SOFTWARE ENGINEERING METHODOLOGIES IN DEVELOPING A RAILWAY CONDITION MONITORING SYSTEM

A Dissertation
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Bhavana Bhardwaj

In Partial Fulfillment of the Requirements
for the Degree of
DOCOTOR OF PHILOSOPHY

Major Department:
Computer Science

July 2022

Fargo, North Dakota

North Dakota State University
Graduate School

Title

SOFTWARE ENGINEERING METHODOLOGIES IN DEVELOPING A RAILWAY CONDITION MONITORING SYSTEM

By

Bhavana Bhardwaj

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

DOCTOR OF PHILOSOPHY

SUPERVISORY COMMITTEE:

Dr. Kendall Nygard
Chair

Dr. Raj Bridgelall

Dr. Simone Ludwig

Dr. Pan Lu

Approved:

08/04/2022 — Date

Dr. Simone Ludwig — Department Chair

ABSTRACT

With the continuous growth of rail track geometry irregularities due to aging, environmental factors, and wheel loads, rail track requires frequent maintenance. Railroads often rely on the precise and correct localization and identification of track irregularities that significantly destroy infrastructure and create life-threatening environments. Therefore, monitoring the conditions of the railroad tracks is vitally essential for ensuring safety, reliability, and cost-efficiency of operations. Consequently, agencies inspect all tracks twice a week per federal track safety regulations. However, their existing methods of track inspection are expensive, slow, require track closure, and pose a high risk to workers. The technical constraints of these methods impede network-wide scaling to all railroads. More frequent, continuous, and network-wide monitoring to detect and fix irregularities can help to reduce the risk of harm, fatalities, property damages, and possible financial losses.

This work introduces and develops a generalized, scalable, affordable inspection and monitoring system called Railway Autonomous Inspection Localization System (RAILS). In particular, the study aims to detect, locate, and characterize track-related issues. The research focuses on designing RAILS architecture, implementing data collection, and building algorithms that include inertial signal feature extraction, data processing, signal alignment, and signal filtering.

Case studies validate and characterize system accuracy by estimating the position of detected irregularities based on a linear referencing system. In one case study, the estimated position of the irregularity is compared with the actual position of ground truth data (GTA) observed by a railroad inspector. In another case study, a railroad inspector verifies the estimated

position of the irregularity to demonstrate the system's effectiveness and affordability for practical applications.

Therefore, railroad agencies employing the developed methods will benefit from reliable track and equipment conditions to make informed decisions that will lead to resource optimization. The conclusion of this research outlines the significant potential of the proposed system, including limitations and future work for practical, real-time, and autonomous implementation.

ACKNOWLEDGMENTS

I would like to thank those who contributed and helped me complete my study by writing a few words. First, I am deeply grateful to my advisor, Dr. Kendall Nygard, for the guidance, support, encouragement, and advice throughout my graduate studies. And I am thankful for his flexibility and understanding at critical moments in my life.

I would like to sincerely thank my co-advisor and supervisor, Dr. Raj Bridgelall, for his confidence, expertise, support, and inspiration during this work; his experience in this field and advice has been invaluable throughout the process of this study. It would be hard to overstate how much I have learned and benefited from his support, knowledge, experience, and expertise.

I would like to thank Dr. Pan Lu for her constant support and valuable assistance throughout this journey. She always motivated and helped me work on this study and being a member of my committee.

My gratitude is extended to Dr. Simone Ludwig for her personal and professional guidance and for supporting my graduate studies at NDSU. I would also like to thank Dr. Oksana Myronovych for being a true mentor. Dr. Oksana Myronovych helped and supported me to flourish in my career.

I would also like to express my appreciation to Upper Grate Plains Transportation Institute for providing me the opportunity to work on this project.

My heartfelt gratitude goes out to my family for their encouragement and the sacrifices they have made over the years. Finally, I am thankful, especially to my husband, Neeraj, for his patience and for never giving up on me and our daughter, Eshana Dhingra.

DEDICATION

To my parents

Late Mr. V.S. Sharma, and Mrs. Sushila Sharma

To my In-Laws

Late Mr. M.M. Dhingra, and Mrs. Niru Dhingra

To my brother and his families

Vineet Bhardwaj, Raina Bhardwaj, and Prisha Bhardwaj

To my brothers-in-laws and their families

Gaurav Dhingra, Ruchi Dhingra, and Anay Dhingra

Sourabh Dhingra, Aarti Malhotra, Zoe Dhingra, and Kai Dhingra

To my Husband

Neeraj Dhingra

To our daughter

Eshana Dhingra

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The railroad is the crucial sector of transport that significantly affects the economic growth of a nation. Railroads of each country are very vulnerable, and its continuous deterioration due to aging, environmental factors, and wheel loads on the rails can trigger possible damages [1]. Therefore, monitoring the conditions of the railroad is vitally essential for ensuring railroad safety, reliability, and cost-efficiency of railroad operations. For the railroad industry, monitoring is to identify faults, understand their causes, and predict their occurrence by identifying and characterizing track irregularities. Rail track irregularity or abnormality is an interchangeable term that is widely used in the railroad industry.

Track irregularities affect ride quality and the safety of vehicle equipment and train operation [2]. For that reason, railroad companies strive to attain the highest quality possible track geometry standard [3]. The basic parameters of track geometry include gauge, cross-level, vertical profile, horizontal alignment, the warp (twist or cross-level deviation) [4] [5]. Deviations from the designed track geometry increase the risk of derailment by intensifying angular movements and linear accelerations, which can lead to accidents, results in traffic delays and financial losses. According to the FRA's current safety regulation and rulemaking proceeding report, track issues are the second major cause of train accidents. Consequently, most serious events involving train derailments result in releasing hazardous material or harm to rail passengers [5]. Figure 1.1 shows track-related issues are the second highest cause of rail accidents followed by human error.

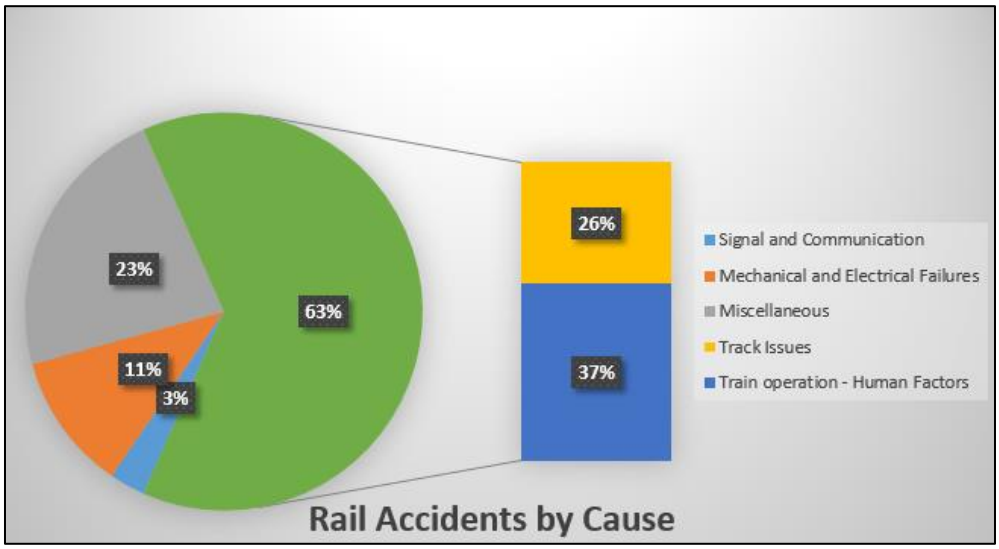


Figure 1.1. Rail accidents due to different causes from 2009-2018

Derailment is the most prominent and significant type of train accident in the United States and preventing them has been a focus area of the rail industry and the government [6]. Consequently, railroads deploy traditional inspection and monitoring methods to search for possible abnormalities. However, these existing methods are laborious, relatively slow, and decrease rail productivity by track closure to search for possible track defects. For that reason, federal track safety regulation requires railroads to inspect all tracks in operation as often as twice weekly [7]. Therefore, railroad maintenance systems integrated with advanced technologies are used to help inspectors detect and localize anomalies that can cause casualties and financial losses. However, railroads cannot afford to use these complex technologies to monitor and inspect all railroad tracks or important routes more often. Subsequently, railroads hire trained inspectors to find and fix anomalies.

Consequently, some track anomalies often go undetected during the monitoring period until they cause significant destruction. Correspondingly, many track sections are only monitored on an annual basis, leading to significant gaps in the knowledge of the current state of many of the lines [8]. Moreover, such resources are generally not available to short lines. Hence, those

lines rely on hi-rail vehicles to identify track irregularities subjectively and to verify them by measuring parameters manually.

Therefore, more frequent, continuous, and network-wide monitoring to detect and fix irregularities can help to reduce the risk of harm, fatalities, property damages, and possible financial losses. Recently, the federal government also mandated railroad to implement positive train control (PTC) technology in their respective systems [9]. PTC system is not a single technology, but it integrates various components such as the locomotive computer, wayside device, communication network, and back office. The locomotive computer, an onboard equipment that accepts speed restriction and movement information to compared against the train's location to ensure compliance. The wayside device on the side of the track is capable of monitoring and reporting switch position and signal status to locomotive computers and the back office. The back office is a centralized office system consists of the back- office server, the geographical information system (GIS), and the dispatch office which interface with other components of the PTC systems.

With these integrated components, PTC system uses a combination of communication networks, GPS (or transponder), and a fixed wayside signal device to send and receive data about the location, direction, and speed of trains. Back offices process these data in real time and provide movement authority and speed restriction information to locomotive computers. Then locomotive computers accept the information and compare it against the train's condition to ensure safety compliance. So, whenever a train crew fails to properly operate within specified safety parameters, PTC systems automatically apply the brakes and bring the train to a stop.

Regardless of preventing accidents due to human error, PTC also designed to prevent train-to-train collision, derailments caused by excessive speeds, unauthorized incursions into

work zones, and movements of trains through misaligned railroad switches. However, due to functional restrictions, there are some challenges in developing and deploying PTC system. Such challenges are the operating system does not track the real-time train location and can only communicate when a train passes the wayside infrastructure; Wi-Fi is not practical in the PTC [10], which restricts greater data transferability, railroad cyber-security risk, en-route failure of PTC, grade-crossing protection, and resource constraints prevents many freight and commuter railroads from meeting the PTC deadline.

A more expansive version of PTC, Communication-based train control (CBTC) is a more sophisticated computer-aided dispatching framework that uses GPS to track the location and speed of the train that sent to a central location, which then disseminates the information to all entities in the network. The central control then automatically sends speed restrictions and movement authorities to individual trains and checks for potential derailment and collisions. However, the system requires seamless communication coverage along the entirety of PTC-equipped track, as temporary communication loss can pose safety risks. The need for constant communication also requires significant investment in either radio towers or fixed transponders. These requirements raise the capital cost, making CBTC more expensive [10]. Irrespective of functional capabilities and allied costs, both PTC and CBTC limited to focusing on improving communication between the system and railroad engineers to reduce specific train accidents and their severity. Nonetheless, these technologies do not focus on detecting track related issues that causes most accidents. Therefore, this work focuses on introducing and developing a generalized and cost-effective system that focuses on detecting track related issues.

1.1. Research Objective

The primary goal of this research is to develop and introduce a low-cost track condition monitoring system called Railway Autonomous Inspection Localization System (RAILS) to reinforce condition monitoring to provide safe and seamless operation. The system will locate and characterize the possible track irregularity by analyzing regular service vehicles' inertial dynamics.

This research develops an approach that assesses the inertial events or responses of low-cost inertial sensors aboard revenue service trains or locomotives, which could indicate the presence of irregular railroad track geometry. Much of the previous research in this area has explored the concept of using on-board sensors to continually monitor the roughness of road pavement [11] [12] [13]. This method has the potential to reduce the monitoring cost and provide more accurate results. The same technique applies to railroads that presents an opportunity to meet all expectations and requirements. The technique utilizes low-cost inertial sensors that are ubiquitous in smartphones to provide unimpeded and continuous monitoring of the entire network. Consequently, this research uses an Android app called “Railroad Infrastructure and Vehicle Evaluation Technology” (RIVET), which can access all the required sensors in smartphones. Thus, the capability of frequent monitoring will enhance the efficiency of traditional track inspections by focusing inspection resources on high-risk locations and helping the assets manager to possess reliable track and equipment condition information to make informed decisions, leading to resource optimization.

1.1.1. Workflow

This section describes the research plans and objectives associated with the development of the RAILS.

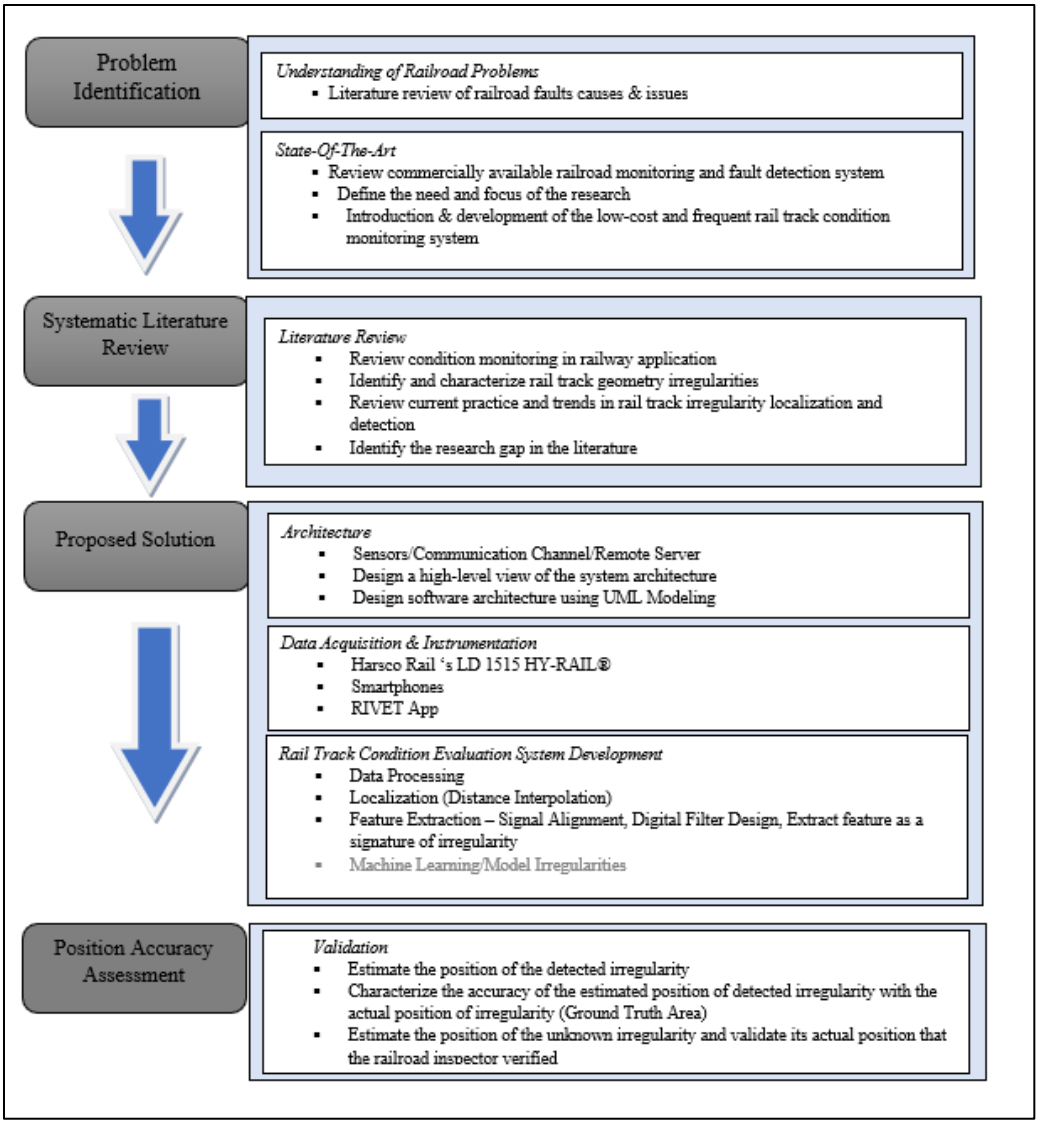


Figure 1.2. Research plan workflow

Based on the research plan workflow, following are the outlines of the research objectives:

1) Design of the RAILS architecture and description of its implementation.
2) Develop a method of processing large scale railroad condition monitoring data.
3) Develop methods that enhances irregularity localization due to GPS error and sample rate variation of the accelerometers. The method also includes the signal position alignment that will enhance signal quality and reduce signal detection error. The

technique leverages the large volume of sensor signals from multiple traversals of a road or rail segments by ensemble average.

4) Design a digital filter by measuring the appropriate cutoff frequency to maximize the signal-to-noise ratio for subsequent feature extraction.
5) Develop a statistical model to extract and combine features of the inertial signals from multiple traversals of a road or track segment. The technique will enhance the detection and localization accuracy of road or track irregularities.
6) Validate the accuracy of estimated position of detected track irregularity.

1.2. Research Contribution

The widespread use of smartphones embedded in all required sensors and network connectivity has become available for network-wide condition monitoring applications. However, the geospatial position estimates from the low-cost GPS receivers are inaccurate due to satellite signal blockage because of tall buildings or narrow streets, bridges, trees, or tunnels called non-line-of-sight (non-Los) condition, causing non-uniform update rates [14]`. Other reasons are: First, low-cost GPS receivers provide position updates approximately once per second. That means, that when the inertial sensor samples at 90 hertz on an average, the GPS coordinate will update after every group of 90 inertial samples. To be precise, the system will tag every block of 90 inertial samples with the same GPS coordinate. Second the GPS updates from individual traversals will be at different geospatial points along the path. Consequently, some position updates will not tag with some signal peaks. Also, the standard deviation (STD) of the position estimates from GPS receivers is three to five meters along the travel direction [15]. Additionally, non-uniform sample rate of accelerometer causes a problem in signal detection that results in additive signal noise and position alignment errors. In another words, the non-uniform

sampling caused each signal to have a different number of samples, and the distances associated with each sample was also non-uniformly distributed. Therefore, it reduces the signal-to-noise ratio (SNR) and pose a significant challenge in detecting, localizing, and characterizing irregularities at high accuracy and high precision. Specifically, it increases the false positives and false negatives.

Therefore, the primary contribution of this research is to develop data processing, signal processing algorithms and models that will transform the on-board sensors data into track geometry irregularities equivalent.

This research utilizes a mathematical transform called “Road Impact Factor” (RIF Transform) that will be applicable for all facility types and at all speeds [12] and produce continuous assessments of irregularities for the entire network called Total Ride Quality (TRQ). The Total Ride Quality (TRQ) is defined as the resultant vector magnitude of the RIF derived by integrating accelerations in the lateral and vertical directions. The developed method also visualizes data by using color-coded TRQ values onto maps of the rail routes using a suitable geographical information system (GIS) platform.

Another contribution is the ensemble averaging the inertial signals from the same position along multiple road or rail segment traversals at any speed. Consequently, it will improve the signal quality as the signal is correlated and reduce noise because of their un-correlation. Therefore, this research focuses on design RAILS architecture and implementation, and building the statistical model that includes inertial signal feature extraction, data processing, signal alignment, and signal filtering. The following are the list of publications from this research work:

Journal Articles:

1. Bhardwaj, B., Bridgelall, R., Lu, P., & Dhingra, N. (2021). "Signal Feature Extraction and Combination to Enhance the Detection and Localization of Railroad Track Irregularities." *IEEE Sensors,* 21(5). DOI: 10.1109/JSEN.2020.3041652.
2. Bhardwaj, B., Bridgelall, R., Chia, L., Lu, P., & Dhingra, N. (2020). "Signal Filter Cut-off Frequency Determination to Enhance the Accuracy of Rail Track Irregularity Detection and Localization." *IEEE Sensors,* 20(3). DOI: 10.1109/JSEN.2019.2947656, pp. 1393-1399.
3. Bridgelall, R., Chia, L., Bhardwaj B., Lu, P., Tolliver, D., & Dhingra, N. (2019). "Enhancement of Signals from Connected Vehicles to Detect Roadway and Railway Anomalies." Measurement Science and Technology. DOI: 10.1088/1361-6501/ab5b54.

