even at low speeds, only the loose track in the curved sections was detectable.

This method, however, is spatially imprecise as each section represents 300 feet of rail. In an attempt to increase the spatial precision of the data processing, the data was analyzed with regards to the frequency domain.

Chapter 5 Frequency Domain Analysis

Processing and analyzing the data in the frequency domain allows for inspection of patterns in the rail motion or changes in vibrational characteristics of the rail. It can be used to analyze the data for information that would not be observable via direct inspection.

5.1 Implementation of the Short Time Fourier Transform

After the sorting algorithm, the next data processing technique which was applied to the TTC data was the use of a Fourier Transform. A Fourier transform is a technique used to convert data from the time domain to the frequency domain, or in our case, since we are interested in the location of the points of interest, from the spatial domain to the frequency domain. The Fourier transform accomplishes this through decomposing a signal into its component sine or cosine signals and returning its power across the frequency spectrum. The classic Fourier transform, however, is a summation from negative infinity to positive infinity, and is operated on a continuous data set. The data for each run was obviously bounded within finite limits and used discrete data rather than the continuous data as assumed in the Fourier transform. In order to accommodate for this, the use of the Fast Fourier Transform (FFT) was implemented on the data.

5.1.1 Background

The purpose of analyzing our data through the use of harmonics can be illustrated by treating the track as a spring-mass-damper system. Previous analysis of track vibration through the use of the Fourier transform uses this model, based on a Timoshenko beam. These previous studies, however, tend to focus on higher frequency responses, in the range of 500 Hz or greater [23] to analyze the squeal effects of the rail, which our system was incapable of measuring. Additionally, these previous studies tend to focus primarily on just the wheel and rail interaction, as the body suspension causes the natural frequency of the rest of the train to be significantly lower than the frequencies of interest. Our experiment, however, was interested in the lower frequency deflection motions of the rail rather than the surface roughness and squeal effects prior Fourier inspections analyzed [24]. In this consideration of the track structure, the track itself acts as a mass while the mechanics of its stiffness and damping comes from its connection to the ties as well as the mechanics involved in the tie's ballast's interaction. The track is effectively a beam with an infinite length compared to the localized vibrations, with a traveling point of vibration. Loose connections to the ties or ballast could produce a much greater vibration or a change in the vibration frequency. Figure 5-1 shows a simplified representation of the stiffness and damping relationships between the ballast, the ties, and the beam as a series of spring-mass-damper systems.

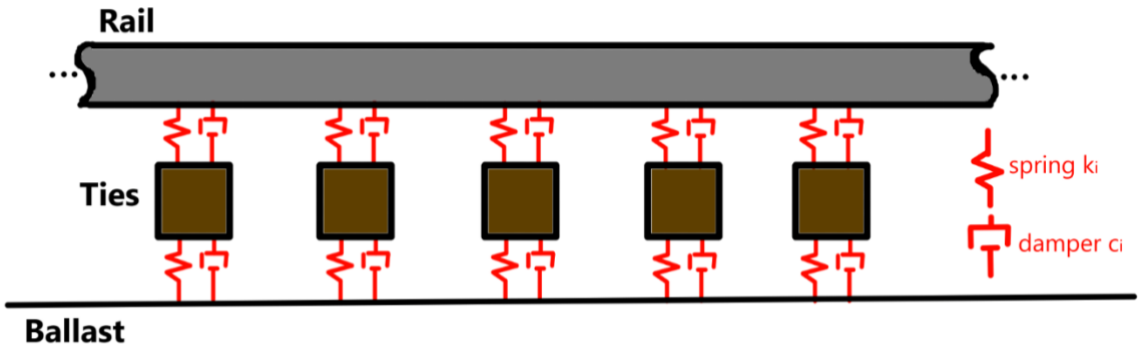


Figure 5-1: Railroad track and ties represented by a series of springs and dampers in the vertical direction

Under normal operation, the damping and stiffness of the rail is quite high, and as such acts without much in the way of a harmonic resonance, and with a high natural frequency. If, however, the damping and stiffness was reduced, the excitation and forced movement of the train onto the track could cause the response to grow. It would also cause a different frequency response as compared to the non-" loose" sections of the track. This apparent lack of damping could come from a bad connection to the ties or through issues with the ballast and its strength and damping capabilities. As such, while we do not know what form the looser rails will take in terms of their presentation on future tracks, it seems appropriate to attempt to analyze its movement through the use of frequency analysis, and in specific the implementation of a Fourier Transform.

In order to use the Fast Fourier Transform, the data must be of a number of points that is a power of 2. As such, due to the data not directly being a power of two when collected, two options remain when attempting to conduct an FFT. The first was to trim the data until it reached the largest power of two that was smaller than the number of data points collected. The second was to pad the remaining data with zeros until it reached the next largest power of two. Both of these methods had issues, however, for practical implementation. With regards to trimming the data, with a large data set, this could involve cutting out a significant portion of the recorded data since up to half of the data set could fall outside this range. For example, with the second 10 mph Corsair data we would lose 43% of the data set, which would remove a large portion of the track response.

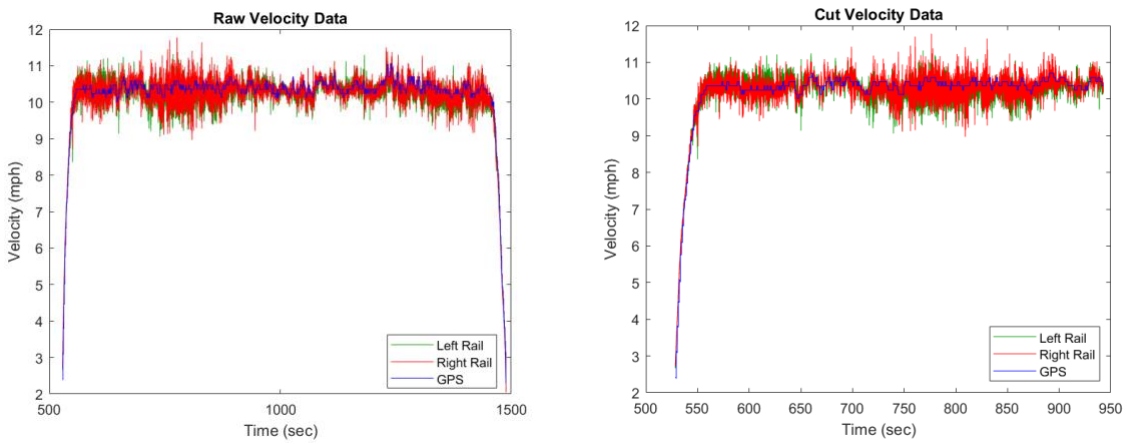


Figure 5-2: Second 10 mph Corsair full data (left) and cut down data (right)

As such, cutting down the data was not a viable solution. In contrast, padding the data set with zeros also produced undesired results. This would have greatly altered the frequency distribution of the resulting Fourier transform. To mitigate some of these issues, padding the remaining numbers with the average value of the data set, or even the average of the initial/final values could serve to reduce the sharp drop off that would happen at either the beginning or the end of the data. The results, however, would still be unrepresentative of the overall characteristics of the track.

Additionally, using a Fourier or Fast Fourier Transform on the entire data for a loop provided very little useful information. The goal of this data processing was to identify the locations on the track where there were changes in conditions, and more specifically, to identify the locations of the track where we purposefully removed vertical tie spikes. The Fourier transform retains no time or spatial information, and as such a single transformation cannot be used to identify the locations of differing track dynamics. One method to account for this issue is to partition the data into smaller sequences and to conduct a Fourier transform on each of the smaller sections.

The method of conducting a Fourier Transform on subsections of data in time or spatial order is known as a Short Time Fourier Transform (STFT). This transform allows for the observation of changing harmonic components throughout the data set; however, the issue of uncertainty has not been fully removed. It also can work to solve the previously noted issue with regards to the application of the Fast Fourier transform. When the data is divided into smaller segments, each of which are of a power two, the possible unusable data is significantly smaller than what it would be as a part of a single transform. In this case, the question of padding or cutting can be quickly addressed since the final section of the data was when the train was not moving and therefore the loss of this data was meaningless when used for analyzing the track conditions. Even in cases where the train did not stop or stopped close enough to the end of the recording session that there was still some movement in the final segment, these locations were far enough away from our points of interest that it did not impact our results.

The STFT has been applied in several other fields as a method of detecting damage or changing surface conditions as a function of time. They have been found to be accurate in the modelling and

monitoring of changing conditions regarding the life of bearings and the use of grinding on metal [25,26]. And the monitoring of bearing health has been applied to those on trains in as well [27].

One of the considerations when implementing the STFT was selecting the size of the segment divisions. The larger the segment size, the more precise the representation of the overall frequency content the results will be, however the ability to identify locations of changing conditions, especially those with shorter time effects, will diminish. In the other direction, segment sizes which are too small wouldn't allow for much analysis of the frequency content due to lack of resolution, even though the locations were much better defined. This tradeoff can be pushed to either extreme, the largest of which represents just the Fourier Transform conducted over the entire data set, and the smallest of segment sizes being representative of just the raw time domain data.

Within this experiment, a single choice of segment size was not implemented. Instead, due to the changing sizes of data files especially due to the changing times from faster speeds, the segment size was determined individually for each run. While this might present itself as a difficult decision to make, the trials were similar enough to one another that the changing sizes only varied by a couple powers in either direction. In most cases, the segment sizes would not range smaller than 64 or greater than 1024, though those external sizes were still looked at. Below are examples of the STFT results for the smaller and larger bounds of segment sizes.

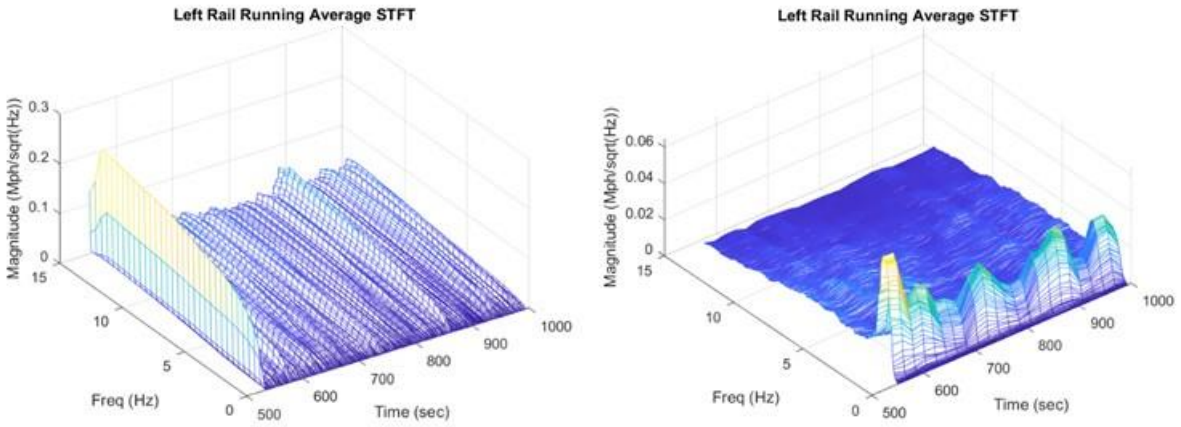


Figure 5-3: Fourth 20 mph Blackbeard data segment size of 64 (left) and segment size of 1024 (right)

These examples were shown on the fourth 20 mph trial Blackbeard data, and it is clearly visible that both extremes lack sufficient information to either make a determination of the frequency content or the location of the content within each segment. The starting segment size was typically around 512, and then be reduced sequentially until a balance of spatial and frequency resolution was found for that speed.

For this data processing technique, we needed the time, GPS speed, and the channel 0 and channel 1 data sets. The same process was repeated for both the lateral and vertical data sets, and so the two channels represent the left with channel 0, and the right with channel 1. The following described process was written and executed through MATLAB. After the refined data set was

imported into the MATLAB workspace, each data set was trimmed to exclude sections of the data where the train was stopped. The data at low speeds around 2 mph or less was both generally inaccurate due to the broad-spectrum input of the returning signal (and therefore lower detectability), and also not of particular relevance to the identification of the points of interest because the train stopped and started well away from these points.

Due to the variable processing times and therefore the inconsistent sample rate of the Lidars, the average sampling rate of the data was collected to serve as the general sampling rate when calculating the cutoff for filters. Due to the mechanics of how the on-board Lidar system processes the incoming data, the processing time was not the same during each evaluation. There was a minimum processing time for the Lidar sensors, in this case limited to a maximum of 34 Hz, however if the time taken to calculate the velocity was longer than this minimum time, it was delivered and saved to the system at that slower rate. As such, the average sampling rate had an upper bound of 34 Hz and a significant portion which was calculated at some slower rate.

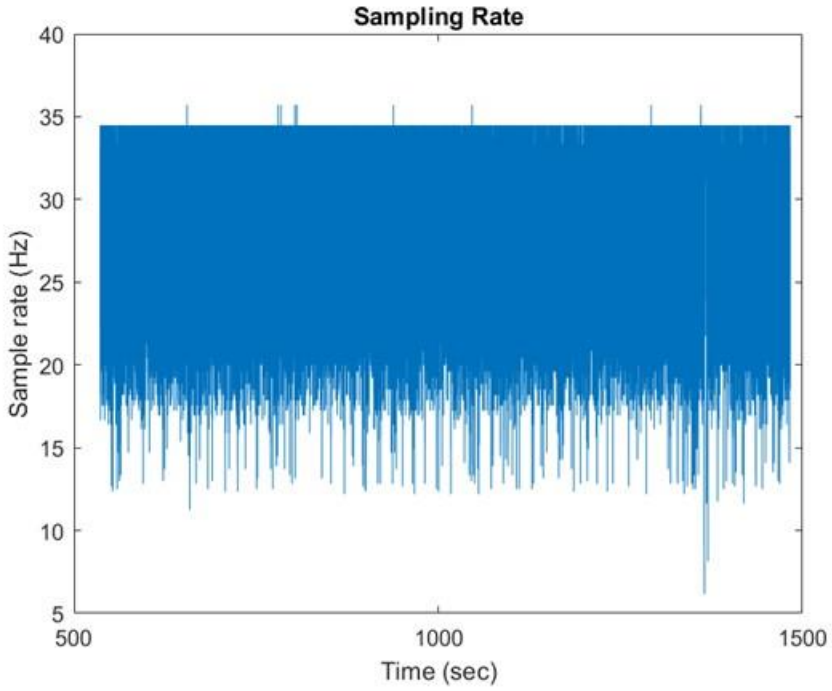


Figure 5-4: Average sampling rate for the second 10 mph Blackbeard data

The above example, from the second 10 mph run on the Blackbeard data, clearly depicts this upper sampling bound with lower frequency samples interspersed throughout the entire data set. The sampling rate, when calculated, shows that the effective sampling rate was lower than the prescribed sampling rate of around 34 Hz, but is actually closer to 27.9 Hz. This is calculated by averaging the inverse of the time difference between consecutive segments.

Due to the 10-degree angle of the Lidar lasers to the track, the data needed to be corrected to only represent the vertical or lateral motion of the rail. The purpose of the forward cant on the Lidar lasers was to impart a velocity offset on the lateral and vertical velocity data. Due to the homodyne

velocity acquisition, the offset imparted by the forward velocity of the train allowed the sensor to measure whether the motion was towards the laser or away from it. Both the lateral and vertical data sets were processed to remove this offset through the use of a high pass filter. Due to the length of time necessary to change speeds rapidly in the longitudinal direction, this high pass filter removed the forward component of the data and resulted in just the lateral or vertical component of that set. As seen in this example, the second 10 mph Blackbeard unfiltered data followed the GPS speed throughout the run.

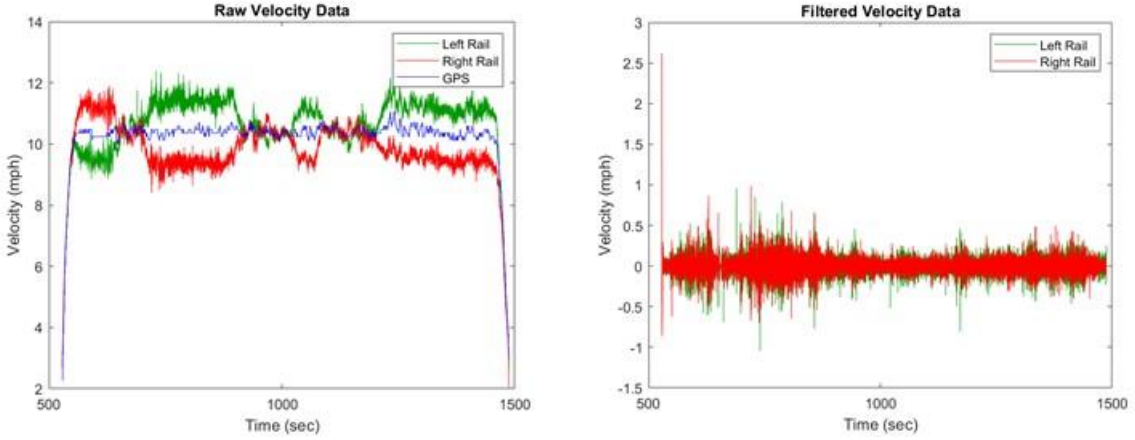


Figure 5-5: Velocity vs time from second 10 mph Blackbeard trial, raw (left) and filtered (right)

A 3rd order Butterworth high pass filter with a cutoff frequency of 0.318 Hz was selected to remove the longitudinal velocity of the train at 10 mph. This frequency corresponds to a track length of around 46 feet. This removed slow changes in speed such as the acceleration and deceleration of the train, and left us with the faster effects of the rail as measured by the lateral and vertical Lidar sensors.

Once the data had been trimmed and filtered, it needed to be divided into the segments necessary for the STFT. One consideration when dividing the segments for analysis was the level of overlap between each segment. This, with the combination of the choice of segment size, determined the data division. In the case of all of the TTC trials, an overlap of 50% was implemented. This overlap choice was settled on because the size of the segments was generally kept small irrespective of speed, and so excess overlap could produce less readable results due to apparently slowly changing conditions. The 50% overlap also ensured that every piece of data was included in two of the segments, and wouldn't be lost if a point of interest was bisected by a segment division. An example of the split data segments can be seen below in Figure 5-6, with the first example using a segment length of 1024 and the second a length of 256 on the lateral data from the second 30 mph run after it was processed through the high pass filter.

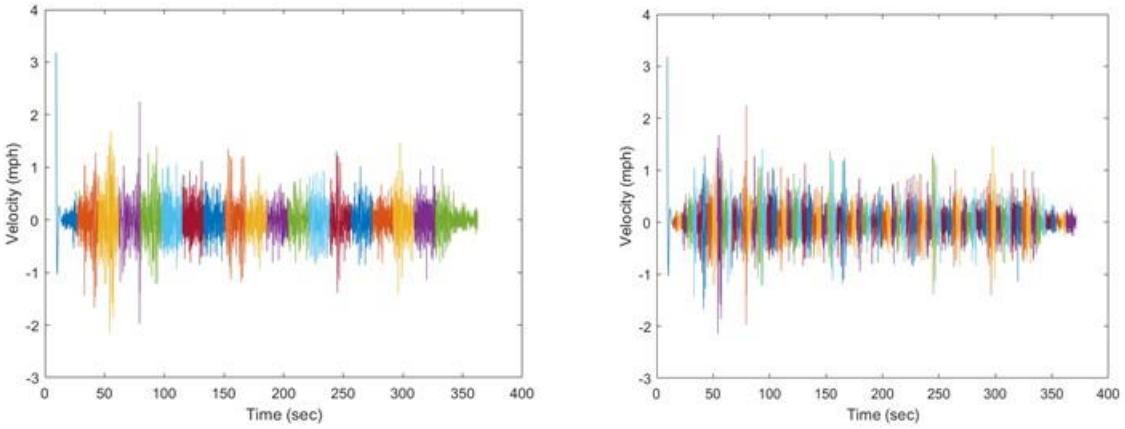
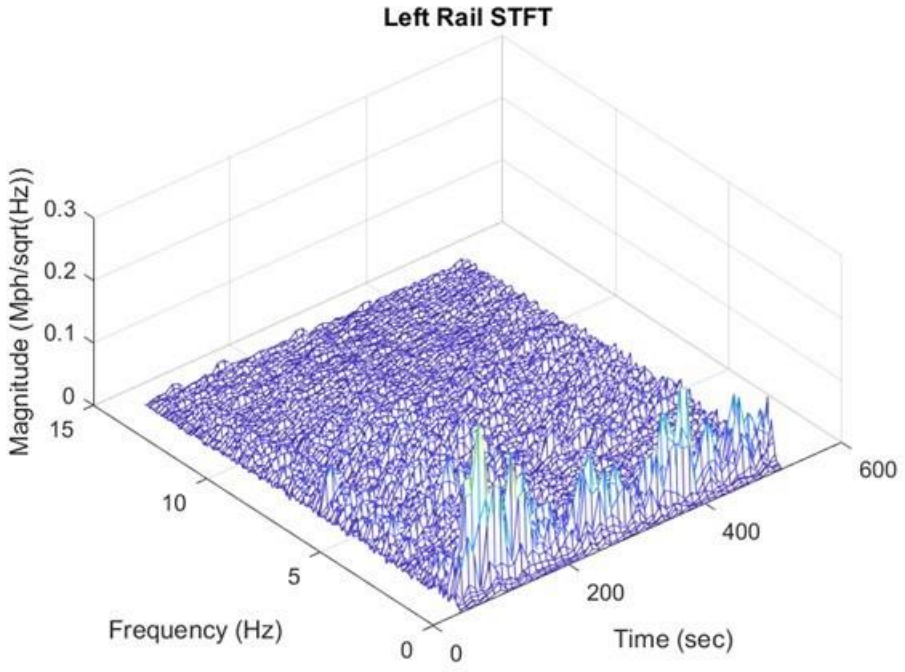


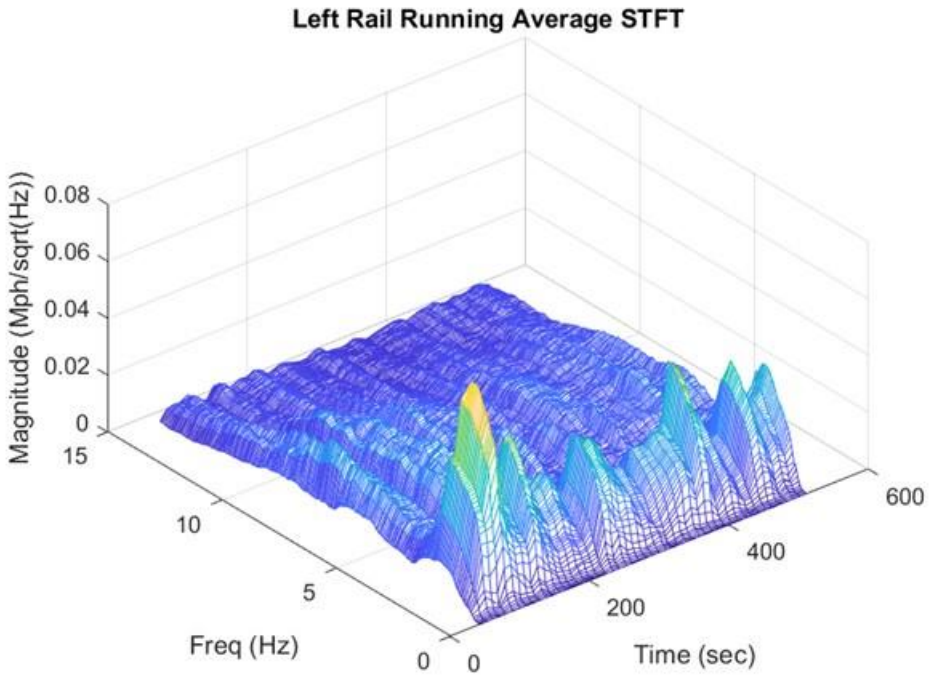
Figure 5-6: Segments division for a size of 1024 (left) and 256 (right)

In order to identify the locations which contained points of interest, a few different methods were implemented. The first of these methods was to create a threshold which would act as a reference to define when a component of the frequency stood out from the noise. In RADAR detection, the threshold for detectability is defined by a 6 dB for a returning signal, and as such was chosen to be the starting threshold to apply for this trial [28]. No standard dB threshold was prescribed for determining whether a component of the signal indicated a resonant frequency, but the 6 dB threshold exceeded the noise of the signal which was about 3 to 4 dB above the linear regression of the data.

This approach identified locations which contained harmonic components that deviated from the normal regression of the FFT output. While this doesn't provide a method of differentiating between locations or selecting for changing conditions, it does illustrate points in the data which contain a clear harmonic component. Several steps were used to get to the point of monitoring whether a signal had a component that cleared the detectability threshold. The first step in this process was to conduct a running average filter over the decoupled data. This running average allowed for a better regression line to be implemented in the further steps, as well as to clarify the shape of the STFT mesh for manual inspection if necessary. The running average filter was a 30-point moving average filter, and acted like a low pass filter, which helped to clean up some of the frequency noise present.



a.



b.

Figure 5-7: Left rail lateral velocity (a) Short Time Fourier Transform and (b) running average STFT

A linear regression line could not be applied directly to the resulting full FFT, since the shape of the power spectrum is clearly not linear. One of the assumptions made in this evaluation was that the shape of the FFT spectrum could considered linear after its "shoulder" location. As seen below, the linear regression line accurately reflects the shape of the frequency output.

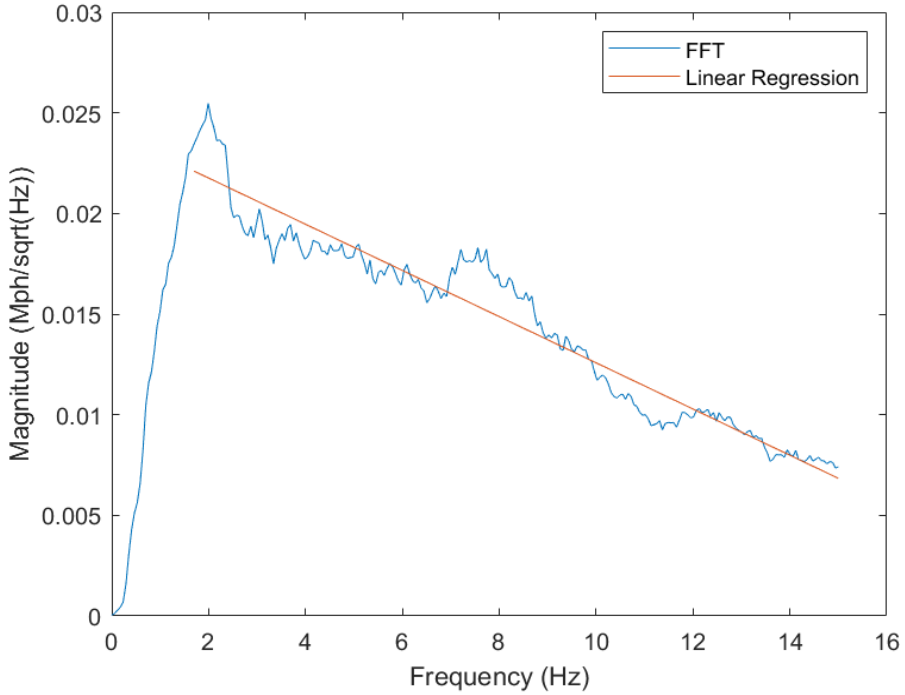


Figure 5-8: Linear regression line for a sample FFT segment

This linear regression line was then raised above the starting position by 6 dB, which reflects the doubled amplitude of the signal. This raised linear regression acted as the threshold to determine whether a response within the power spectrum was sufficiently differentiated from the background frequencies. This threshold detection method works for a signal with a sufficiently narrow frequency range within the spectrum, as too broad of a frequency distribution would impact the linear regression line. With this raised threshold, if a part of the signal is greater than the raised regression line, it can be claimed to be distinct and detectable.

This detectability only refers to one particular form of track motion which was found through the STFT, which is that of a harmonic excitation of the track beam. If an excitation in this form was found, it would indicate that the track was being excited at or about a natural frequency and that the damping was not sufficient to attenuate the motion. While developing this process we used examples from prior data sets collected by the Lidar system. One of the runs, the data collected on our study from the Norfolk Southern 2019 Hyrail test demonstrated a clearly detectable harmonic component. As can be seen in Figure 5-9, the STFT results exceed the threshold developed by the linear regression curves, and even without such a threshold are clearly apparent.