Conference Proceedings & Presentation:

1. Bhardwaj, B., Bridgelall, R., Lu, P., Nygard, K., & Dhingra, N. (2020). "Architecture for an Intelligent Low-Cost Rail Track Condition Evaluation System." In *ASCE International Conference on Transportation & Development*. DOI: 10.1061/9780784483145.020.
2. Bhardwaj, B., Bridgelall, R., Lu, P., & Dhingra, N. (2020, January). "Signal Feature Extraction and Combination to Enhance the Detection and Localization of Railroad Track Irregularities." *The 99th Annual Meeting of the Transportation Research Board*, Washington, D.C.
3. Bhavana, B., Bridgelall, R., Lu, P., & Dhingra, N. (2019). "Railroad Track Irregularities: Position Accuracy Assessments Using Low-Cost Sensors on a Hi-Rail

Vehicle." In *Proceedings of the ASCE International Conference on Transportation & Development* (ICTD 2019). DOI: 10.1061/9780784482575.043.

1.3. Organization of the Dissertation Proposal

Chapter 2 describe a literature review on the general state of condition monitoring and its application in railways. It also illustrates gap and scope for the improvement of established monitoring and inspection techniques. Chapter 3 introduced the architecture of the proposed system. It described the high-level view system architecture and its implementation. It also designs the software architecture of the system using UML modeling. Chapter 4 briefly explains the data processing technique to visualize the irregularities proportional to the amount of peak inertial events. Chapter 5 develop and introduced the methods required to enhance the detection and localization of the track irregularity. Chapter 6 shows the potential use of low-cost sensors aboard the regular vehicle. The demonstrated two case studies that estimate and validate the position of detected irregularity 1) by comparing it with ground truth area (GRA) overserved during the rail inspection and 2) by verifying the estimated position of unknown irregularity. Chapter 7 summarizes the conclusion and future work of the research.

2. LITERATURE REVIEW

Currently, the sector of condition monitoring is proliferating and covers a large scale of industries and applications. Research on railway condition monitoring and inspection techniques has been evolving over the years in the railway industry. Some condition monitoring techniques have been introduced and established into rail track networks and rolling stock.

A literature review describes the general state of condition monitoring and summarizes the research on condition monitoring in railway applications. Conversely, a review discloses gaps in the literature for the enhancement of established monitoring techniques.

2.1. Condition Monitoring

Generally, condition monitoring is a technique to determine a state of a system according to the parameters of a system. Frequently monitoring and taking notes of conditions or any irregularities that would reduce equipment lifespan allows scheduling preventive maintenance tasks to address the issues before they cause major failures and avoid their consequences [16] [17]. The goal is, with the use of collected real-time data, the current condition of a system or equipment can be monitored.

In large-scale industries, the cost of maintenance can be high as 40% of their total budget. Nonetheless, inadequate maintenance could lead to major accidents, which can cause environmental pollution and damage to human lives [18]. Therefore, condition monitoring would help in optimizing the use and efficiency of a system or equipment and significantly increase profitability. In the survey of condition monitoring [19], Neale and Woodley addressed the potential benefits of condition monitoring, such as reduced injuries and fatal accidents to personnel caused by reduces inspection time and speeds up the start of correct remedial action.

There are two main condition monitoring methods, and these are continuous and periodic monitoring [19] [20]. Continuous monitoring is the practice of recording continuous or regular measurements during machine operation at a fixed sampling rate. It is often recommended for those critical components with a high impact on the costs and a short time-to-failure. In other words, when the cost of replacing a component or the complete breakdown of a malfunctioning component is too high. In periodic monitoring, measurements are recorded at scheduled time intervals. It is particularly suitable for non-critical components or where failure either can be predicted easily or where changes occur very slowly, that is, medium-high time-to-failure.

2.2. Condition Monitoring in Railway Application

Condition Monitoring detects and identifies faults in rail structure and infrastructure before they cause a failure or prevents rail operations [21]. In the railway industry, vehicle components, track conditions, and derailments are attractive targets for condition monitoring because of the growing need for safety, reliability, and minimal cost operations. Generally, railroad condition monitoring techniques locate or find degradation of suspension, derailment risks, or track faults such as track irregularities [22]. Therefore, the condition monitoring technique should be selected based on how it can handle severe nonlinear systems, robustness, sensitivity to disturbances, and computation performance [23].

2.2.1. Rail Track Irregularities

Rail track irregularities are the primary concern of the railroad industry because they can lead to accidents, resulting in traffic delays and financial losses. In the railroad industry, railroad irregularities or faults are interchangeable terms. Repeated stress from heavy axle loads increases the risk of deviations from uniform track geometry [24]. Consequently, it may result in heavy tread wear, fatigue cracks, high noise, freight damage, passenger discomfort, and in extreme

cases, derailments [25]. Therefore, railroads spend billions of dollars each year on infrastructure inspections and maintenance [26]. The deviation is an abnormal track surface feature, characteristic, or occurrence in the rail track surface. This research only focuses on track surface irregularities detection.

Track geometry consists of several parameters that describe each rail's position, or the track centerline occupies in space. Track geometry parameters can be specified or grouped by projecting into various planes they reside in. There are horizontal plane parameters, longitudinal vertical plane parameters, transverse vertical plane parameters, and track plane parameters. The main parameters of the track geometry include profile (longitudinal vertical plane), alignment (horizontal plane), cross-level and warp (transverse vertical plane), and gage (track plane) [25], [27]. The profile, cross-level, and warp represent the track surface condition, and the tool used to measurements is called a track level [27], [28]. Figure 2.1 shows the profile, alignment, and warp track geometry parameters.

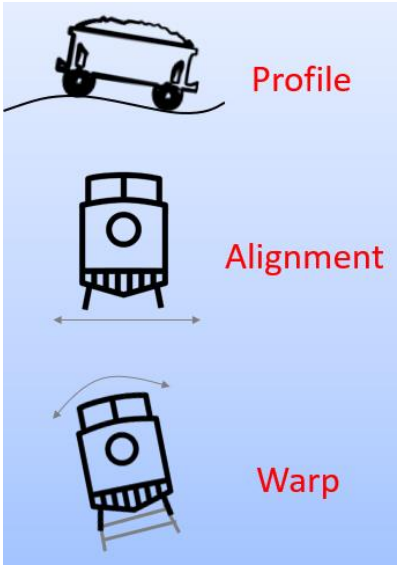


Figure 2.1. Profile, alignment, and warp track geometry parameters
Image Source: (Benefit Cost Analysis of Railroad Track Monitoring Using Sensors Onboard Revenue Service Trains, 2021 [29])

A *profile* is the track geometry of each rail, or the track centerline projected onto the longitudinal vertical plane. Deviation in the *profile* is the change in elevation of the two rails along with the track relative to a designated grade. This type of condition might happen where the track is especially muddy right under an unloading spout [27], [28]. Typically, a *profile* is measured as the vertical deviation in a certain chord length. If the profile is more than 3 inches, take remedial action like tamping fresh ballast [28] . *Alignment* is an indicator of how well-positioned the rails are horizontally along the intended route. *Alignment* deviation is the difference between the designated alignment and what is actual. It is measured as the horizontal distance between the gauge side railhead and the center of a certain chord length (measured at 62-foot string length). *Gauge* is the right-angle distance between the two rails at a given location, measured 16 mm below the top surface of the railhead. *Cross-level* is the deviation between the top surfaces of two rails at a given location. Basically, if the track is straight, elevation should be at track level and have a half-inch of elevation if the track is curved. Moreover, *cross-level* should be notified if it is greater than 3 inches. That means, on a curve, the limit is more than 3 ½ inches and more than 2 ½ inches on the straight line [7], [28]. *Warp* is the difference in cross-level of any two points within the specified distance along the track. The warp parameter is used to specify the maximum in the cross-level difference of the track in any segment. The limit for warp is 3 inches [7], [28]

There are some other track irregularities such as *broken rail, rail joint, and turnouts*. *Broken rail/track* means that either fish plate is removed, or track is broken at any section. *Broken rails* are the leading cause of derailment; therefore, they must be changed before another car or locomotive may pass over it. The *joint* bar holds ends of the rail together. If a joint bar is cracked or broken between the middle two bolt holes, it must be replaced. There must be at least

one bolt in each rail end at a joint. If the rail at a joint is mismatched more than ¼ inch, it must be corrected [28]. A *turnout* is where one-track separates from another. *Turnouts* are the most important piece of the track and require more preventive inspection.

2.2.2. State-of-Arts Inspection and Monitoring Techniques

Railroad supplements many tracks inspection & monitoring techniques to maintain safe operation. These techniques include manual inspection and non-destructive evaluation methods. This section briefly describes and reviews the main NDE methods for rail inspection. However, some monitoring systems exist besides NDE methods that many railroad companies use.

2.2.2.1. Ultrasonic Inspection

Ultrasonic inspection is normally done in a pulsed or a beamed mode [30]. A piezoelectric element produces an ultrasonic energy pulse that is transmitted into a rail. Some portion of the ultrasonic pulse is then reflected or scattered back to a receiving transducer [30], [31] . Here, the information collected from the magnitude of the signals is used to identify defects. The standard ultrasonic inspection equipment or special high-speed testing trains carrying ultrasonic probes. These inspection vehicles use liquid-filled rubber wheels to couple the excitation energy into the rail. The inspection vehicle provides different angles for detecting the defects such as 0 degree, 37 degree, or 45 degree, and 70 degree [31], [32]. The typical ultrasonic frequency is 2.25 MHz. The speed of the inspection testing trains varies from 10 km/h up to 100km/h.

Despite being good at detecting flaws in railhead and web, the technique misses the cracks smaller than 4 mm deep [31], [32]. Also, maintaining the coupling while the test car is moving is a significant challenge. For the 0 degree, the equipment showed a strong back reflection from the bottom of the rail base. However, it is not the case for the other beam angles. As a result, coupling loss becomes intermittent without the operator realizing it.

Nonetheless, residual layers from wheel burns can shadow internal defects. Therefore, the backscattered energy pulse or waveforms require complex signal processing and experts to interpret them. The technique is also relatively poor at detecting surface or near-surface defects where most faults are located. Therefore, most NDE equipment includes electromagnetic probes to compensate for this deficiency. Additionally, multiple probe types are high in power consumption and require large equipment. Thus, there are not suitable for integrating revenue service vehicles [33].

2.2.2.2. Visual Inspection

Visual inspection is one of the traditional methods of the railroad industry. Typically, visual inspection is conducted by specialized inspectors who walk along the track and focus and address all expected track defects. Recently, visual inspection systems based on cameras have been used to measure the rail head profile, percentage of wear, rail gap, moving sleepers, absence of ballast, missing bolts, and surface defects. These systems first capture images and process images using image feature analysis to extract features that would identify and characterize fault type and its severity [33].

The advantage of visual inspection systems based on camera is that it provides more reliable and consistent results than human observation. However, these systems do not provide sufficient information for the existence of internal defects. Nonetheless, they require a large storage capacity to store and process the images. Also, image processing has self-learning and computationally complex algorithms to detect faults in the image frame [33].

2.2.2.3. Eddy Current Inspection

A typical eddy current inspection involves a sensing coil that induces current into the rail track to generate a magnetic field. This magnetic field produced an eddy current as it approached

the test surface area [31]. So, when the defects areas exist, the difference in the magnetic fields produced by the coil and the test surface results in variations in the impedance and electromotive force. Therefore, the information related to defects such as defect existence, size, and depth can be recorded through variation in the electric signals [34]. This approach works very well in practice, performing more reliably for near-surface defect detection [31]. However, they cannot analyze non-conductive materials [33]. Also, maintaining the eddy current inspection equipment is costly and time-consuming. Nonetheless, it requires specialized users of eddy currents to get accurate defects data using inspection programs [34].

2.2.2.4. Magnetic Inspection

Magnetic inspection uses a magnetic field to detect flaws in the rail. The current is put into the rail through direct contact or using an electromagnet. Any disturbance in the magnetic field indicates the existence of a potential defect that is detected by pickup coils. This technique is good at detecting near-surface or transverse defects. But defects in the vertical and horizontal plane often go unnoticed. Due to speed sensitivity, higher speed inspection is not recommended, which results in lower magnetic flux density in the rail head [30], [35].

2.2.2.5. Track Recording Vehicle (TRV)

Track recording vehicles comprise sensor signals, measuring systems, and data management systems to detect track parameters. However, the high cost and complex structure of TRV make it challenging to apply on daily inspection and maintenance operations which results in inconvenient and time-consuming operations [36].

2.2.2.6. Vehicle/Track Interaction (VTI) Systems

VTI uses vehicle responses as a measure of track geometry conditions. VTI systems are originated and evolved as US Federal Railroad Administration (FRA) research & development

and with the joint R&D collaboration between ENSCO, Inc., in the late 1990s. This system uses an array of sensors to accurately measure the dynamic response of a rail vehicle in interaction with the underlying track. The system consists of an onboard monitoring unit (computer), necessary electronics, five accelerometers, and an externally mounted dual-purpose antenna for GPS and cellular communication. The accelerometer measurement is recorded continuously, and an exception is created when a value exceeding a predetermined threshold defined by Federal Track Safety Standard (FTSS) is encountered [37]. However, tests demonstrate that the system could detect 84% of FTSS exception conditions using a neural network to determine optimum shock-level thresholds. Otherwise, it is hard to establish these thresholds by trial-and-error or analytically [33].

2.2.3. Academic and Experimental System in Railroad Condition Monitoring

Research found that the fault detection at variable speed using in-service vehicle has been an attractive and interesting topic in industry and academia [22]. Traversing the same track again and again gives an opportunity for obtaining the track geometry degradation information. The obtained information is then used to take necessary actions.

Tsunashima et al. [38] used the same approach for monitoring the condition of the track during regular operation hours. The study demonstrate that the in-service vehicle equipped with sensors and GPS receivers may serve as probes to detect and analyze real-time vehicle vibration. The Root Mean Square (RMS) values of the vertical acceleration is measured to access the longitudinal level irregularities.

Mori et al. [39] also developed a portable track conditioning monitoring system using in-service vehicles. The system consists of a noise meter to detect corrugation, accelerometers to detect track irregularity, a gyroscope, a GPS receiver for position information, a computer for

analysis, and an analog input terminal for inputting signals from each sensor to the computer. A field test shows that the track faults are detected using the proposed system. Similarly, Bocciolone et al. [40] detects the track corrugation from accelerometers mounted on the axle of a passing train.

We et al. [41] presented an on-board monitoring sensor network system for high-speed trains in China. The proposed system that comprises of sensor nodes, a hierarchal data transporting scheme, and an electronic tag-based addressing method for making maintenance decisions. In another study by Weston et al. [42], sensors are mounted on rail bogie to monitor lateral and vertical irregularities by double integration of acceleration signals and high pass filtering.

A high-speed track-irregularity measurement system is introduced and installed it into an HSR-350x locomotive and analyzed the test run result on a high-speed line [43]. The Korean high-speed train (HSR-350x) is composed of seven cars, two power cars, two motorized cars, and three trailer cars. The system measures longitudinal level, cross-level, gauge, alignment, and twist irregularity. A global positioning system (GPS) and an encoder attached to the axle also measured train speed. The result shows excellent performance of the measurement system.

Nonetheless, some railroad industry focuses on advanced visual inspection techniques include LIDAR and three dimensional (3D) cameras to find and fix the anomaly on the track. Such systems use automated algorithms to extract the features from high-resolution two-dimensional (2D) images or 3D light detection and LIDAR data. This type of system uses automated algorithms for the extraction of designated features from high-resolution 2D images or 3D light detection and ranging (LIDAR) data [44], [45], [46].

2.3. Conclusion

All the discussed techniques & studies in the literature review uncovered some disadvantages. The disadvantages include complex sensor system networks that provide dynamic behavior of vehicles with high cost, require more energy and low robustness [47], [48]. Thus, such systems are not suitable for freight trains [48]. Moreover, it may not be necessary to achieve inspection quality accuracy using in-service system that traverse over the same track frequently [42]. Therefore, the literature review reports the research gap that no one focuses on the error produced in finding and localizing potential anomaly due to non-uniform acceleration sensor sampling and GPS signals degrade from the loss of line-of-sight (LOS) to the satellites. Also, another gap is the approach combining multiple signal streams obtained from multiple traversals at varying speed, frequently and across the network to enhance the detection and localization of irregularity.

3. RAILS SETUP AND ARCHITECTURE

3.1. Objective

This chapter's primary objective is to introduce a low-cost track condition evaluation system that generates real-time information using collected data from the accelerometers and global positioning system receivers that are ubiquitous in smartphones. Another objective is to design a low-cost, robust, and low power consumption system that offers accurate condition estimation of tracks for both passenger and freight trains.

3.2. Contribution

This chapter's main contribution is to describe the system architecture that will autonomously identify track anomalies to focus follow-up inspections. A second contribution is to demonstrate the high-level architecture and system implementation. A third contribution is to design and draw UML diagrams that illustrates the quantifiable aspects of the system.

3.3. System Architecture

This section describes the overall design of a proposed system based on the Intelligent Transportation System (ITS) approach. The system consists of low-cost inertial sensors, remote servers, and a decision support system.

3.3.1. System Overview

Figure 3.1 shows the overall architecture of an intelligent, low-cost rail track condition evaluation system. The architecture is divided into three units: on-board measurements, a communications channel, and data processing with result diagnosis.

The on-board measurement unit utilizes inertial sensors embedded in the smartphone to sense and record multi-directional inertial responses retrofit in service vehicles or any other locomotive during ordinary commercial operation. The measured data are transferred over a

communication channel using a secure standard protocol and uploaded to the remote server. Figure 3.2 represents the inertial sensor data transmission from smart phone using a secure channel to the remote window server.

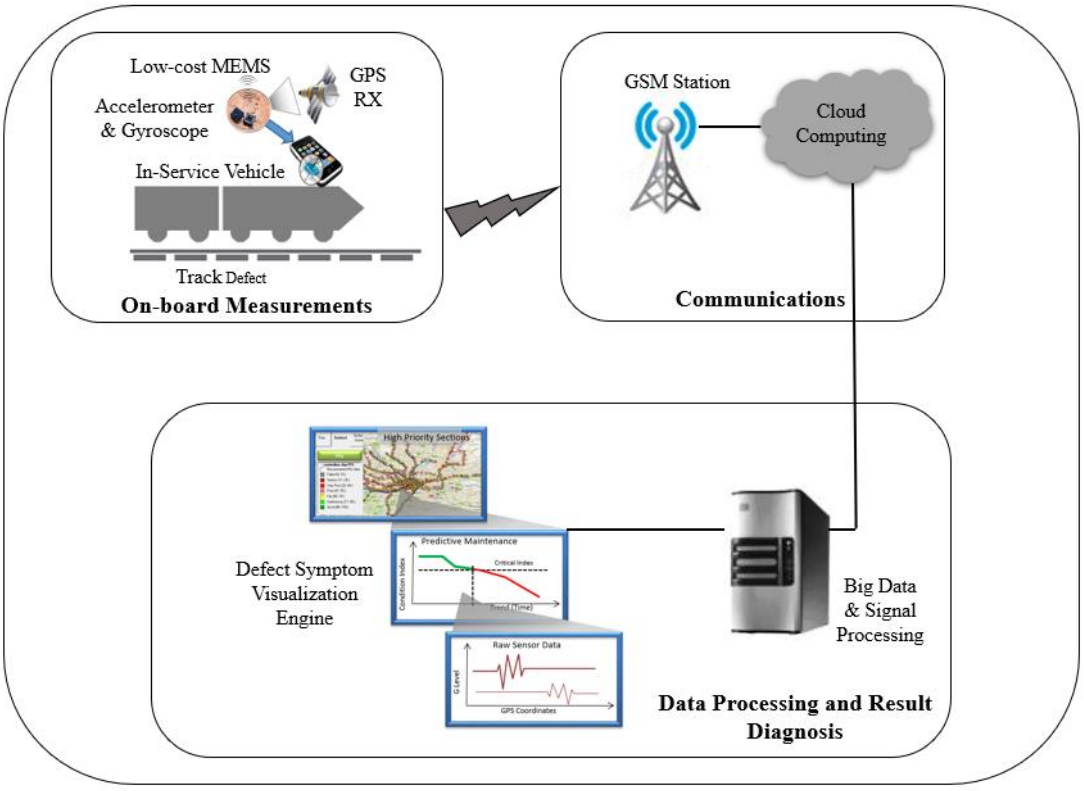


Figure.3.1. Architecture of the track condition evaluation system

The data processing and result diagnosis unit employs big data and signal processing algorithms that will compress three-dimensional linear acceleration, angular acceleration, and geospatial data to identify and localize the signature of track irregularity. The diagnostic results can provide reliable condition information that using color-coded TRQ values onto maps of the rail routes to the railroad inspector via the defect symptom visualization engine. Thereupon, with the enhanced situation understanding, railroad companies will be able to optimize inspection and maintenance practice to minimize track closure and slow order frequency. Therefore, it will

minimize the cost and safety risks while maintaining reliable track and equipment condition information.