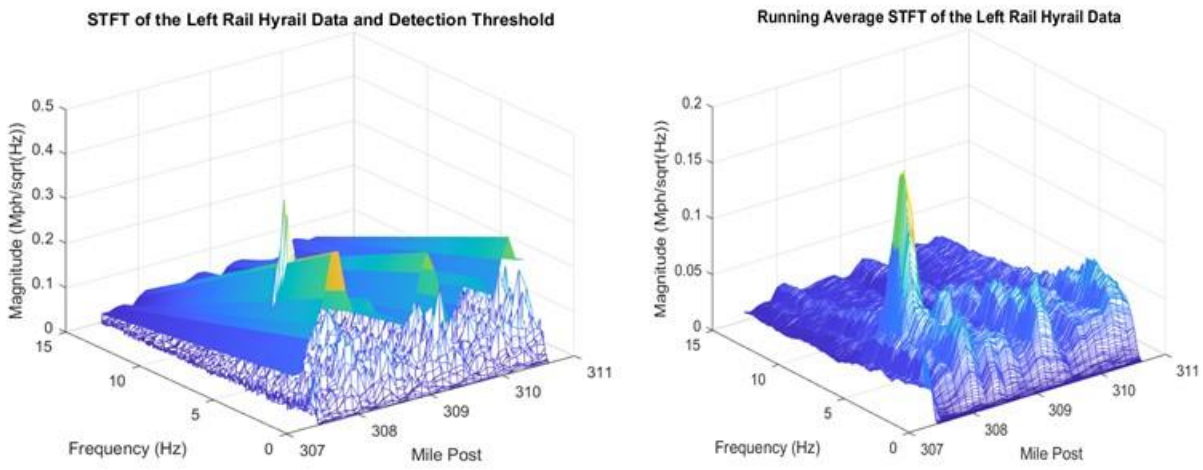


Figure 5-9: 2019 Hyrail test STFT with detectability threshold (left) and STFT running average (right)

This result lines up with the section of track which was previously manually selected from a visual inspection of the data to be loose or unstable. In addition to demonstrating the purpose of this approach, the example above can also provide a clear demonstration of the tradeoff between the frequency and positional specificity when determining segment size. If, for example, the segment size chosen was doubled, the harmonic would still exceed the threshold, but its location would decrease in precision. From a manual inspection of the raw velocity data, the harmonic at this location persisted from mile 307.946 to mile 308.06. At the segment size of 1024, the threshold was exceeded from mile 307.86 to 308.066, while at the doubled segment size of 2048, the threshold was pierced from mile 307.676 to mile 308.132. The segment size of 1024, therefore, overestimated the location by around 1.8 times while the segment size overestimated by 4 times. These expand the range of detected track from 600 feet to around 1090 and 2410 feet respectively. This approach was intended to be used in conjunction with adjusting the segment sizes to find the optimal clarity of deviating frequencies.

Of course, if there did exist some effect which occurred over a broad range of frequencies, or that the effect of note was transient, analysis of a single segment of FFT data would not produce any useful evidence towards that fact. For example, the broadest frequency input would be that of an impulse, which when transformed into the frequency domain would convert a unit impulse into a frequency response of unity. Hence, a large impulse or several smaller impulses in the segment would result in a broad raising of the FFT response in comparison with its surrounding segments, but would not change the distribution of that FFT curve. As seen in this example from the 2012 data collected from the Wayne to Chicago test, a pair of broad frequency responses can be clearly seen in Figure 5-10.

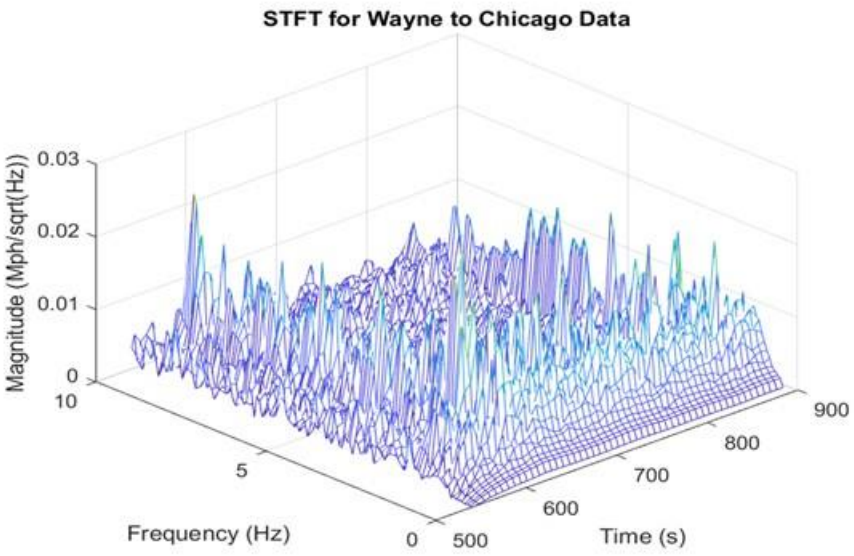


Figure 5-10: Wayne to Chicago STFT presenting the result of impulses

Additionally, if a harmonic result is found within a segment, but that harmonic remains relatively unchanged throughout the run, as in this example from our 2011 gauge corner testing in Figure 5-11, this could be the result of a persisting mechanical effect present on the entire track or on the train itself.

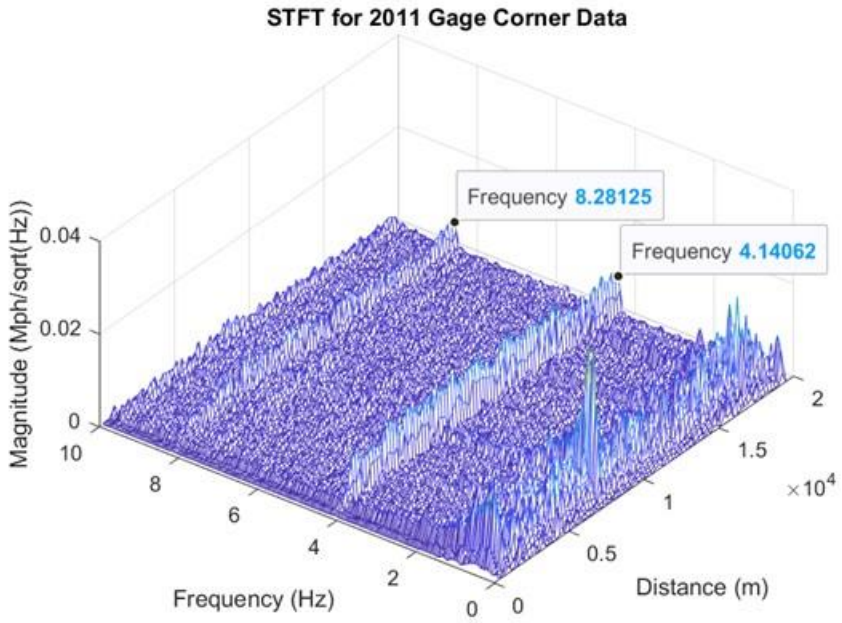


Figure 5-11: 2011 Gauge corner testing STFT depicting a persistent pair of elevated frequency responses

If the train were to rock, or to have uneven wheels, these could be conditions of dynamics on the train which would result in this effect. We can also see that these excitations are at 4.14 and 8.28 Hz respectively, and, given that one is an exact multiple of the other, they might indicate that there is an aliased effect presenting itself as harmonics.

5.1.2 Non-Point-of-Interest Locations detected by Rail Inspection

The processes described above were then conducted on the collected data from the testing at the Transportation Technology Center on their High Tonnage Loop. From these trials, we did not see any results which contained harmonic components in excess of the detectability threshold of 6db above the linear regression. Additionally, reducing this threshold value did not yield useful results as no set of data contained a response narrow enough in frequency. While this was the case, it still proved useful to manually review the results and observe and record their effectiveness at identifying the loose rail. From this process, several locations stood out as consistently differentiated throughout the HTL. Several of these locations were seen equally through both the Blackbeard and Corsair systems, while some of them were more clearly apparent with regards to their intensity and distinction in just one. The clarity in one direction of detection over the other came from the interaction between vertical, lateral, and orbital translations of the rail head.

Of the sections identified which were not the “loose” rail, the most reliably detected location was the switch occurring after the curve coming out of section 40. This switch could be easily seen in the STFTs of both Blackbeard and Corsair data for all speeds and for each trial. It was one of the most prominent peaks in all of the STFT meshes, and could be found at any reasonable tradeoff between segment size and frequency precision. The frequency distribution also was relatively broad, indicating there may have been impulses in that section, which would raise the energy without exciting any particular frequency. This explanation would also help explain how the location was clear at all speeds, since the higher speeds may impact with higher force, and therefore raise the energy to match, along with the impacts of impulses not being affected to the same extent from the spatial sampling rate. In addition to the clear broad-frequency response at the switch, and most apparent in the Blackbeard data, there seemed to be an additional low frequency response at around 3.4 Hz.

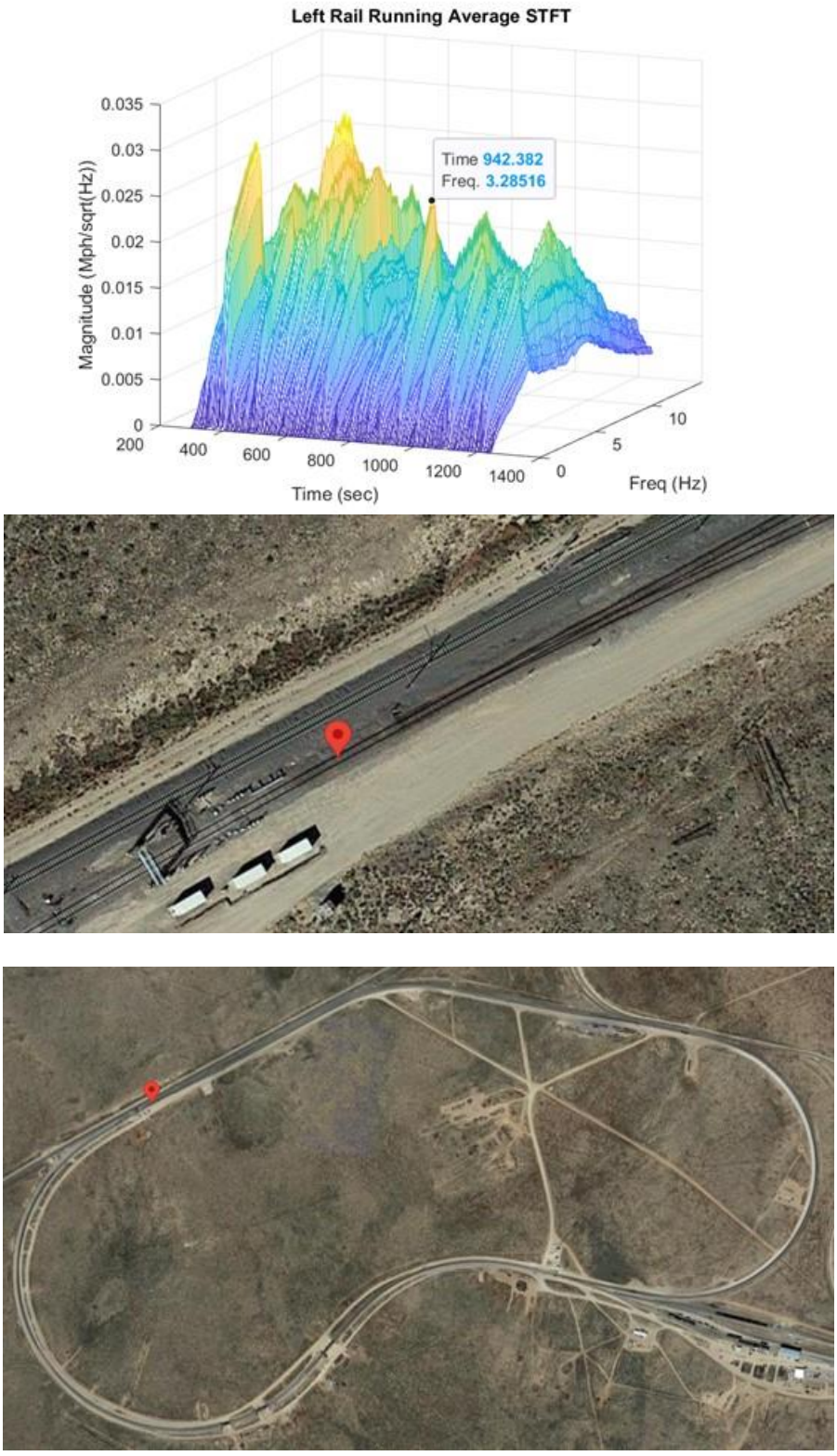


Figure 5-12: Lateral vibrations STFT for the third 10 mph trial showing the location for the switch in section 35 along with the corresponding GPS location of that point

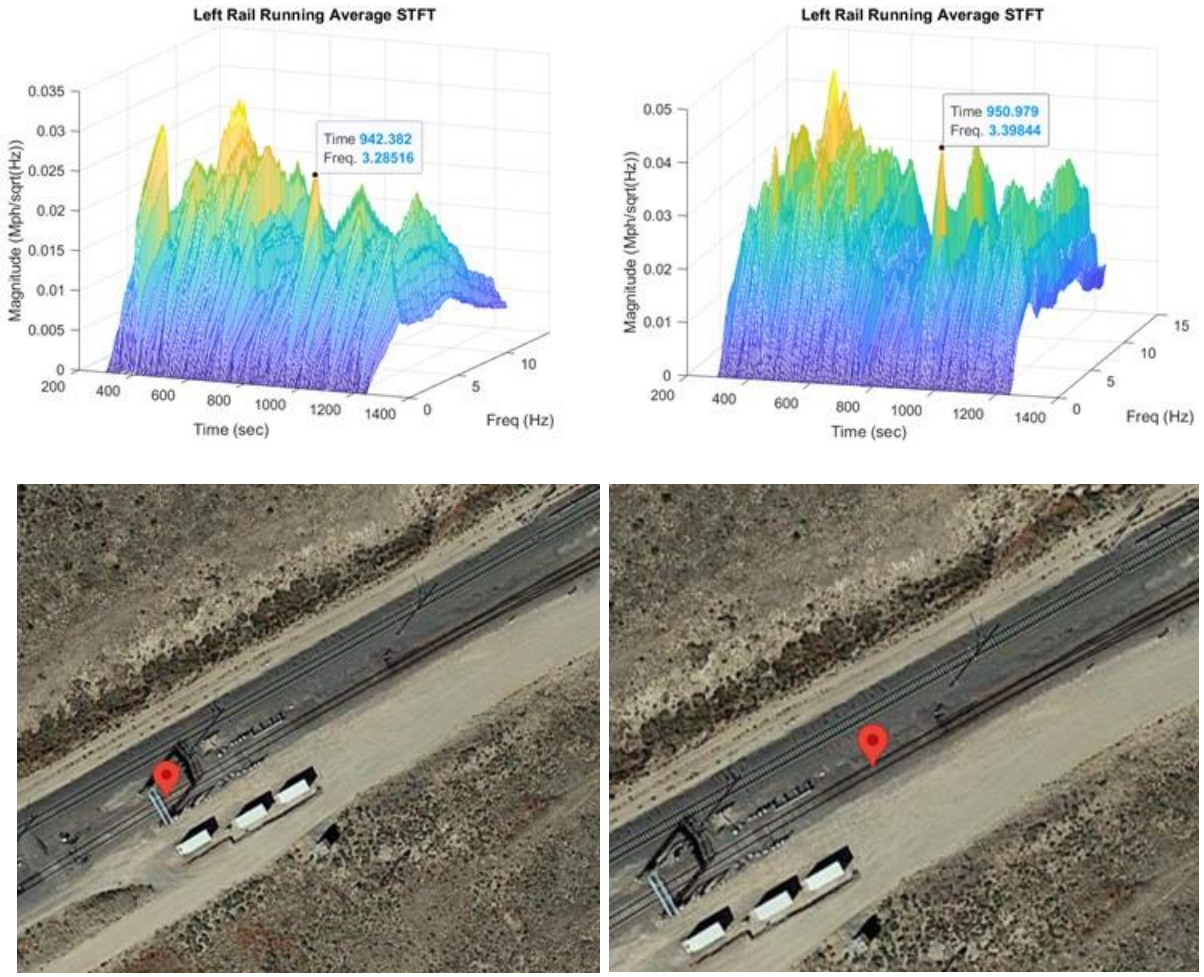


Figure 5-13: Blackbeard (left) and Corsair (right) switch STFT of the left rail data for the third 10 mph trial along with their GPS locations

In addition to the switch in section 35, several other features of the track were detected with relative consistency. The lubricator was detected in all trials of the Blackbeard data and in the majority of trials in the Corsair data. As with the switch, the response of the signal to the lubricator involved a broad-frequency elevation of energy. The lubricator is located in section 24 of the HTL and the same explanations as with the switch apply to the form the responses took here. The largest distinction was that the intensity of this signal was not as significant as the switch, and was therefore not detected in every trial, but was found in around 90% of trials in the lateral data, and around 60% of trials in the vertical data, with a sharp drop off at higher speeds.

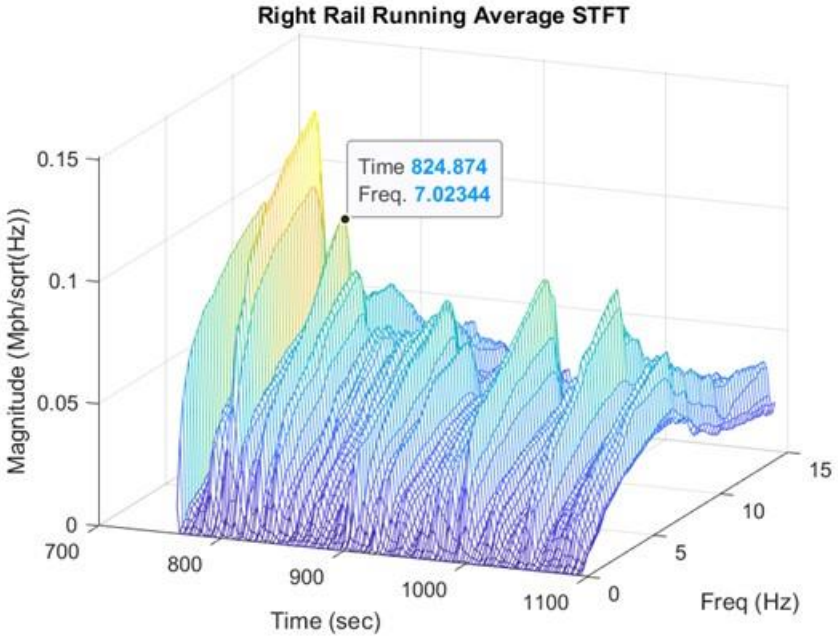


Figure 5-14: Lateral lubricator STFT response in the right rail of the third 30 mph trial

Figure 5-15: GPS location of the Blackbeard lubricator response

Figure 5-16: Location of the lubricator on the HTL

The final non-point-of-interest locations that were found when analyzing the data were the beginnings and ends of the curve spanning section 25, the curve in section 7, and at the midpoint of section 3. All of these presented relatively similarly to each other. Taking the second 20 mph Blackbeard run as an example, we can see how all of these curves did not extend as broadly across the frequency spectrum as the other signatures, and tended to fall off around 7 Hz. Figure 5-17 through Figure 5-20 shows how in the Blackbeard data, this seemed to persist as the cutoff frequency for each speed, and the 10, 20, and 30 mph results each show this relatively steep drop-off. In contrast, the 40 mph runs did not depict this same drop at higher frequencies, but this may be explained by the much lower frequency resolution that comes with the smaller segments. In comparison to the 256 and 128 segment sizes of the slower speeds, the 40 MPH runs were analyzed at segment sizes of 64 to maintain distinction in the locations, but as a result it was much more difficult to identify changes in the frequency distribution.

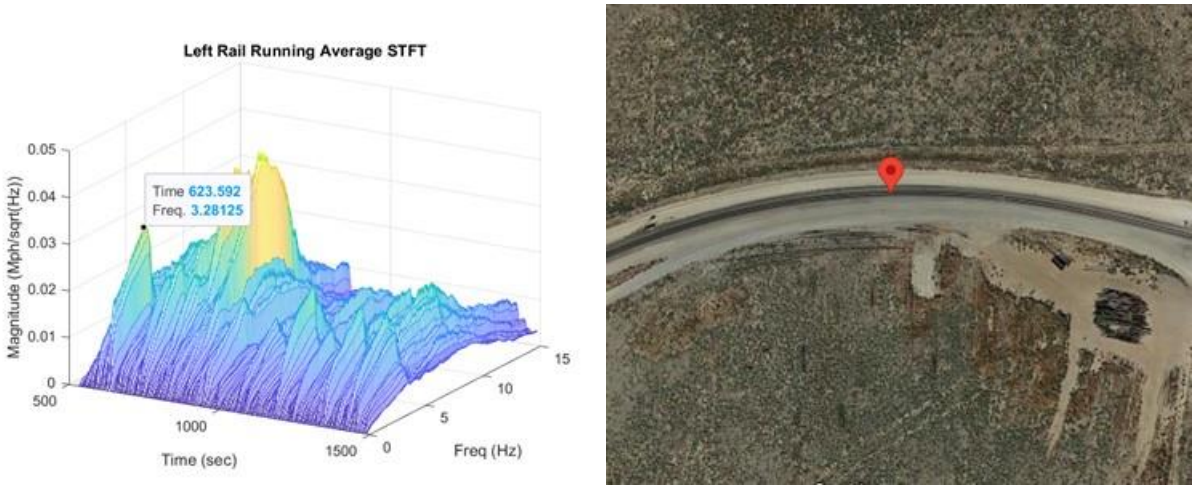


Figure 5-17: Section 7 lateral left rail STFT response and location in the second 10 mph trial

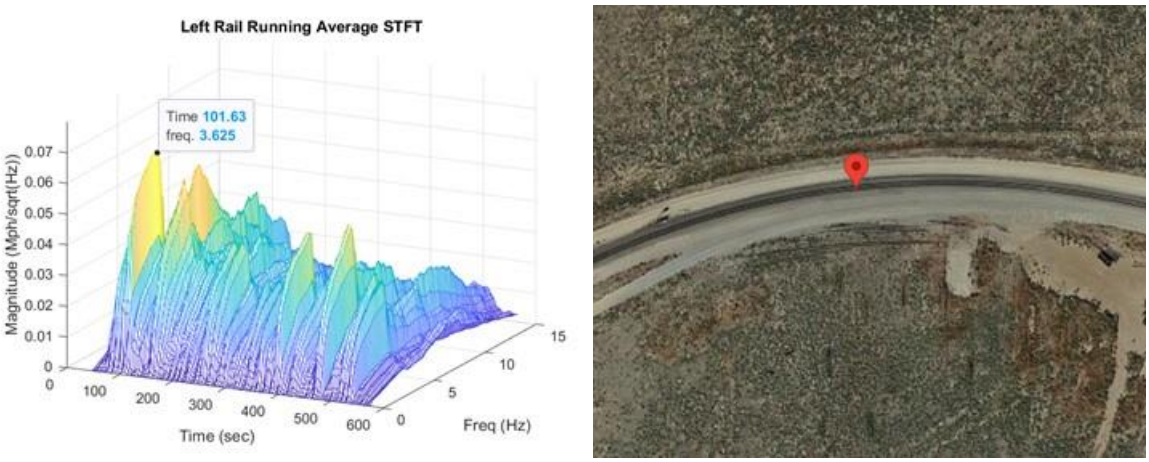


Figure 5-18: Section 7 lateral left rail STFT response and location in the second 20 mph trial

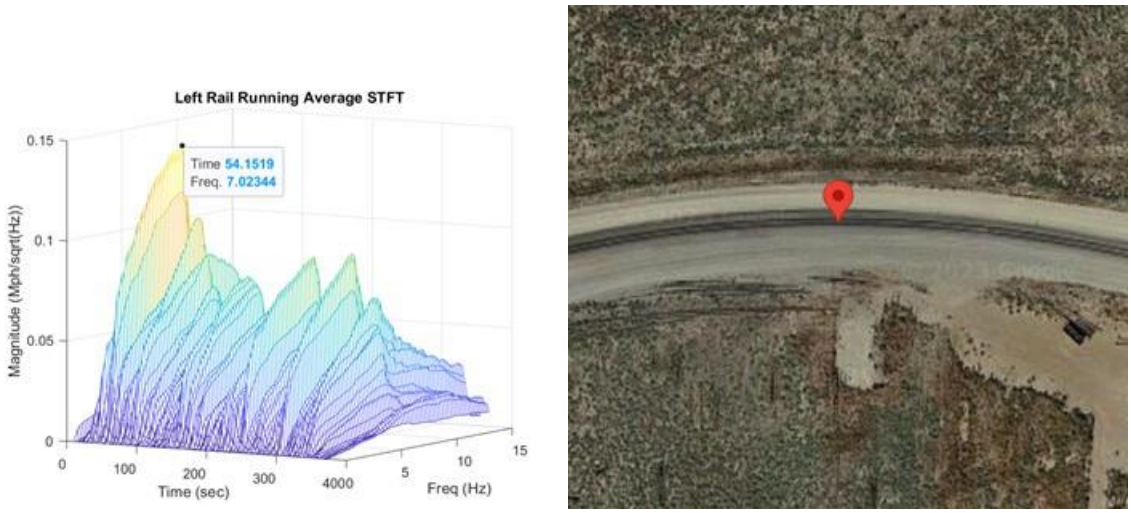


Figure 5-19: Section 7 lateral left rail STFT response and location in the second 30 mph trial

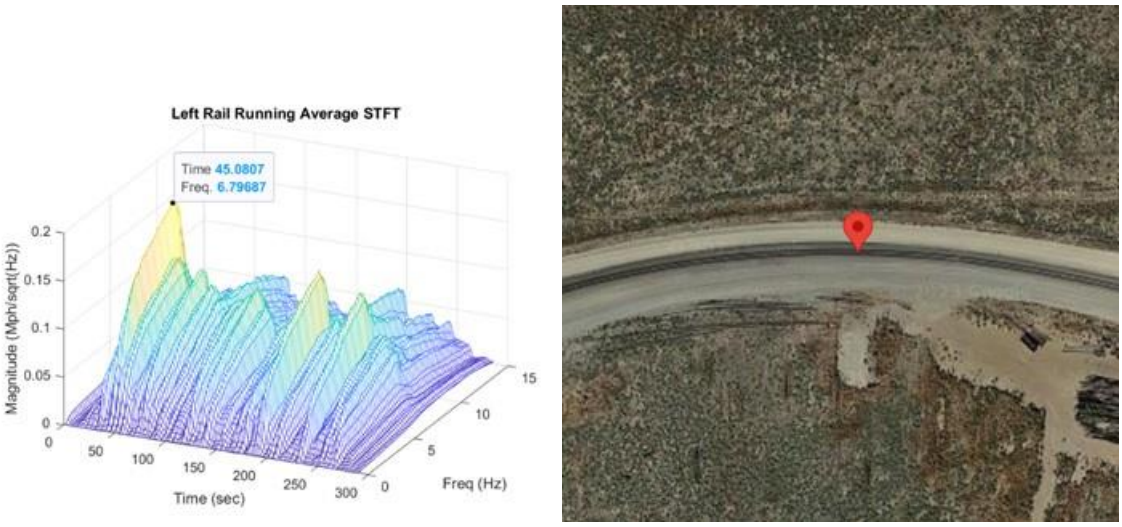


Figure 5-20: Section 7 lateral left rail STFT response and location in the second 40 mph trial

Figure 5-21: Location of these section 7 STFT responses in the HTL

In the prior figures, the emphasis was placed on the curve within section 7, as it was the most discernible. The remaining curves of note have been made clear in
Figure 5-22 where the responses in sections 24 and 3 have been labeled along with their corresponding locations.

Figure 5-22: Locations of the section 3 (lower left) and section 24 (lower right) lateral responses

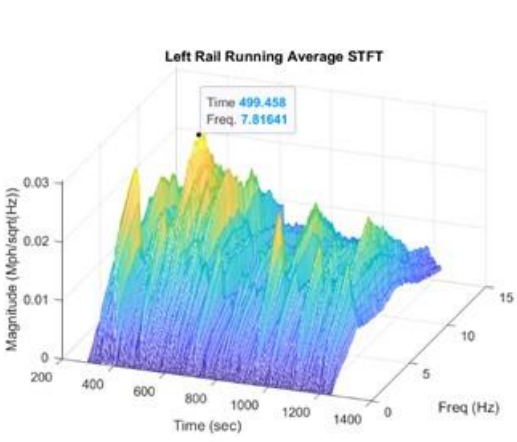


Figure 5-23: Section 24 left rail lateral response and location in the third 10 mph trial

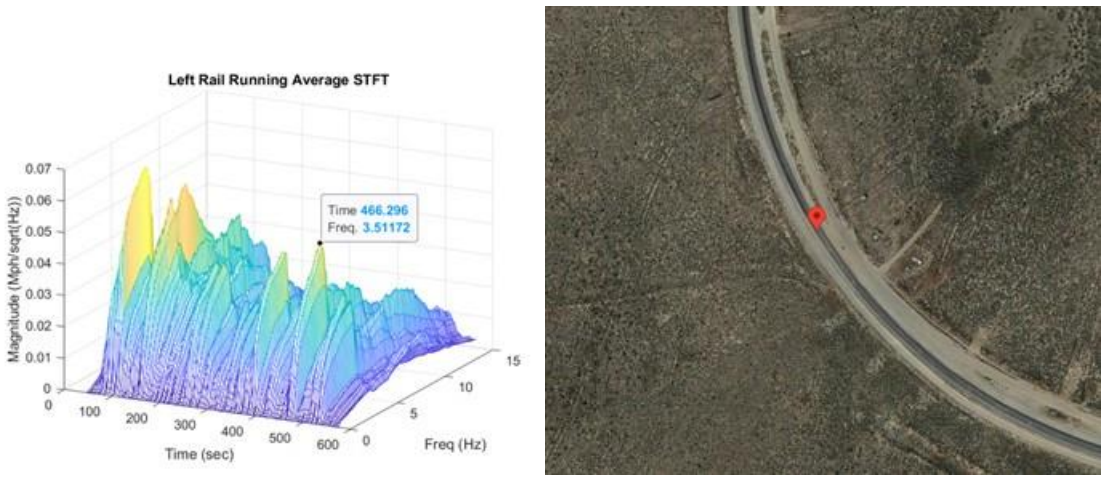


Figure 5-24: Section 3 left rail lateral response and location in the second 20 mph trial

5.1.3 Curved Point-of-Interest Detected by Lateral Rail Inspection

Finally, with regards to the identification of the points of interest, the first result of note is the lack of detection for either the control or the emulated loose track in the tangent section 40. There was no frequency response change, either in the distribution or on the elevation which was noticeable in the STFT mesh. The section 25 response, on the other hand, provided some evidence that the use of the Fourier Transform could identify looser sections of rail. While not found in every trial, the center of the section 25 curve, which corresponded to the ties with removed vertical spikes, was identified in the majority of both the Corsair and Blackbeard data sets.