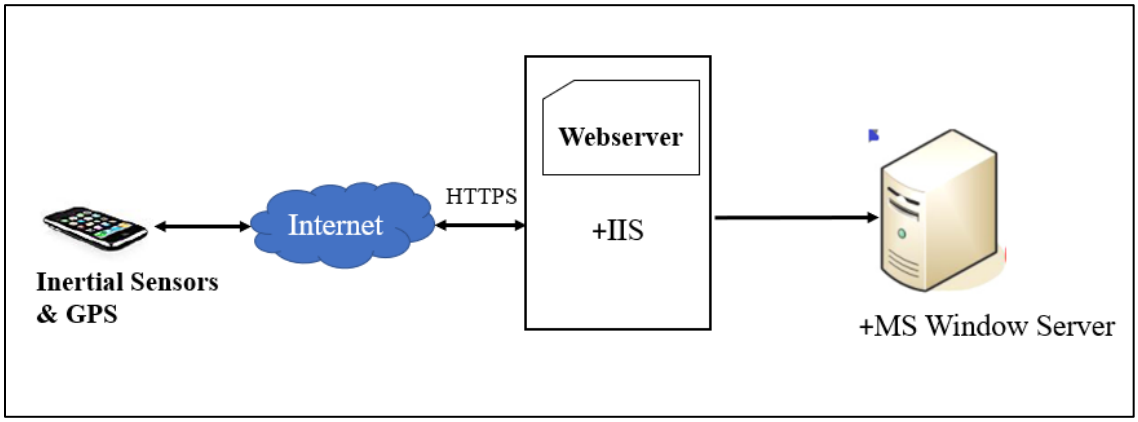


Figure 3.2. Inertial data communication diagram

3.3.1.1. High-Level System Architecture

Figure 3.3 Briefly illustrate the high-level view of the track condition evaluation system architecture.

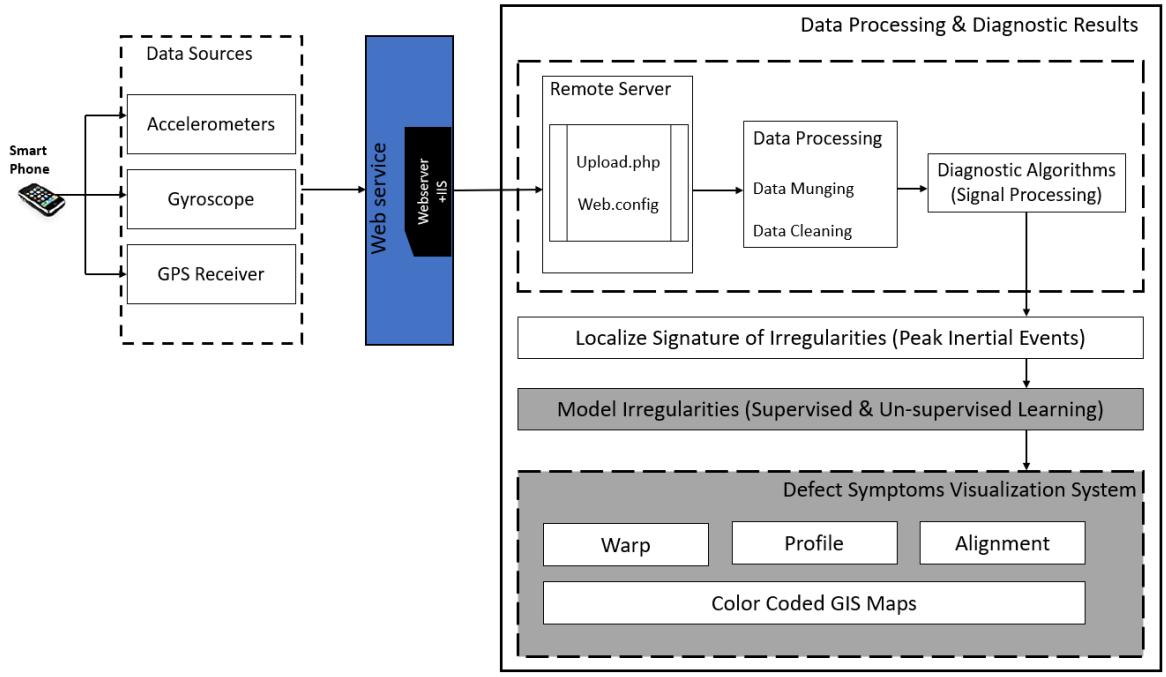


Figure 3.3. High-level view of the system architecture

The core of the high-level system architecture is the smart phones that can be maneuvered as a provenance of data. The smartphones have powerful motion sensors such as accelerometers, gyroscopes, and magnetometer, and GPS receiver. With all these integrated sensing technologies, we can utilize smartphones as a data source. Web services is selected as a method of communication between the data source and data processing & diagnostic results, due to their ability to support interoperable machine to machine interaction and to manage the securely transfer of the data [49]. Data storage is implemented using NTFS, known as New Technology File System, to store and retrieve files on a hard drive that attached to the remote server effectively as shown in figure 3.4. NTFS is strong and fault-tolerance, high performance logging file system. It offers data security and reliability and supports large volume and file size. The data is stored in their original format as .CSV files. Each file is about 4600 KB each minute. The data processing algorithm is then implemented on the raw data to processed and transformed into new datasets. Thereupon, the signal processing algorithms, called diagnostic algorithms are employed to extract the features that fulfill the requirements of analytical models.

Figure 3.4. Hard drive attached to the server

3.4. Software Architecture: System Modeling Using UML

The focus of this section is to design and describe an extensible software architecture using the Unified Modeling Language (UML). The architecture leverages and provides flexibility to accommodate other potential condition monitoring applications. The Unified Modeling Language (UML) is a well-known, recognized, influential, and foremost diagrammatic modeling language. It offers a variety of diagrams that express and represent the different views of a system. For example, UML provides a static view of the system through the class diagram and component diagram, and a dynamic view of the system through the sequence diagram [50].

3.4.1. Component Diagram for RAILS System

The UML component diagram of the RAILS is shown in figure 3.5. The Android mobile device package consists of a RIVET application [51], Android I/O, and an HTTP client. An HTTP client connects to the web server that communicates and transfers data to the Windows

server. The web-sever carries out the request of the HTTP client of the Android app. The data handling component consists of data processing and signal processing applications. The model irregularity uses the input values given by the data handling applications to fuel the analytical model (machine learning). The Defect Symptoms Visualization Systems (DSVS) component consists of a DSV Engine and a DSVS user interface (UI). The compound results have classified irregularities with annotated GIS map that can be viewed using DSVS UI.

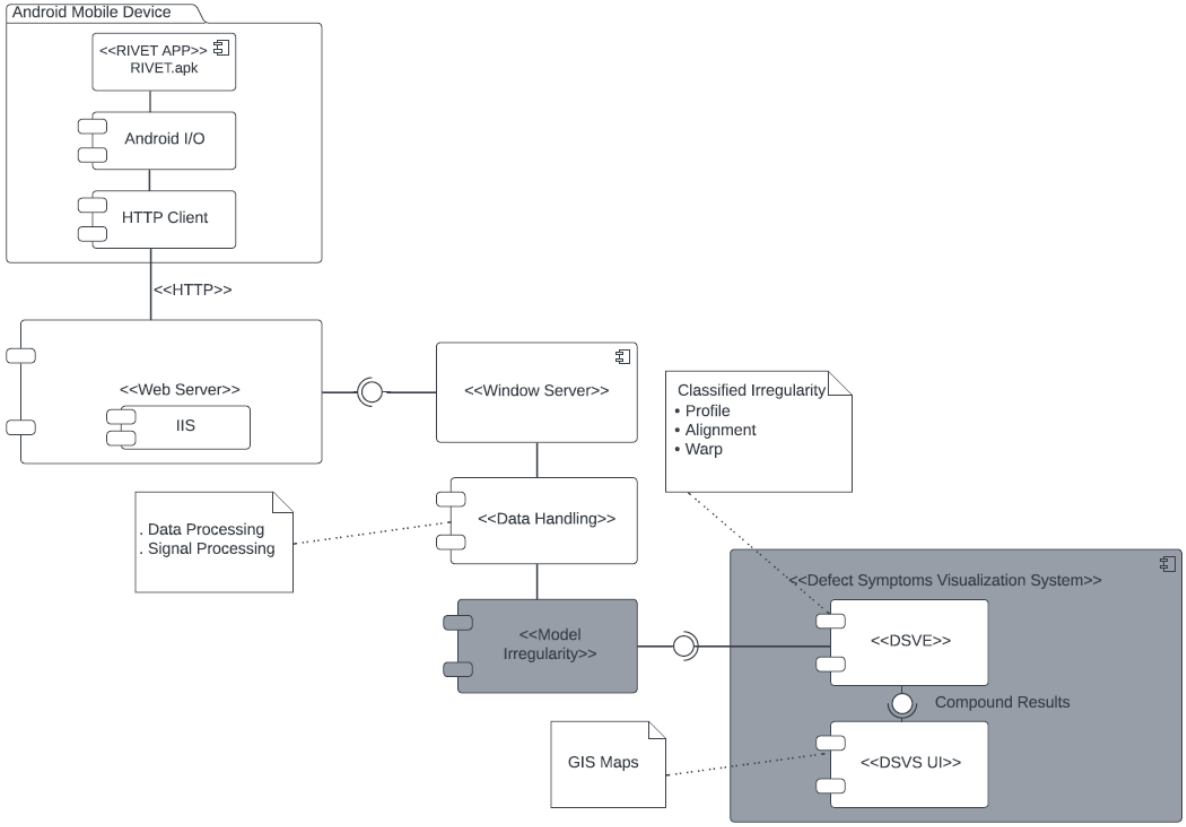


Figure 3.5. UML component diagram of RAILS

3.4.2. Class Diagram for RAILS System

The elements of the UML class diagram are limited to the data processing, signal processing, and analysis shown in figure 3.6. UML class diagram shows the methods and properties of classes and their relationships. Aggregation, composition, association, dependency,

and inheritance are examples of relationships. A generic data source class is defined in figure 3.6, from which motion sensors and GPS receiver data source take their interfaces. The data source class has six attributes: Queue:RIVETDataQueue, SampleRateArray, CalibrationConstant, Accelerometer, Gyroscope, and GPS. It has five operations: DataSource, getData, clearData, uploadData, and setCalibrarion.

"RIVETDataQueue" stores an n-ary signal data on the remote server (i.e., Window server). RIVETDataQueue class has two attributes: DataQueue and Time. It also has six operations: RIVETDataQueue, getData, getAccelerationData, getGyroscopeData, getSpeed, and getLocation.

RemoteServer and DataProcess classes are associated with each other. The DataProcessStrategy class describes the generic interface for the different data processing strategies such as data munging and data cleaning. Subclasses such as Sort, RemoveBadGPS, UnitConversion, and GenerateRIF are generalized by DataMunging class. The DataCleaning class uses DataMunging to define Threshold and ColorCodes subclasses.

The SignalProcessStrategy class describes the generic interface for the signal processing strategies to extract the signal features and localize the irregular rail track. The SignalProcessStrategy class has two operations: SignalProcessStrategy and getProcessedData. Subclasses such as SignalAlignment, FilterDesign, and FeatureExtarction are generalized by SignalProcessStrategy class. LocalizeIrregularity subclass includes peak inertial events (PIEs) that is the high magnitude of RIFs.

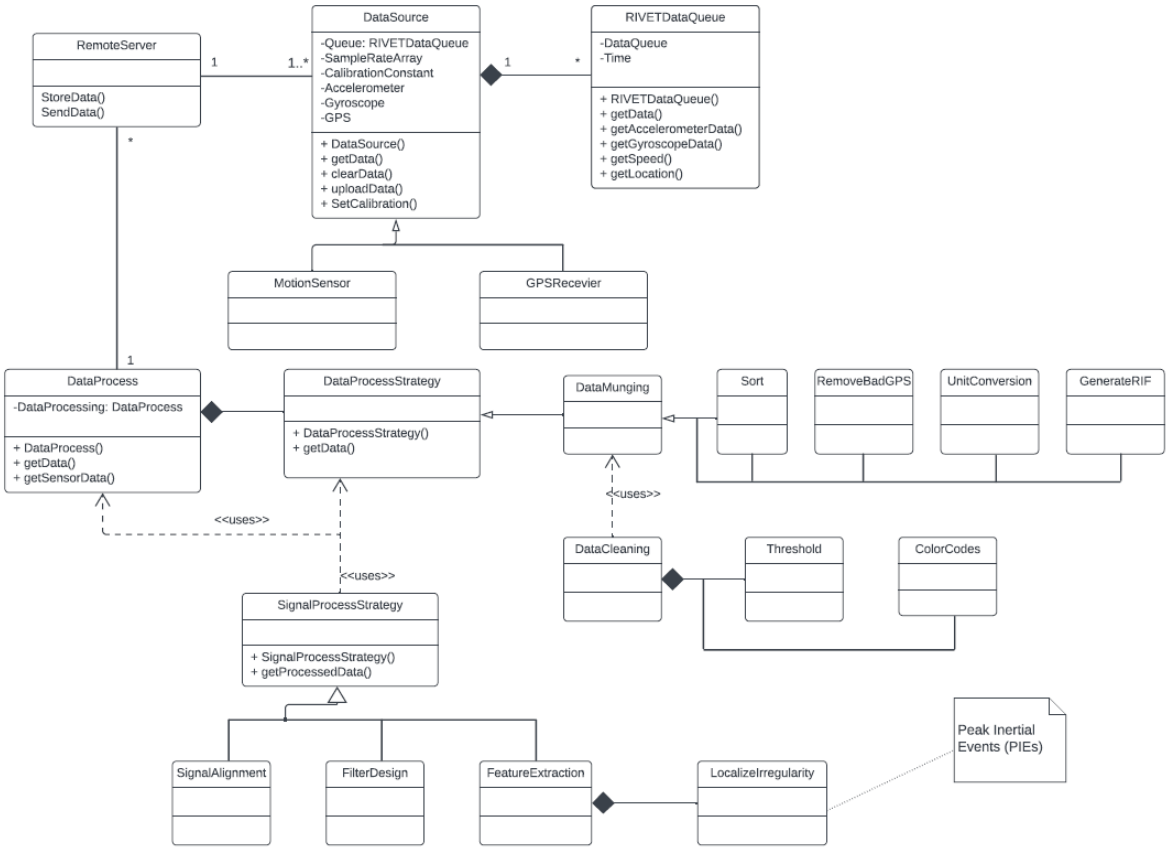


Figure 3.6. Principal components of the design for a data processing & signal processing of RAILS system

3.4.3. Sequence Diagram for RAILS System

The UML sequence diagram of the RAILS system is shown in figure 3.7. The main objects are the RIVET app, Windows server, data processing, localization irregularity estimation, model irregularity and defect symptoms visualization system (DSVS). The process starts with the “RIVET app module” generating and logging data. The logged data is uploaded to the server. Alternatively, “Window server module” sends an acknowledgment to the “RIVET app module.” Once the data stream is available to the server, the “data processing module” reads the data from the corresponding data store (Windows server). The data is processed by the data processing strategies that produce cleaned data. The “localization irregularity estimation module” uses the cleaned data for signal processing activities; it identifies isolated features. Subsequently, the

“model irregularity module” read and uses these identified features as input for supervised/unsupervised learning activities.

Consequently, it classified the features such as profile, alignment and warp. Once the features are classified, the “model irregularity module” passes the optimized irregularity with annotated GIS maps to the “DSVS module.” The DSVS mirrors this action.

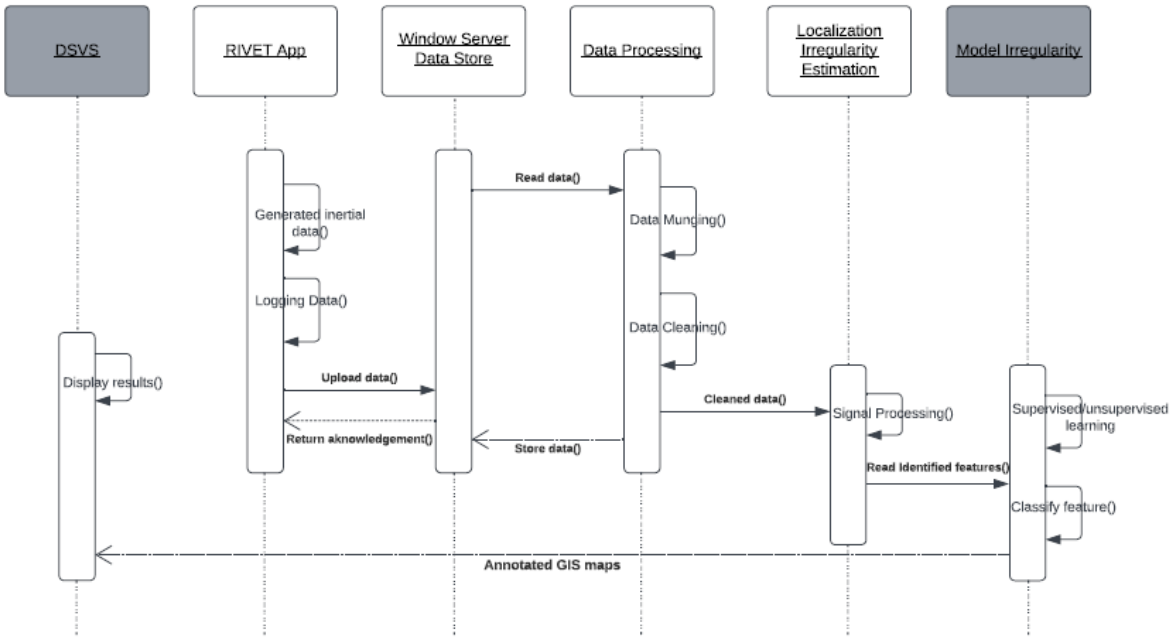


Figure 3.7. UML sequence diagram for RAILS

3.4.4. Software/Hardware

This research uses software: SAS 9.4 software version to build and design data processing algorithms. The minimum system requirements are as follows:

Operating system: Windows 10 (64-bit version)

- Hardware: Intel or Intel-compatible Pentium 4 class processor or above
- Memory: 4GB RAM or more.

MATLAB R2018a for signal processing. The minimum requirements are as follows:

- Operating system: Windows 10 (64-bit version)

- Hardware: Any Intel or AMD x86-64 processor
- Memory: 4GB RAM or more.

External Hard Drive: G-RAID 20TB

3.5. System Implementation

This study primarily starts with smart phone application (app) called “Railroad Infrastructure and Vehicle Evaluation Technology” RIVET that is capable of autonomously collecting and uploading data from rolling stocks such as hi-rails, locomotives, flat cars, freight cars, and other vehicle where power is available.

3.5.1. Data Acquisition and Collection App

The RIVET app was developed as a part of Mountain Plains Consortium (MPC) project at the North Dakota State University, Fargo, North Dakota, USA[30]. The app currently works only as a data collection tool. Figure 3.8 displays the RIVET app screen. The app utilizes all embedded sensors in the smartphone and any of its available wireless connection to communicate with a server. The universal resource locator (URL) for the server is entered in the app setup screen. After collecting the data, the app starts uploading logged files automatically onto a server. If in case of any pending logged file, researcher can tap “stop logging” which uploads any queued file to the server. The app logged inertial and geospatial position data from multiple traversals of railroads.