The form this frequency response took was that of a generally elevated response with a peak frequency around 5-8 Hz. This response in the Blackbeard data, however, was most found in the Channel 0 response, which corresponded to the lateral movement of the left rail. The Channel 1 Blackbeard responses did not contain the same Fourier response, which indicates a difference in the motion of the rails. If the result was a function of the freight car vibrating or moving, the responses should have been reflected in both tracks.

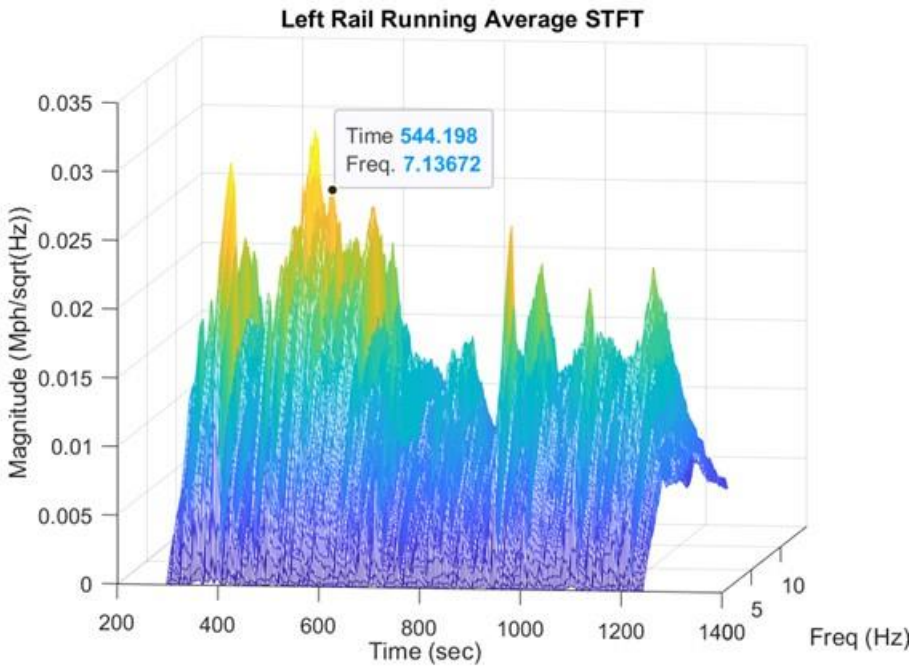


Figure 5-25: Third 10 mph section 25 lateral loose rail response in the left rail

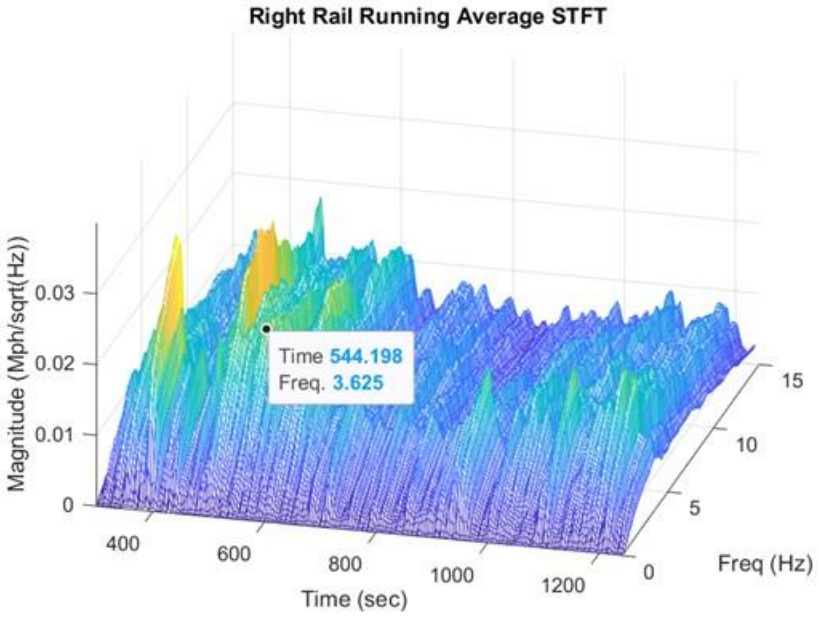


Figure 5-26: Third 10 mph section 25 lateral loose rail response in the right rail

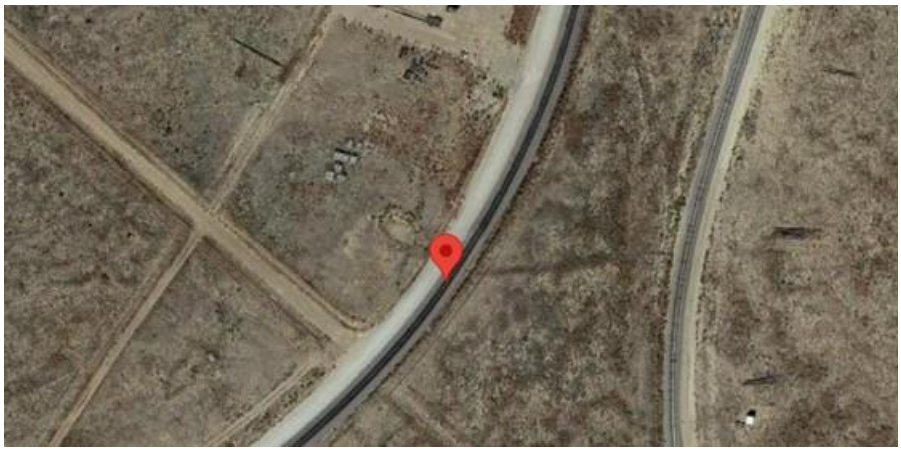


Figure 5-27: GPS location of the section 25 loose rail for the third 10 mph lateral response

One of the challenges of using the Blackbeard frequency response to identify the loose rail was that the level to which its location stood out compared to the surrounding effects varied. In some cases, such as the fourth 10 mph run on the left rail, the largest response was that of the point of interest. In others, such as the fourth 20 mph run, the response of the loose rail did not stand out, or was overshadowed by other phenomena.

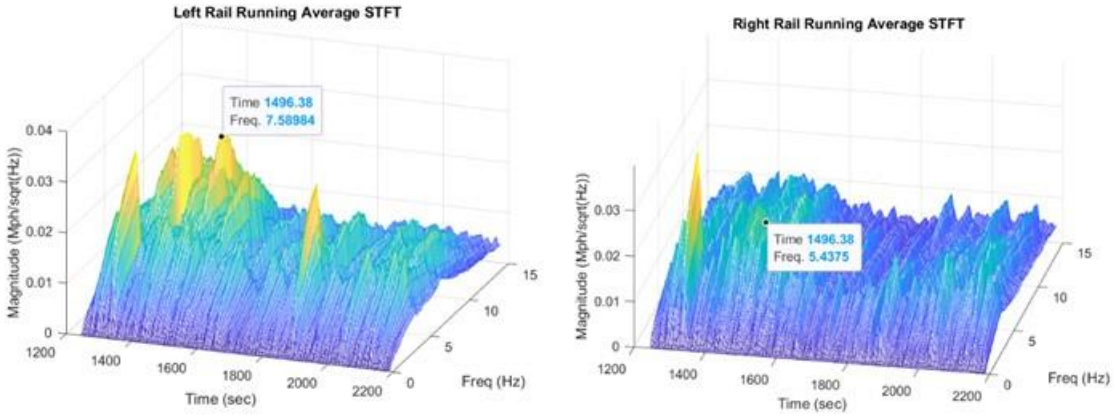


Figure 5-28: Fourth 10 mph section 25 loose rail lateral response in the left rail (left) and the right rail (right)

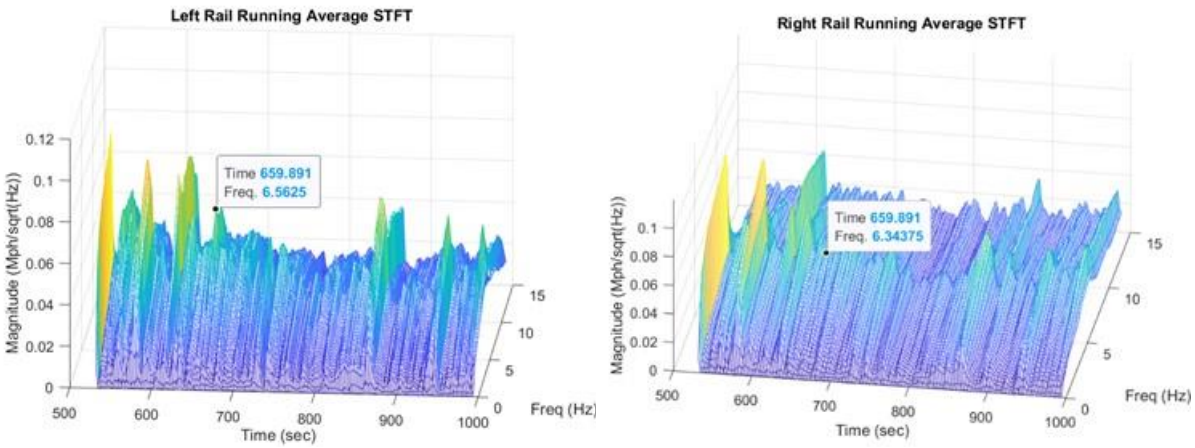


Figure 5-29: Fourth 20 mph section 25 loose rail lateral response in the left rail (left) and the right rail (right)

5.1.4 Curved Point-of-Interest Detected by Vertical Rail Inspection

By comparison, the Corsair data could identify the point of interest with difficulties only at slower speeds, while at higher speeds it lost its distinction. This indicated that there was some speed dependency on the detection of the points of interest for the Corsair data. In contrast to the Blackbeard data, the trend among Corsair results was that the point of interest was identified in the Channel 1 data as opposed to Channel 0. The greatest vertical response for this section, then, was found in the right rail. Combining this with the previous Blackbeard observations suggests that there might be a counteracting vibration with its response dominating the orthogonal vector.

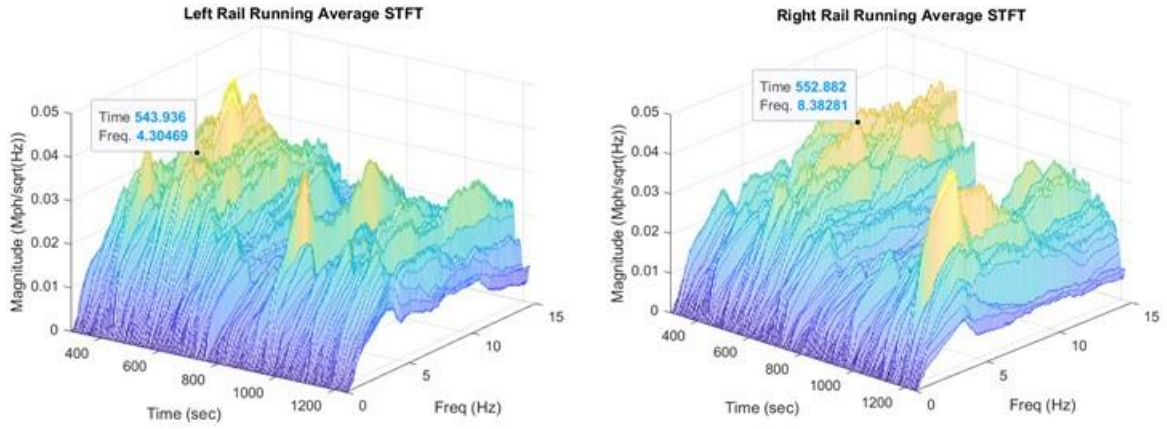


Figure 5-30: Third 10 mph trial section 25 loose rail vertical response in the left and right rails

Figure 5-31: Locations of vertical response in the left and right rail for the third 10 mph trial

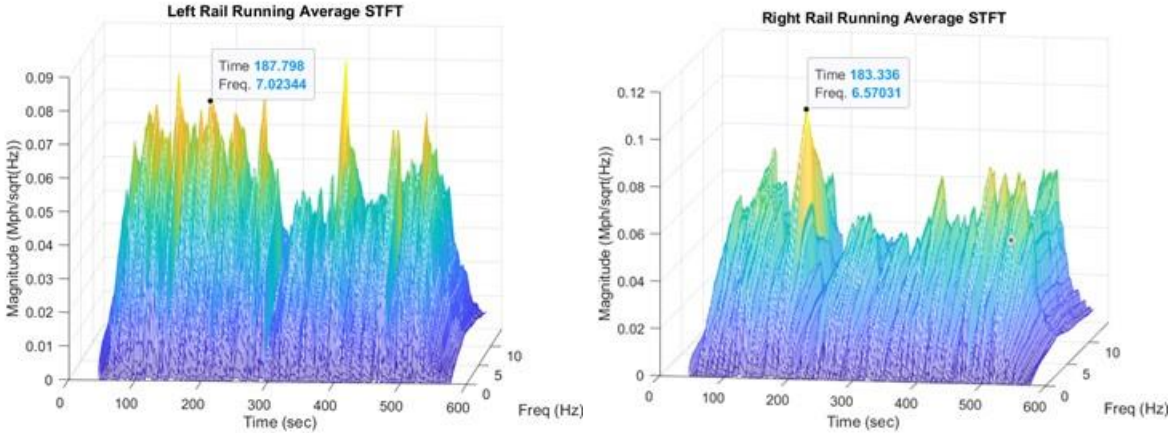


Figure 5-32: Second 20 mph trial section 25 loose rail vertical response in the left and right rails

Figure 5-33: Locations of vertical response in the left and right rail for the second 20 mph trial

The vertical data was unable to detect the loose track at a speed between 20 to 30 mph. The 30 mph results did not provide enough distinction to select even a set of larger peaks which contained the point of interest. Selection of the loose rail could only occur with prior knowledge of its location, and as such has little use for the purpose of using the Doppler Lidars to detect loose track.

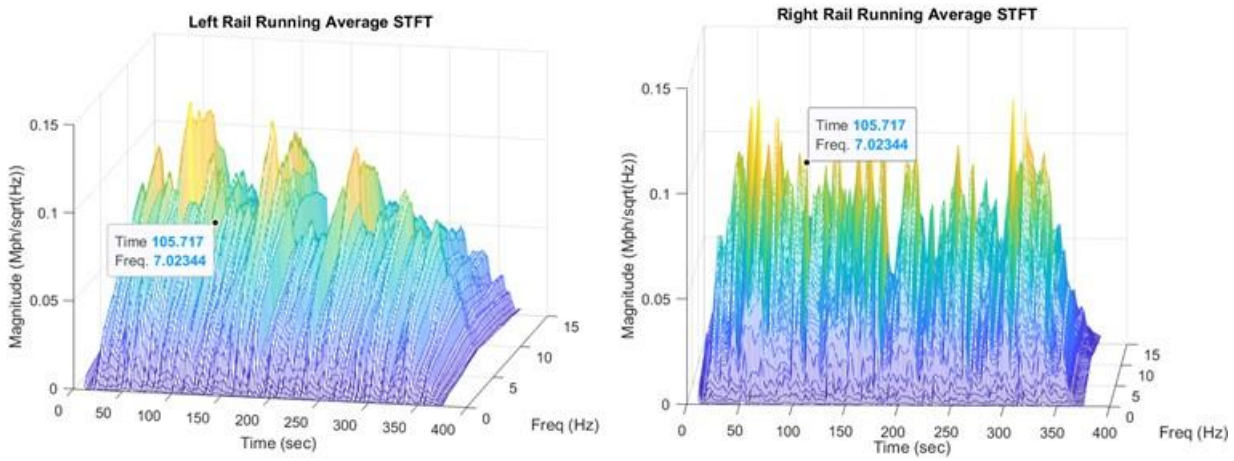


Figure 5-34: second 30 mph vertical rail response in the left and right rail

Overall, while not consistent in both channels or at all speeds, there is potential for the identification of loose vertical tie connections in curved rail. It did not meet an automated detection threshold, and many other locations on the track produced equivalent or larger responses in the STFT. Additionally, there appeared to be a connection between the lateral and vertical motions of the track for loose vertical tie connections around curves. Many issues still exist with the data collection and identification of track conditions to reach the point of unsupervised sorting and

detection, but this current test at least proved that curved “loose” track could be distinguished through the use of non-contact Doppler Lidar sensors.

The issues presented could be a result of several factors in the data collection process. One of these factors could be indicated by the elevated frequency response across the entirety of section 25. This section had a greater vibrational response in general, and as such the distinction of the point of interest within it could be a result of the variances in that elevated response. The counterpoint to this explanation, however, is that the train traveled around several curves on the track, and only this section seemed to result in this phenomenon. The other curves are also of a similar degree of curvature being 5-degree curves and were traveled on at similar speeds. From the comparison of the curved track containing section 25 and the curved section after the tangent track around section 3, we see that the speed difference was less than half a percent for the second 30 mph trial. Therefore, it seems unreasonable that the change in frequency response was due solely to the travel around the curve.

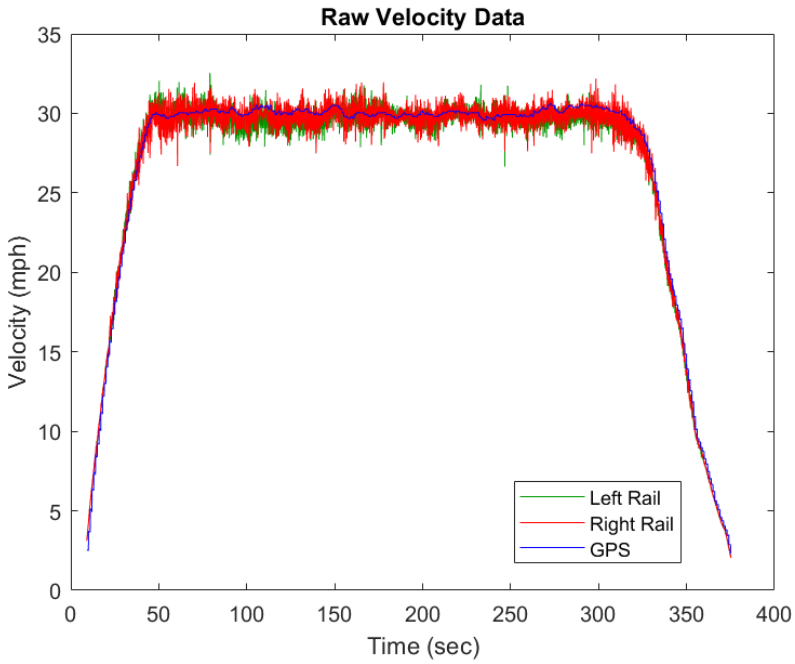


Figure 5-35: Left and right raw lateral velocity from the second 30 mph trial

At higher speeds, however, the clarity of the distinction was reduced. Specifically, the clarity of the location of the elevated frequency response was reduced by both a broadening and shifting of the overall elevation in the frequency curve as well as by the increased frequency responses of the segments surrounding our point of interest. Given the faster speeds, the time which the sensors measured the points of interest on the track was reduced. These fewer points served a twofold purpose in the difficulty of identifying the sections of loose track. The first of these is that, with the same segment size as the slower speeds, a significantly smaller proportion of the segment contained the loose track, which would manifest itself by suppressing the harmonic responses. In order to counteract this first issue, the segment sizes for the higher speeds were reduced, once again returning the proportion of the segments containing loose track to a similar level. The second

apparent issue is derived from the solution to this first; which is the reduction in the precision of the frequency content. With a smaller data set due to the shorter segment sizes, it became harder to identify the frequencies at which the responses were occurring, and at the highest levels, it became difficult to differentiate from the low frequency content.

In addition to these data processing difficulties arising from the reduced time of measurement, the track conditions themselves might change and produce their own negative impacts of the detectability of the loose sections of rail. Depending on the mechanism which excites the loose rail, the speed of the train may not play a large factor in the vibration or displacement of the track. Especially in the vertical direction, as the Corsair systems measures, the speed of the train around the track should mostly impact the speed at which the track is allowed to return to its uncompressed form. In comparison, other track motions and responses may present themselves more prominently with the greater energy being put into the tracks from the higher speeds. Since the exact method by which the rail was excited is unknown, the source of the apparently changing frequency response of the track remains inconclusive.

Overall, the Short Time Fourier Transform allowed for the inspection of the track's condition through a manual review of the collected data. At lower speeds, the curved point of interest was able to be selected from the resulting frequency mesh, but only so in some of the channels of the data. At higher speeds, while possible, detection relies more greatly on the prior knowledge of the section of interest and would require some way to remove the responses indicative of other conditions. Overall, this method works to identify the section of loose rail in the curve, but also produces a large number of false positives.

One limitation of analyzing the Lidar data through the use of an STFT is that it mostly detects harmonic motion, and so repeated patterns of motion that do not fall within that shape of motion would not be found. As an alternate exploration of the data in the frequency domain, the application of the Wavelet transform allows for the inspection of different shapes of motion than from sinusoidal motion.

5.2 Implementation of the Discrete Wavelet Transform

5.2.1 Background

Another way unstable track could present itself is as a transition from "soft" to "firm" track or vice-versa. Rather than looking at the STFT representation of the data, which involved shifting from shorter and longer segments to tradeoff between the spatial and frequency resolutions, a wavelet transform was used to approach from a multiresolutional perspective. The wavelet transform is bounded by the same uncertainty principle as found in the STFT, however it approaches the problem through a different decompositional lens. In the STFT approach, with a given segment size, the temporal and frequency resolutions are the same across all time periods and frequencies. With the wavelet transform, this tradeoff is manipulated to reflect a greater location resolution for high frequency motion and a greater frequency resolution for low frequency motion, as illustrated in Figure 5-36 [29].

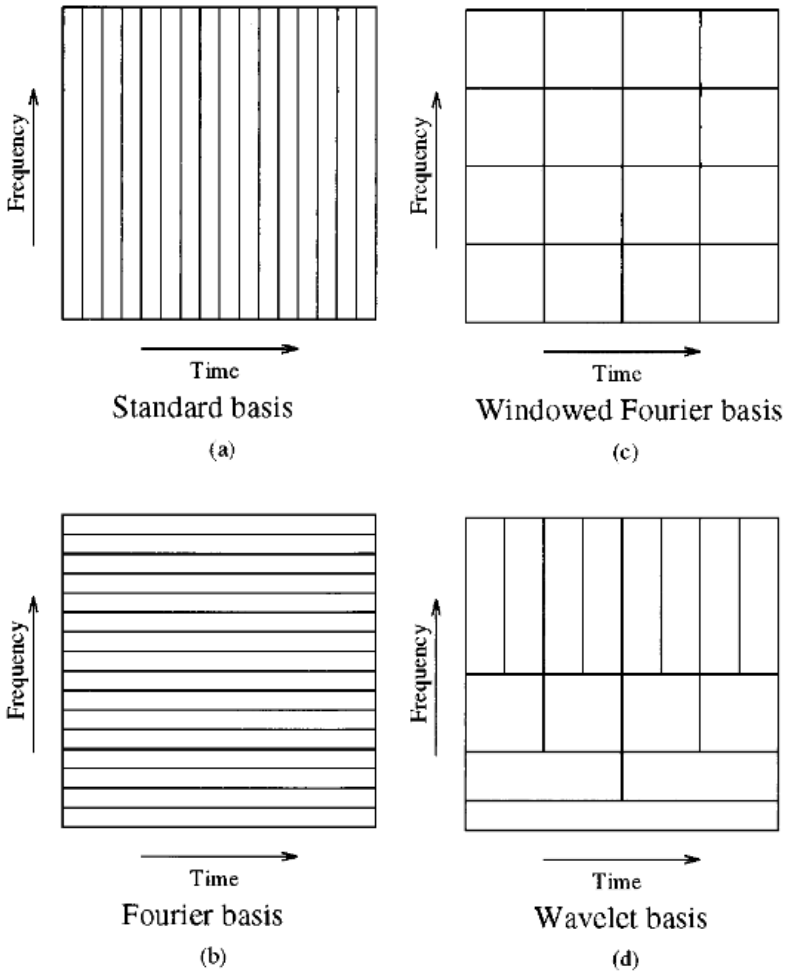


Figure 5-36: Time-frequency decomposition for: (a) standard basis, (b) Fourier basis, (c) Short Time Fourier basis, and (d) wavelet basis [30]

Wavelets have been investigated for their use in structural analysis with regards to detecting locations of damage as a function of time [31], and can be used to reflect the characteristics of both longer lasting low frequency content and shorter lasting high frequency content. This exploration into the evaluation of structural damage on the railroad tracks seems appropriate for investigation.

The wavelet transform operates through the filtering of the signal with a convolution of a wavelet. This convolution works through the selection of a mother wavelet which acted as the base of the transformation. In the case of the MATLAB implementation of the Discrete Wavelet Transform (DWT), the chosen mother wavelet must be either orthogonal or biorthogonal. The MATLAB database includes several examples which could be used in this process. Once the mother wavelet has been selected, the wavelet decompositions use the mother wavelet to determine the coefficients used in the filtering of the data.

Each of the Lidar data signals each act as an input data set for this transform in the form of a data vector. This data vector was split via a pair of filters based on the wavelet coefficients of the mother

wavelet. This pair of filters acts as a quadrature mirror filter, which works to either smooth over the data like a low pass filter or to isolate the details of the signal similar to that of a high pass filter [29]. The selection of the mother wavelet determines the coefficients of the high pass filter component, h, and the low pass filter component, g, which must be spatially localized and orthogonal to each other. The coefficients for g are based on those for h and are given by the relationship of:

$$g_k = (-1)^k h_{n-k-1}, k \in \{0, \dots, n-1\} \tag{5-1}$$

The length of these sets of coefficients is a function of the chosen mother wavelet. In all of the mother wavelets capable of performing a discrete wavelet transform, their lengths are a multiple of 2, as the discrete wavelet transform iterates across the data vector by incrementing by two points at a time. An example of the coefficients present in one of these mother wavelets is that of the Daubechies 2 (db2) coefficients. These sets of coefficients are:

$$h = [\frac{(1-\sqrt{3})}{4\sqrt{2}} \frac{-(3-\sqrt{3})}{4\sqrt{2}} \frac{(3+\sqrt{3})}{4\sqrt{2}} \frac{-(1+\sqrt{3})}{4\sqrt{2}}]$$
$$g = [\frac{(1+\sqrt{3})}{4\sqrt{2}} \frac{(3+\sqrt{3})}{4\sqrt{2}} \frac{(3-\sqrt{3})}{4\sqrt{2}} \frac{(1-\sqrt{3})}{4\sqrt{2}}]$$

These low pass and high pass filter coefficients are used to transform an equivalent number of samples as there are coefficients, and then increment through the raw data set by two. The result of the low pass filter component generates the approximate results and subsamples the data vector to half of its original length. The high pass filter follows the same process but generates the detailed results of the transformation, still subsampled to half its original length. The result of both of these filters is that both the low and high-passed results are down-sampled by a factor of two. In Figure 5-37, an example data vector, x, has 8 terms, and a wavelet with a length of four is used to demonstrate the process of a one level wavelet transform.

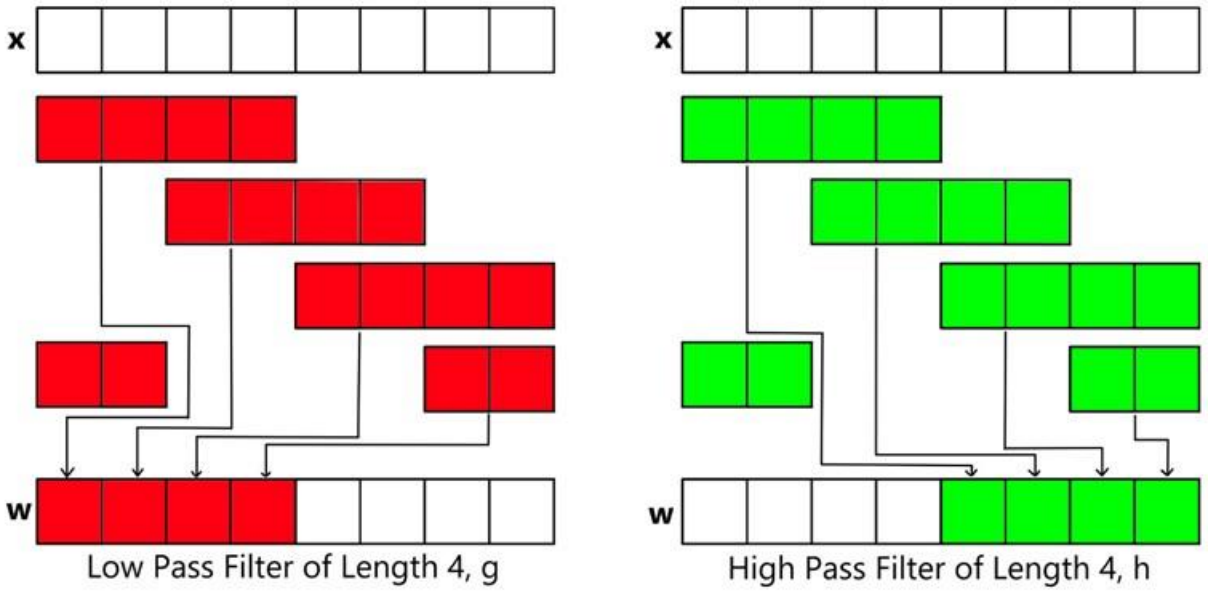


Figure 5-37: The process of the discrete wavelet transformation on a signal length of 8 with a low pass filter (left) and a high pass filter (right) to produce the resulting groups of the two

This process can be repeated to further levels by using the results of the low pass filter as the new starting data vector. Each time, the results are down sampled by a factor of two, and as such the effective frequency of the data analyzed is halved each time. This process can only be repeated up to n times if the initial data set has a length of two to the power of n.