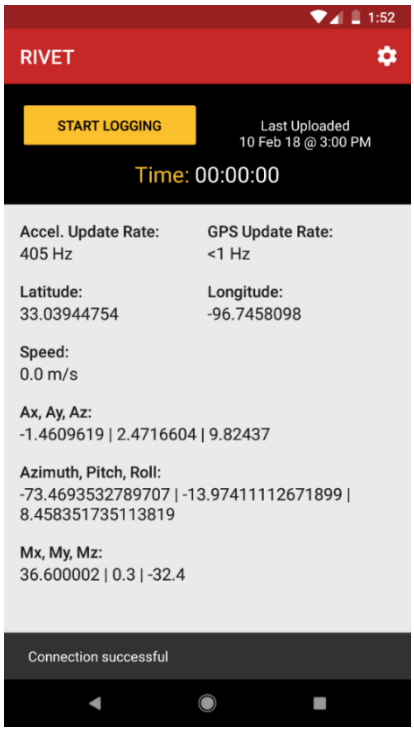


Figure 3.8. RIVET app screen shot

3.5.2. Data Collection Instrumentation

This study uses Harsco Rail ‘s LD 1515 HY-RAIL® shows in figure 3.9 is a light duty track and utility vehicle for railway applications requiring travel on the highway and on the rail. Road/rail units are equipped for track inspection, crew and material transportation, and specialized track maintenance assignments.

Figure 3.9. Harsco rail's light duty 1515 HY-RAIL

For the data collection purpose, three google pixel phones are used. Careful consideration was given to the smartphone location setup, including different kinetic energy responses, power supply resources, and the GPS signal strength. Figure 3.10 shows all three android phones location setup in a light duty 1515 Hy-Rail. First phone is mounted on the dashboard of the vehicle and the other two phones are fixed securely under the driver and passenger seat. Therefore, phones recorded the intensity of the roughness that seated driver and passenger may have experienced. All three phones can access the power supply resource and collect the data from multiple traversals over the same segments irrespective of the direction of travel. The primary area of interest in this study are the locations where geometry issues such as warp, profile and alignment are known and exist.

Figure 3.10. a) Phone installed under the passenger seat b) phone installed on the dashboard c) phone installed under the driver seat

3.5.3. Smart Phones Sensors and Specifications

Modern smartphones have several kinds of sensors. However, there are three main sensors that embedded in the smart devices for the motion detection: accelerometer, gyroscope, and magnetometer. Google pixel uses Bosch Sensortec BMI160 IMU chip (IMU = inertial measurement unit) utilizes the fascinating MEMS technology (Micro-electromechanical systems) that integrates a triaxial accelerometer and a triaxial gyroscope [52]. The smartphones have built-in GPS receiver that trilaterates the device's position using data from at least three GPS satellites and the receiver. The accelerometer of the smartphone was set to sample at a rate of 400 Hz and the GPS receiver updates location in approximately 1 Hz.

The *accelerometer* detects changes in the device displacement (change in velocity), orientation, and tilt around three axes by measuring acceleration forces or inertial force. Such accelerations are popularly measured in terms of g-force. On the other hand, the *gyroscope* measures orientation of the device, based on the principal of angular momentum. Basically, a *gyroscope* is a spinning wheel or disk in which the axle is free to assume any orientation and returns three-directional angular velocity [53], [54]. Figure 3.11 shows the orientation of the mobile device around three axes.

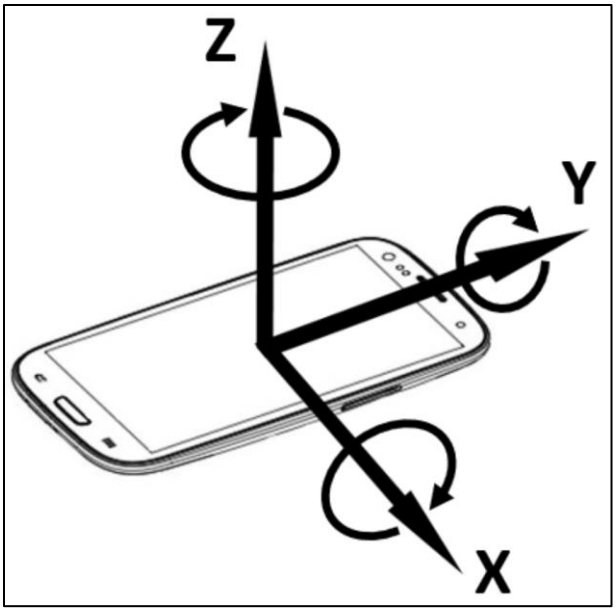


Figure 3.11. Orientation of the smartphone device
Image Source: (Nguyen, Wang, Li, Luo, and Watkins, 2019 [55])

A *magnetometer* measures the strength of the magnetic field around the phone from which the phone can obtain its absolute direction related to the earth's geomagnetic field. The *magnetometer* also returns the three-dimensional values.

4. RAILS DATA PROCESSING

4.1. Objective

The first objective of this chapter is to remove likely erroneous signal features so that the remaining data enable best estimates of the positions and magnitudes of railroad track irregularities. Another objective develops a method to visualize the estimated geospatial positions and relative magnitudes of railroad track irregularities. Also, to eliminate the minimum number of outliers such that the remaining distribution fits a classical distribution that is typical of measurements from physical characteristics in nature.

4.2. Contribution

The contribution is to introduce a method of cleaning large-scale rail data to visualize the condition of the railroad tracks spatially.

4.3. Data Processing

The amount of data and data sources is rapidly growing and expanding and becoming more diverse in nature and structure. So, it is essential to organize the available data for analysis. However, data processing and preparation takes a lot of time in any project or research [56]. Data is often merged or fused from different sources, and the quality of data needs to be checked and ensured [56], [57].

Bad data quality adversely affects the decision-making process and negatively impacts the results. Also, an explorative analysis will be carried out to identify the abnormalities in the data. For this purpose, data munging can be performed, also referred to as data wrangling. It integrates heterogeneous data sources, finds dirty or unclean data and extreme outliers, and transforms and cleans it. The data munging prepares and transform raw data into a more readily used format that can be used for further analysis and visualization. The exact methods differ

from project to project depending on the data you are leveraging and the goal you are trying to achieve [58]. The below subsections explain the raw data and methods of data munging for this research.

4.3.1. Data Munging

The inertial and geospatial position data collected from sensors embedded in the smartphones aboard a hi-rail vehicle. A portion of the collected data and its format shows in table 4.1. The first row contains a header with labels for each column of data collected from inertial and GPS sensors. Each data row updates with each sample of the accelerometer data. The DateTime variable is in epoch time format for milliseconds. The GPS data is latitude (Lat) and longitude (Lon) in decimal degree format. The ground speed (Speed) is in units of m s-1. The Ax, Ay, and Az, signals are the linear accelerations in m s-2 along lateral, longitudinal, and vertical directions, respectively. The Rx, Ry, and Rz signals are the angular rotations (angular velocity) in degrees-per-second around the X, Y, and Z axis, respectively. The integrated gyroscope produces the “Azimuth”, “Pitch”, and “Roll” for the sensor orientation angles in degree. Azimuth is angle around the z-axis in degrees; Pitch is angle around the x-axis in degrees; Roll is angle around the y axis in degrees. The Mx, My, and Mz are the geomagnetic field strength along the X, Y, and Z axis in micro-Tesla.

Table 4.1. Data sample and format

DateTime	Lat	Lon	Speed	Ax	Ay	Az	Azimuth	Pitch	Roll	Rx	Ry	Rz
1.55E+12	48.17365	-96.2332	10.47	-1.06	-0.73	10.04	-45.9	4.79	8.38	-0.01	0.02	-6.15E-04
1.55E+12	48.17365	-96.2332	10.47	-1.06	-0.73	10.04	-45.9	4.79	8.38	-0.01	0.02	-6.15E-04
1.55E+12	48.17365	-96.2332	10.47	-1.06	-0.73	10.04	-45.9	4.79	8.38	-0.01	0.02	-6.15E-04
1.55E+12	48.17365	-96.2332	10.47	-1.06	-0.73	10.04	-45.9	4.79	8.38	-0.01	0.02	-6.15E-04
1.55E+12	48.17365	-96.2332	10.47	-1.06	-0.73	10.04	-45.9	4.79	8.38	-0.01	0.02	-6.15E-04
1.55E+12	48.17365	-96.2332	10.47	-0.85	0.04	9.561	-45.9	4.79	8.38	-0.01	0.02	-6.15E-04

Once the raw data is accessible, the data processing algorithm fuse all the transmitted data and sort it via timestamps. The algorithm then removes all the bad GPS data (zero values) and converts linear acceleration Ax, Ay, and Az from ms-2 to g-force units by dividing each value by 9.80665 m s-2. After combining, sorting, and cleaning bad GPS data, the algorithm generates the road impact factor (RIFs) using equations 1 and 2. The data munging flow shows the sequence in brief in figure 4.1.

The Road Impact Factor (RIF) "a measure of intensity that is proportional to resultant acceleration rate along the x- and z-axis such that

$$RIF_{\mathrm{Gi}} = \sqrt{\frac{1}{L}\sum_{n=0}^{N-1}|Gi_n\, v_n|^2\, \delta t} \quad (1)$$

where RIF_{Gi} is the average g-force magnitude experienced per unit of distance L traveled. The sample period is δt and v_n the instantaneous traversal speed at sample n for N total samples produced by a speed sensor. Gi is a resultant acceleration signal sample in units of g-force calculated as

$$Gi = \sqrt{(G_{\mathrm{x}})^2 + (G_{\mathrm{z}})^2} \quad (2)$$

where Gx and Gz are acceleration signals about lateral and vertical directions, respectively. Appendix A shows the detail procedure of RIF calculation. When a locomotive or rolling stock crossing broken rail or ties on the track, it produces a vertical acceleration. Lateral acceleration always cornering or diagonally to the direction of travel of a vehicle shown in figure 4.2 as a reference. Hence, both indicate the existence of track irregularities.

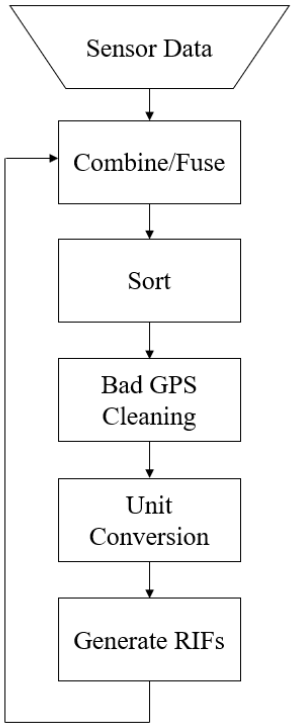


Figure 4.1. Data munging flow chart

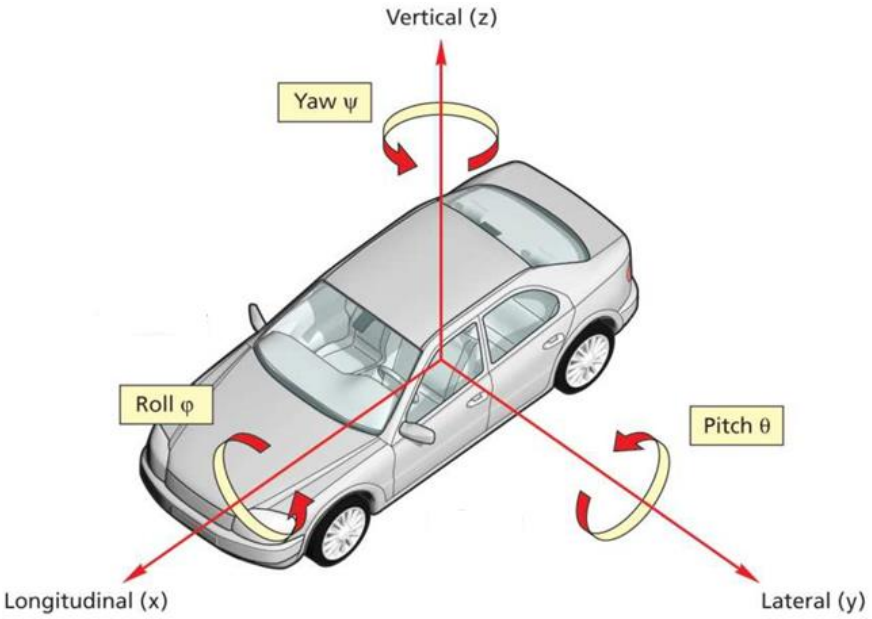


Figure 4.2. Three-axis accelerometer orientation
Image Source: European new car assessment program [58]

4.3.2. Data Cleaning – Removing Noise

After generating RIFs, the approach next clean large-scale data by removing noise such as invalid values, skewed values, and outliers to visualize the estimated geospatial positions and relative magnitudes of railroad track irregularities. It is significantly important to remove outliers as it adversely affects the accuracy of the algorithms. The approach first removes the invalid values and outliers in the dataset and check the distribution of the data. Based on the distribution, the method applied some heuristic tactics and use normality tests to ensure that data follow the normal distribution. The below subsection shows each applied steps and the flow diagram (figure 4.3) that explains the process more precisely.

4.3.2.1. Cleaning Steps

- Eliminate all RIF indices of zero, because g-force values of zero do not produce any inertial events.
- Remove RIF indices above 1.0, as it represents airborne values i.e., flying off your seat [59]. Moreover, accelerometer produced very large g-force values that resulting in very high outlier RIF values in case of some disturbance such as manual handing of the devices, network transmission issues, spurious powering issues, or Android OS issues (perhaps downloading updates). Therefore, based on these observations, algorithm remove outliers RIF values above 1.0.
- Start trimming either of the tail (higher or lower tails) based on the use of some heuristics in knowing how much to trim in which tail depending upon the shape of the resulting histogram.
- Do normality tests with chi-squared goodness-of-fit test with binning variations until all tests fail to reject the hypothesis that the data came from a normal distribution.

- If at least one of those tests does not reject the null for at least one distribution type, then stop. Go to step 3
- If the algorithm never stops and still rejecting the null, then randomly generate the normally distributed data with the same mean and standard deviation as a RIF data.
- Repeat all the normality tests with chi-squared goodness-of-fit test and seeing if the tests fail to reject.

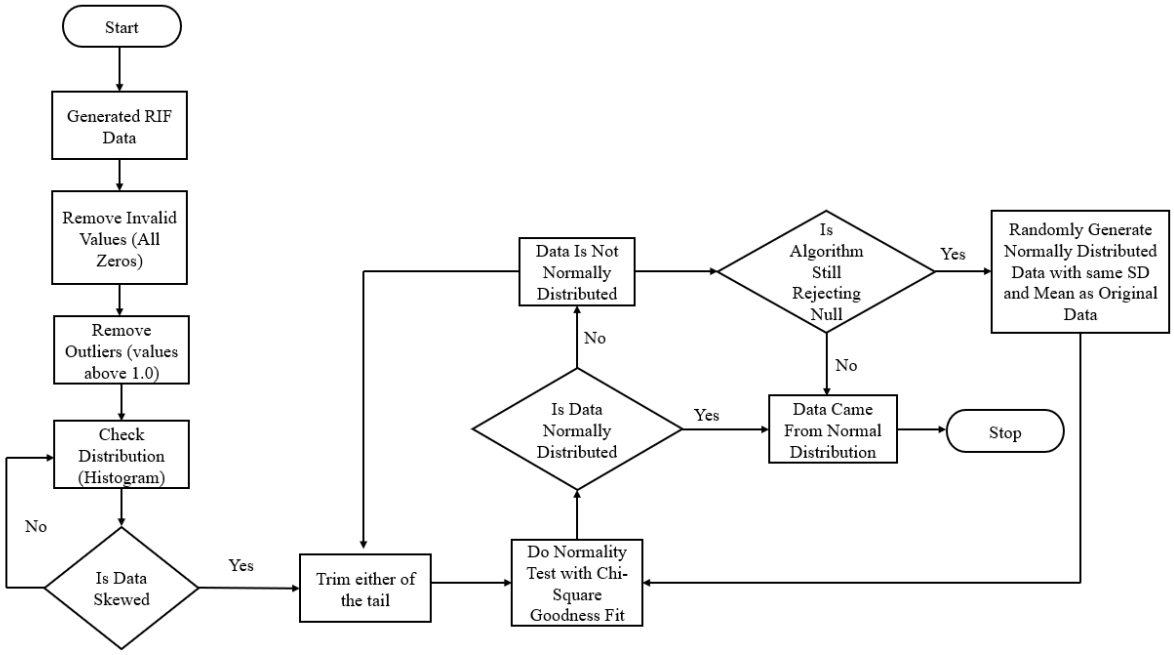


Figure 4.3. Data cleaning process flow chart

4.3.3. Results

RIF Data values (RIF-Gi) [0.05-0.451] (figure 4.4) shows the best fitting curve to data with truncated left tail. Except chi-squared (p = 93.48%), all the normality tests reject the hypothesis.

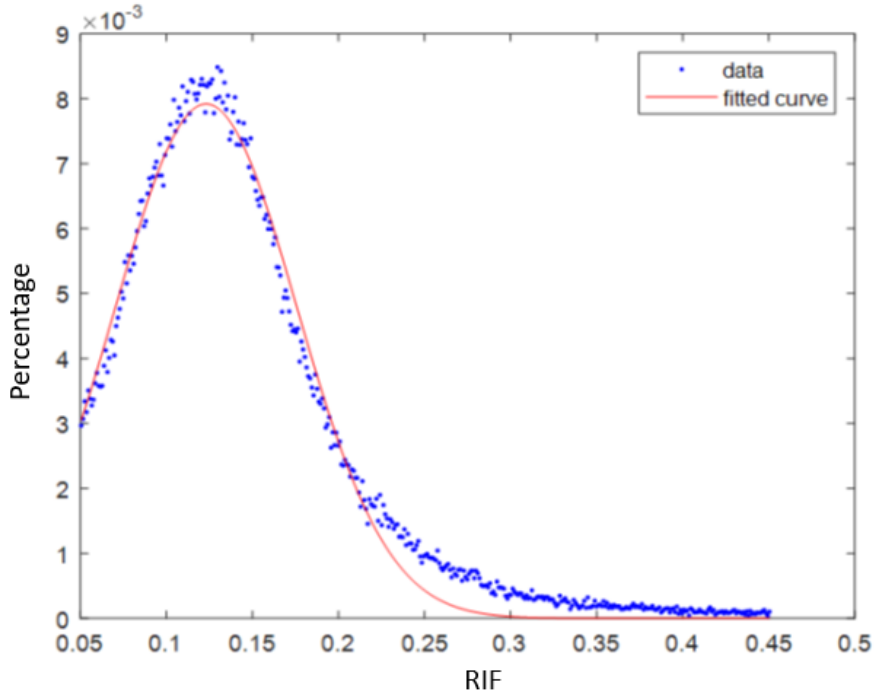


Figure 4.4. Best fitting curve to RIF data values with a truncated left tail

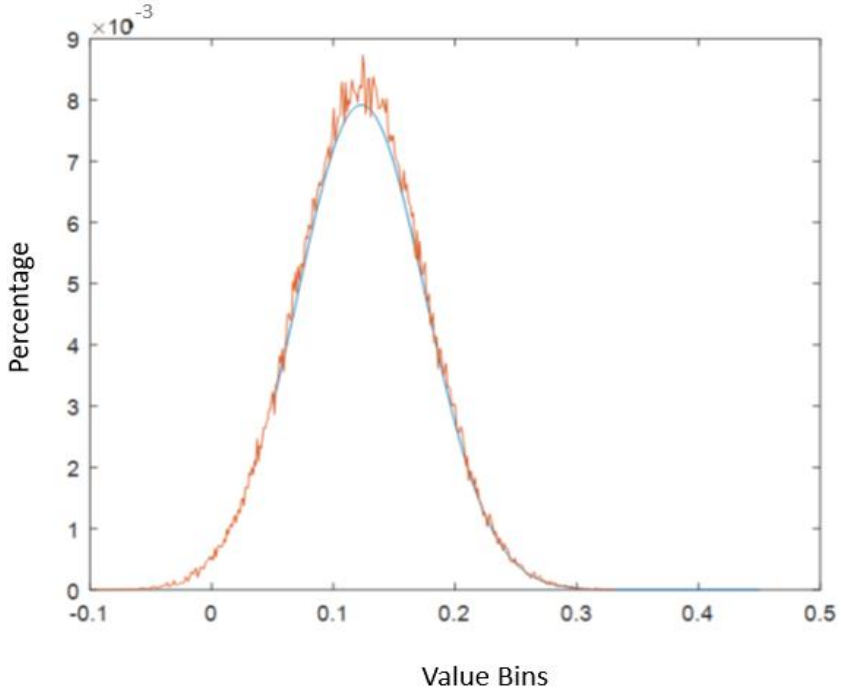


Figure 4.5. RIF data fit curve (Blue) relative to normal random data (Red)

However, randomly generated data with same mean and STD has a longer left tail that include negative values shown in figure 4.5. Test of normality (Gaussian) does not reject the null

if we extrapolate the left tail to negative values. However, our RIF data has positive values. Thus, we conclude that RIF data follows a normal distribution (Gaussian fit) with truncated left tail. Chi-square with Log-Normal fit fails to reject the hypothesis due to similarity. However, normal with truncated left tail fits better than Log-Normal (figure 4.6). In case of Weibull distribution fit, data values [0.02-0.43], chi-square also fails to reject the hypothesis. Log-Normal doesn't seem to be as good a fit as the Weibull. Weibull (figure 4.7) requires the measured values to be positive, which it is in our case.