Overall, the result of this process is that the low frequency decomposition of the data has a higher frequency precision, but is spatially imprecise. In comparison, the high frequency decomposition results in a reduced detection of specific high frequency effects, but the spatial resolution is much higher. This multiresolution analysis allows for a similar exploration as the STFT, however the tradeoff between spatial and frequency resolution is conducted as a function of the frequency rather than a set value for the entire operation. This particular multiresolutional analysis tends to reflect the real-world vibrational phenomena, as low frequency events persist much longer than high frequency events. Due to this tradeoff, it was determined that our implementation was appropriate for a number of reasons. The first of these reasons was that the railway vibration phenomena should reflect this pattern of both long and persisting low frequency content combined with short high frequency content. The second of these reasons was that the particular motions of interest can be reflected in these restrictions; namely the “loose” track segments are relatively short and most likely dominated by a high frequency response. Difficulties may arise from an insufficient length of track to detect low frequency responses, or insufficient sampling rate to distinguish high frequency responses. Hence, for the purposes of this analysis, this attempt to use wavelet analysis may indicate if the process itself is viable for our current implementation of the sensor system and track configuration.

5.2.2 Wavelet Selection

The first step in conducting the discrete wavelet transform is to select an appropriate mother wavelet. For the ease of application, we selected from the orthogonal wavelets listed in the MATLAB documentation database. The ones selected for initial inspection were the Symlets 2 and 4, the Daubechies 1 (AKA: Haar), 2, and 4, and the Coiflet 1. These wavelets were chosen as a way of identifying points in the run which contained movement indicative of loose rail. One of the issues with using wavelets for the identification of an unknown movement is that the wavelet decomposition is much stronger when the mother wavelet reflects the shape of the motion it is being used to identify. As such, we could not be sure that the wavelets chosen represented the best option for this type of analysis, and in fact, a further exploration and analysis into optimizing the selection of wavelets may be useful if a defined motion of interest is identified. For now, the selection of test wavelets, as seen in Figure 5-38 below, are meant to cover a variety of possible motions.

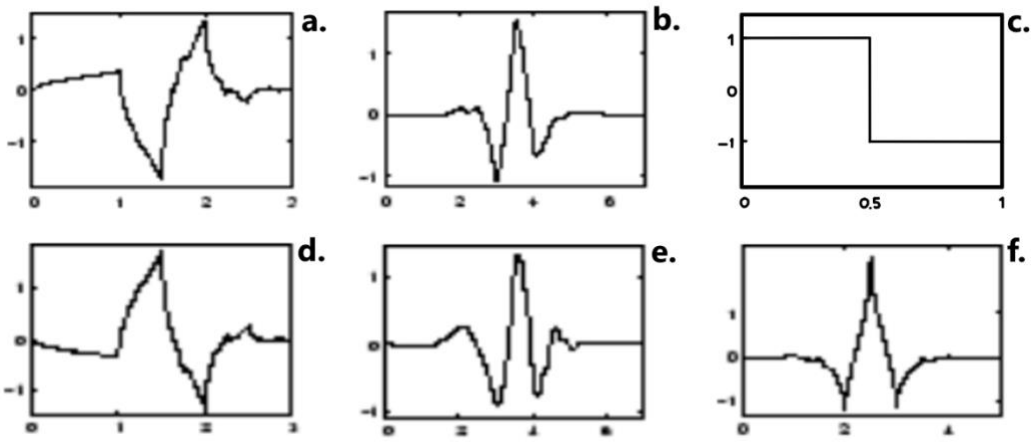


Figure 5-38: Selected mother wavelets of: (a) 'sym2', (b) 'sym4', (c) 'Haar', (d) 'db2', (e) 'db4', (f) 'coif1' [32]

While the motion of the loose track in response to the train was unknown, a few estimations were made as to what would be the most likely, and therefore what would be best represented in the wavelet form. From our understanding of the possible motions which might represent loose track, an initial outward or inward motion followed by a roughly equivalent response in the other direction was possible and worth an attempt at analysis, hence the selection of the 'sym2' and 'db2' mother wavelets. Another motion which could be found within the movement of the track as a function of the loose rail is the slight deflection as a lead-in to a much larger opposite direction of motion, followed by the recovery back in the first direction. To represent this possibility, the 'sym4', 'db4' and 'coif1' mother wavelets were selected for initial analysis. The Daubechies 1 mother wavelet was selected as a comparison, even though it did not present a realistic path of motion, it still provided another method of movement analysis. If the movement of the affected rail repeated, this would present itself in the data as a series of larger amplitudes of convolution, and further analysis through the use of higher order mother wavelets could be viable. If the motion is more complex than what is presented in these mother wavelets, the wavelet decompositional analysis may not be an effective form of "loose" track identification.

Once the wavelets were selected, they were used to filter the data into progressively higher and higher levels of decomposition, resulting in the same frequency/spatial tradeoff that was mentioned before. An example of one of these decompositions can be seen in Figure 5-39 with the use of a db4 mother wavelet for the second 10 mph trial in the Corsair data for the left rail.

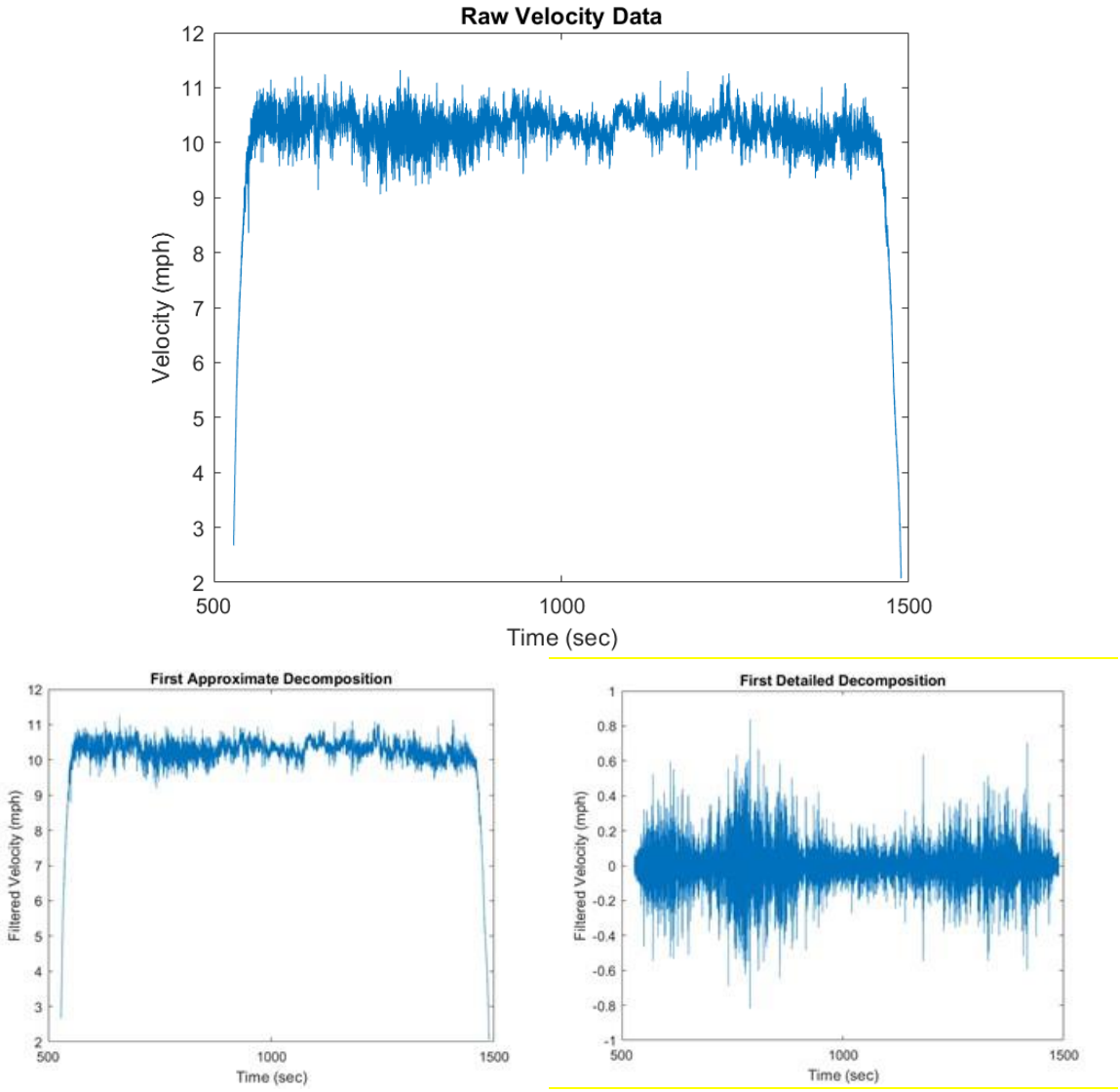


Figure 5-39: The first step of the discrete wavelet transform applied to the second 10 mph vertical data (top) resulting in the approximate components (bottom left) and the detailed components (bottom right)

Again, as a function of not knowing in which particular frequency the effect should appear, the decomposition was conducted several times and stopping after the effective wavelength of the wavelet was no longer appropriate or useful for analysis. Given that each step in the process decomposes the data into the higher and lower halves of frequency content, each progressive level in the detailed data would have twice the effective wavelength as the prior level. In this case, given

the starting sample rate of 29 Hz, for 10 MPH the 7th level of detail produced an effective frequency of 0.23 Hz and wavelength of 30.1 feet. This was larger than the distance of the sections of interest by over twice their length and produced a resultant signal which was ineffective at distinguishing differences in the track motion or conditions.

Figure 5-40 presents a decomposition which was filtered nine times. No one particular decompositional level could be selected with confidence to be containing the locations of interest at this point in the analysis, even if there were elements which stood out from the rest. These responses needed to be taken together as a whole with regards to their different levels and used to look for any patterns or relations across them that might reliably indicate where the loose rails are.

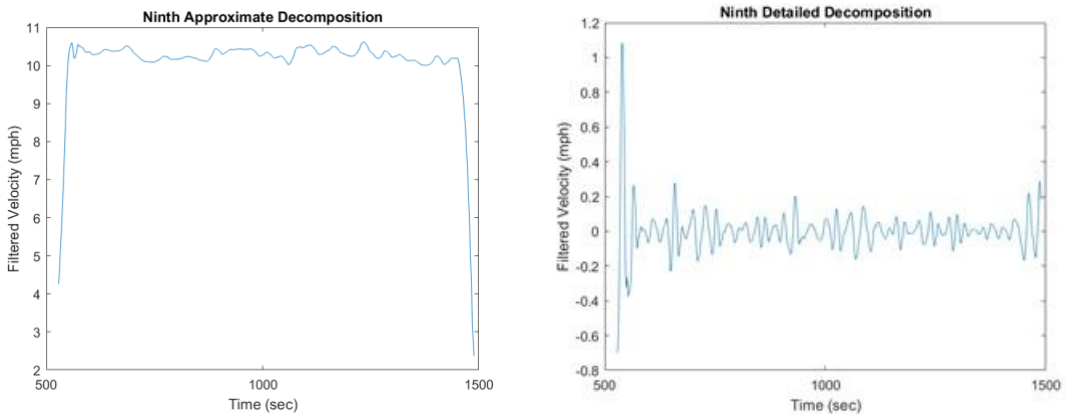


Figure 5-40: The ninth level of the discrete wavelet transform with the approximate results (left) and the detailed results (right) for the second 10 mph vertical velocity

Much like the STFT, the purpose of the wavelet transform was to identify locations with loose rail via their apparent responses. The decomposition of this data into detailed and approximately wavelet filtered data, with its variance in wavelengths, allowed for the detection of outstanding motions associated with each segment of the track. This direct form, however, did not simultaneously allow for this comparison between the different decomposition levels. In the same form as the STFT, a 3D mesh representing the decomposition levels is aligned by the time variable. Both forms produced a comparison between the time/location, the frequency/wavelet decomposition, and the magnitude of the response. In the case of the STFT, each data set was split along the time/space variable and so each part displays a full FFT frequency spectrum. By comparison, each wavelet was distinct between decomposition levels and therefore each component is a vector in the time/space domain. Hence both of these approaches attempt to reach a similar goal through different avenues.

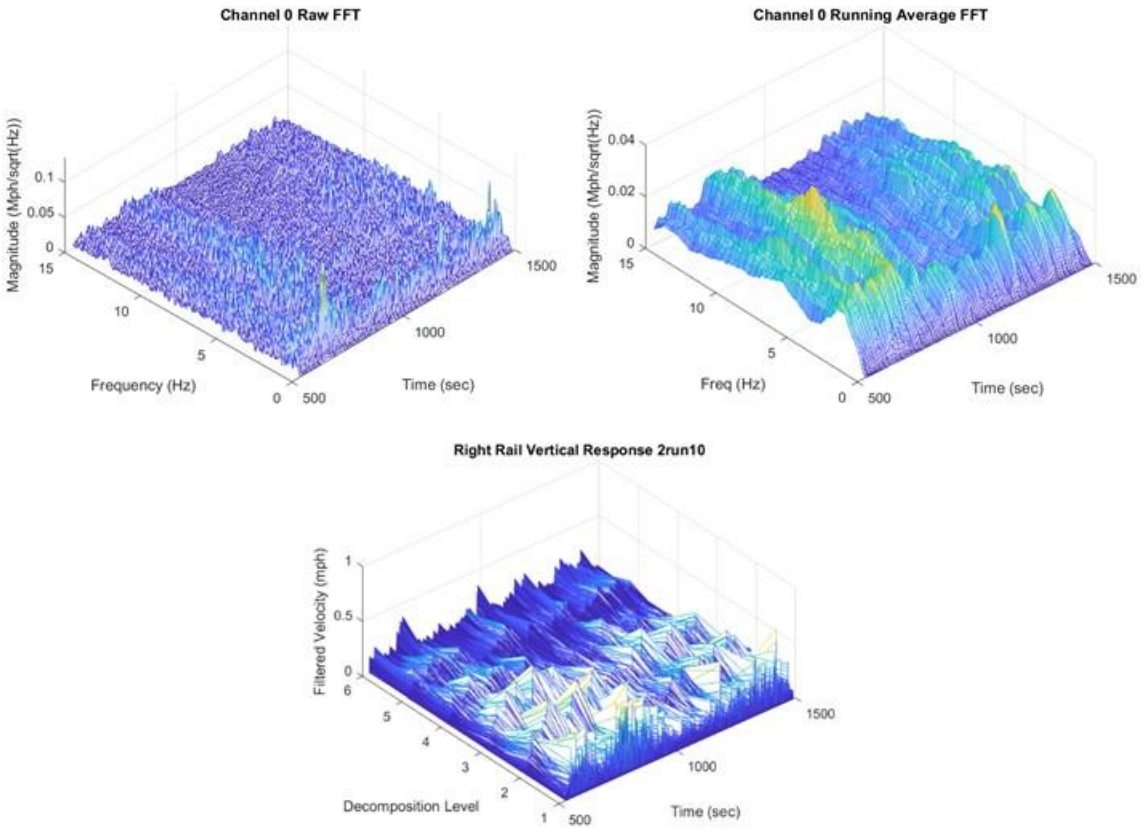
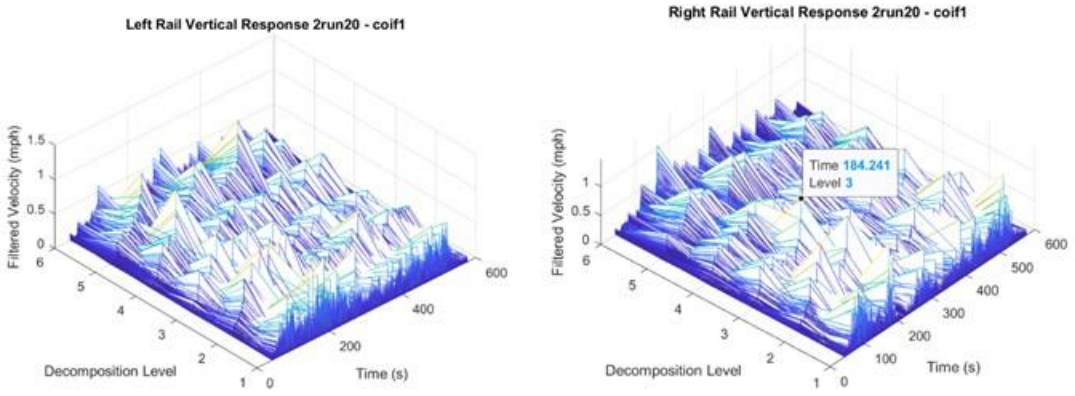


Figure 5-41: Comparison of the second 10 mph vertical vibrations as presented by the raw STFT (top left), the running average STFT (top right), and the wavelet decomposition (bottom)

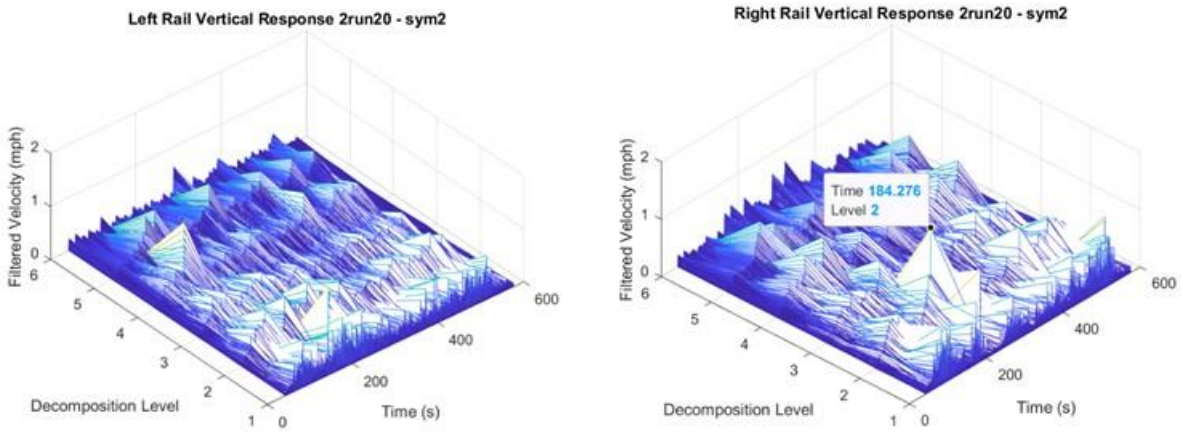
In
Figure 5-41, the left rail motion captured by the Corsair system for the second 10 mph run is represented in both the form of an STFT and its corresponding version of the wavelet decomposition. For the wavelet analysis, this process produced a mesh which was inspected manually for sections which indicated a notable response. Due to the lack of known rail responses, it would be impractical to attempt to design an automated detection method, and as such these initial reviews relied on the researcher to make assumptions about which parts of the graphs were significant. The main criteria used for the selection of locations of interest through the wavelet decomposition was the distinction of the particular location's response compared to the locations around it, along with the magnitude of the response itself. Some sections of the runs include areas with several peak amplitudes close together, and within these sections this method of detection would most likely present itself as unreliable. With external knowledge from the GPS information, however, peaks within these raised sections that correspond to "loose" sections may still be indicated for the purposes of further analysis.

With regards to the selection of the optimal mother wavelet, they were evaluated by using them to filter the data and generate a graph, followed by a manual inspection of the results. This was used to determine the usefulness of the particular wavelet inspected. From an initial inspection of the

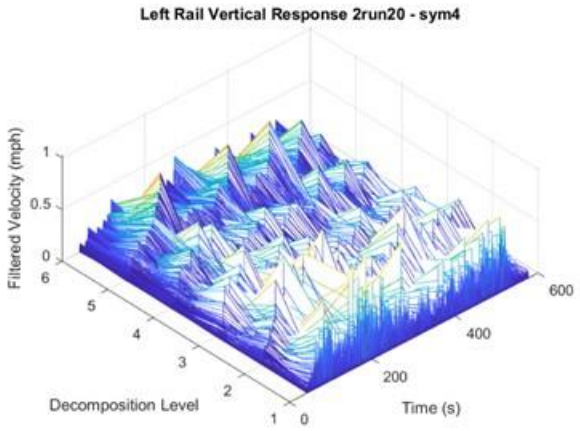
second 20 mph run lateral responses, the most promising of the mother wavelets was the Daubechies 4, or ‘db4’, wavelet, and as such was the wavelet used for the initial battery of examinations on the trial runs. This selection was determined by the use of comparing the represented data and the ease of selecting the time(s) in the run which aligned with when the train would be passing over a point of interest. As seen in
Figure 5-42, the results indicated a more favorable illustration of the loose rail in the curve through the ‘db4’ on the Corsair data.

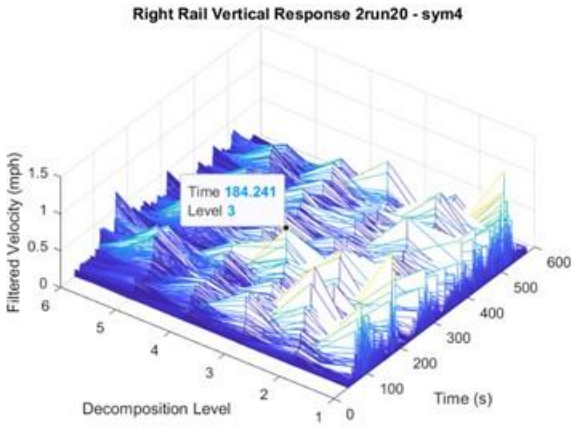


a.

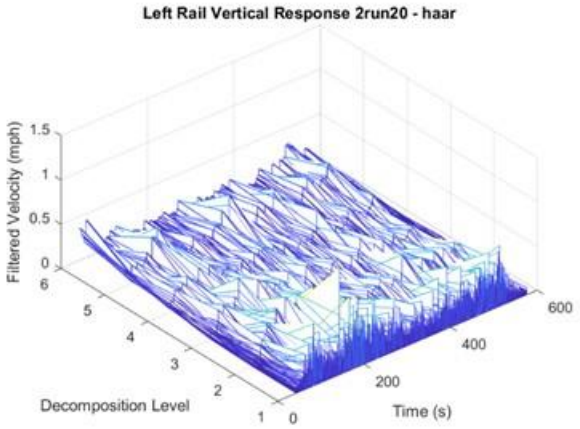


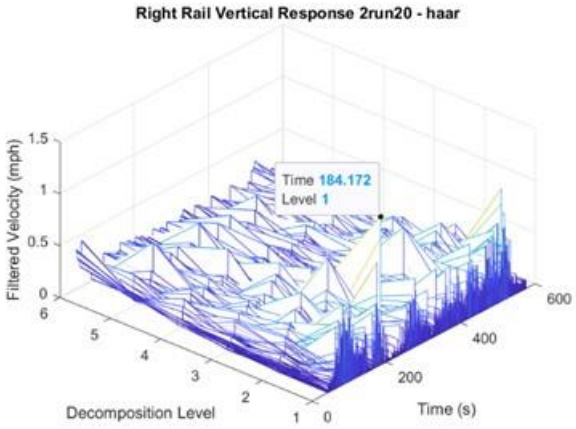
b.


Left Rail Vertical Response 2run20 - sym4
Filtered Velocity (mph)
Decomposition Level
Time (s)

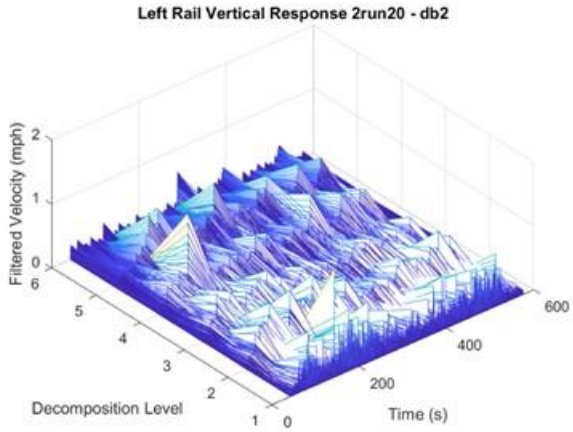

Right Rail Vertical Response 2run20 - sym4
Time 184.241
Level 3
Filtered Velocity (mph)
Decomposition Level
Time (s)

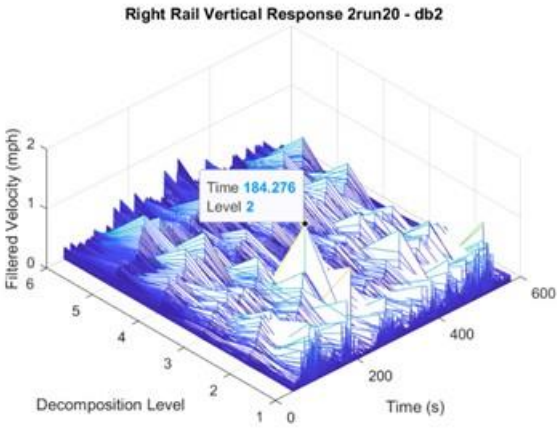
c.


Left Rail Vertical Response 2run20 - haar
Filtered Velocity (mph)
Decomposition Level
Time (s)

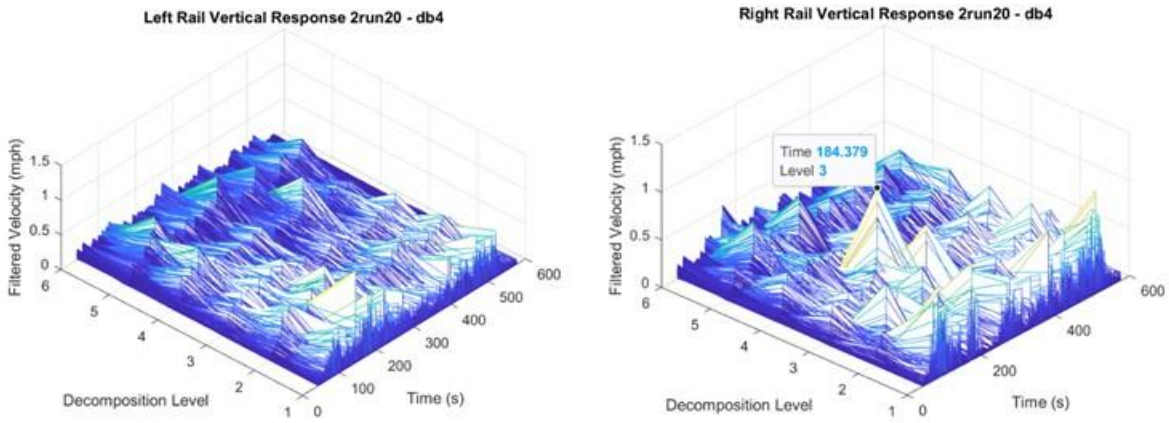

Right Rail Vertical Response 2run20 - haar
Time 184.172
Level 1
Filtered Velocity (mph)
Decomposition Level
Time (s)

d.


Left Rail Vertical Response 2run20 - db2
Filtered Velocity (mph)
Decomposition Level
Time (s)


Right Rail Vertical Response 2run20 - db2
Time 184.276
Level 2
Filtered Velocity (mph)
Decomposition Level
Time (s)

e.



f.

Figure 5-42: Vertical rail wavelet decomposition with the (a.) 'coif1', (b.) 'sym2', (c.) 'sym4', (d.) 'Haar', (e.) 'db2', and (f.) 'db4' mother wavelets

In order to further validate the use of the 'db4' wavelet, the GoPro aligned locations of interest and their velocities were visually inspected and compared to the wavelet transformed data. From the inspection of the 2nd 20 mph trial's vertical data, which was selected due to the clarity of the response tied directly to the timing for the location of interest found by aligning the GoPro video with the velocity data, the results indicated that a Daubechies 4 mother wavelet would be most appropriate. The shape appears to reflect the movement identified at that location much better than when identified using the Daubechies 2. Both mother wavelets could identify the location of interest, however the 'db2' wavelet identified the location through effectively overfitting to the shape of the signature, while the db4 wavelet aligned strongly without over or underfitting to the shape. As seen in Figure 5-43, in addition to both wavelets identifying the apparent large response at the point of interest, they do so at different decomposition levels. The db4 wavelet was most aligned with the signal at the detail level of 3, which would produce an effective frequency and wavelength at 20 MPH of 1.06 Hz and 3.76 feet. The db2 wavelet was able to reflect the data but was conducted at a detail level of 2, in which it repeated itself to achieve the effect and therefore had an effective frequency and wavelength of 0.53 Hz and 1.88 feet and occurred twice. From this, the db4 appeared to be a stronger indicator of the motion found in the second 20 MPH run on the Corsair data.