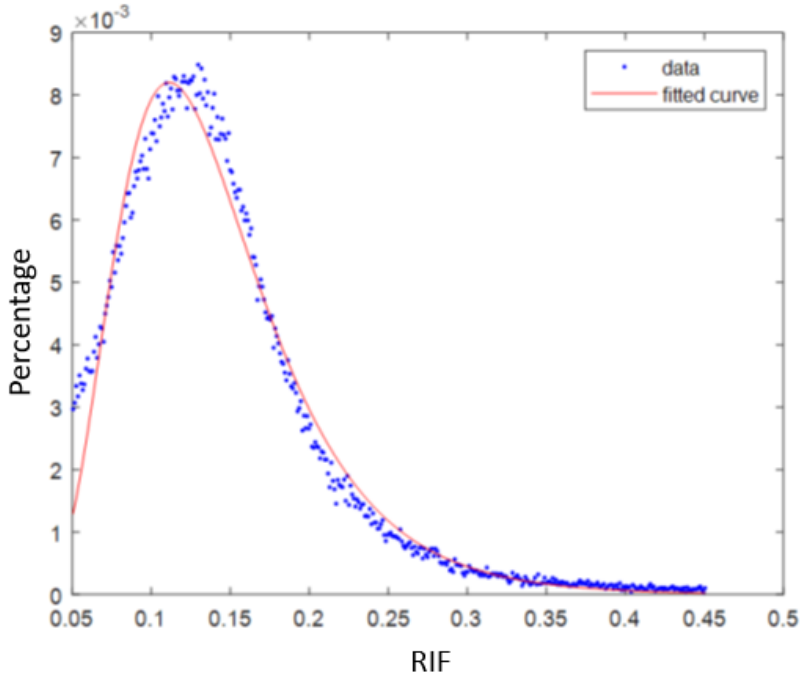


Figure 4.6. Log-normal distribution fit curve to RIF data values with truncated left tail

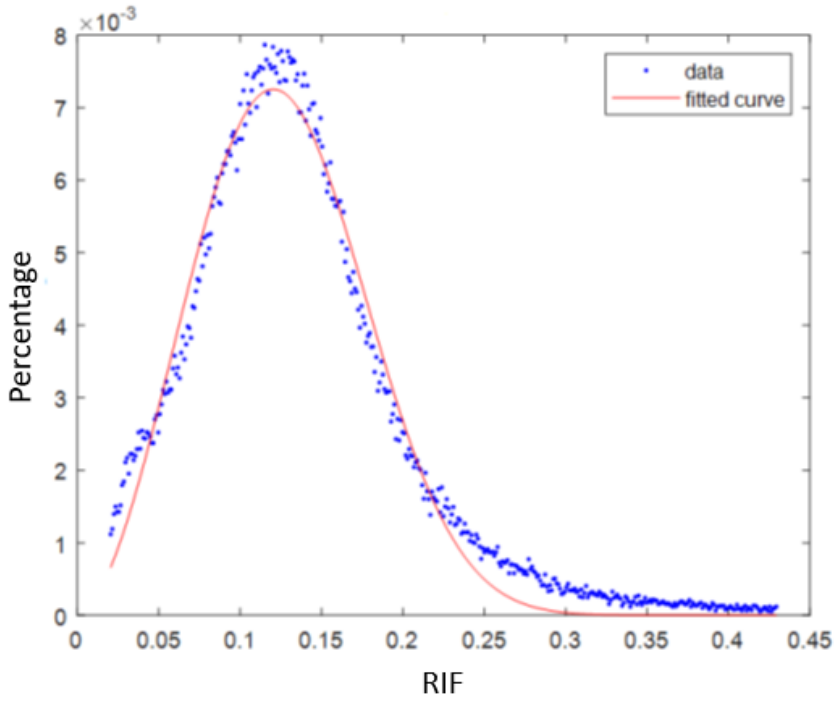


Figure 4.7. Weibull distribution fit to RIF data values

Hence, the normality tests with chi-squared goodness-of-fit test with binning variations shows that the data came from a normal distribution. Finally, it is concluded that RIF feature values follows a normal distribution shown in figure 4.8. Therefore, depending on the inferred distribution, a percentile threshold and color-codes are defined to visualize the peak inertial events (PIEs) that is high magnitude RIF values as represents in table 4.2. Peak inertial events are also defined as the peak accelerometer signals in the unit of g-force values produce by isolated roughness or irregularity.

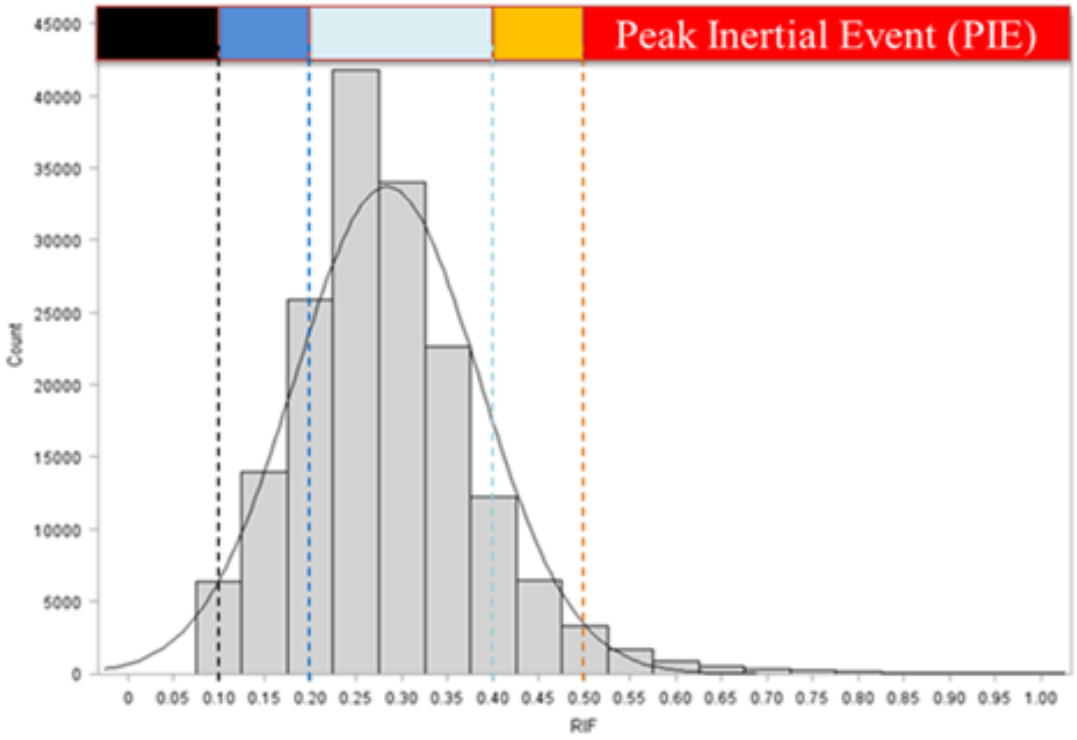


Figure 4.8. RIF data distribution along with define thresholds & color-codes to visualize the PIEs

Table 4.2. Color-coded scheme to visualize the PIEs

Color	Percentile
Black	Lower 2.5-percentile
Blue	Within the lower 1 to 2 standard deviations from the mean
Light Blue	Within 1 standard deviation of the mean (68%)
Orange	Within the upper 1 to 2 standard deviations from the mean
Red	Upper 2.5-percentile

5. IRREGULARITY POSITION LOCALIZATION AND DETECTION ENHANCEMENT[1]

5.1. Objective

The objective of this chapter is developing a procedure that provide robust detection and enhanced accuracy in the localization of irregularities. Another objective is to demonstrate the effectiveness of the proposed approach by employing data used in the literature [60].

5.2. Contribution

The main contribution of this chapter are the signal processing methods that improve the low-resolution and low accuracy of GPS receivers, and the non-uniform sample rate of the accelerometer signals. Another contribution is an algorithm that can enhance the accuracy of identifying and locating track irregularities by combining the data from multiple traversals of the same or different trains, at various speeds. The method also enhances the signal-to-noise ratio (SNR), which reduces both false positive and false negative feature detection errors.

5.3. Procedure

Sensors, when placed on a locomotive or a rail vehicle, do not always perform perfectly. With the advent of MEMS technology, the inertial sensors such as accelerometer and gyroscope have become smaller, use less power, have better performance, and are lower cost. Moreover, the Global Navigation Satellite System (GNSS) is a generic term for a satellite-based positioning

[1]The material in this chapter was co-authored by Bhavana Bhardwaj and Dr. Raj Bridgelall. Bhavana Bhardwaj had primary responsibility for data processing and method evaluation. Bhavana Bhardwaj was also contributed to the development of the conclusion that is advanced here. Bhavana Bhardwaj also drafted and revised all versions of this chapter. Dr. Raj Bridgelall and Dr. Kendall Nygard served as a proofreader and checked the math in the statistical analysis conducted by Bhavana Bhardwaj.

system called GPS (Global Positioning System). It is also assumed that attaching the antenna to the roof of a vehicle and connect it to a GNSS module, will provide accurate position information. However, this is often not the case. Conversely, for monitoring the development of the geometry faults over time, accurate location information is critical. Thus, systems such as unattended geometry measurement system (UGMS) and automated track geometry monitoring system requires accurate localization.

Due to non-sufficient quality and low accuracy of GPS receivers and the non-uniform sample rate of the inertial accelerometer sensors, affects the signal position alignment and adds noise to the signal. As a result, it is hard to detect and localize the track irregularities. In other word, it reduces the signal-to-noise ratio that can increase the false positive and false negative. Therefore, the proposed procedure could improve the SNR and reduces the significant challenges in detecting and locating the irregularities. The proposed procedure includes distance interpolation technique called *Localization* that interpolates the distance between the GPS positions to increase the position resolution and *heuristic* alignment. The primary goal of the alignment algorithm is to improve both accuracy and precision of the traversal alignment. That means, the algorithm aligns the position of inertial signals collected from multiple traversals so that the extracted features among traversals can be ensemble average (EA) to increase the SNR. Therefore, ensemble averaging is not only enhancing the signal quality but also improving the information and reduce noise in the signals. Nonetheless, both precision and accuracy increase with the addition of each traversal during ensemble averaging the aligned signals.

Moreover, the procedure designs an appropriate digital signal filter that filter the noise from the signal, necessary to maximize the signal-to-noise ratio (SNR) of each signal prior to feature extraction. The following subsections represents the steps of the proposed procedure.

5.3.1. Localization- Distance Interpolation

High distance resolution can be attained by interpolation using the instantaneous speed and sample periods obtained from the sensors aboard regular vehicle. Fundamentally, the approach is to position aligns the inertial signals by replacing the GPS position tags with higher resolution distance tags and to identify the starting position along the traversal path. Therefore, the inertial sample at the starting position y_0 of a traversal is set to distance zero, and subsequent samples are positioned at the accumulated distance

$$y_n = y_{n-1} + v_n \times \Delta t_n \quad (3)$$

where n is the sample number, v_n is the instantaneous speed for n samples, and Δt_n is the associated sample period.

This study uses simulated data shown in table 5.1 from railroad track traversals by driving a paved road segment (Cell 40) of the Minnesota road (MnROAD) research facility in the United States. A smartphone application (app) named pavement analysis via vehicle electronic telemetry (PAVVET) logged the inertial and geospatial data from a sedan.

Table 5.1. Stimulated data and its format

Time	Gz	Lat	Lon	GSpeed	Pitch	Roll	Yaw	Gx	Gy	RotX	RotY	RotZ
44.142	-1.057	45.263	-93.71	9.586	6.693	4.886	-0.319	-0.088	-0.152	3.177	0.629	-0.452
46.768	-1.216	45.263	-93.71	9.586	6.693	4.886	-0.319	0.047	-0.241	3.177	0.629	-0.452
50.26	-1.087	45.263	-93.71	9.586	6.693	4.886	-0.319	0.026	-0.272	3.177	0.629	-0.452
62.927	-0.854	45.263	-93.71	9.586	6.741	4.903	-0.329	-0.002	-0.212	1.246	-0.13	-0.332
73.909	-0.912	45.263	-93.71	9.586	6.752	4.907	-0.332	0.022	-0.161	1.865	0.214	-0.258
86.754	-0.942	45.263	-93.71	9.586	6.776	4.908	-0.341	0.038	-0.144	2.005	-0.67	-0.189

The data consists of 53 traversals at nearly uniform speeds. The authors verified that the inertial signal patterns observed from traversing isolated road bumps are similar to those observed when a train traverses track irregularity. Traversing the bump produced a consistently large peak inertial events (PIE) that is used to evaluate the SNR as a function of the number of

traversals combined. In table 5.1, the time variable is in milliseconds from the start of the data collection. The GPS data is latitude (Lat) and longitude (Lon) in decimal format. The ground speed (GSpeed) is in units of m s−1. The g-force values of Gx, Gy, and Gz are accelerations in the lateral, longitudinal, and vertical directions, respectively. The integrated gyroscope provides the pitch, roll, and yaw angles of the sensor orientation in degrees. The RotX, RotY, and RotZ signals are the angular rotations in degrees-per-second around the X, Y, and Z axis, respectively. The accelerometer sample at approximately 90 hertz and GPS updates at approximately 1 hertz, the latitude and longitude remained unchanged for blocks of inertial samples, or GPS blocks. Figure 5.1 illustrates how the positions of the GPS update, and the accelerometer updates distribute nonuniformly both within and across traversal datasets [48].

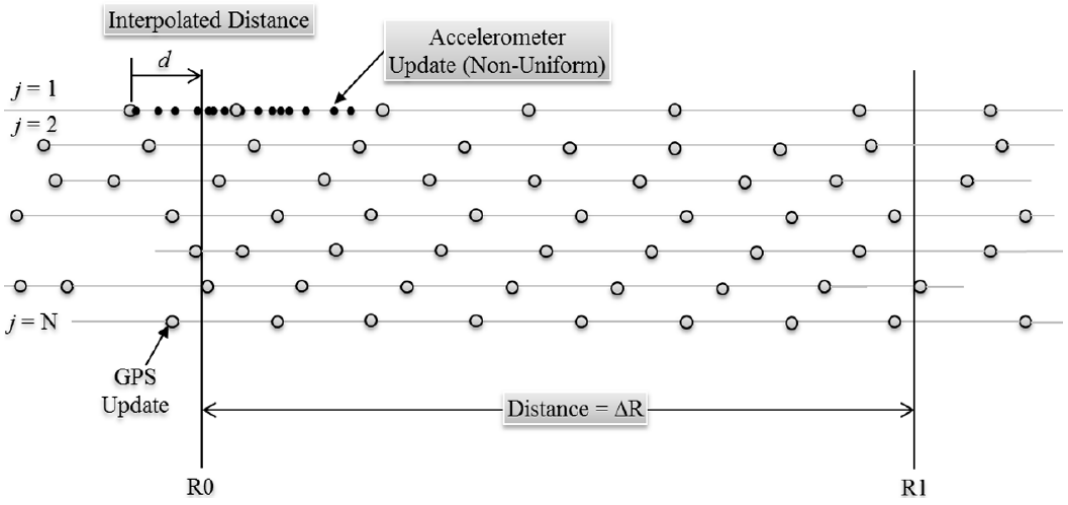


Figure 5.1. GPS position variation and non-uniformly distributed accelerometer sample

The geospatial processing technique contracted a starting and ending geofences to indicate the interested segment of road or railroad. The data then extracted from all traversals contained within the geofences R0 and R1. These geofences are perpendicular to the traversal path and jointly served as a distance reference.

The reference distance d0 = 0 is the first sample of the GPS block that is closest to the left of R0. In Sedan dataset, the GPS updated continuously after every second. That means, there is no missing GPS updates in this dataset. Consequently, the variation in the relative distance of the peak inertial event (PIE) indicates the misalignment. After interpolation, the procedure extracts approximately equal length traversals. The relative position of the two PIEs in figure 5 shows the position misalignment of the two signals. It also shows two extracted traversals starting from the interpolated zero distance position [48]. As shown in the figure 5.2, the positive peak is the local maxima and negative peak is the local minima of the vertical acceleration profile from traversing the bump.

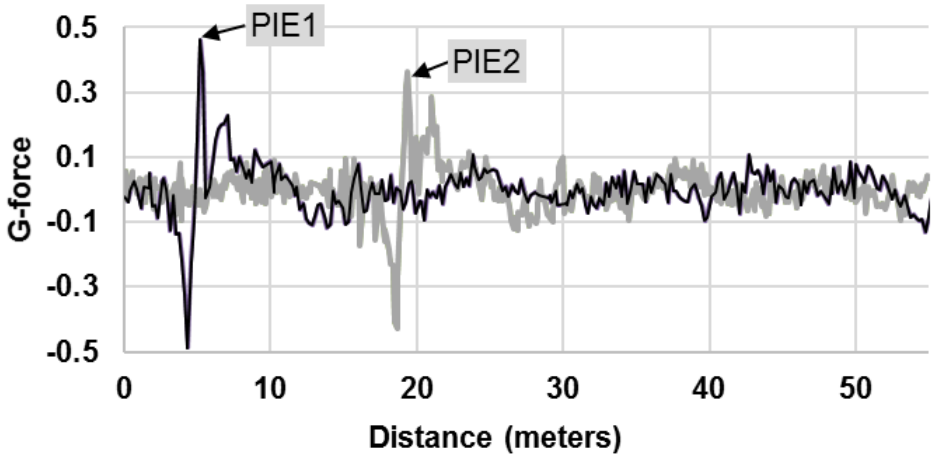


Figure 5.2. Misalignment of the PIE position between the two signals

Most importantly, the distance interpolation benefits ensemble averaging in computing signal means along the small and fixed window sizes along the traversal, regardless of vehicle speed and sampling rate. Now, the next step discussed about the signal alignment algorithms using the same dataset.

5.3.2. Signal Position Alignment

This section utilizes the *heuristic* technique to align the misaligned signals from multiple traversals. The signals are aligned when the peaks of the signals are at the same distance

position. The authors [48] developed some heuristic methods include *Dual Geofence, Midway Midpoint, Centroid Asymmetric, and Centroid Symmetric* to position align the signals. It is better to mention that next subsection "Design Digital Signal Filter", uses the same data set that is the outcome of the best performance alignment method. Therefore, for this research, we only explain the best performed alignment method and the statical method that evaluate the performance of all these heuristic methods. After distance interpolation, the alignment methods identify the first and last inertial samples of each traversal to produce maximally aligned signals.

5.3.2.1. Heuristic Alignment Methods

These alignment algorithms used extracted dataset obtained from the interpolated traversals. Before introducing these methods, this study developed a reference method that remove GPS position estimation error to calculate the residual or remaining error as global methodology quality reference in the article [60]. The reference method is attained by equalizing the position of the known isolated PIE for each traversal.

The *reference method* is known to be able to align miss-positioned data almost perfectly because the method relies on a known ground truth anomaly and its actual true location in the data. However, in practical application, the railroad agencies don't have a known anomaly location to start with. The anomalies need to be determined from the misaligned data. Therefore, this reference method labels the isolated PIEs on all traversals as position zero. Here, the known ground truth generates the isolated PIE for each traversal shown in figure 5.2. The residual errors are from estimates of the speed and sample intervals.

The *Centroid Asymmetric* method first calculates the centroid at the center of the maximum overlapping segments and draws a center line between the geofences R0 and R1 that is perpendicular to the traversal path. The center line is then spatially joined to the traversal

layers to obtain a distance. The centroid calculates from the geospatial positions of that set of coordinates as

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}, \bar{Y} = \frac{\sum_{i=1}^{n} y_i}{n} \quad (4)$$

where xi and yi are the coordinates for GPS block i and n is the total number of GPS blocks.

Subsequently, the algorithm identifies the immediate left GPS blocks closest to the center line for all traversals and then interpolate the distance from the first sample of that GPS block. All traversals are truncated based on the recorded shortest distance. The first sample of all truncated traversals is marked as the new starting zero position. Starting from the new zero position, the distance is interpolated to the new ending position.

5.3.2.2. Statistical Performance Evaluation

The performance of the alignment methods is tested using the empirical distribution test that include Kolmogorov –Smirnov (KS), Anderson-Darling (AD) and Cramer-von Mises tests (CVM) [61]. The PIE distributions resulting from each alignment method is the bases of the performance evaluation. All three tests use an empirical distribution function (EDF) that defined for a set of n independent observations X1…….,Xn that have a common distribution function F(x). Under the null hypothesis, F(x) is the normal distribution. The tests reject the null hypothesis if the p-value of the test-statistic is less than 0.05. The empirical distribution function (EDF), Fn(x), takes a step of height 1/n at each observation such that

$$F_n(x) = \begin{cases} 0, & x < X_{(1)} \\ \frac{i}{n}, & X_i \le x \le X_{(i+1)}, i = 1, \ldots, n-1 \\ 1, & X_{(n)} \le x \end{cases} \quad (5)$$

At any value x, Fn(x) is the proportion of observations less than or equal to x, while F(x) is the probability of an observation less than or equal to x. EDF statistics measure the

discrepancy between Fn(x) and F(x). The EDF tests use the probability integral transformation U = F(X) such that if F(X) is the distribution function of X, then the random variable U is uniformly distributed between 0 and 1, where U(i) = F[X(i)] given n observations X(1), ..., X(n) are inputs for the EDF test statistics.

KS Test:

The Kolmogorov-Smirnov statistics (D) is

$$D = sup_x |F_n(x) - F(x)| \tag{6}$$

The Kolmogorov-Smirnov statistics belong to the supremum class of EDF statistics, which is based on the largest vertical difference between F(x) and Fn(x). The Kolmogorov-Smirnov statistic is the maximum of D+ and D-, where D+ is the largest vertical distance between the EDF and the distribution function when the EDF is greater than the distribution function, and D- is the largest vertical distance when the EDF is less than the distribution function. That is,

$$\begin{aligned} D^+ &= max_i \left(\frac{i}{n} - U_{(i)}\right) \\ D^- &= max_i \left(U_{(i)} - \frac{i-1}{n}\right) \\ D &= max(D^+, D^-) \end{aligned} \tag{7}$$

AD Test:

The Anderson-Darling statistic and the Cramer-von Mises statistic belong to the *quadratic* class of EDF statistics, which is based on the squared difference $(F_n(x) - F(x))^2$. Quadratic statistics have the following general form:

$$Q = n \int_{-\infty}^{+\infty} \left(F_n(x) - F(x)\right)^2 \varphi(x) dF(x) \tag{8}$$

The function φ(x) weights the squared difference $[F_n(x) - F(x)]^2$. The Anderson-Darling statistics (A^2) is

$$A^2=n\int_{-\infty}^{+\infty}[F_n(x)-F(x)]^2[F(x)(1-F(x))]^{-1}dF(x) \quad (9)$$

Here the weight function is $\varphi(x) = [F(x)(1-F(x))]^{-1}$. The discrete form for computing the Anderson-Darling statistic is

$$A^2=-n-\frac{1}{n}\sum_{i=1}^{n}\left[(2i-1)\log U_{(i)}+(2n+1-2i)\log\left(\left\{1-U_{(i)}\right\}\right)\right] \quad (10)$$

CVM Test:

The Cramer-von Mises statistic (W^2) is

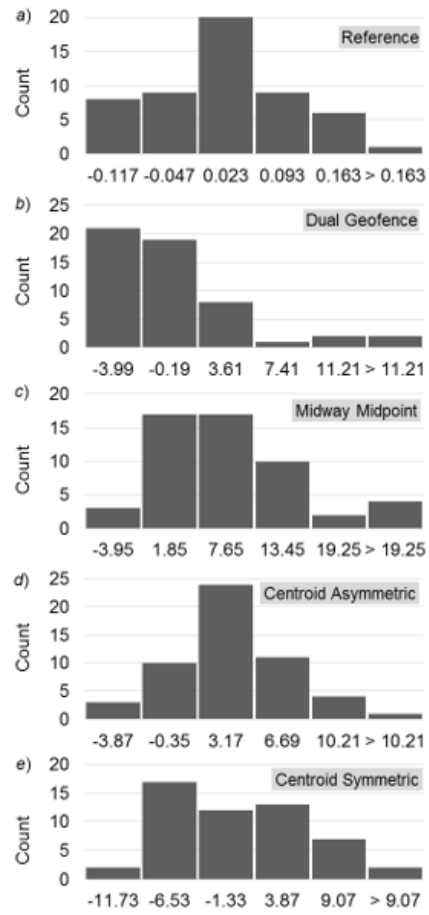
$$W^2 = n\int_{-\infty}^{+\infty}[F_n(x)-F(x)]^2dF(x). \quad (11)$$

Here the weight function is $\varphi(x) = 1$. The discrete form for computing the Cramer-von Mises statistic is

$$W^2 = \sum_{i=1}^{n}\left(U_{(i)}-\frac{2i-1}{2n}\right)^2+\frac{1}{12n}. \quad (12)$$

5.3.2.3. Findings and Conclusion

1.



2.

Method	PIE STD	Skew	P-Value KS	P-Value CVM	P-Value AD	Length STD
Reference	0.09	0.159	>0.150	>0.250	>0.250	0.09
Centroid Asymmetric	4.04	0.335	>0.150	0.128	0.149	0.13
Dual Geofence	4.84	1.546	<0.010	<0.005	<0.005	5.59
Centroid Symmetric	6.32	0.522	0.101	0.099	0.080	0.24
Midway Midpoint	7.24	0.690	0.125	0.072	0.045	0.20

Figure 5.3. 1) PIEs distance distribution of reference method and four heuristic alignment methods. 2) statistical summary of all methods

A normal distribution is a symmetric distribution with no skew that means skewness is zero and standard deviation is 1. Fundamentally, skewness represents the degree and direction of the asymmetry of the distribution about its mean. Conversely, if the distribution has long left tail, it means the distribution is skewed to the left and has a negative skewness. The statistics that use to compare for the PIE distributions and the length of the aligned traversals are mean, STD, and skewness. Figure 5.3 1) shows the distribution of the reference and all four alignment methods. It is found that the large offset is present at the beginning of the aligned traversals. And removing

such offsets will improve the resolution of the distribution for statistical testing. Therefore, the distribution of the PIE distances relative to the PIE distance of the reference traversal provides a measure of the alignment spread and removed the offsets. Interestingly, the traversals with the most samples for approximately equal distances selected as reference traversal for reference PIE. This is because having more samples achieves the highest resolution in distance estimates.

The reference method overtakes all other method in figure 5.3 1) and 5.3 2). So, the performance ranking of all four heuristic methods is decided through the statistics closest to the reference method. Here, centroid asymmetric is performed closely to the reference method. The centroid asymmetric and reference methods are the methods in which PIE distribution follows the normal distribution pattern. That means, in both the methods, none of the normality tests could reject the hypothesis. Thus, the study selected centroid asymmetric as a signal alignment method that generate aligned signals with the lowest STD of their lengths. Moreover, this alignment step prepares the signal for the further process of filtering that explain in the next subsection.

5.3.3. Signal Filter Design

Appropriate signal filtering, alignment, and combination from multiple traversals can enhance the signal-to-noise ratio. However, it is not straightforward in determining the best cut-off frequency for the filter. This section introduces a method that is suitable for any signal filtering approach. A method informs the cutoff frequency selection for a FIR low-pass filter that maximizes the SNR, with low computational complexity and high stability for practical use. The method achieves this by combining the FFTs from multiple signal runs of the same traversal to produce an ensemble average FFT (EA-FFT). Figure 5.4 illustrates the application of EA-FFT.

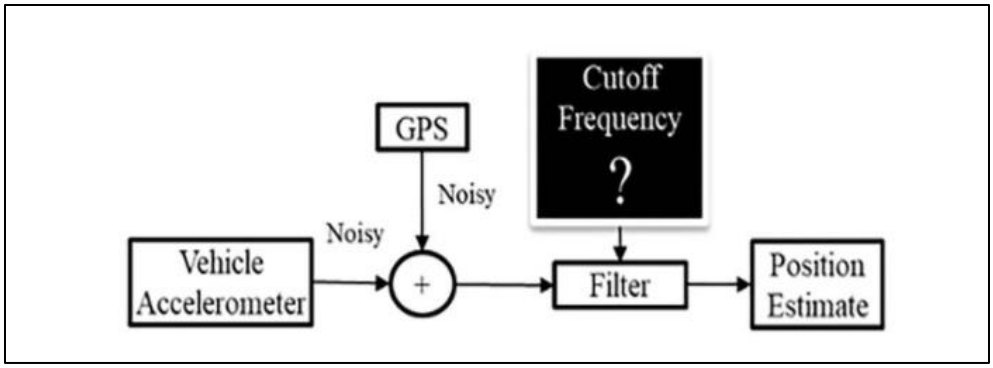


Figure 5.4. Application of EA-FFT technique

Additionally, this technique benefits to overcome some significant limitations in the literature. First, the absence of prior knowledge of frequency response parameters, it is hard to design standard digital filters due to their changing nature of the filter's requirements [62]. Therefore, they are not suitable for the general application [63]. Second, some authors recommended adaptive filtering approach. However, adaptive filtering also has some shortcomings, which the previous studies did not highlight. The adaptive process requires the use of a more complex cost function and bank of filters to evaluate the SNR in a closed-loop manner. Hence, it has high computational complexity and numerically unstable. Nonetheless, it works on single signal stream or run and does not guarantee convergence to a local minimum that would maximize the SNR, which limits their use in practice.

5.3.3.1. Method

The individual FFTs of the inertial signals from each traversal produce patterns in the spectral domain that do not clearly distinguish between information and noise. The EA-FFT enhances the clarity of the patterns by canceling randomness and boosting the correlated portions of the patterns. Subsequently, the sharpened pattern revealed a clearer breakpoint between information and noise. Figure 5.5 briefly explains the individual section of the overall approach.

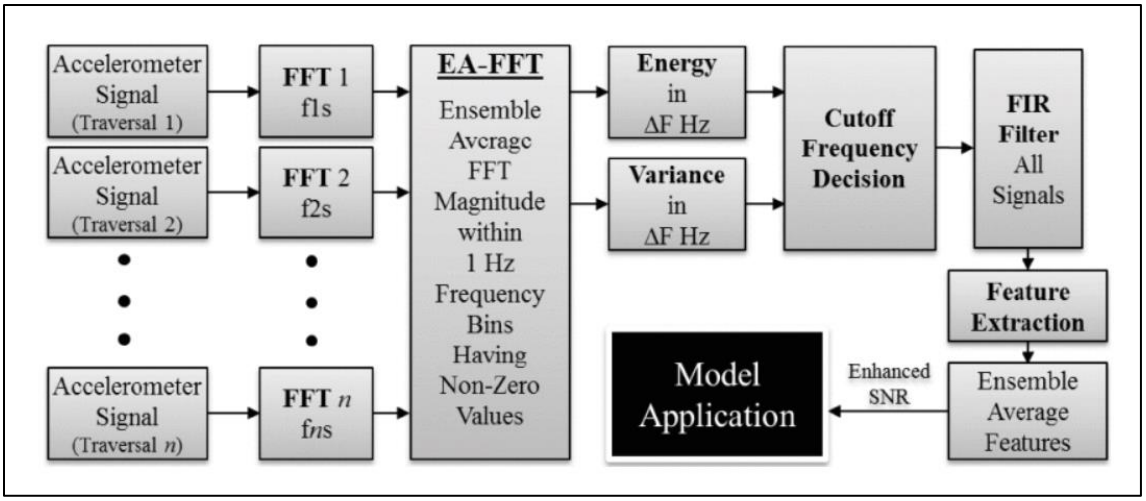


Figure 5.5. Method of determining the signal filter cut-off frequency in preparation for feature extraction and their application

Before calculating the FFTs and EA-FFTs of the inertial signals, the signals are extracted from approximately equal length heuristically aligned traversals. The remaining procedure are to 1) compute the EA-FFT from the individual FFTs of the inertial signals, 2) compute energy and variance windows along the frequency range, 3) determine the appropriate cut-off frequency, and apply a low-pass digital filter to each inertial signal shown in workflow figure 5.6. Each FFT used the mean sampling frequency (Fs) of its individual signal. However, the mean sample frequency varies among traversals. Therefore, the EA-FFT algorithm aligned the frequency bins before averaging the non-zero magnitudes. A frequency bin size of 1 Hz provided sufficient resolution to compute 5 Hz energy and variance windows for cut-off frequency selection. In particular, the observed transition of energy and variance from high to consistently lower values pointed to the appropriate cut-off frequency for a low-pass noise filter.

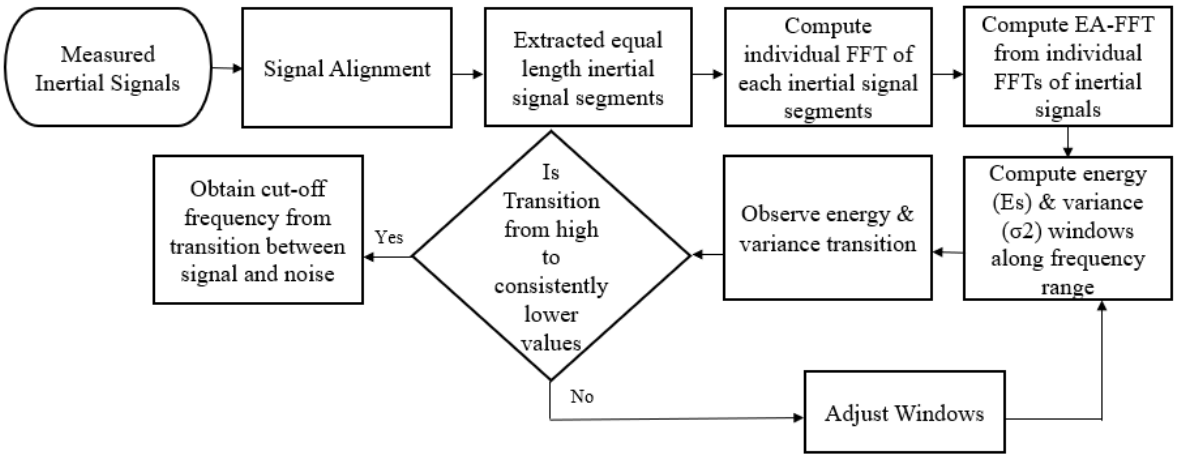


Figure 5.6. Determination of digital filter cut-off frequency workflow

- Fast Fourier Transform (FFT)

  The Fast Fourier transform (FFT) is an efficient method of computing the discrete Fourier transform (DFT) X_j of x_l as follows [64]:

$$X_j = \sum_{l=0}^{N-1} x_l W^{-jl} \qquad W = exp\frac{2\pi\sqrt{-1}}{N}, \tag{13}$$

$$j = 0, \ldots, N\text{-}1$$

where each x_l is an equally spaced sample of a function $x(t)$, l represents the time index, and j represents the frequency index. N is an integral power of 2.

- Ensemble Average FFT (EA-FFT)

  The EA-FFT is the mean value of all spectral magnitudes within a frequency bin *x* such that

$$X_{Nx} = \frac{1}{N}\sum_{i=1}^{N} X_{ix} \tag{14}$$

where X_{Nx} is the ensemble average of *N* FFTs across bin *x*, and X_{ix} is the magnitude of the FFT of an individual signal within frequency bin x. Figure 5.7 shows the EA-FFT of the signals extracted from the approximately equal length segments after position alignment. Fig. 4 also illustrates the difference between the FFT of two individual signals, and the EA-FFT.)

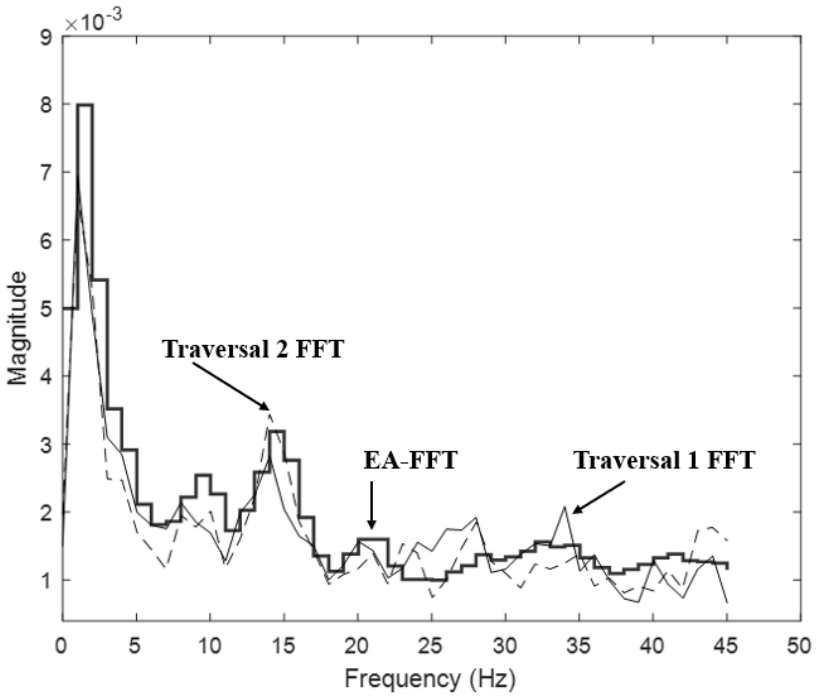


Figure 5.7. EA-FFTs of the signals extracted from equal length traversals

- Signal Energy

  The energy [65] of a finite signal x(n) defines as

$$E_s = \sum_{n=0}^{N-1} |x(n)|^2 \tag{15}$$

where Es is the energy of a signal, n is the sample number, and *N* is the total number of samples.

- Signal Variance

  The variance of a signal is defined as

$$\sigma^2 = \frac{1}{N-1} \sum_{i=0}^{N-1} [x_i - \mu]^2 \tag{16}$$

where the signal samples are x_i, μ is the signal mean, *N* is the number of samples and σ^2 is the signal variance [65]. The variance of the signal signifies how much a signal varies about the average value. Figure 5.8 a, and 5.8 b shows the calculated energy and variance of the EAFFT along the resolution window 5 Hz.