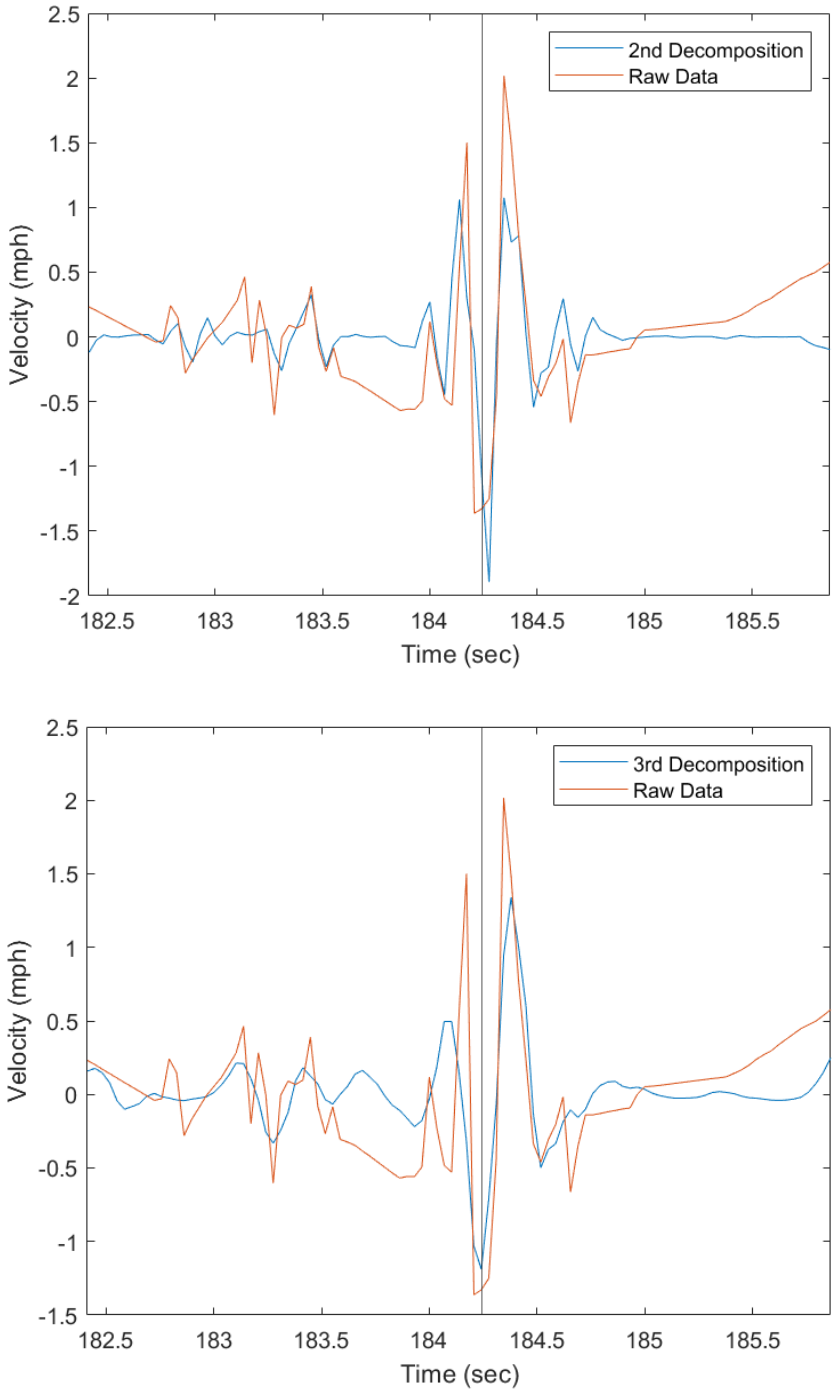
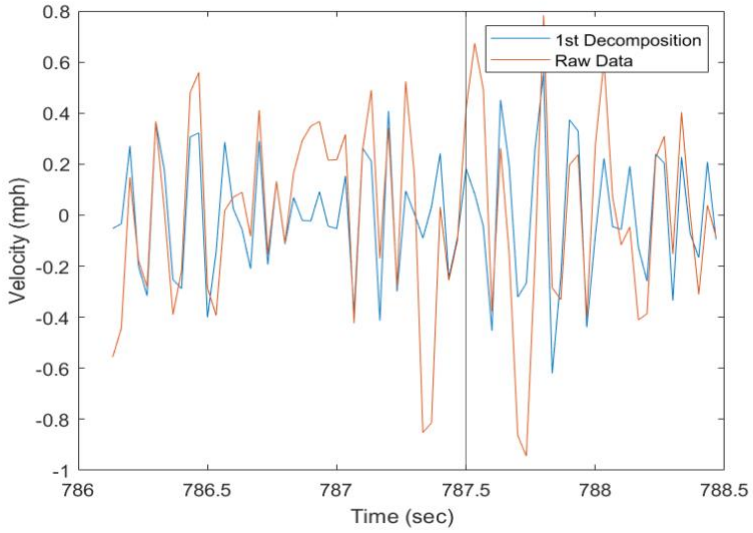


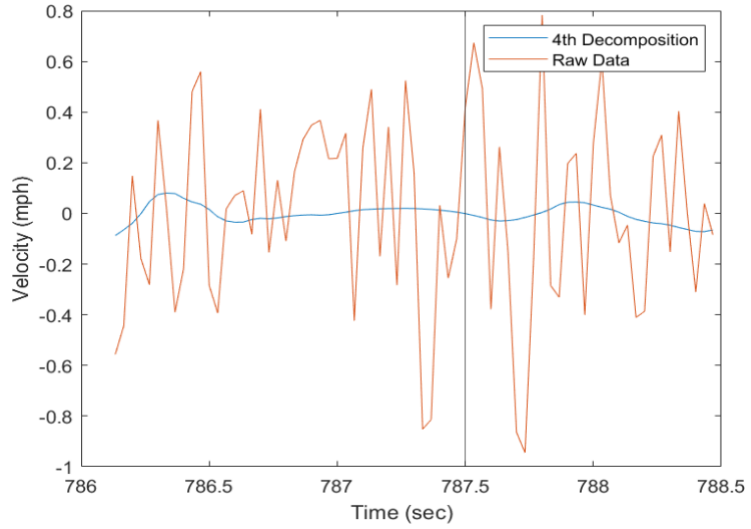
Figure 5-43: The vertical vibrations of the loose right rail in section 25 represented by the second level 'db2' wavelet (top) and the third level 'db4' wavelet (bottom)

In addition to the comparison between the db2 and db4 mother wavelets, a brief example of the differences in effect of the levels of detail could be pertinent to the following explanations. At low levels of detail decomposition, i.e., at the first or sometimes second level, the wavelets tend to

overfit to the data. In this case, they provide little useful information for identifying large movements of the rail, especially at slower speeds where their effective spatial wavelength along the track is as low as 5.6 inches. As a result, the outputs of this process would reflect a response that would happen over distance shorter than the length of a tie. In Figure 5-44, the effects of overfitting on the representation of the data for the vertical right rail movement in the second 10 mph run can be seen and compared to that of an underfitted filtered data set produced by decomposing the signal past the point of returning useful motion information.



a.



b.

Figure 5-44: The vertical vibrations at the loose right rail in section 25 represented by the (a) overfitted first level and the (b) underfitted fourth level of wavelet decomposition with the 'db4' wavelet

To continue with the validation of the use of the ‘db4’ wavelet over the ‘db2’, the same process was repeated for the 2nd 10 MPH Corsair run, which provided another strong example for the preference of the use of the Daubechies 4 over the Daubechies 2 especially when looking at the point of interest defined in the curve of the track in section 25. The signature centered around the first loose tie of that section is once again better reflected by the shape prescribed by the db4 as opposed to the db2 wavelets, as seen in Figure 5-45. This identification, however, was more present in the signal for Channel 1 than for Channel 0, which did not contain much of a strong signature anywhere, and as such did not assist in the selection of the appropriate mother wavelet.

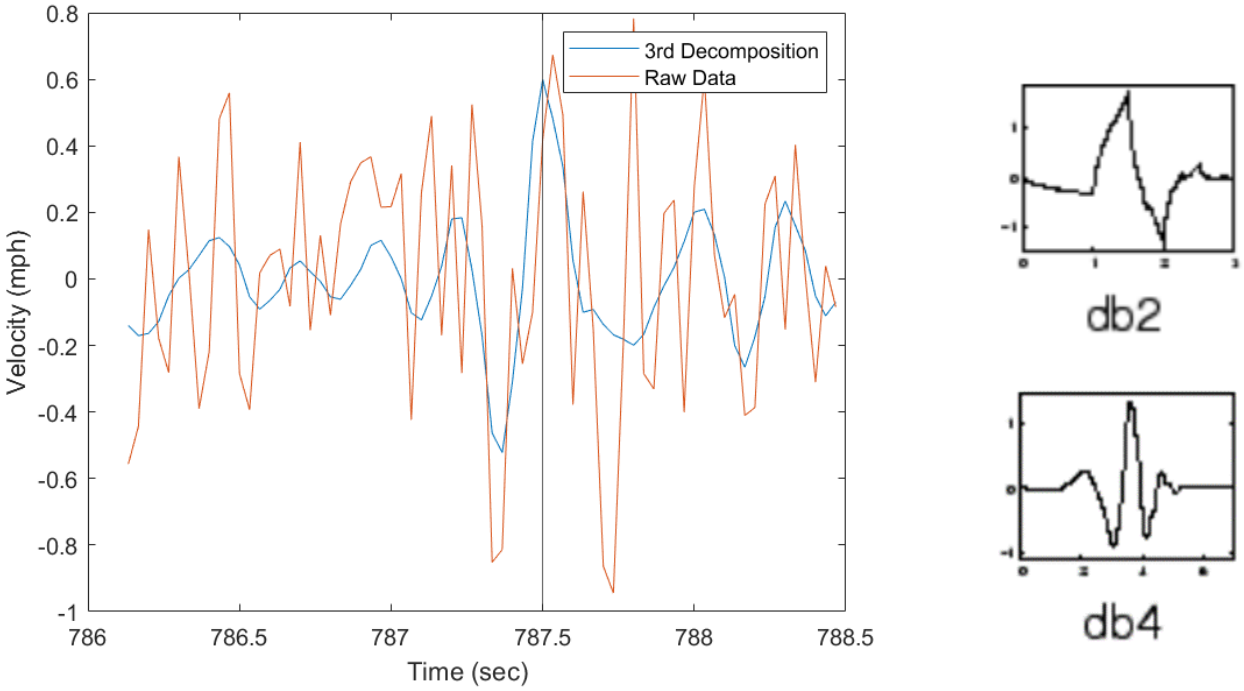


Figure 5-45: Vertical vibrations of the loose right rail represented by a wavelet transform [32]

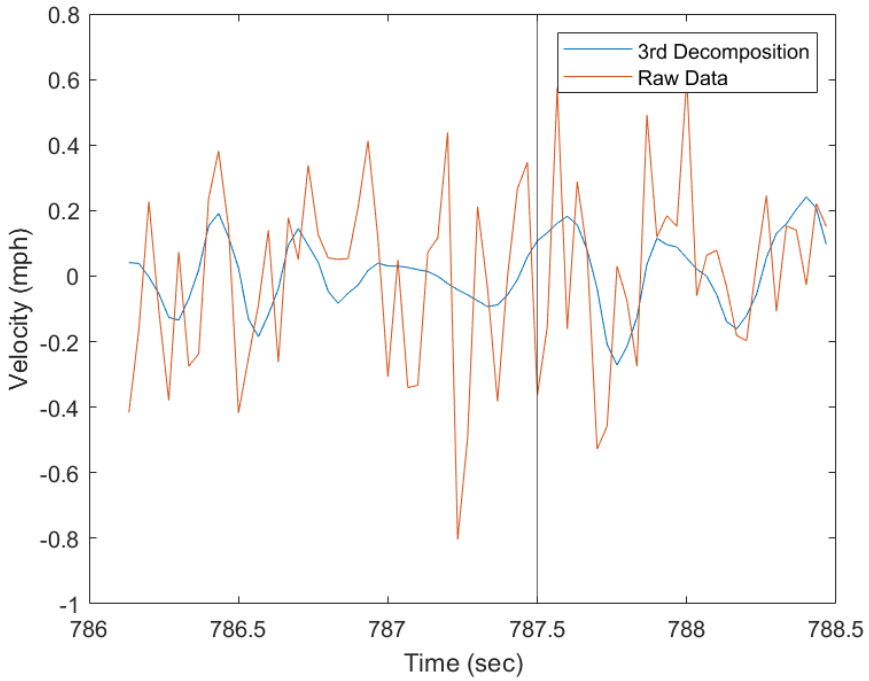
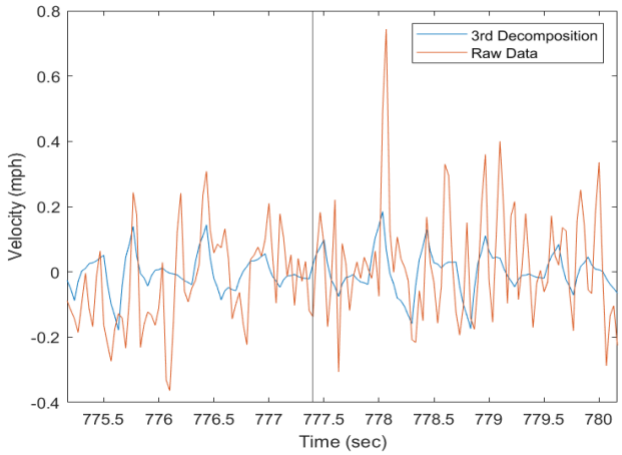
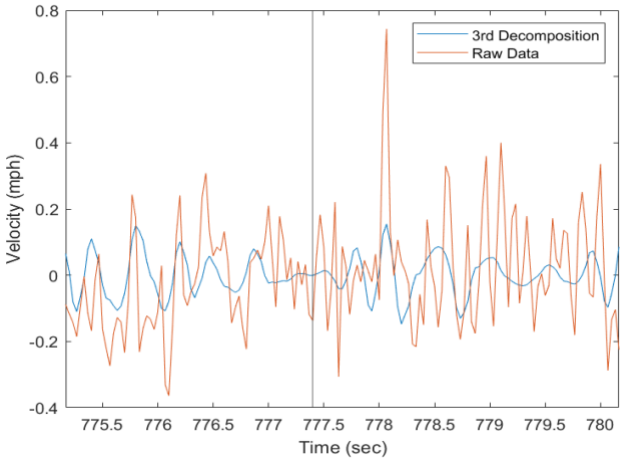


Figure 5-46: Vertical vibrations of the loose left rail represented by a wavelet transform

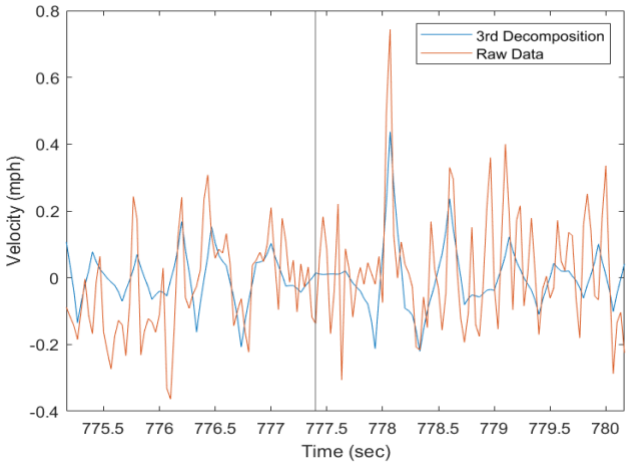
For both of these inspections, the vertical data was selected, and this was mostly due to the strength with which the signals in the vertical direction indicated our points of interest. Even from the initial inspections of the data, the Corsair data was better at identifying the point of simulated loose rail than the Blackbeard data, which reflects the removal of the vertical spikes and the still retained lateral spikes. The Blackbeard data still could be used for the detection of loose track, but the clarity and reliability with which it could be used for was reduced compared to the Corsair data, especially for defining the representative type of wavelet for implementation. In this example from the 2nd 10 mph run on the Blackbeard system, the peak response happened a short time after the first tie loose, but within the loose track since it was positioned four and a half feet after the fish scale on tie 552. This peak was not defined clearly by either the Daubechies 2 or the Daubechies 4, and in fact seemed best represented by the Coiflet, or the ‘coif1’ wavelet.



a.



b.



c.

Figure 5-47: Lateral vibrations of the loose rail in section 25 represented by a (a) 'db2', (b) 'db4', and (c) 'coif1' mother wavelet

This result, however, did not seem to be consistent throughout the Blackbeard data, as inspection of the 20 MPH and 30 MPH runs did not reveal this same Coiflet motion at the location of the point of interest in the curve, nor did they provide a clear enough depiction of any particular signature to indicate if another wavelet would be appropriate. In both cases, for both the left and right rails, the Blackbeard data did not produce any clear and definable motion which could be best represented by wavelets. This is not to indicate that other locations within the Blackbeard data were undetectable, but that it was inconsistent for both the curved section 25 and in the tangent section 40. As such, the emulated loose rail was not able to be represented as cleanly by a wavelet, and therefore not as reliably detectable via manual inspection of the results of the discrete wavelet transform. One particular feature of note was found in the 2nd 20 mph Blackbeard results; the largest response in the section of interest was in phase for its 3rd level detail response rather than the normal out of phase motion seen in the Blackbeard data. The second 30 mph run, however, did not have as noticeable of an in-phase response in its most notable response around the point of interest. From all of these results, it is clear that the Blackbeard results produce inconsistent responses around the section 25 point of interest, and as such might not be as productively analyzed through the wavelet transform.

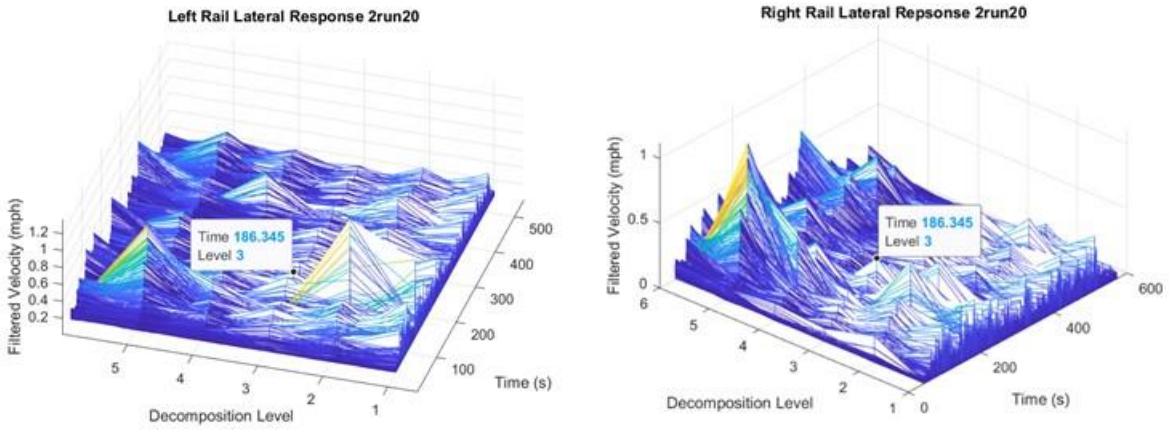


Figure 5-48: Lateral vibrations of the loose rail in section 25 wavelet decomposition for the left and right rails at 20 mph

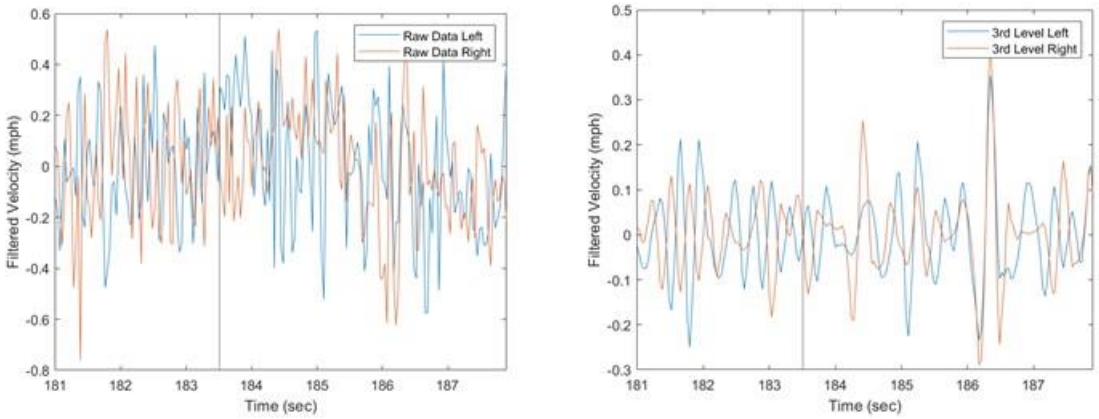


Figure 5-49: Raw data (left) and wavelet representation (right) of the section 25 loose rail at 20 mph

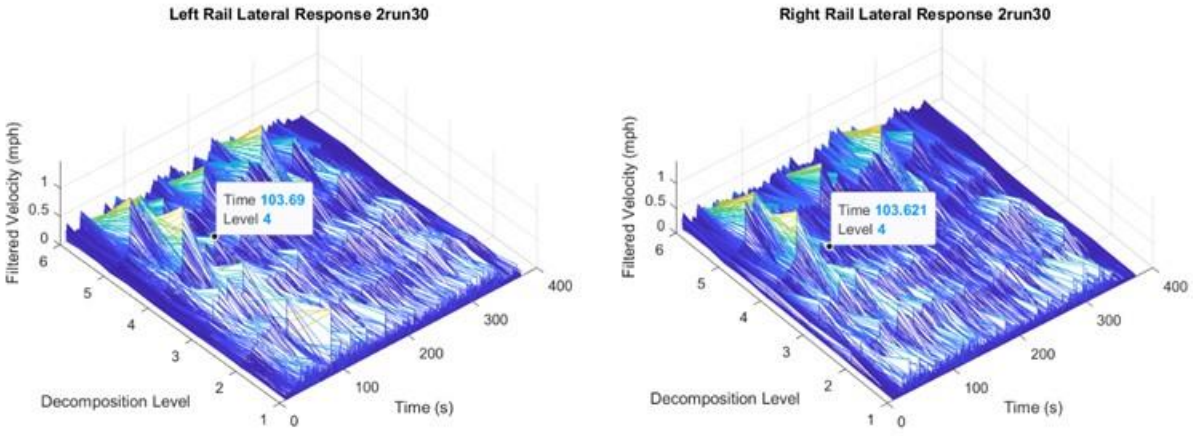


Figure 5-50: Lateral vibrations of the loose rail in section 25 wavelet decomposition for the left and right rails at 30 mph

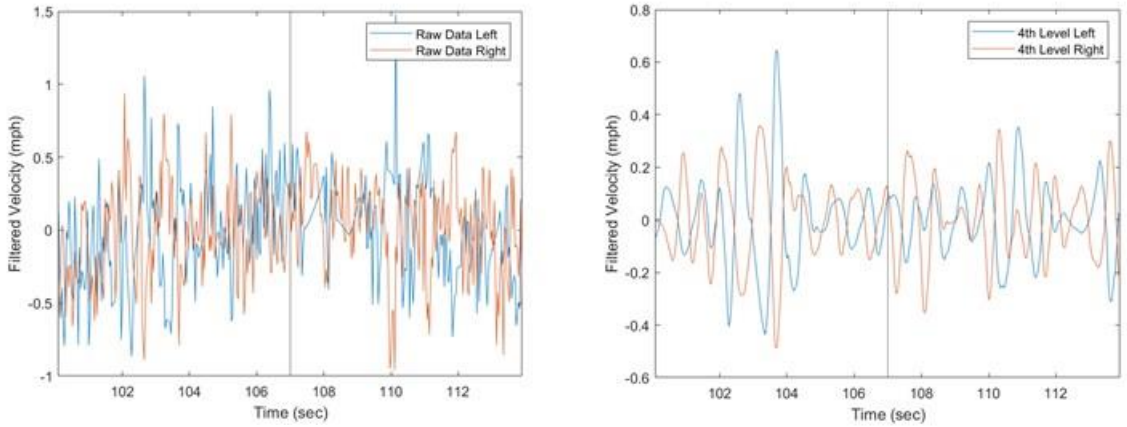


Figure 5-51: Raw data (left) and wavelet representation (right) of the section 25 loose rail at 30 mph

5.2.3 Non-Point-of-Interest Locations Detected by Lateral Rail Inspection

Even after determining that the Blackbeard data was unreliable in terms of defining the points of interest in sections 25 and 40, it was still worth analyzing if it was able to reliably detect any other phenomena within the track. From the visual inspection of the data from the 'db4' wavelet, several aspects appeared relatively consistently within the data. Through the use of Google Earth images and the GPS outputs of our system, the locations determined by this inspection were assigned to the different track structures. Given that there was an inconsistent GPS lead/lag time in the data and that it was a function of speed, the locations determined via this method were considered to indicate the closest object.

The first of the locations indicated by the Blackbeard wavelet analysis was the inside of the curve near the end of section 7. This occurrence was apparent in both the left and right channel data sets

and across all speeds. In at least two of the three trials for all four speeds, this location was selected via manual analysis. This event was characterized by the higher level of detail decomposition required for this location to stand out, typically occurring between the 4th to 6th level of decomposition, and hence having a spatial wavelength of between 3.8 to 15 feet at 10 mph and 15 to 60 feet at 40 mph

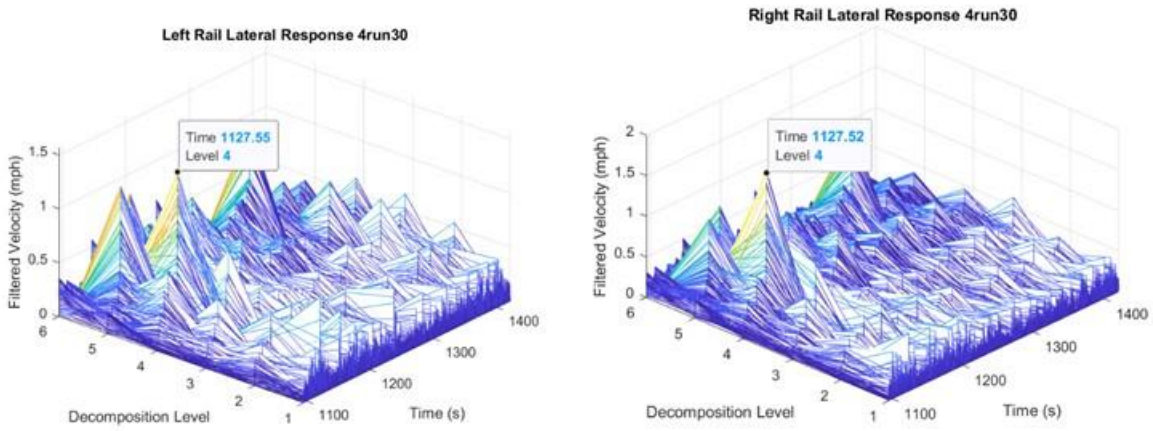


Figure 5-52: Section 7 left and right wavelet decomposition at 30 mph

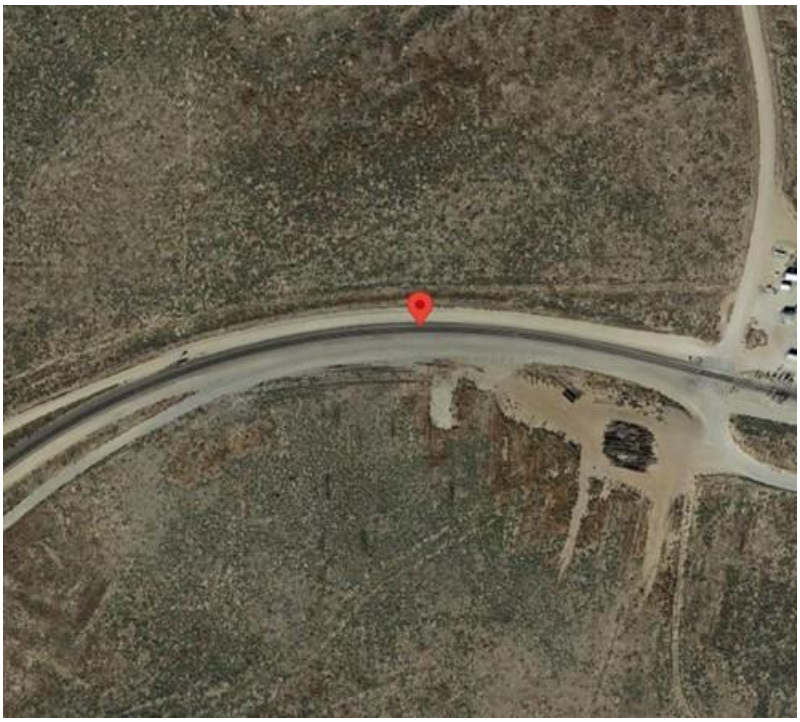


Figure 5-53: Location of the greatest lateral response in section 7 at 30 mph

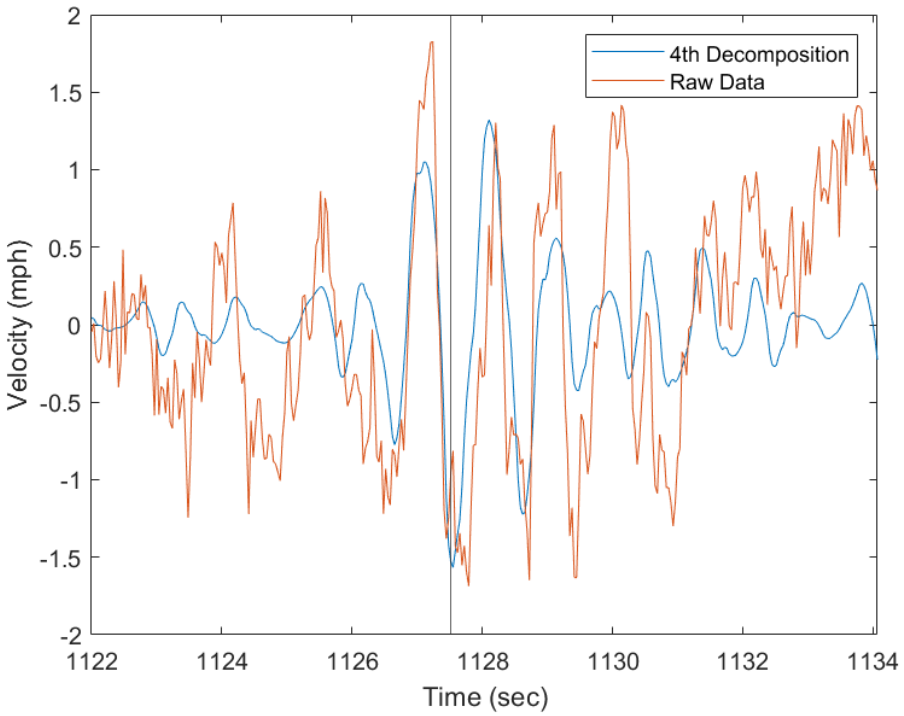


Figure 5-54: Raw data vs the fourth level wavelet of the lateral response at 30 mph

In the above figure, the Google Earth view depicts the location indicated by the wavelet inspection. The wavelet convolution mesh also indicates additional features of this location, namely the level to which it is differentiated from the surrounding data along with its broad and high decompositional levels. From this, along with a more detailed inspection of the location, we were able to identify the signature of the location to have what appears to be a harmonic resonance, which corresponds to the STFT analysis which shows a large magnitude response at the same location.

Another element which the Blackbeard data wavelet analysis consistently identified was the "lubricator". This element can be identified by a large magnitude response across a short section of track. This response occurred around the second decomposition and most prominently in the left rail. As the speed of the train increased, the response of the lubricator section, as with most of the other responses, became harder to identify.

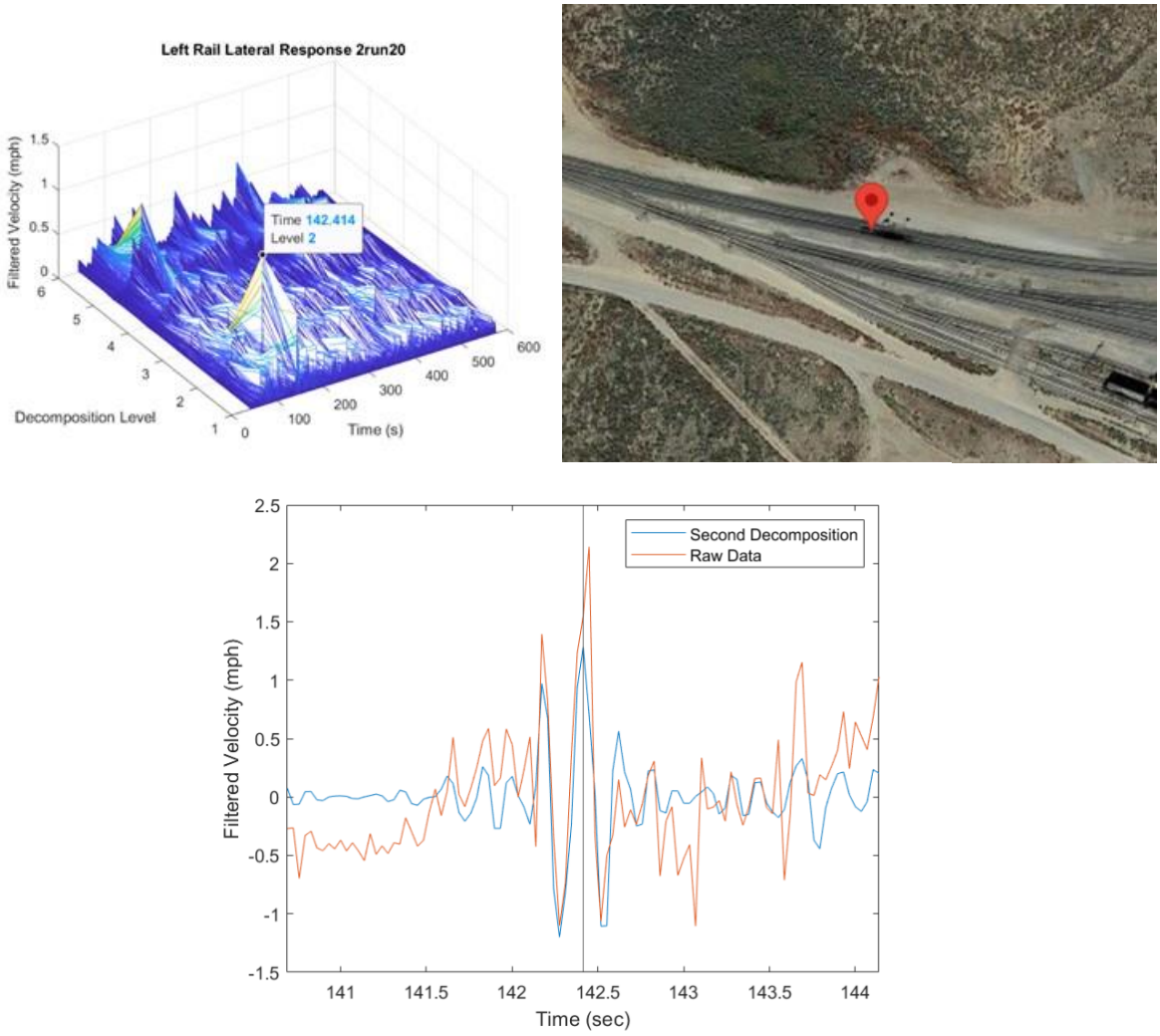


Figure 5-55: The left rail lateral response as depicted by the wavelet decomposition (top left), the GPS location (top right), and the raw vs second wavelet (bottom) for 20 mph

At this lubricator, both the left and the rail signals moved in similar ways, which was counter to the typical result of the lateral data oscillating 180 degrees out of phase. This could indicate that either the rail's gauge was narrowing/widening or that the train's speed was rapidly increasing/decreasing for a short time. The GPS speed data could not confirm or reject either of these possibilities, as the update rate did not appear to be fast enough in this section to have captured if the train had a rapid change in speed. This in phase response appeared at the higher speeds as well, as found in the second 30 mph data. In both of these, the data indicates that the two channels are 180 degrees out of phase in general, most likely due to the rocking or shifting of the car, but at the location of the lubricator, they both respond in phase.

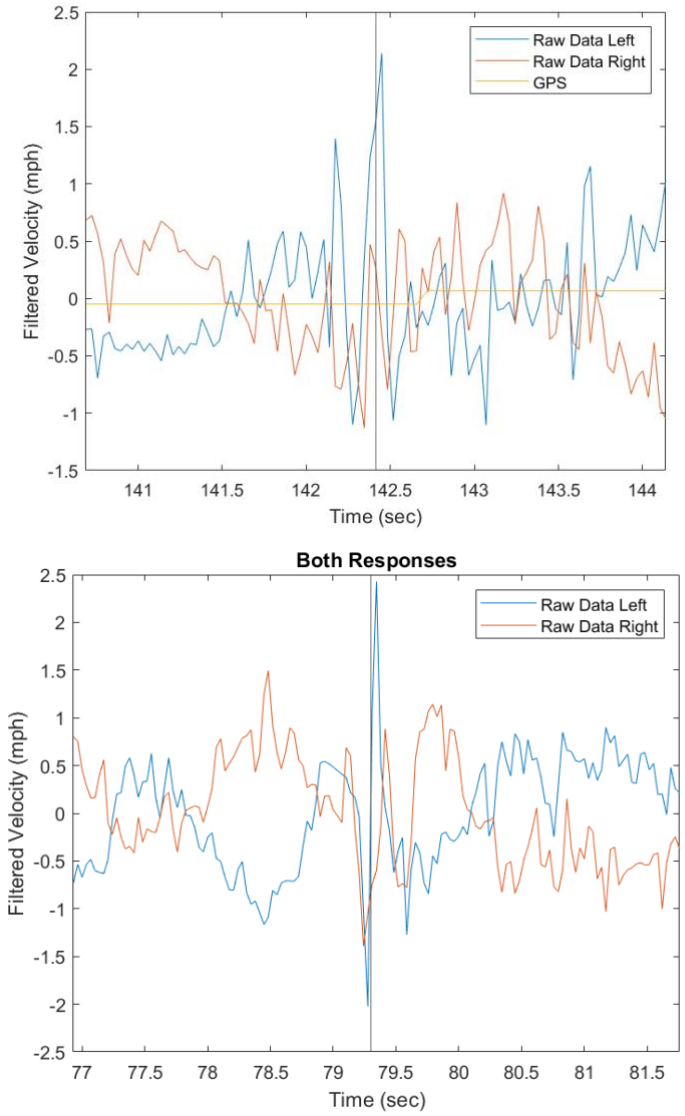


Figure 5-56: Lateral velocity of the left and right rail at the lubricator 20 mph (top) and for 30 mph (bottom)

In addition to these locations, another that stood out within the Blackbeard data was the switch reconnecting the tangent section containing the control and loose rail with section 35. The motion of these sections also seemed defined by the rocking of the freight car back and forth, and at the location of the switch, the data indicated a much larger deflection, maintaining the out of phase motion. Given that the rate of this deflection was relatively slow, and that the motion stayed out of phase, it would appear that this resulted from an increased deflection of the rail from the rocking of the train. As such, it seems that the switch was detected via either the rocking of the train or the shifting of the bogeys changing the angles of the Lidars to the rails.

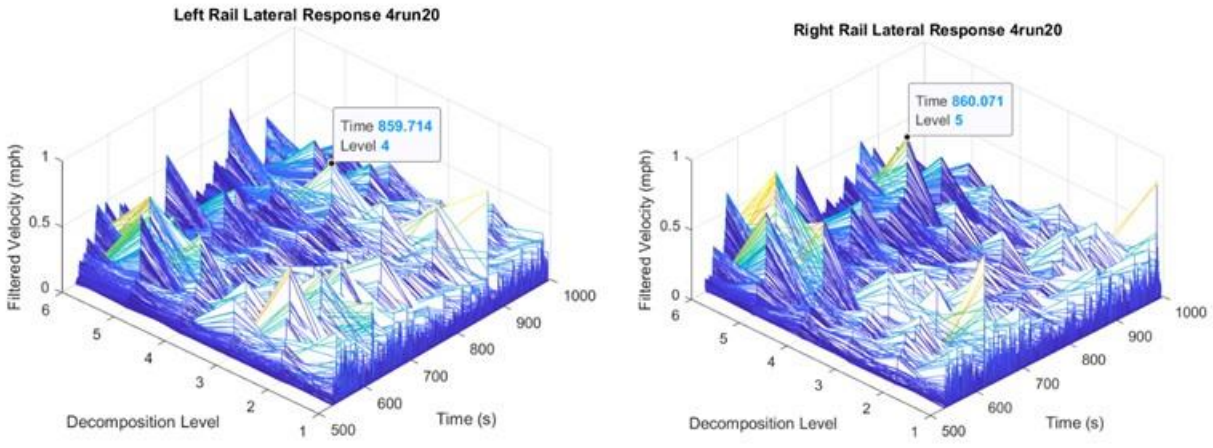


Figure 5-57: The wavelet decomposition depiction of the left and right rail lateral responses at the switch in section 34

Figure 5-58: The GPS location of the switch indicated by the lateral velocity wavelet decomposition

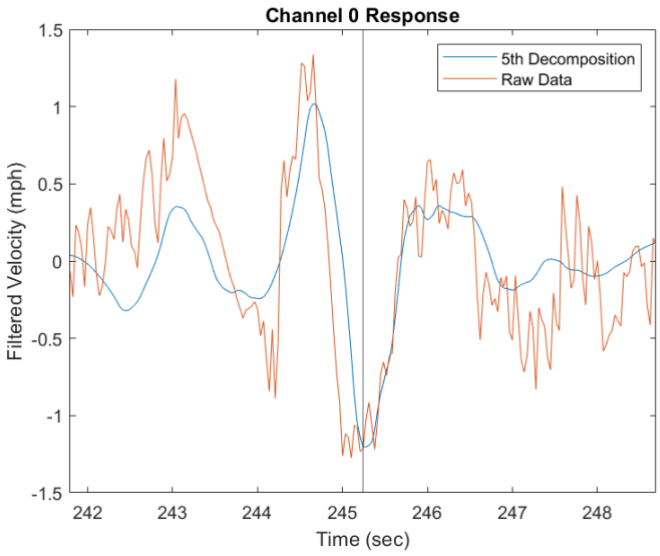


Figure 5-59: Left rail lateral raw and 5th level wavelet response of the section 34 switch

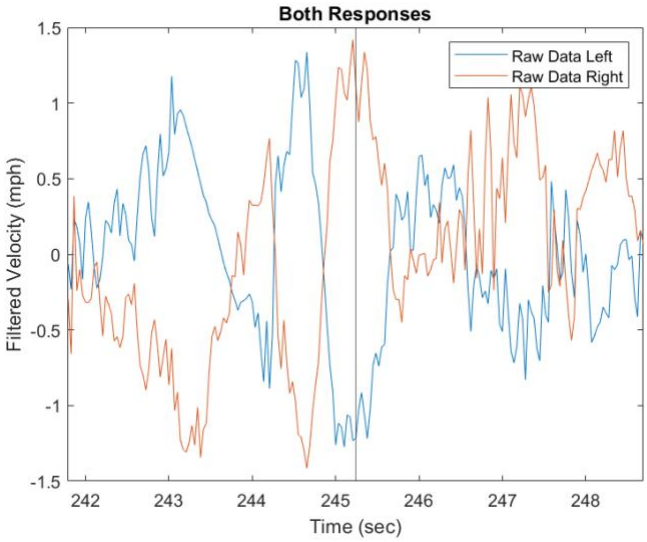


Figure 5-60: Left and right rail lateral velocity at the section 34 switch

5.2.4 Non-Point-of-Interest Locations Detected by Vertical Rail Inspection

As compared to the Blackbeard data, the Corsair data indicated a much greater potential from our testing to find the location of rails which were "loose". Although similarly to the Blackbeard responses, the Corsair data also indicated locations which did not occur concurrently with the targeted points of interest. These locations and responses would therefore prove to be an obstacle which prevents the implementations of a form of automated or unsupervised detection algorithm, as without a way to filter these other signatures out, there would be too many false positives to provide any real utility.

Of the locations which provided false positives for the identification of the simulated loose rail, the lubricator stood out the most. As with the Blackbeard data, it seems that the wavelet analysis was most adept at identifying the brief but large effects of the lubricator on the velocity data. In Figure 5-64, the motion of the left and right rails is clearly distinct from the Blackbeard movement in that they appear to be out of phase, rather than in phase. This effect is even more apparent in the 3rd level of detail, occurring around an effective wavelength of around 8 feet.

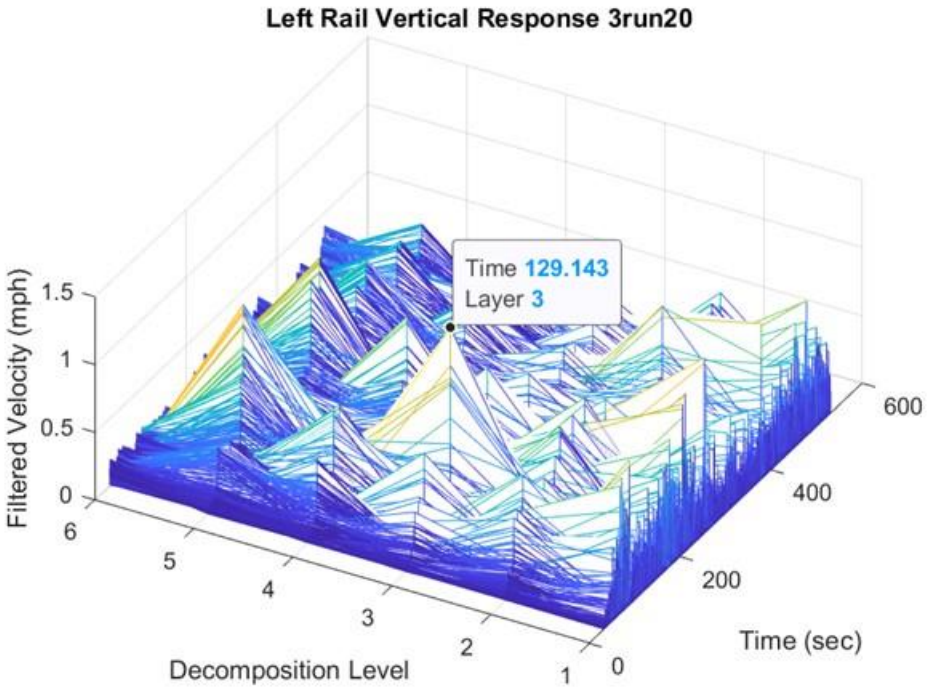


Figure 5-61: Left rail vertical wavelet decomposition for the lubricator

Figure 5-62: Lubricator GPS location indicated by the vertical response

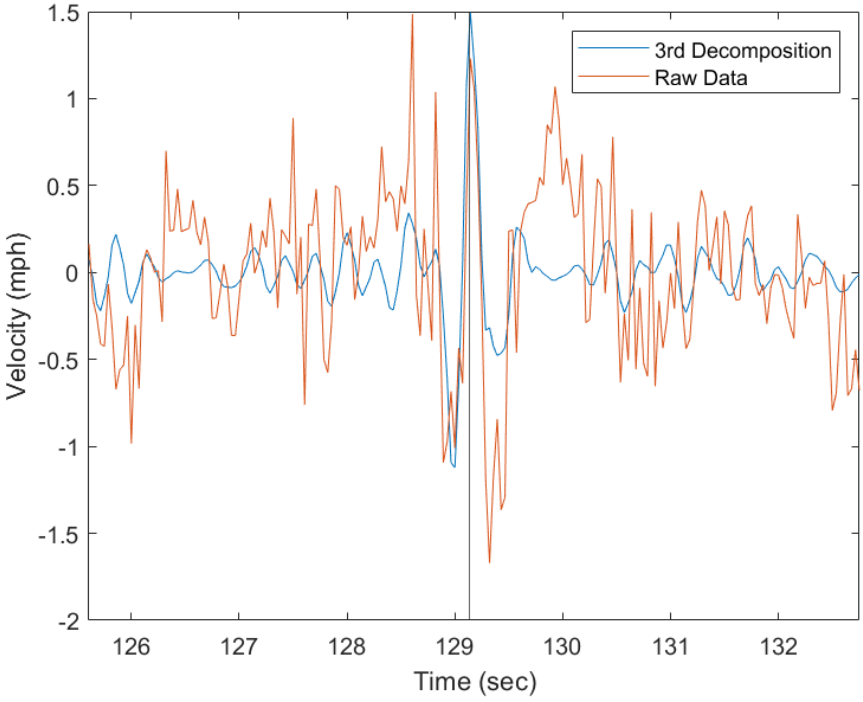


Figure 5-63: Lubricator raw and wavelet transformed data of the vertical response at the lubricator

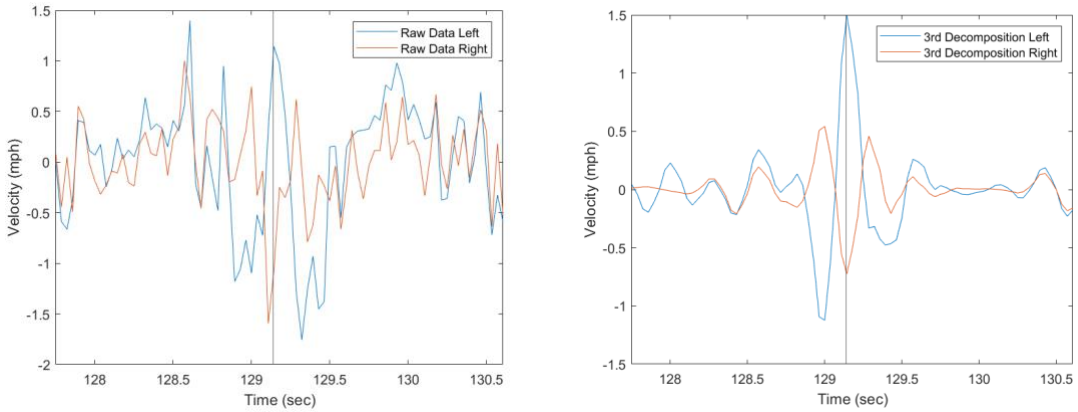


Figure 5-64: Left and right rail raw (left) and wavelet transformed (right) vertical responses at the lubricator

Additionally, the switch after section 40 was readily apparent in the Corsair data, just as in the Blackbeard data. As seen in
Figure 5-68, the motion still reflects the out of phase movement found in the Blackbeard data, and hence provides further justification for the effect to be a result of the rocking train.

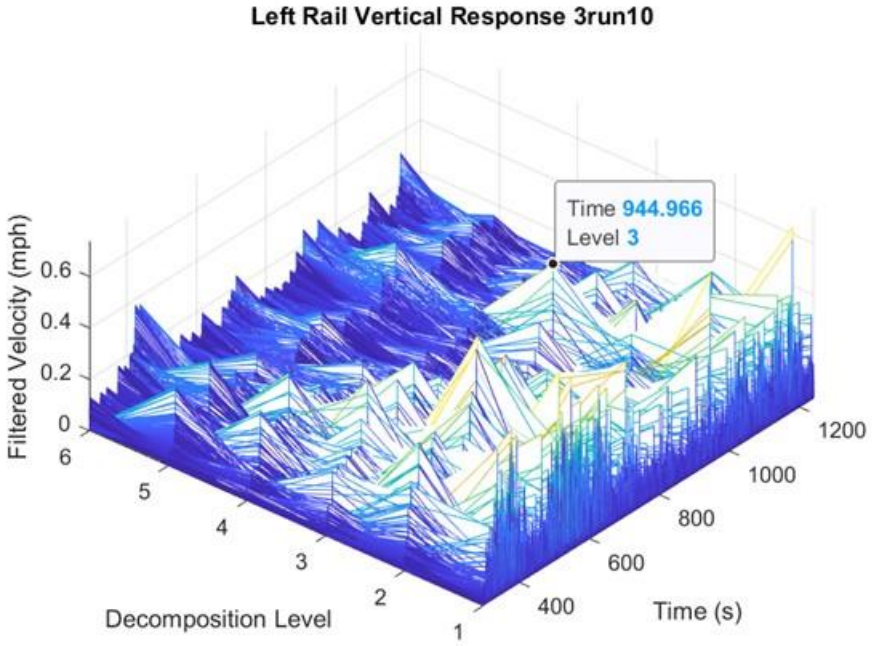


Figure 5-65: The vertical left rail wavelet decomposition of the switch at 10 mph

Figure 5-66: GPS location of the switch as indicated by the wavelet decomposition

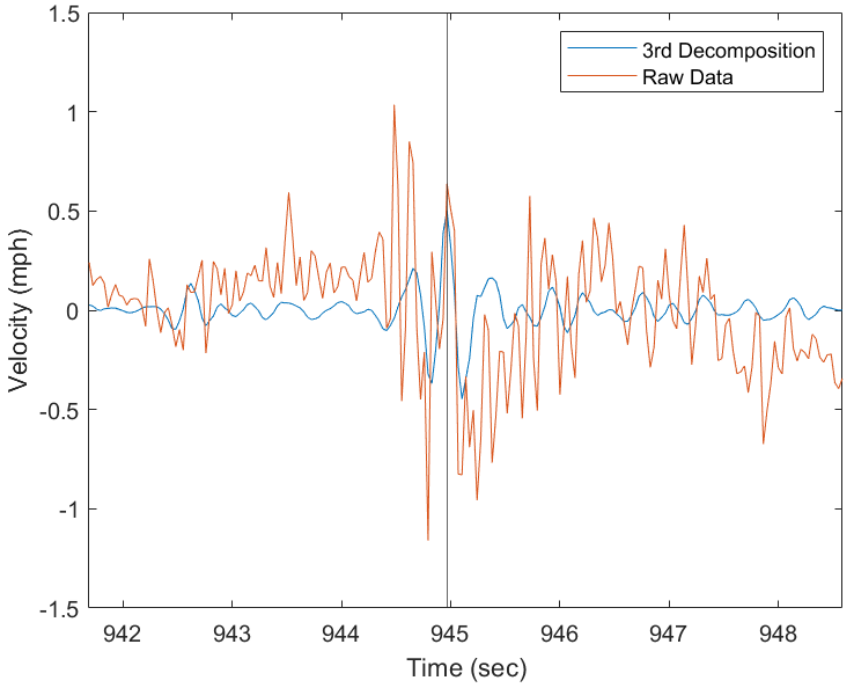


Figure 5-67: left rail raw and 3rd level wavelet transform vertical velocity for the switch

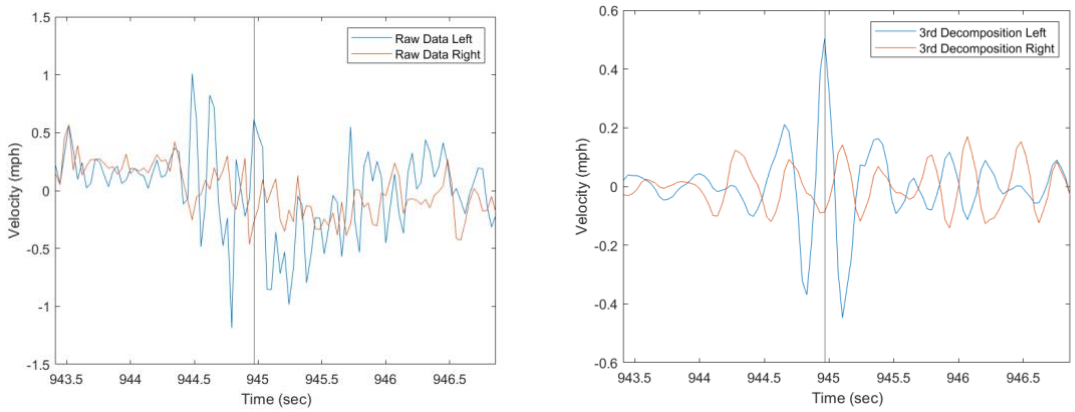


Figure 5-68: raw (left) and 3rd level DWT (right) for the switch at 10 mph

5.2.5 Point-of-Interest as Detected by Vertical Rail Inspection

With regards to the main purpose of the Wavelet analysis, namely the identification of the "loose" rail, the Corsair data able to be used to determine the location, but was inconsistent. From the 2nd 10 and 20 MPH runs, where we can see that there was an elevated wavelet response directly at the start of the section 25 point of interest, we compared the corresponding GPS locations found on Google Earth. Unfortunately, we can see via direct comparison that the pinned location was not the same across both of these runs. These locations occurred at a similar point along the curve, but it was clearly imprecise. As a result, even though the use of the GPS positioning was implemented to determine if a wavelet response was indicative of the loose rail via comparison to the known location, error in the location should be expected. We could also observe that the higher speed position appears earlier in the track, which could be explained by the GPS delay in comparison to the true location of the train.

a.

b.

c.

Figure 5-69: Location of the first fish scale in the section 25 loose rail as correlated by manual video analysis and GPS dad for the (a.) 2nd 10 mph, (b.) 2nd 20 mph, and (c.) 2nd 30 mph trials

Through using the range of detected locations as the bounds of identification for the point of interest via the GPS, we identified that the Corsair data was capable of identifying the emulated loose rail of section 25 while at slower speeds. From the 10 and 20 MPH runs, of the three runs of each speed, all three 10 MPH trials were able to select a peak corresponding to the loose rail. Of the 20 mph runs, 2 of the locations fell squarely within our defined bounds held by the GPS lag, however one of them lied just on the line for the 30 MPH run. Because the GPS lag delay was inconsistent across runs, it would be prudent to consider this result a partial success, since we cannot be as confident in its location as at or around the “loose” rail.