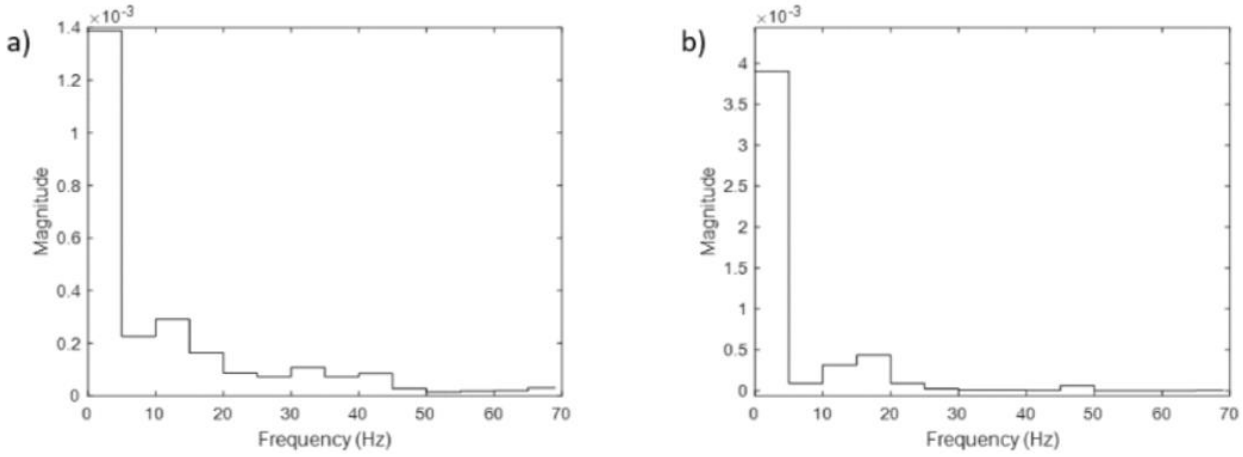


Figure 5.8. a) Energy windows of the EA-FFT, and b) variance windows of the EA-FFT along 5 Hz window

- Finite Impulse Response (FIR) Low-Pass Filtering

  A low-pass FIR filter of order N [66] is defined as:

$$y[n] = b_0 x[n] + b_1 x[n-1] + \cdots + b_N x[n-N] \tag{17}$$

Or

$$y[n] = \sum_{i=0}^{N} b_i x[n-i] \tag{18}$$

where $x[n]$ is the input signal, $y[n]$ is the output signal, and the filter order symbolizes as N. Each coefficient b_i is the value of the impulse response at the i[th] instant for $0 \leq i \leq N$. The impulse response of the filter is:

$$h[n] = \sum_{i=0}^{N} b_i \delta[n-i] = \begin{cases} b_n & 0 \leq \text{n} \leq \text{N} \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

The phase response of the FIR filter causes a delay in the filtered signal. Hence, the algorithm estimates and corrects the distance offset accordingly as the filter will delay all signals similarly. Figure 5.9 compares the unfiltered (black) and filtered (gray) signals.

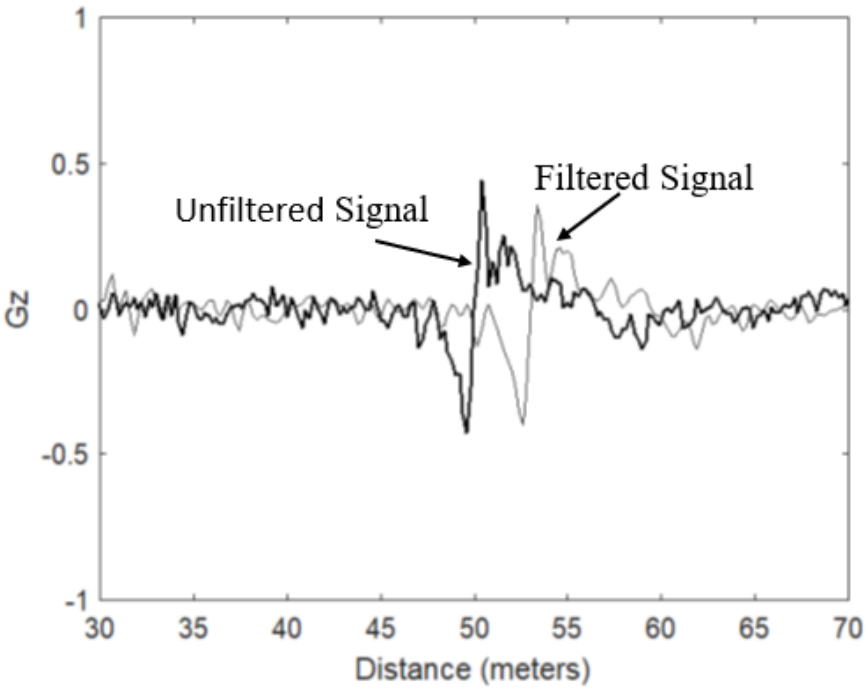


Figure 5.9. Unfiltered and filtered signals comparison

5.3.3.2. Result Analysis

The SNR analysis demonstrates the effectiveness of the proposed method by comparing the filtered and unfiltered signal to validate improvements in SNR. The results show that the method of EA-FFT with statistical decision criteria is effective for informing a cut-off frequency to digitally low-pass filter the inertial signals. From figure 5.7, it is evident that the windowed ensemble averaging provides a clearer picture of the frequency transition from signal to noise. Also, it shows that Ensemble Average (EA) reduces noise and enhances correlated energy features in the spectral representation of the signal. Subsequently, there is agreement among the EA-FFTs on the frequency at which both the energy and variance in the spectral windows reach a minimum before rising again slightly. In particular, both the energy and variance windows of the EA-FFT shows such a transition at 25 Hz.

- Filtered Signal vs. Unfiltered Signal

Figure 5.9 shows that the filter effectively removed noise while preserving the quality and strength of the inertial signature. The procedure can use any low-pass filter, including infinite impulse response (IIR) filters and is independent of the data collection apparatus and their operating systems.

- SNR Improvement Validation

The filter order determines the degree to which the filter attenuates noise. A poor choice of cut-off frequency could result in the attenuation of both signal and noise, and subsequently no improvement in SNR. Therefore, the observation of SNR improvements with increasing filter order validates the cut-off frequency selection because the filter reduces noise while preserving the strength of the desired signal.

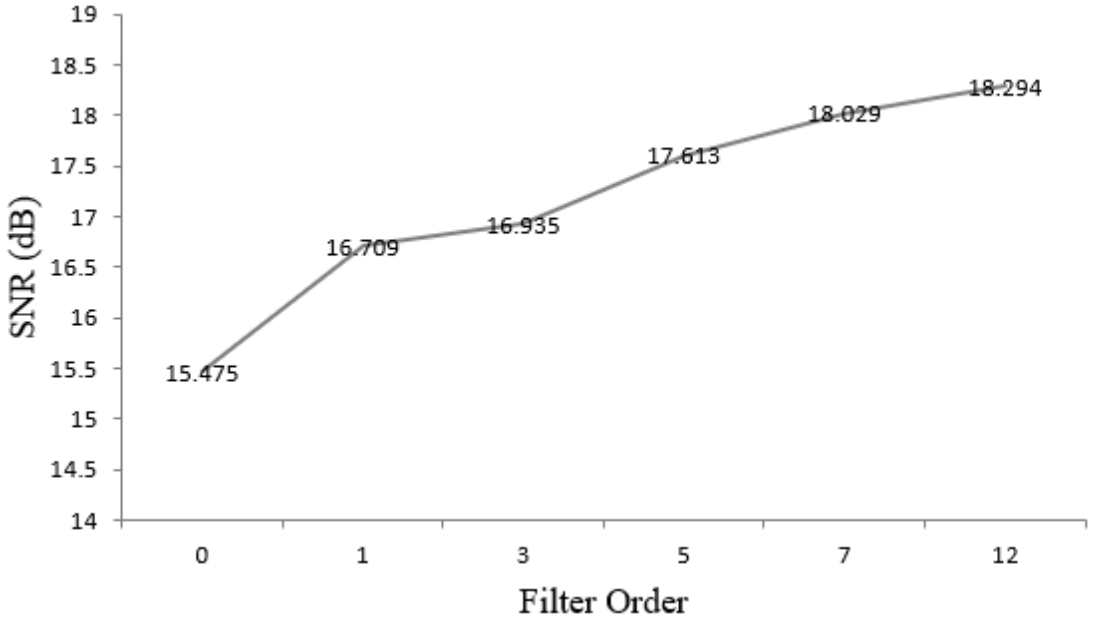


Figure 5.10. SNR as a function of filter order (zero order is the unfiltered signal) demonstrates the effectiveness of the cutoff frequency selection

Figure 5.10 shows that after applying a low-pass FIR filter with the selected cut-off frequency, the SNR steadily increased with the order of the filter. The SNR shown at order zero

is for the unfiltered signal. This procedure prepared the inertial signals for the remaining stages of an application that involves signal position realignment, distance interpolation, and truncation to obtain signals from approximately equal length traversals for feature extraction and feature ensemble averaging.

5.3.3.3. Conclusion

Resolving issues such as non-uniform sample rate that adds noise and decreases the signal strength enables the placement of sensors in vehicles to monitor road or railroad condition. This work showed that appropriate signal filtering and alignment for ensemble averaging could be effective in improving the SNR for subsequent feature extraction. However, an objective determination of the filter cut-off frequency is necessary. From the individual FFTs, it is difficult to see a clear pattern that separates signal from noise. Hence, method ensemble averages the individual FFTs (EA-FFT) from the approximately equal length and position aligned inertial signals to enhance the clarity of the underlying pattern. The subsequent application of two statistical methods to frequency windows of the EA-FFT provides an objective means for selecting a frequency threshold where signal transitions to noise.

This proposed EA-FFT method to identify the cut-off frequency of a low-pass noise filter can be generalized to any sensor data that meet at least two criteria. The first is that the signal contains both signal and noise. The second is that the data source is from multiple traversals of a segment of some transportation network such as a rail track or a roadway. Such data produces distinct inertial signal patterns or features in common time windows. Moreover, variable train speed, GPS position registration errors, relatively slow update rates of the GPS receivers, and the uneven sampling rate of their inertial sensors result in feature alignment errors that can degrade

the signal-to-noise ratio. Therefore, it is necessary to align and equalize the position of similar features or responses across multiple traversals.

5.3.4. Signal Feature Extraction

This section introduces an analytic framework that includes mathematical and statistical methods to enhance the detection and localization accuracy of track or road irregularities by extracting and combining features of the inertial signals obtained from multiple traversals of a track segment. Figure 5.11 shows the workflow framework of the feature extraction and localizing irregularities application. The framework explains each step that enhance the signal quality for subsequent feature extraction discussed in previous sections. The RIF transform is employed to extract the features within a selected window size.

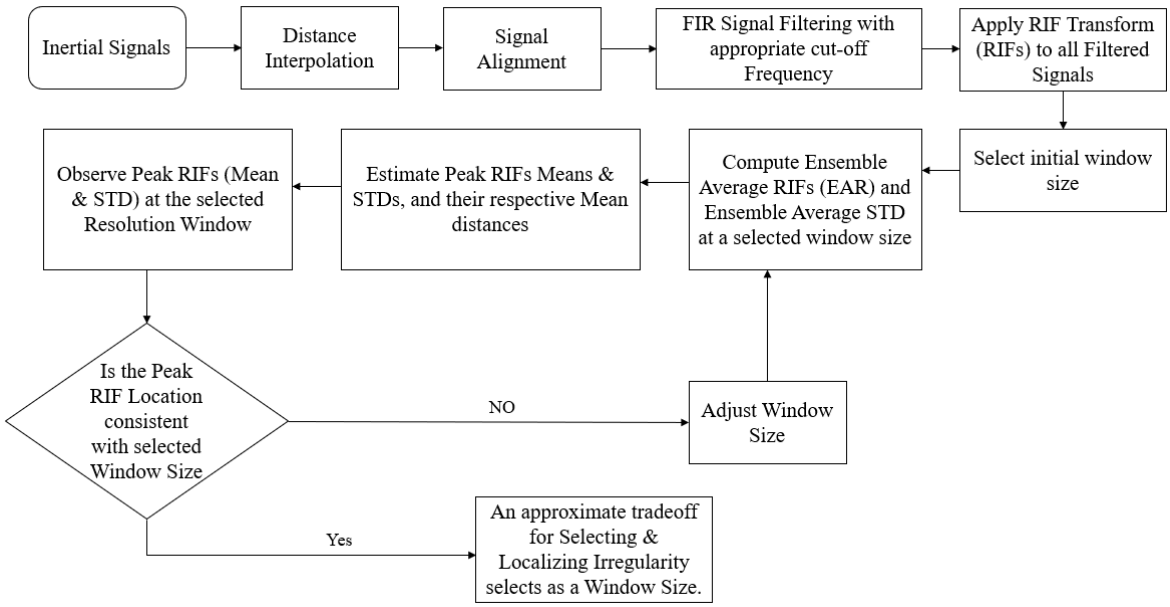


Figure 5.11. Signal feature extraction & localizing road or track irregularities workflow framework

The experiment has also been conducted on local railroad data to show the effectiveness of the procedures. The collected data was from the traversal of a rail grade crossing (RGC)

because it reliably simulated the presence of a rail anomaly. The study concluded that the method applies identically on actual railroad detect and localize railroad track anomalies.

5.3.4.1. Method

The method inherits a scalable technique that not only extract roughness or irregularity features of the inertial signals obtained from multiple traversals of a road or track segment but also transform and combine features into a reliable single-index summary of road or railway roughness called EAR-index. The value of each peak EAR is the effective estimator for both the roughness intensity and position. The workflow framework shown in figure 5.11, the method primarily includes distance interpolation from geo-spatial reference position to align the signals with approximately equal length segments. The algorithm then applies a finite impulse response (FIR) low-pass filter to filter the aligned signals as described in the previous subsection. Afterword, the EAR is computed across the small and fixed distance windows of multiple traversals. The following are the mathematical computations of the algorithm.

- Multi-Resolution RIF Feature Indices

The RIF transform reduces the data into features that are proportional to the roughness by replacing the segment length L in Equation (1) with a distance resolution window ΔL such that

$$\breve{R}^{\Delta L} = \sqrt{\frac{1}{\Delta L}\sum_{n=0}^{N-1}\left|G_{z[n]}\overline{v}_n\right|^2\delta t_n} \tag{20}$$

where the RIF index $\breve{R}^{\Delta L}$ is the average g-force magnitude per unit of distance ΔL travelled. $G_{z[n]}$ is the vertical acceleration for signal sample n, sample period instant δt_n and, the instantaneous traversal speed is $\overline{v}_n$. The window position varied with GPS error. The RIF intensity varied with traversal speed. Hence, an *ensemble average of the RIF-indices* (EAR) within a selected distance

window along the traversal path and across N traversals produced an estimate of the average roughness $\bar{E}^w$ experienced in window w, at any speed such that

$$\bar{E}^w = \frac{1}{N}\sum_{x=1}^{n} \breve{R}_w^{\Delta L}[x] \tag{21}$$

where $\breve{R}_w^{\Delta L}[x]$ is the RIF index from traversal x within distance window w.

The peak of EAR shows the amount of roughness produced when a vehicle traverses any road or railway anomalies. The roughness represents the non-uniform track geometry that affect vehicle dynamics and ride quality.

- Ensemble STD RIF Indices

The ESR indices represents as $\tilde{\sigma}_{RIF}$ is a measure of the variability of the roughness intensity across traversals, within a distance window. It is defined as:

$$\tilde{\sigma}_{RIF} = \sqrt{\frac{\sum_{x=1}^{N}\left(\breve{R}_w^{\Delta L}[x] - \bar{E}^w\right)^2}{N}} \tag{22}$$

- Margin-of-Error

The margin-of-error denoted as $\breve{R}_{1-\alpha}^{\Delta L}$ within a (1-σ)% confidence interval defined as

$$\breve{R}_{1-\alpha}^{\Delta L} = \mathrm{z} * \frac{\sigma^{\breve{R}^{\Delta L}}}{\sqrt{n}} \tag{23}$$

where $\sigma^{\breve{R}^{\Delta L}}$ is the standard deviation of the intensity of the peak RIF within the selected distance window sizes and n is the total number of traversals within the window size. Z is a critical value of 1.96 at 95% confidence interval (CI). The $\breve{R}_{0.95}^{\Delta L}$ indicates that 95% of the RIF data values are likely to be within that percentage of the ensemble average RIF-index.

5.3.4.2. Results Analysis

The experiment has been conducted on both roadway and railway data that demonstrate the reliability and accuracy of the employed method. Also, tradeoff analysis found an optimal window size of anomaly detection that minimizes the false positives and false negatives.

- Road Experiment

Figure 5.12 displays the *EAR* for the 1-, 5-, 15-, and 20-meters distance resolution window sizes.

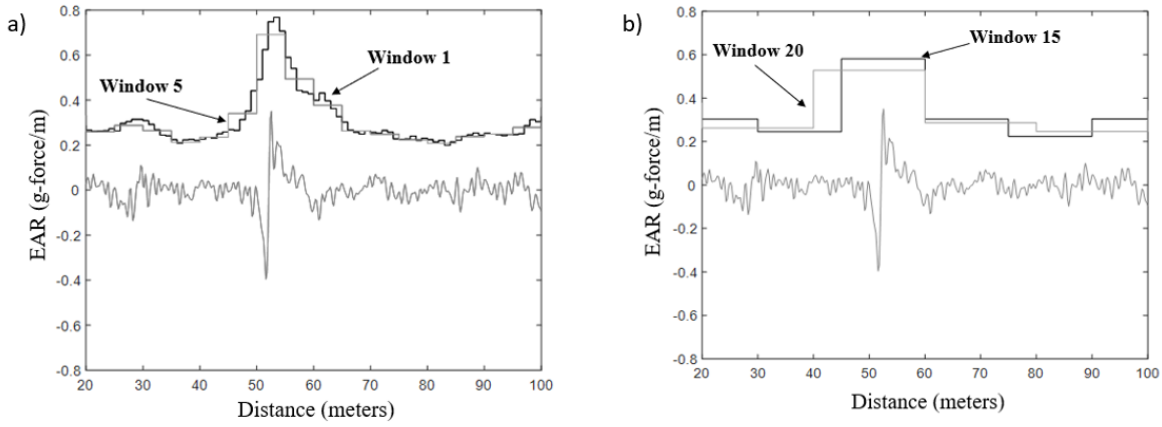


Figure 5.12. a) Ensemble average RIF Indices at the resolution window 1 and 5-meters.
b) ensemble average RIF Indices at the resolution window 15 and 20-meters

The chart shows the unprocessed inertial signal from one of the traversals for reference. The position of the maximum EAR is an estimate of the position of the track irregularity. The precision of position estimate increases with the number of traversals. Larger window sizes provide greater data reduction but reduce the precision of estimating the position of an irregularity.

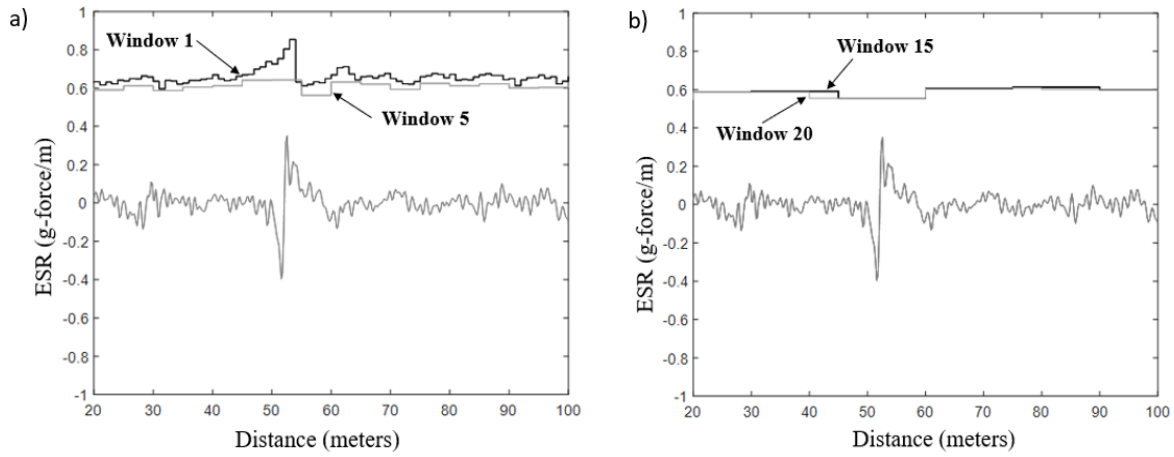


Figure 5.13. a) Ensemble STD RIF Indices at the resolution window 1 and 5-meters.
b) ensemble STD RIF Indices at the resolution window 15 and 20-meters

Figure 5.13 a) and b) displays the ESR for the 1-, 5-, 15-, and 20-meters distance resolution window sizes. The *ESR* generally declines as the window size increases because the EAR also decreases. This represents a tradeoff in data size for accuracy and precision in estimating the position of anomalies.