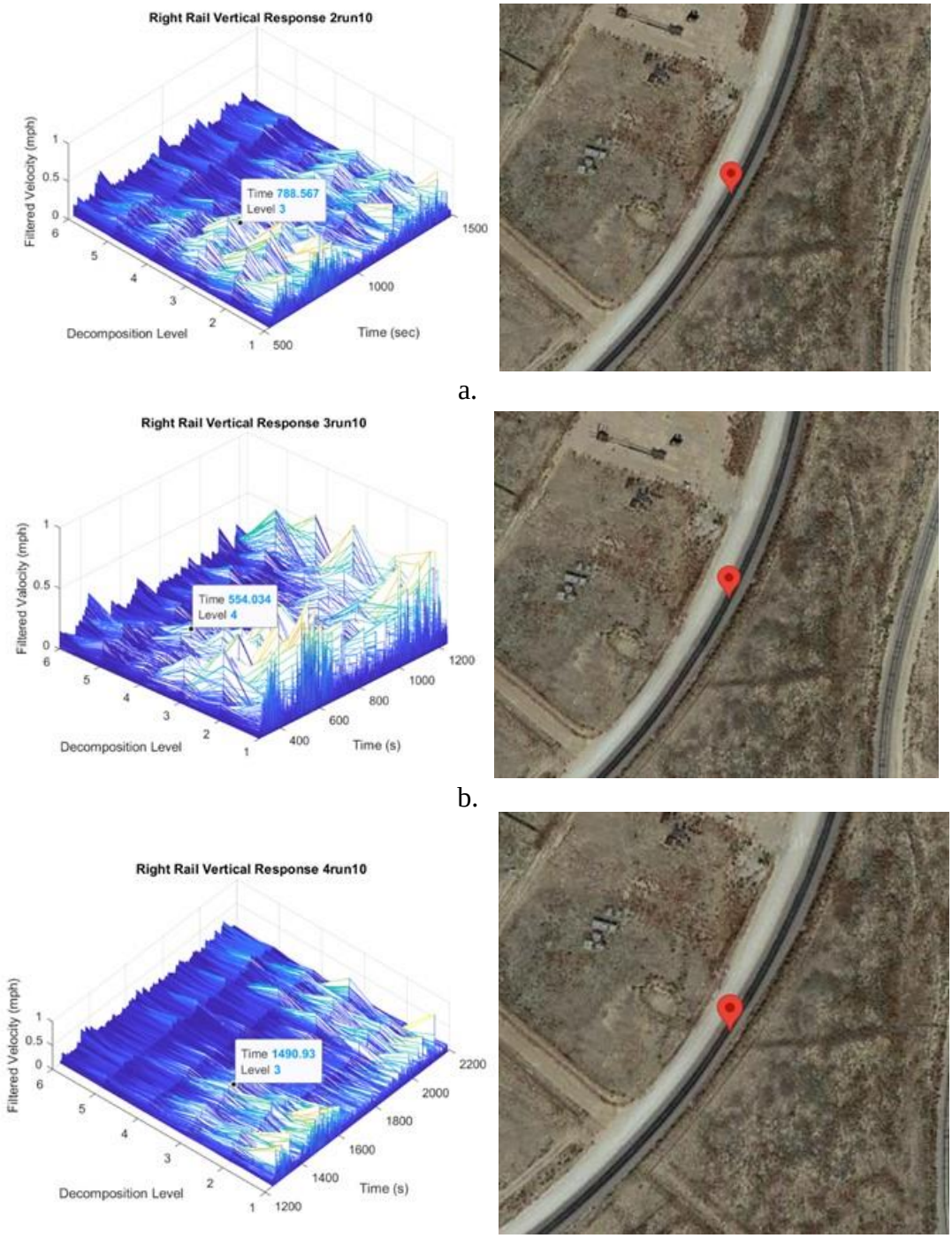


Figure 5-70: Vertical wavelet decompositions and corresponding GPS locations for the section 25 loose rail in the (a.) 2nd, (b.) 3rd, (c.) and 4th 10 mph trials

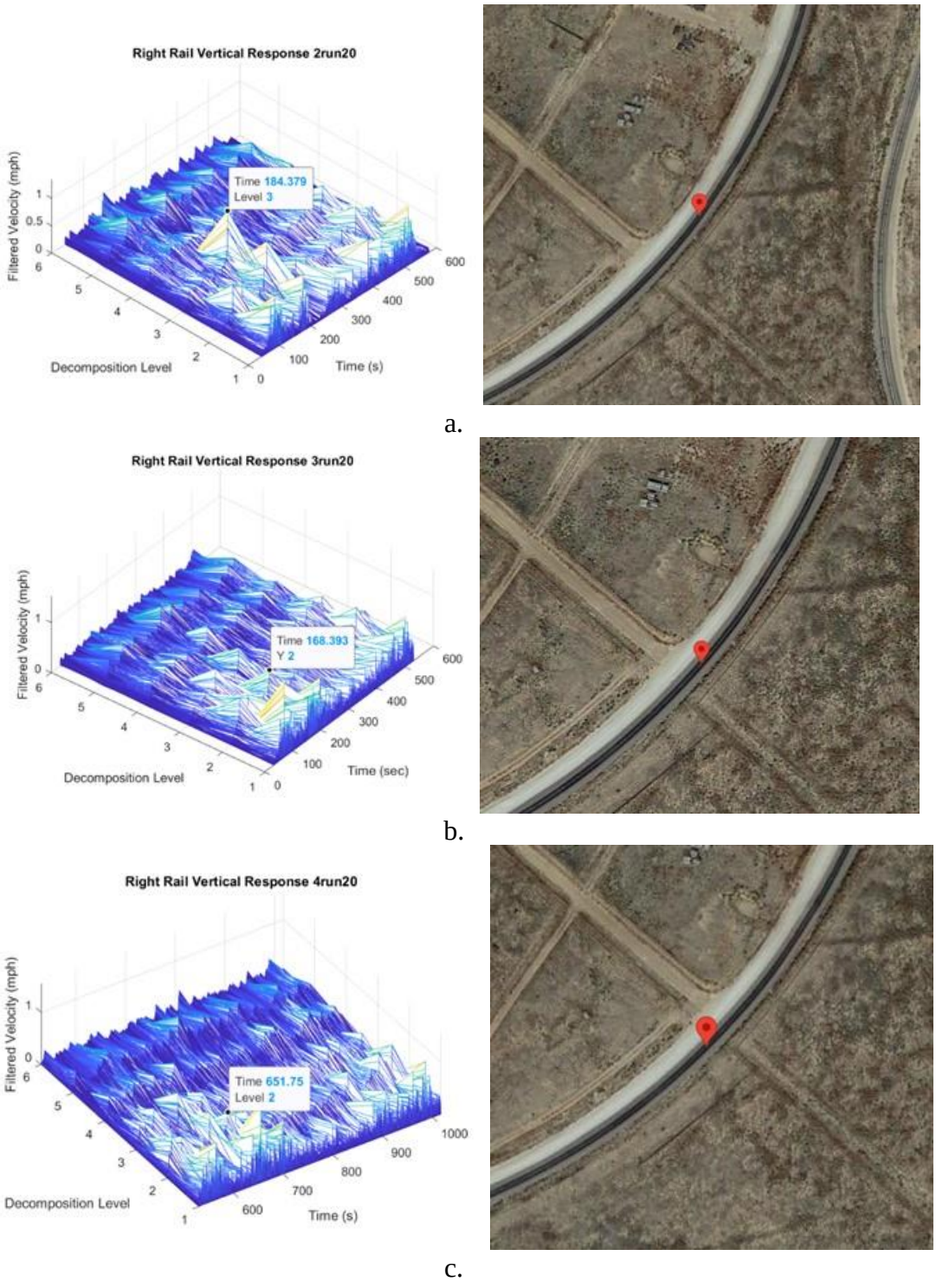


Figure 5-71: Vertical wavelet decompositions and corresponding GPS locations for the section 25 loose rail in the (a.) 2nd, (b.) 3rd, (c.) and 4th 20 mph trials

Upon further inspection of the data at these locations, it appeared that both channels were in phase, however the right channel produced a much stronger response. While the train rounded the curve,

the outer rail would serve to push the train around, but due to the superelevation, the weight of the train would be most loaded on the inside left rail. This might have prevented the outside rail from being as fully damped and therefore generating a stronger signal.

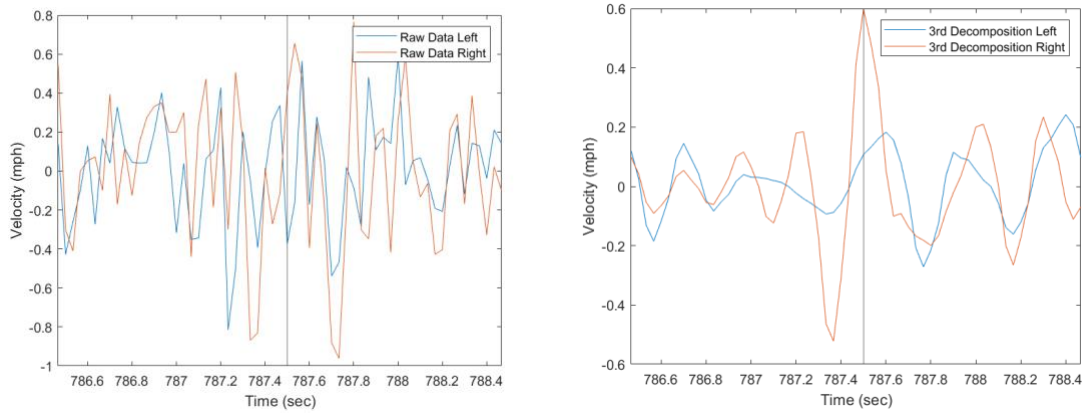


Figure 5-72: The raw (left) and 3rd level wavelet transform (right) for the left and right rail vertical responses for the emulated loose rail in section 25 at 10 mph

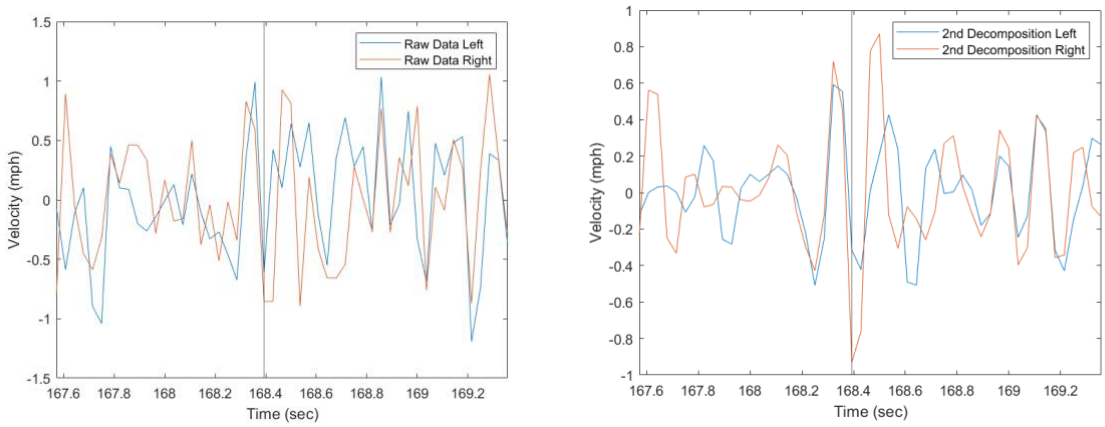


Figure 5-73: The raw (left) and 3rd level wavelet transform (right) for the left and right rail vertical responses for the emulated loose rail in section 25 at 30 mph

As described before, the method of identification of these locations occurred through manual inspection of the wavelet detailed decompositional mesh. Of the left and right rails, the channel 1 signal, which corresponded to the right rail, was more prominent for the purposes of displaying the point of interest location. One of the difficulties with this process was the surrounding responses of the signal. Compared with the other structures in the trials, such as the lubricator and switches, the response of the loose rail was a smaller effect. Surrounding the response of the loose rail in section 25 was also a series of other large deflections, which made selecting the representative spike a challenge. This particular issue did not express itself in every trial, however this example from the second 30 mph run Corsair data depicts how, without prior knowledge of the estimated location, the selection of the response of interest would be infeasible.

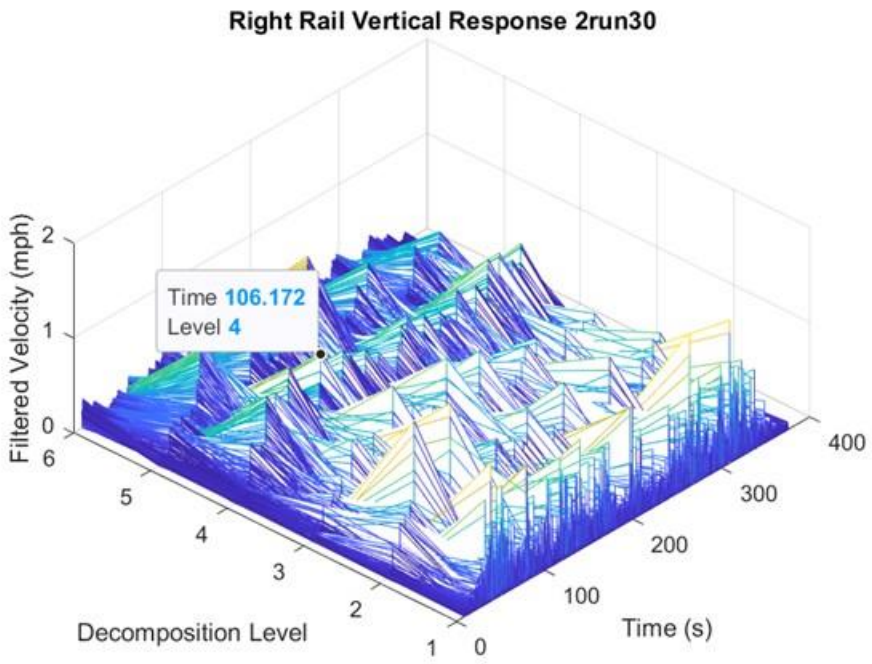


Figure 5-74: Vertical rail wavelet decomposition not clearly indicating the loose left rail

One of the aspects which helped with this low-speed identification was the selection of the level of detailed decomposition that best described the motion at the points of interest. In the majority of cases, the level of detail in the Corsair channel 1 data which indicated the point of interest most accurately compared to the surrounding motion was the 3rd and 4th detail levels. Given the speeds of travel and the sampling rate, these correspond with the effective wavelengths of between 1.8 to 7.6 feet from between 10 to 20 mph, which would span from between one to four railroad ties and hence could indicate that the effect occurs as a direct response of the changed condition of those ties.

As the speeds increased, however, the distinction of the locations corresponding to the points of interest decreased drastically. Both 30 and 40 MPH trials in Corsair and Blackbeard had an increased ability to indicate other locations such as the lubricator and switches. This can be seen in Figure 5-75, which shows the clarity with which the lubricator is found in the vertical data of the left rail from 2nd 30 MPH Corsair trial. Even without the increased false positives at higher speeds, the distinguishability of the "loose" rail is still significantly diminished.

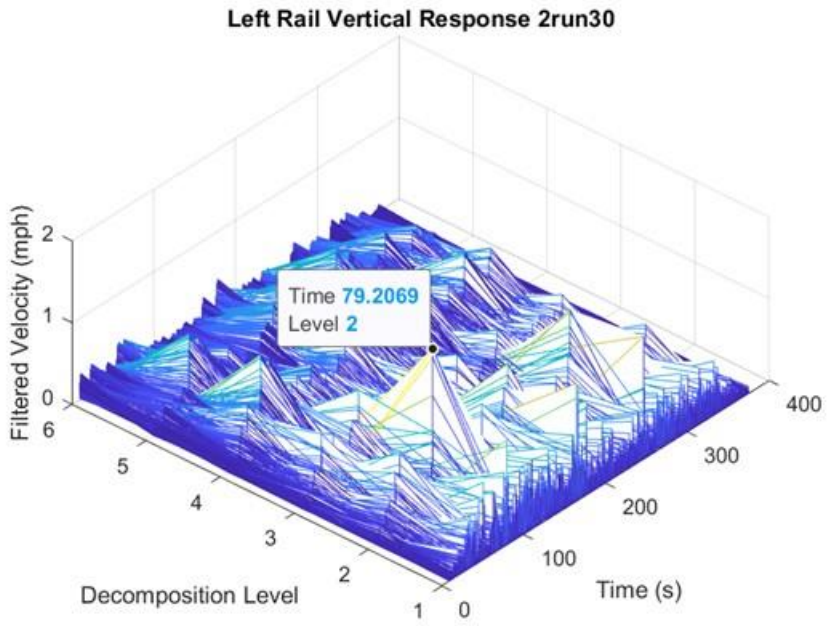


Figure 5-75: Wavelet decomposition clearly indicating the location of the lubricator at 30 mph

As with the same issues presented in the STFT, the faster speeds reduced the number of samples present over any section of the track, and with short sections of interest around 15 ft long, this made the analysis harder. The representation of the results also made the manual inspection of the meshes somewhat difficult. The responses changed rapidly enough that it was hard to distinguish between events which occurred within a short time from each other.

Overall, the method of wavelet analysis, when successful, could identify the location of interest much more precisely that from the STFT analysis. Unfortunately, it was less reliable at distinguishing the different locations around the track. Both systems were designed to identify different aspects of track movement, with the STFT looking for harmonic movement and specifically resonances or changes in resonance, while the wavelet inspection focused on reflecting the particular movement as defined by the mother wavelet. With a high enough sampling rate, and further inspection of the track movement at the points of loose rail, it could be possible to allow for the detection of these different phenomena at the higher speeds and with greater reliability at the lower ones. As of now, these processes do, however, show that it is possible to identify sections of rail with loose vertical spikes that are located within curves.

Neither the frequency domain nor the sorting algorithm methods of data analysis, however, were able to distinguish the loosened rail located in the tangent section of the track. As such, an alternative analysis method needed to be implemented to specifically target the detection of the loose track in the tangent. To do so, a manual analysis of the data was implemented to specifically inspect the data found in that tangent.

Chapter 6 Manual Analysis

While the sorting algorithm, STFT, and wavelet decomposition styles of analysis had some success with regards to identifying the loose rail in the curved section 25, none of the techniques were particularly successful with regards to the detection of the loose track in the tangent section 40. As a way to address this issue, manual inspection of the data was implemented in order to identify if there was a change in the dynamics of the signal as a result of the looser connection between the rails and the ties.

Similar to the analysis conducted using the Gaussian Mixture Model, the manual analysis and inspection of the data was also conducted by other members of CVeSS for this study. As before, the results of this analysis were presented in a previous report [21].

While manual analysis of the raw Lidar velocity data does not allow for the inspection of frequency content, it may be prudent to review the signal for changes in motion and dynamics. Without removing the low-frequency content, the responses of the rail and the train were still coupled in the data set. Analysis of this data, which was dominated by the low frequency content, provided a method of viewing changes in the track condition due to the response of the train itself. The rocking of the train car back and forth provided potential problems when collecting and analyzing the data, however, in this case the rocking provided an opportunity to observe any changes in the pattern due to changing track conditions. While it is not a direct method of track analysis, it may still provide a method of distinguishing which sections of the data indicate loose track.

Through the use of the video aligning techniques as described earlier, the locations of the points of interest were identified in the cleaned Lidar data. On the 2nd 10 mph run, the start and the end of the tangent loose section were illustrated in the data for
Figure 6-1. From manual inspection, it is clear that the location containing the loose track has a higher apparent velocity variance in the vertical data. This could be explained by the removal of the vertical tie spikes, which would reduce the resistance of the track and train system to vertical motion. As seen in the vertical data, there seems to be a pair of triple peaks which exist over the length of the section of interest, starting from the control section and bridging to the emulated loose track. The distance between the tangent control section and emulated loose track was around 150 feet. Given the length of the train consist being about 159 feet in length, this allowed for the front of the train to be starting its contact with the “loose” track while the rear was still in the control section. As such, this elevated response could be a result of the reverberations present in the “loose” track traveling back through the rail and being picked up before reaching the point itself. At the moment, however, we do not have an explanation as to the existence of this double set of peaks, but it could be possible that the three train cars each contribute to a specific peak. The overall range of these peaks do have the potential to represent the train itself, as they are small enough to be fully encapsulated by its length.

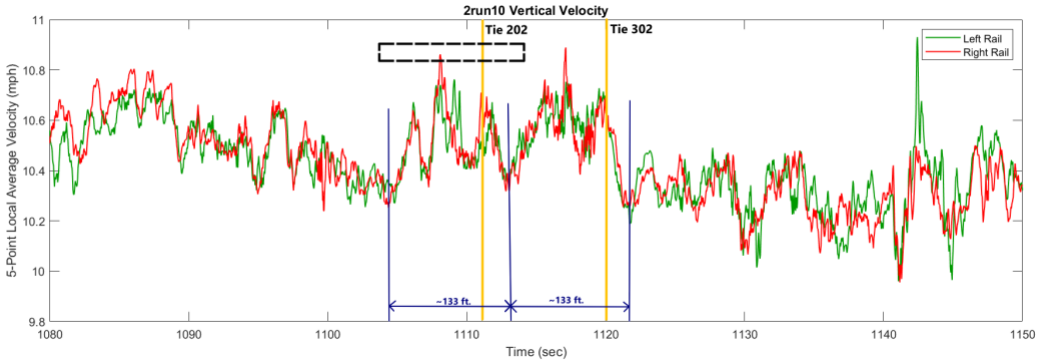


Figure 6-1: Vertical motion manual inspection of section 40

As for the lateral Blackbeard data, the rocking of the freight car is clearly demonstrated through the oscillatory nature of the two signals. As the car leans or drifts towards one direction or the other, the two velocity measurements will be moving opposite to each other i.e., they are shifted in phase by 180 degrees. During the section leading up to the rear car passing the "loosened" point of interest, the oscillations in the rocking seem to shift in phase and increase in amplitude. This increase in amplitude could correspond once more to the "loosened" rails. Both the lateral and vertical responses are aligned with one another. Since both effects are present, and correspond to the distance with which the train would be passing over the emulated loose rail, it could be inferred that this is due to the track condition.

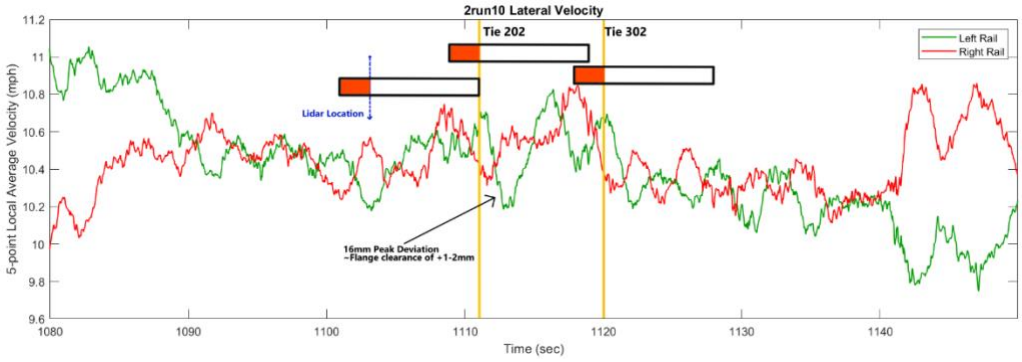


Figure 6-2: Lateral motion manual inspection of section 40

While a change in track-train dynamics is able to be identified via manual analysis, it falls under a similar condition as the STFT and wavelet forms of analysis. In all of these cases, the overall goal of implementing a data processing method which is able to distinguish between loose and non-loose track is not fully met. These methods validate that the prior sensing techniques were capable of identifying loose rail, but that there would be a risk of falsely identifying a section of track as loose. This prevents automated systems from being viable with these current techniques. Even with manual inspection, without prior knowledge of the locations of interest, differentiating the effect of track conditions on the consist would not be feasible. Further exploration of the interactions between the track structure and the resulting body motion of the train might allow for a greater specificity with which the loose rail could be detected.

In the case of manual analysis of the tangent section, one of the issues regarding the comparison between the control and track with removed ties is that the train consist was large enough to bridge the gap between them, and as such effects occurring due to the locomotive approaching the loose track could be propagated back to the control section. A larger gap between the two sections might have been necessary to prevent this feedback.

Overall, the manual inspection of the data allowed for the identification of different motion around the loose rail than that of the stiffer rail sections. This process was not able to automatically detect the changes in the rial condition, but these results indicate that the sensor platform is able to capture the changes in train and rail dynamics that reflect a change in the track structure. As such, while the tangent section proved to be much more difficult to detect vertically loose rail, the non-contact Doppler Lidar lasers are able to measure the changes in motion. While more time consuming than the other methods of data analysis, and less reliable, the manual inspection of the velocity data allowed for the validation of the system capabilities in tangent track. This difficult to acquire and unreliable result, however, does not show much future in larger scale rail monitoring systems, and so improvements will need to be made in order to detect rail with removed vertical tie spikes.

Chapter 7 Conclusion and Future Study

7.1 Summary of the study

Our purpose for this experiment was to validate the capability for our non-contact system to detect instability in railroad tracks. In previous experiments, the Doppler Lidar was used to measure the lateral dynamics of in-revenue railroad tracks, and as such, we wanted to confirm the effectiveness of those analyses. To achieve this, we conducted an experiment to compare the results of our analysis with prescribed locations of instability on the track. The first analysis method implemented was the same as was used in prior experiments, and was based on an unsupervised sorting algorithm to identify sections containing abnormal rail motion. We also incorporated several other analysis techniques in this study to compare the effectiveness of the different approaches and to expand our understanding of the capabilities of the system. These additional methods of analysis were the Short Time Fourier Transform, the Discrete Wavelet Transform, and manual forms of analysis.

To achieve this comparison between techniques, we conducted runs around the Transportation Technology Center's High Tonnage Loop at speeds of 10, 20, 30, and 40 miles per hour. These speeds allowed us to review the effect of increasing speed on the capabilities of our system to detect track abnormalities. At all of these speeds we measured the lateral and vertical track translation instantaneous velocities for both of the rails. Using GPS receivers and an accelerometer, we were able to calibrate and isolate the motions of interest during our trials. Each trial consisted of having the train travel around the full length of the track, and after passing the locations of interest, also recording the maximum lateral displacement of the track heads.

Before analyzing the data, it was cleaned through a multi-step process. This cleaning involved removing data points which fell outside of the reasonable range of speeds and replacing any missing data with data generated through the use of linear regression. The processing also decoupled the longitudinal speeds from the lateral and vertical data sets through a high pass filter.

The first technique used to analyze the data was the implementation of a Gaussian Mixture Model in order to sort the data into clusters. This process used the vertical and lateral absolute velocities as well as the variances in those velocities as inputs. The trials were split into small segments and sorted by these inputs to four clusters. Combining the result of this sorting, along with previous experience in using this process, the loose rail in section 25 was identified, though spatially imprecise due to the 300-foot segment size used in the sorting. This process also sorted the surrounding locations of that curve into its own cluster. It did not, however, identify the emulated loose rail in the tangent section 40.

The second technique to review the data was the implementation of a set of Short Time Fourier Transforms to identify sections containing larger or shifting frequency responses. The frequency content of the vibrating rail indicates whether the damping or stiffness components of the rail have changed for that segment of the track. This process required a balance between the spatial and frequency precisions, as imprecise frequency analysis would not allow for observing the different track characteristics, while imprecise spatial analysis would not allow for proper identification of the location. From this analysis, like the prior processing technique, the "loose" rail in the curved

section 25 was detected with relative reliability at speeds lover than 30 mph, but neither tangent track element generated a notable response. This process also reliably indicated the locations of the lubricator, starts and ends of curves, and track switches. Unfortunately, none of these responses were distinctly different in the data from the loose rail and so could not be filtered through an automated process. Therefore, a combination of manual analysis and GPS review was used to determine the effectiveness of this technique. One point of favorability for future studies of track dynamics through the use of Fourier analysis is that different elements tended to have different characteristics regarding dominant frequencies, but these tests did not produce sufficient evidence to determine which frequencies were most important or what their causes may have been.

Based on the tradeoff between location and frequency resolution, a multiresolutional analysis was then implemented through the application of the Discrete Wavelet Transformation. The data was processed through a discrete Daubechies 4 wavelet transform, but there was some evidence that further exploration into generating a different wavelet that might produce a more robust response. Like the Fourier transform, the dominant wavelet transform responses corresponded to a frequency of approximately 5 Hz, with variations across the speed. Overall, the characteristic wavelength was between around 1.6 to 3.2 feet in length, which would seem to indicate the results of loose rail can be found within the length of track covering one to two railroad ties. In the same fashion, these results were able to detect the curved loose rail with relative accuracy but were not able to detect the tangent track conditions. For this method and all the prior ones, there is a clear difference in the capability to detect loose rail in curved as opposed to tangent track.

To account for the lack of tangent rail detection, the next processing technique implemented was a manual review of the Lidar velocity data in the tangent section of the track. From the rocking motion generated on the freight car, we were able to identify a correlation between the increased variation in the vertical velocity and a phase shift in the lateral rocking. These results persisted across speeds, and as such suggest that the track conditions caused this effect. While this method did not allow for any form of detection of loose track without prior knowledge of its location, it did demonstrate that some elements of the track dynamics measurably change in response to emulated “loose” rail, even in tangent sections. Even though this method does not serve the purpose of creating an automated detection method, it does validate the capabilities for the system to at least detect the changes in rail conditions.

Overall, the results of this experiment and its corresponding data processing techniques show a clear preference for the identification of this particular type of loose rail in curved sections. The tangent sections, while not discernable without prior knowledge of the track conditions, do at least indicate that future analysis may be possible with greater knowledge of the reactions in tangent sections of track. If this test were to be expanded upon, the control in the tangent section of track should be placed significantly far away from the “loose” track as this experiment allowed for the train to bridge the two sections, possibly muddling the results.

7.2 Recommendations

Additionally, for all of these results, this experiment only examined the capabilities for the non-contact sensors to detect the effect of removed vertical spikes. This should allow for a greater motion of the rail, but still retain the lateral motion of the base of the rail. If the lateral retaining

spikes were removed instead, or if both sets were removed, the results could have presented themselves quite differently. Even with regards to the removal of spikes, several other forms of track instability have yet to be inspected and validated for their ability to be detected by this non-contact system. These could include earlier or more reliable detection of the other major causes of derailment, such as cracked track, which would change its vibrational characteristics.

Even within the investigation of loose rail connections and gage widening, further exploration into the use of wavelets could provide useful information. Testing different track structures for characteristic shapes of motion could, along with an appropriately selected or constructed mother wavelet, allow for the automated detection of much more specific motions within the track.

Another consideration which was not explored in these experiments, was the impact of the choice of location of the Lidar on the measured response of the track. While we knew positioning the target of the Lidar lasers either too close or too far from the point of contact of the wheel would produce bad results, we did have a specific offset selected for the pair of lasers behind the wheel. A study specifically on the results caused by different offsets from the wheel could improve the effectiveness of future Lidar studies. In the same vein, and exploration of the difference between placing the sensors in front of or behind the wheel, as well as if the wheel is on the front or the rear of the car, could be useful.

REFERENCES

[1] Y. Ren, P. Lu, C. Ai, L. Gao, S. Qiu, and D. Tolliver, "Review of emerging technologies and issues in rail and track inspection for local lines in the United States," *Journal of Transportation Engineering, Part A: Systems*, vol. 147, no. 10, 2021.

[2] X. Liu, C. P. Barkan, and M. R. Saat, "Analysis of derailments by accident cause," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2261, no. 1, pp. 178–185, 2011.

[3] C. P. L. Barkan, C. T. Dick, and R. Anderson, "Railroad Derailment Factors Affecting Hazardous Materials Transportation Risk," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1825, no. 1, pp. 64–74, Jan. 2003, doi: https://doi.org/10.3141/1825-09.

[4] "49 CFR Ch. II Part 213--Track Safety Standards," Jan. 10, 2021.

[5] L. El-khateeb, "Defect-based Condition Assessment Model of Railway Infrastructure," Thesis, Concordia University, 2017.

[6] I. A. khouy, H. Schunnesson, U. Juntti, A. Nissen, and P.-O. Larsson-Kråik, "Evaluation of track geometry maintenance for a heavy haul railroad in Sweden: A case study," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 228, no. 5, pp. 496–503, Mar. 2013, doi: https://doi.org/10.1177/0954409713482239.

[7] W. Ebersohn and E. Selig, "Use of Track Geometry Measurements for Maintenance Planning," *Transportation Research Record*, vol. 1470, pp. 84–92, 1994.

[8] L. Ferreira and M. H. Murray, "Modelling rail track deterioration and maintenance: current practices and future needs," *Transport Reviews*, vol. 17, no. 3, pp. 207–221, Jul. 1997, doi: https://doi.org/10.1080/01441649708716982.

[9] Q. Li, Z. Zhong, Z. Liang, and Y. Liang, "Rail Inspection Meets Big Data: Methods and Trends," *IEEE Xplore*, Sep. 01, 2015. https://ieeexplore.ieee.org/abstract/document/7350636?casa_token=--WAdo2I_YMAAAAA:Vj-2utda3KZKddg8xD2JAq5Wp34smyFbfPab-iCI4J2vFyMjtjr450zIVK-gZZJL0IBDcl0wg7lI (accessed Dec. 29, 2020).

[10] P. Chandran, M. Rantatalo, J. Odelius, H. Lind, and S. M. Famurewa, "Train-based differential eddy current sensor system for rail fastener detection," *Measurement Science and Technology*, vol. 30, no. 12, p. 125105, Sep. 2019, doi: https://doi.org/10.1088/1361-6501/ab2b24.

[11] Q. Li and S. Ren, "A Real-Time Visual Inspection System for Discrete Surface Defects of Rail Heads," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 8, pp. 2189–2199, Aug. 2012, doi: https://doi.org/10.1109/tim.2012.2184959.

[12] M. Karakose, O. Yaman, M. Baygin, K. Murat, and E. Akin, "A New Computer Vision Based Method for Rail Track Detection and Fault Diagnosis in Railways," *International Journal of Mechanical Engineering and Robotics Research*, pp. 22–27, 2017, doi: https://doi.org/10.18178/ijmerr.6.1.22-27.

[13] Y. Li and J. Ibanez-Guzman, "Lidar for Autonomous Driving: The Principles, Challenges, and Trends for Automotive Lidar and Perception Systems," *IEEE Signal Processing Magazine*, vol. 37, no. 4, pp. 50–61, Jul. 2020, doi: https://doi.org/10.1109/msp.2020.2973615.

[14] C. I. Rablau, “Lidar: a new self-driving vehicle for introducing optics to broader engineering and non-engineering audiences,” *Fifteenth Conference on Education and Training in Optics and Photonics: ETOP 2019*, Jul. 2019, doi: https://doi.org/10.1117/12.2523863.
[15] Y. Ma, J. Anderson, S. Crouch, and J. Shan, “Moving Object Detection and Tracking with Doppler Lidar,” *Remote Sensing*, vol. 11, no. 10, p. 1154, May 2019, doi: https://doi.org/10.3390/rs11101154.
[16] T. Wei, H. Xia, B. Yue, Y. Wu, and Q. Liu, “Remote sensing of raindrop size distribution using the coherent Doppler Lidar,” *Optics Express*, vol. 29, no. 11, p. 17246, May 2021, doi: https://doi.org/10.1364/oe.426326.
[17] J. P. Goit, A. Yamaguchi, and T. Ishihara, “Measurement and Prediction of Wind Fields at an Offshore Site by Scanning Doppler Lidar and WRF,” *Atmosphere*, vol. 11, no. 5, p. 442, Apr. 2020, doi: https://doi.org/10.3390/atmos11050442.
[18] B. Yang and L. Fang, “Automated Extraction of 3-D Railway Tracks from Mobile Laser Scanning Point Clouds,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 12, pp. 4750–4761, Dec. 2014, doi: https://doi.org/10.1109/jstars.2014.2312378.
[19] M. Taheri Andani, A. Peterson, J. Munoz, and M. Ahmadian, “Railway track irregularity and curvature estimation using doppler LIDAR fiber optics,” *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 232, no. 1, pp. 63–72, Aug. 2016, doi: https://doi.org/10.1177/0954409716660738.
[20] S. Mace, “EVALUATION OF RAIL ROLLOVER DERAILMENT STUDY,” US Department of Transportation, Springfield, Virginia, May 1993.
[21] M. Ahmadian and C. Holton, “In-Motion Track Stability Assessment using Doppler LIDAR Velocity Methods - Phase III Annual Progress Report,” Dec. 2021.
[22] K. Krishnan Nair and A. S. Kiremidjian, “Time Series Based Structural Damage Detection Algorithm Using Gaussian Mixtures Modeling,” *Journal of Dynamic Systems, Measurement, and Control*, vol. 129, no. 3, pp. 285–293, Aug. 2006, doi: https://doi.org/10.1115/1.2718241.
[23] X. Sheng, M. Li, C. J. C. Jones, and D. J. Thompson, “Using the Fourier-series approach to study interactions between moving wheels and a periodically supported rail,” *Journal of Sound and Vibration*, vol. 303, no. 3–5, pp. 873–894, Jun. 2007, doi: https://doi.org/10.1016/j.jsv.2007.02.007.
[24] A. Pieringer, “A numerical investigation of curve squeal in the case of constant wheel/rail friction,” *Journal of Sound and Vibration*, vol. 333, no. 18, pp. 4295–4313, Sep. 2014, doi: https://doi.org/10.1016/j.jsv.2014.04.024.
[25] M. T. Pham, J.-M. Kim, and C. H. Kim, “Accurate Bearing Fault Diagnosis under Variable Shaft Speed using Convolutional Neural Networks and Vibration Spectrogram,” *Applied Sciences*, vol. 10, no. 18, p. 6385, Sep. 2020, doi: https://doi.org/10.3390/app10186385.
[26] R. Thomazella, W. N. Lopes, P. R. Aguiar, F. A. Alexandre, A. A. Fiocchi, and E. C. Bianchi, “Digital signal processing for self-vibration monitoring in grinding: A new approach based on the time-frequency analysis of vibration signals,” *Measurement*, vol. 145, pp. 71–83, Oct. 2019, doi: https://doi.org/10.1016/j.measurement.2019.05.079.
[27] H. Liu, D. Yao, J. Yang, and X. Li, “Lightweight Convolutional Neural Network and Its Application in Rolling Bearing Fault Diagnosis under Variable Working

Conditions," *Sensors*, vol. 19, no. 22, p. 4827, Nov. 2019, doi: https://doi.org/10.3390/s19224827.
[28] M. I. Skolnik, *Radar handbook*. New York: Mcgraw-Hill, 1990.
[29] A. Graps, "An introduction to wavelets," *IEEE Computational Science and Engineering*, vol. 2, no. 2, pp. 50–61, 1995, doi: https://doi.org/10.1109/99.388960.
[30] P. Kumar and E. Foufoula-Georgiou, "Wavelet analysis for geophysical applications," *Reviews of Geophysics*, vol. 35, no. 4, pp. 385–412, Nov. 1997, doi: https://doi.org/10.1029/97rg00427.
[31] A. Silik, M. Noori, W. A. Altabey, R. Ghiasi, and Z. Wu, "Comparative Analysis of Wavelet Transform for Time-Frequency Analysis and Transient Localization in Structural Health Monitoring," *Structural Durability & Health Monitoring*, vol. 15, no. 1, pp. 1–22, 2021, doi: https://doi.org/10.32604/sdhm.2021.012751.
[32] "Introduction to Wavelet Families - MATLAB & Simulink," *www.mathworks.com*. https://www.mathworks.com/help/wavelet/gs/introduction-to-the-wavelet-families.html (accessed Mar. 27, 2023).

APPENDIX

MATLAB Implementation of the STFT

```
%2nd Run - 10 mph - Corsair_uniform (1)
clc
clear all
close all

%%%Take the basic methods from the single FFT and split the overall data
%%%into subvectors to look through

%Load in the file as a table
T=readtable(['2nd Run - 10 mph - Corsair_uniform (1).csv']);

%%% CORSAIR DATA

%2 run 10 Cor Un 44666
Time_Raw=T(15854:44666,5);       %Time
Ch0Vel_Raw=T(15854:44666,6);     %Velocity of channel 0
Ch1Vel_Raw=T(15854:44666,7);     %Velocity of channel 1
GPSSpeed_Raw=T(15854:44666,8);  %GPS speed

% Convert the data from table to a vector
Time_Raw=table2array(Time_Raw);
Ch0Vel_Raw=table2array(Ch0Vel_Raw);
Ch1Vel_Raw=table2array(Ch1Vel_Raw);
GPSSpeed_Raw=table2array(GPSSpeed_Raw);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
                %Mile marker for specifit run%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
location_1=Time_Raw(1);
location_2=Time_Raw(end);
Mile=linspace(Time_Raw(1),Time_Raw(end),length(Time_Raw));

%Display the velocity data for both channels and the GPS
f=0;
f=f+1;
figure(f)
plot(Mile,Ch0Vel_Raw,'Color',[0 0.6 0])
hold on
plot(Mile,Ch1Vel_Raw,'Color',[1 0 0])
plot(Mile,GPSSpeed_Raw,'b')
legend('Left Rail','Right Rail','GPS')
```

```
xlabel('Time (sec)')
ylabel('Velocity (mph)')
title('Raw Velocity Data')

%cut length
cut=length(Mile)-(2^14);
%%CUt down data
f=f+1;
figure(f)
plot(Mile(1:cut),Ch0Vel_Raw(1:cut),'Color',[0 0.6 0])
hold on
plot(Mile(1:cut),Ch1Vel_Raw(1:cut),'Color',[1 0 0])
plot(Mile(1:cut),GPSSpeed_Raw(1:cut),'b')
legend('Left Rail','Right Rail','GPS')
xlabel('Time (sec)')
ylabel('Velocity (mph)')
title('Cut Velocity Data')

%Graph the instantaneous sampling rate of the data
for z=1:length(Time_Raw)-1
    Sample(z)=1/(Time_Raw(z+1)-Time_Raw(z));
end

%Plot the sampling rate as a function of time
f=f+1;
figure(f)
plot(Time_Raw(2:end),Sample)
title('Sampling Rate')
ylabel('Sample rate (Hz)')
xlabel('Time (sec)')

%Find the average sampling rate of the data
AvgSample=mean(Sample);
disp('The average sampling rate of the data:')
disp(AvgSample);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
              %User selected segment size and cutoff frequency%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
segment=256;       %Segment Size
fc=0.318;          %frequency cutoff at 0.318 Hz
fs=AvgSample;      %sampling frequency at average sampling rate

%impliment a 3 power high-pass butterworth to clean out velocity as a
```

```
%function of longitudinal motion
[Regression_Length,i]=butter(3,fc/(fs/2),'high');

Ch0Vel_Filtered=filter(Regression_Length,i,Ch0Vel_Raw);
Ch1Vel_Filtered=filter(Regression_Length,i,Ch1Vel_Raw);

f=f+1;
figure(f)
plot(Mile,Ch0Vel_Filtered,'Color',[0 0.6 0])
hold on
plot(Mile,Ch1Vel_Filtered,'Color',[1 0 0])
legend('Left Rail','Right Rail')
title('Filtered Velocity Data')
xlabel('Time (sec)')
ylabel('Velocity (mph)')

%Split the full data vectors into segments of equal length that are
%overlapped with each other by 50%
y1=buffer(1:numel(Ch0Vel_Filtered),segment,segment/2);
Ch0Vel_Split=Ch0Vel_Filtered(y1(:,all(y1)));

y2=buffer(1:numel(Ch1Vel_Filtered),segment,segment/2);
Ch1Vel_Split=Ch1Vel_Filtered(y2(:,all(y2)));

y3=buffer(1:numel(Time_Raw),segment,segment/2);
Time_Split=Time_Raw(y3(:,all(y3)));

% % %Can Graph the new data as split vectors if needed
% % f=f+1;
% % figure(f)
% % plot(Time_Split,Ch0Vel_Split)

%Implement the FFT transformation of each segment of the data and store in
%a new matrix
Ls=segment; %Length of each data segment
Ts=1/fs; %Time segment
tv=(0:Ls-1)*Ts; %Time vector

%%Image for the splittign of the data into its component segments
f=f+1;
figure(f)
plot(Time_Split,Ch0Vel_Split)
xlabel('Time (sec)')
ylabel('Velocity (mph)')
hold off
```

```
                %%%%----Channel 0----%%%%

for i=1:length(Time_Split(1,:))
   %Apply rectangular window when converting to dB
   Ch0_FFT(:,i)=fft(Ch0Vel_Split(:,i));

   PS2_0(:,i)=abs(Ch0_FFT(:,i)/Ls);
   PS1_0(:,i)=PS2_0(1:Ls/2+1,i);
   PS1_0(2:end-1,i)=2*PS1_0(2:end-1,i);
   freq_0(:,i)=fs*(0:(Ls/2))/Ls;
end

%Generate location/Time vector for graph
segCount=linspace(1,length(Ch0Vel_Split(1,:)),length(Ch0Vel_Split(1,:)));

location_Segment=linspace(location_1,location_2,length(Ch0Vel_Split(1,:)));

%decibel magnitude of PS1_0
PS1_0_dB=mag2db(PS1_0);

%Graph the 3d plot of the power spectrum for each location down the track
f=f+1;
figure(f)
%Option to switch between decibel and PSD graphs
% % % mesh(location_Segment,freq_0,PS1_0_dB)
mesh(location_Segment(:,2:end),freq_0(:,2:end),PS1_0(:,2:end))
xlabel('Time (sec)')
ylabel('Frequency (Hz)')
zlabel('Magnitude (Mph/sqrt(Hz))')
title('Left Rail STFT')

                %%%%----Channel 1----%%%%

for i=1:length(Time_Split(1,:))
   Ch1_FFT(:,i)=fft(Ch1Vel_Split(:,i));
   PS2_1(:,i)=abs(Ch1_FFT(:,i)/Ls);
   PS1_1(:,i)=PS2_1(1:Ls/2+1,i);
   PS1_1(2:end-1,i)=2*PS1_1(2:end-1,i);
   freq_1(:,i)=fs*(0:(Ls/2))/Ls;
end

%Generate location/Time vector for graph
segCount=linspace(1,length(Ch0Vel_Split(1,:)),length(Ch0Vel_Split(1,:)));
location_Segment=linspace(307.448,310.639,length(Ch0Vel_Split(1,:)));
location_Segment=linspace(Time_Raw(1),Time_Raw(end),length(Ch0Vel_Split(1,:)));
```

```
%decibel magnitude of PS1_0
PS1_1_dB=mag2db(PS1_1);%decibel magnitude of PS1_1

%Graph the 3d plot of the power spectrum for each location down the track
f=f+1;
figure(f)
% % % mesh(location_Segment,freq_1,PS1_1_dB)%%-------------------------7 A
mesh(location_Segment,freq_1,PS1_1)%%----------------------------------7 B
xlabel('Mile Post')
ylabel('Frequency (Hz)')
zlabel('Magnitude (Mph/sqrt(Hz))')
title('Channel 1 Raw FFT')

            %%%%----Channel 0 Regression----%%%%

%%%For each FFT:
%%%take the low pass filter for the data, I.e. a running average (1.5 pole
%%%filter)
%%%Add a regression line of the filtered data
%%%Compare each point to the corresponding regression curve to look for
%%%large enough differences

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
                %Running Average Filter Settings%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
B=1/30*ones(30,1);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
                  %Frequency Shoulder Location%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
shoulder_location_0=31;

for i=1:length(PS1_0(1,:))
   PS1_0_Temp=PS1_0(:,i);
   freq_0_Temp=freq_0(:,i);
   Power_0_Avg(:,i)=filter(B,1,PS1_0_Temp);

   %Determine the linear regression line coefficients for each of the segments
   regress_line_temp_vectors=[ones(length(freq_0_Temp(shoulder_location_0:end)),1)
freq_0_Temp(shoulder_location_0:end)];
   Regression_Length=regress_line_temp_vectors\Power_0_Avg(shoulder_location_0:end,i);
```

```
   Channel_0_FFT_Regression(:,i)=regress_line_temp_vectors*Regression_Length;
   Channel_0_Freq_Regression(:,i)=freq_0_Temp(shoulder_location_0:end);
end

%Plot the average power of Channel 0 from using the Running Average filter
f=f+1;
figure(f)
mesh(location_Segment(:,2:end),freq_0(:,2:end),Power_0_Avg(:,2:end))
%removed first segment term since it was too large to see other
% details in the graph
xlabel('Time (sec)')
ylabel('Freq (Hz)')
zlabel('Magnitude (Mph/sqrt(Hz))')
title('Left Rail Running Average STFT')

%Plot the averaged FFT along with the regression line to confirm accuracy
f=f+1;
figure(f)
mesh(location_Segment,Channel_0_Freq_Regression,Channel_0_FFT_Regression)
xlabel('Time (sec)')
ylabel('Frequency (Hz)')
zlabel('Magnitude (Mph/sqrt(Hz))')
title('Left Rail STFT + Linear Regression')
hold on
mesh(location_Segment,freq_0,Power_0_Avg)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
                %Decibels Over Regression Line%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Decibels_0=6;

%ratio generated by target decibel change
ratio=10^(Decibels_0/10);

Channel_0_FFT_Regression=Channel_0_FFT_Regression.*ratio;

%Plot the raw data to find any points that pierce the surface generated by
%the regression line to indicate detectability
f=f+1;
figure(f)
mesh(location_Segment,Channel_0_Freq_Regression,Channel_0_FFT_Regression)
xlabel('Time (sec)')
ylabel('Frequency (Hz)')
zlabel('Magnitude (Mph/sqrt(Hz))')
```

```
title('Channel 0 Raw Data vs Detectability Threshhold')
hold on
mesh(location_Segment,freq_0,PS1_0)

f=f+1;
figure(f)
mesh(location_Segment,Channel_0_Freq_Regression,Channel_0_FFT_Regression)
xlabel('Time (sec)')
ylabel('Frequency (Hz)')
zlabel('Magnitude (Mph/sqrt(Hz))')
title('Channel 0 Running Average vs Detectability Threshhold')
hold on
mesh(location_Segment,freq_0,Power_0_Avg)


            %%%%----Channel 1 Regression----%%%%


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
                %Frequency Shoulder Location%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
shoulder_location_1=31;

for i=1:length(PS1_1(1,:))
   PS1_1_Temp=PS1_1(:,i);
   freq_1_Temp=freq_1(:,i);
   Power_1_Avg(:,i)=filter(B,1,PS1_1_Temp);

   %Determine the linear regression line coefficients for each of the segments
   regress_line_temp_vectors=[ones(length(freq_1_Temp(shoulder_location_1:end)),1)
freq_1_Temp(shoulder_location_1:end)];
   Regression_Length=regress_line_temp_vectors\Power_1_Avg(shoulder_location_1:end,i);

   Channel_1_FFT_Regression(:,i)=regress_line_temp_vectors*Regression_Length;
   Channel_1_Freq_Regression(:,i)=freq_1_Temp(shoulder_location_1:end);
end

%Plot the average power of Channel 0 from using the Running Average filter
f=f+1;
figure(f)
mesh(location_Segment(:,2:end),freq_1(:,2:end),Power_1_Avg(:,2:end))
xlabel('Time (sec)')
ylabel('Freq (Hz)')
zlabel('Magnitude (Mph/sqrt(Hz))')
```

```
title('Right Rail Running Average STFT')

%Plot the averaged FFT along with the regression line to confirm accuracy
f=f+1;
figure(f)
mesh(location_Segment,Channel_1_Freq_Regression,Channel_1_FFT_Regression)
xlabel('Time (sec)')
ylabel('Frequency (Hz)')
zlabel('Magnitude (Mph/sqrt(Hz))')
title('Channel 1 Averaged FFT + Linear Regression')
hold on
mesh(location_Segment,freq_1,Power_1_Avg)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
            %Decibels Over Regression Line%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Decibels_1=4;

%ratio generated by target decibel change
ratio_1=10^(Decibels_1/10);

Channel_1_FFT_Regression=Channel_1_FFT_Regression.*ratio_1;

%Plot the raw data to find any points that pierce the surface generated by
%the regression line to indicate detectability
f=f+1;
figure(f)
mesh(location_Segment,Channel_1_Freq_Regression,Channel_1_FFT_Regression)
xlabel('Mile Post')
ylabel('Frequency (Hz)')
zlabel('Magnitude (Mph/sqrt(Hz))')
title('Channel 1 Raw Data vs Detectability Threshhold')
hold on
mesh(location_Segment,freq_1,PS1_1)

%Demonstration FFT:
Test_PS1_0=PS1_0(:,12);
Test_freq_0=freq_0(:,12);

B=1/30*ones(30,1);
out3=filter(B,1,Test_PS1_0);

X=[ones(length(Test_freq_0(shoulder_location_1:end)),1)
Test_freq_0(shoulder_location_1:end)];
```

```
b=X\out3(shoulder_location_1:end);

yCalc=X*b;

f=f+1;
figure(f)
plot(Test_freq_0,out3)
hold on
plot(Test_freq_0(shoulder_location_1:end),yCalc)
xlabel('Frequency (Hz)')
ylabel('Magnitude (Mph/sqrt(Hz))')
legend('FFT','Linear Regression')
```

MATLAB Implementation of the DWT

General DWT code:

```
clear all
close all
clearvars
clc

%% Read data
NAME='2nd Run - 10 mph - Corsair_uniform (1).xls'
D=xlsread(NAME);

%%
Ch0_Vel_temp=D(:,6);
Ch1_Vel_temp=D(:,7);
time_temp=D(:,5);

Ch0_Vel=Ch0_Vel_temp(15854:44666);
Ch1_Vel=Ch1_Vel_temp(15854:44666);
time=time_temp(15854:44666);

f=1;
figure(f)
plot(time,Ch0_Vel)

xlabel('Time (sec)')
ylabel('Velocity (mph)')
title('Raw Velocity Data')

% Wavelet
nLevel=6;
TYPE = ['db4'];
[Ax, Dx, f_temp] = FilterUsingWavelet(Ch0_Vel,nLevel,TYPE,'Acc_x',f);
f=f_temp;
[Ay, Dy, f_temp] = FilterUsingWavelet(Ch1_Vel,nLevel,TYPE,'Acc_y',f);

f=f_temp;

Matrix_A_Ch0=cell2mat(Ax);
Matrix_D_Ch0=cell2mat(Dx);

Matrix_A_Ch1=cell2mat(Ay);
Matrix_D_Ch1=cell2mat(Dy);

k=1;
for i=1:nLevel
   for j=1:length(Ch0_Vel)
```

```
        finalMat_A_Ch0(i,j)=Matrix_A_Ch0(k);
        finalMat_D_Ch0(i,j)=Matrix_D_Ch0(k);
        finalMat_A_Ch1(i,j)=Matrix_A_Ch1(k);
        finalMat_D_Ch1(i,j)=Matrix_D_Ch1(k);
        k=k+1;
    end
end

y=linspace(1,nLevel,nLevel);

f=f+1;
figure(f)
plot(time.',finalMat_D_Ch0(1,:))
title('First Detailed Decomposition')
xlabel('Time (sec)')
ylabel('Filtered Velocity (mph)')

f=f+1;
figure(f)
plot(time.',finalMat_A_Ch0(1,:))
title('First Approximate Decomposition')
xlabel('Time (sec)')
ylabel('Filtered Velocity (mph)')

f=f+1;
figure(f)
mesh(time.',y,abs(finalMat_A_Ch0))
title('Left Rail Approximate Vertical Response 2run10')
xlabel('Time (s)')
ylabel('Decomposition Level')
zlabel('Filtered Velocity (mph)')

f=f+1;
figure(f)
mesh(time.',y,abs(finalMat_D_Ch0))
title('Left Rail Vertical Response 2run10')
xlabel('Time (sec)')
ylabel('Decomposition Level')
zlabel('Filtered Velocity (mph)')

f=f+1;
figure(f)
mesh(time.',y,abs(finalMat_A_Ch1))
title('Right Rail Approximate Vertical Response 2run10')
xlabel('Time (s)')
ylabel('Decomposition')
```

```
zlabel('Filtered Velocity (mph)')

f=f+1;
figure(f)
mesh(time.',y,abs(finalMat_D_Ch1))
title('Right Rail Vertical Response 2run10')
xlabel('Time (s)')
ylabel('Deomposition Level')
zlabel('Filtered Velocity (mph)')
```

FilterUsingWavelet – function

```
function [A, D, f] = FilterUsingWavelet(x,nLevel,TYPE,Title,f)

[A, D]=GetDWT(x,nLevel,TYPE);

f=f+1;
figure(f);
subplot(nLevel+1,2,[1 2]);
plot(x,'r');
ylabel('x');
title(Title)

c=2;
for i=nLevel:-1:1
   c=c+1;
   subplot(nLevel+1,2,c);
   plot(A{i},'b');
   ylabel(['a_{' num2str(i) '}']);

   c=c+1;
   subplot(nLevel+1,2,c);
   plot(D{i},'g');
   ylabel(['d_{' num2str(i) '}']);
end
end
```

GetDWT – function

```
function [A, D]=GetDWT(x,N,wname)

   [C, L]=wavedec(x,N,wname);

   A=cell(N,1);
   D=cell(N,1);
   for i=1:N
      A{i}=wrcoef('a',C,L,wname,i);
      D{i}=wrcoef('d',C,L,wname,i);
   end

end
```

ACKNOWLEDGEMENTS

The authors wish to thank and acknowledge the US Department of Transportation, University Transportation Center Program (RailTEAM UTC) for funding support for this research.

ABOUT THE AUTHORS

Mehdi Ahmadian, J. Bernard Jones Chair and Director

Dr. Mehdi Ahmadian is a Dan Pletta Professor of Mechanical Engineering at Virginia Tech, where he also holds the position of Director of Center for Vehicle Systems and Safety (CVeSS), and the Railway Technologies Laboratory (RLT). Dr. Ahmadian has authored more than 130 archival journal publications and more than 250 conference publications, including a number of keynote lectures. He has served as Editor or Editor-in-Chief for four journals on Vehicle System Dynamics, Vibration and Control, Shock and Vibration and Automobile Engineering. Dr. Ahmadian is Fellow of American Society of Mechanical Engineers of the American Institute for Aeronautics and Astronautics (AIAA). He has received many distinguished scholar awards.

Ian A. Larson

Mr.Ian Larson was a Graduate Research Assistant for his Masters' degree in mechanical Engineering. He combined his interest in mechanical design with the study of dynamic systems. He was researching railroad track vibration mechanics and the instrumentation and processing associated with it. He has his bachelor's degree in mechanical engineering from Virginia Tech as well.