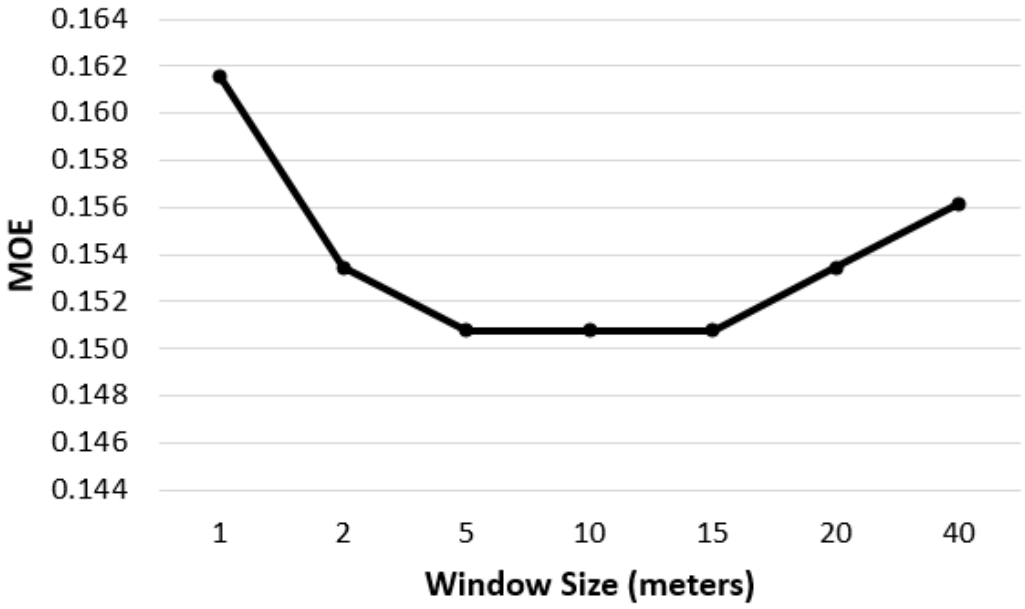


Figure 5.14. The MOE of the peak RIF for varying distance windows

Figure 5.14 shows the MOE of the peak RIF for distance windows of 1, 2, 5, 10, 15, 20 and 40 meters. The chart shows that the MOE declines most substantially between window sizes of 1 and 5 meters. This suggests that that the reliability of the estimate is best for the window sizes in that range.

- Findings

Table 5.2. Statistical summary of peak RIF within the selected distance resolution window size

Window Size	Peak RIF Mean	Peak RIF STD	Peak RIF Dist Mean	Peak RIF Dist STD
1	1.48	0.60	53.38	4.02
2	1.21	0.57	53.59	4.02
5	0.87	0.56	54.47	8.50
10	0.70	0.56	56.79	15.91
15	0.60	0.56	56.15	20.15
20	0.53	0.57	51.98	5.91
40	0.43	0.58	101.22	56.57

From the road experiment, it is found that the position of the maximum EAR is an estimate of the position of the irregularity and the ESR declines as the window size increases, which is the expected outcome. As a function of window size, Table 5.2 summarizes the means and standard deviations (STD) of the peak RIF intensity and the center position of the window for the peak RIF relative to the beginning of the traversals.

A smaller window size increases the precision of locating an irregularity but also increases the variability of that estimate. The results show that an optimum window size exists that minimizes the STD of estimating the RIF intensity. A lower STD improves the consistency of detecting an anomaly and, therefore, reduces the false positive and false negative detection errors. From the table 5.2 and figure 5.15 a, and 5.15 b, it is observed that that keeping the distance window size below 5 meters maintains a stable estimate of the position of the peak RIF relative to the beginning of the traversal (approximately 53 meters for the experiment) and minimizes the STD of that estimate.

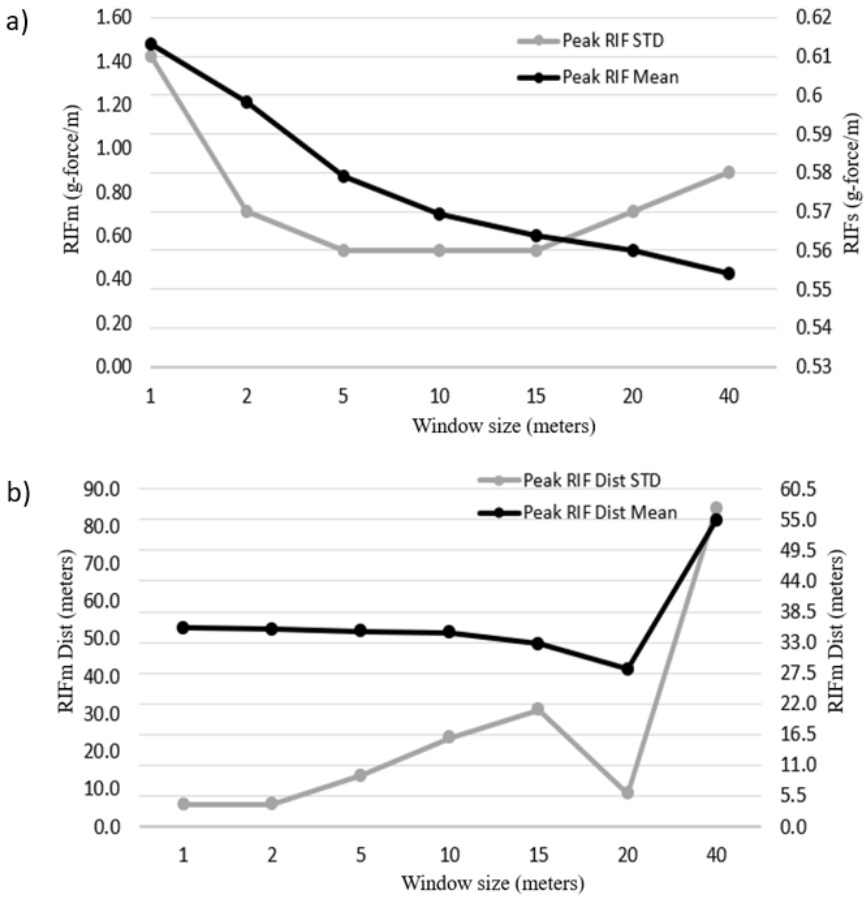


Figure 5.15. a) The mean and standard deviation of the intensity of the peak RIF within the small and fixed distance window sizes. b) the mean of peak RIF distance relative to STD of peak RIF distance at different window size

Additionally, figure 5.14 indicates that MOE provides the confidence of the measurements with an indication that there is an optimum window size selection. Table 5.3 shows the margin-of-error (MOE) of peak RIF within the selected window size.

Table 5.3. MOE of the peak RIF within the selected window

Window Size	Total Traversals	MOE Peak RIF
1	53	0.162
2	53	0.153
5	53	0.151
10	53	0.151
15	53	0.151
20	53	0.153
40	53	0.156

- Rail Experiment

From the rail experiment, Table 5.4 represents the mean and STD of the peak RIF intensity relative from the beginning of each traversal.

Table 5.4. Statistical summary of peak RIF within the selected resolution window using rail

Window Size	Peak RIF Mean	Peak RIF STD	Peak RIF Dist Mean	Peak RIF Dist STD
1	0.764	0.187	90.114	43.780
2	0.635	0.145	86.713	30.112
5	0.498	0.097	87.263	29.889
10	0.424	0.074	91.513	21.590
15	0.387	0.063	88.512	28.958
20	0.370	0.054	86.013	27.988
40	0.335	0.059	98.013	8.945

Figure 5.16 shows that the RIF transform presents a stable trade-off between resolution and confidence as a function of window size with an asymptotic decrease of the mean value and a consistent decrease in the STD of the estimate. Figure 5.16 reveals that a window size of 5 meters provides a stable estimate for locating rail anomalies relative to a linear reference at the beginning of the traversals, which was approximately 90 meters in these experiments.

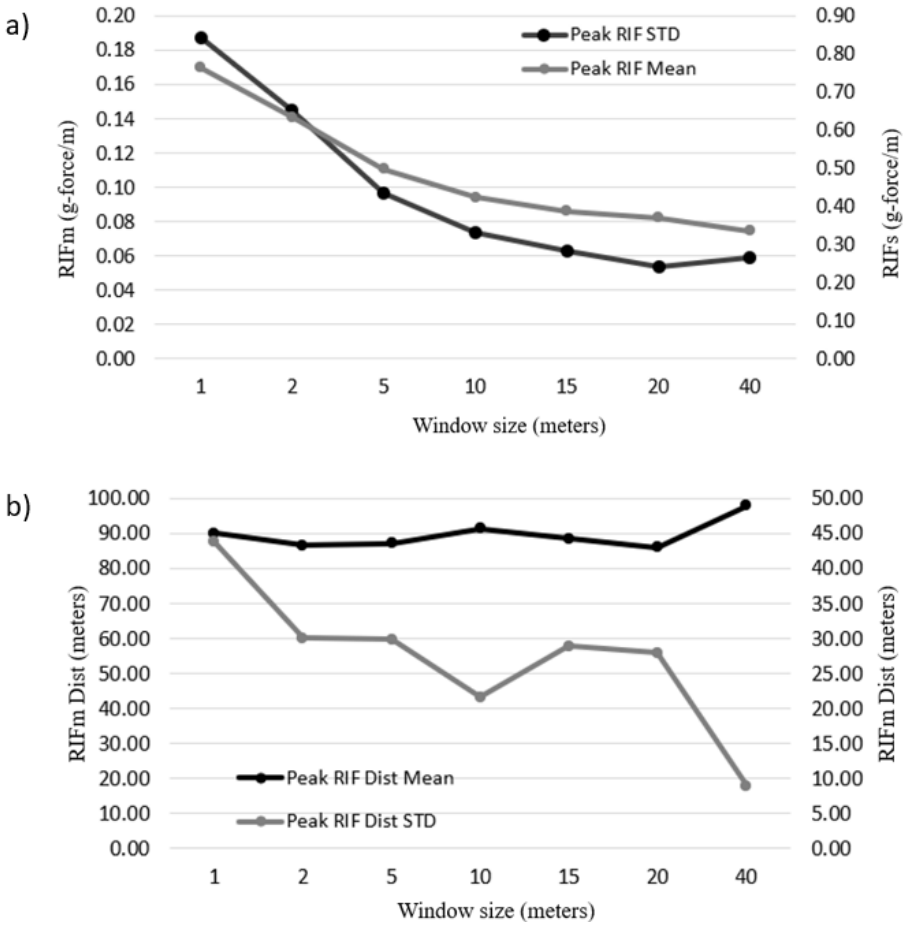


Figure 5.16. a) The mean and STD of peak RIFs of varying window size using rail data. b) mean of peak RIF distance relative to STD of peak RIF distance at different window size using rail data

Therefore, a window size of 5 meters in both the road and rail experiments balances the tradeoff in data size, precision, and accuracy of locating road or rail irregularities with sensors. The best window size is also within the visual range for inspectors to locate the anomaly.

5.3.4.3. Conclusion

The ensemble average of RIF-indices within each window provides an estimate for the intensity and position of an irregularity. Increasing the window size decreases the variability of the intensity estimate and decreases the data size. However, a larger spatial window size also reduces the accuracy and precision of position estimates. Both roadway and railway experiments were conducted to demonstrate the reliability and accuracy of the employed method. In both

experiments, tradeoff analysis found that a window size of 5 meters provided a good balance between data reduction, accuracy, precision, and the consistency of anomaly detection while minimizing the potential for false positives and false negatives. A significant benefit of the ensemble averaging approach is that both precision and accuracy increase with the number of traversals. Furthermore, the MOE of peak RIF values also validates confidence in the measurements and points to an optimal window size selection.

6. POSITION ACCURACY ASSESSMENTS

6.1. Objective

This chapter's primary objective is to investigate the potential use of low-cost sensors aboard the hi-rail vehicle to monitor automatically and continuously for inertial events caused by irregular track geometry. Second objective is to characterize and validate its accuracy by comparing the estimated positions of detected irregularities with the actual positions of irregularities that the railroad inspector observed.

Another objective is to estimate the position of the unknow irregularity and validate its actual position that the railroad inspector verified.

6.2. Contribution

The chapter incorporates a signal processing and statistical method to estimate the position of peak inertial events from multiple traversals. The method reduces the inertial signals associated with irregularity to features called Road Impact Factors (RIFs) [60]. The method also utilizes a geographical information system (GIS) platform to visualize peak inertial events (PIEs), which are RIF values in the high five percentiles of its distribution.

Another contribution is the case studies that validate and characterize the detected irregularities. It demonstrates the practical utilization of RIF-indexes by using a smartphone app to record the required sensor data from a regular service vehicle or locomotive. This will help railroad agencies identify and catalog irregularities that demand critical repair.

6.3. Case Studies of Irregularity Position Assessment

The following sections demonstrate and estimate the position of irregularities called linear references. The study first estimates the position for known irregularity called ground truth areas (GTA), and the second is for unknown irregularity. The case study uses inertial and

geospatial position data collected from sensors aboard a hi-rail vehicle that regularly monitors the condition of a local rural railroad. A portion of the data and its format shows in table 4.1 in the previous chapter 4.

6.3.1. Position Estimation of Known Irregularity (GTA)

Track irregularity not only degrades the ride comfort but also increases the risk of derailments [67]. When a train traverses an irregularity, it produces a Peak Inertial Event (PIE). So, the magnitude of PIEs is an effective means of extracting track condition patterns. A Road Impact Factor (RIF) transform extracts track condition patterns as features from inertial sensor signals to reveal the signature of a track irregularity. In another words, the magnitude of a RIF feature is proportional to the amount of irregularity reported by the railroad inspection supervisor such as warp or cross level. The cross-level is the elevation difference between rails at a given position along the track. For a fixed amount of cross leveling, the intensity of the RIF feature varies as a function of traversal speed.

6.3.1.1. RIF Transform Indexes

The RIF-transform produces an intensity RIF_{Rt} that is proportional to the resultant rotation rate about x and y-axis. The RIF-transform is defined as:

$$RIF_{\mathrm{Rt}} = \sqrt{\frac{1}{L}\sum_{n=0}^{N-1}|Rt_n\, v_n|^2\, \delta t} \tag{24}$$

where RIF_{Rt} is the average magnitude of rotation rate per unit of distance L traveled. The sample period is δt. A speed sensor produces the instantaneous traversal speed v_n at sample n for N total samples. Rt_n is the resultant rotation signal sample n in units of the radians-per second such that

$$Rt = \sqrt{(R_{\mathrm{x}})^2 + \left(R_{\mathrm{y}}\right)^2} \tag{25}$$

where Rx and Ry are angular rotations around X and Y axis of the phone.

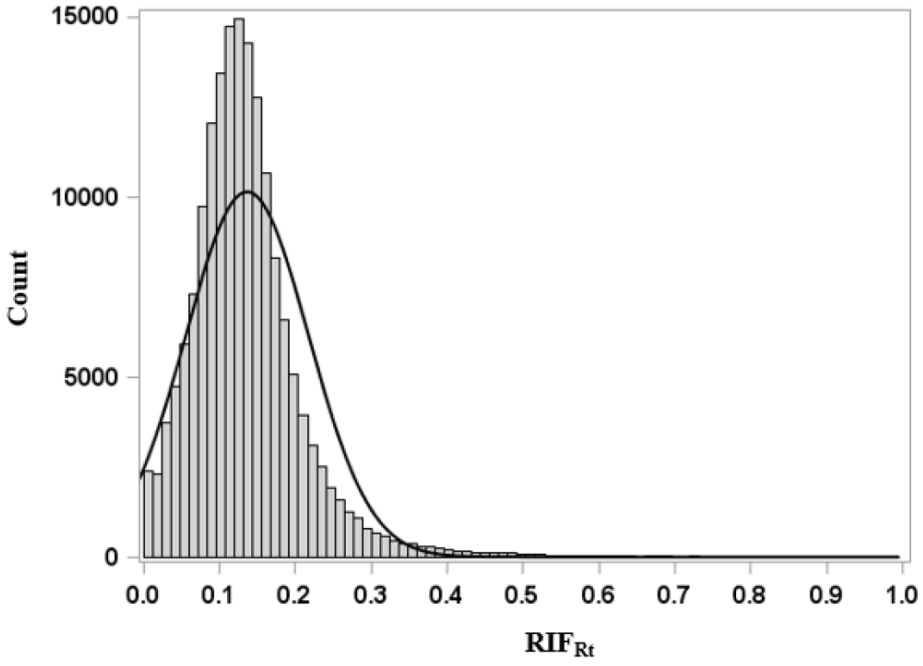


Figure 6.1. Distribution of RIF_{Rt}.

As discussed in the chapter 4, RIF features follows a normal distribution. Hence, after data cleaning, a color-coding scheme (shown in table 4.2 in the chapter 4) based on percentile thresholds of the RIF_{Rt} distribution (figure 6.1) provided a means to visualize the PIEs using a GIS package. Figure 6.2 shows the color-coded RIF_{Rt} values for the visualization of PIEs that correspond to severe track irregularity.

6.3.1.2. Position Accuracy Estimation

The method of estimating the position of an irregularity involved accumulating PIE clusters across multiple traversals and then determining a centroid for each cluster. Estimating the position accuracy with this method required accumulating from all available traversals the PIE that was closest to the position of the reported irregularity. The method built a cluster from PIEs with intensity levels in the ‘orange’ or ‘red’ range. That is, track irregularities produce PIEs in the ‘orange’ range when the vehicle moves more slowly across them.

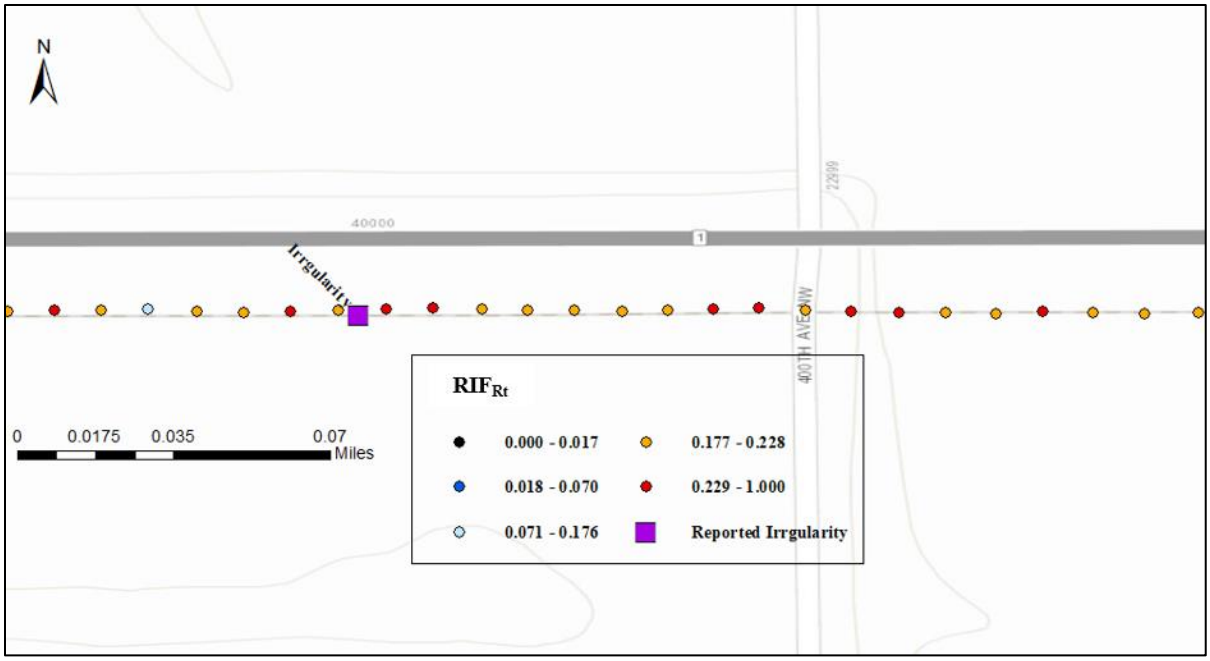


Figure 6.2. A map of PIEs that correspond to severe track irregularity.

6.3.1.3. Results

Figure 6.3 is a GIS map that shows the PIE cluster near the reported irregularity. Figure 6.4 shows the centroid position relative to the actual irregularity. The centroid was offset 13.1 feet (3.9m) from the position of the reported irregularity.

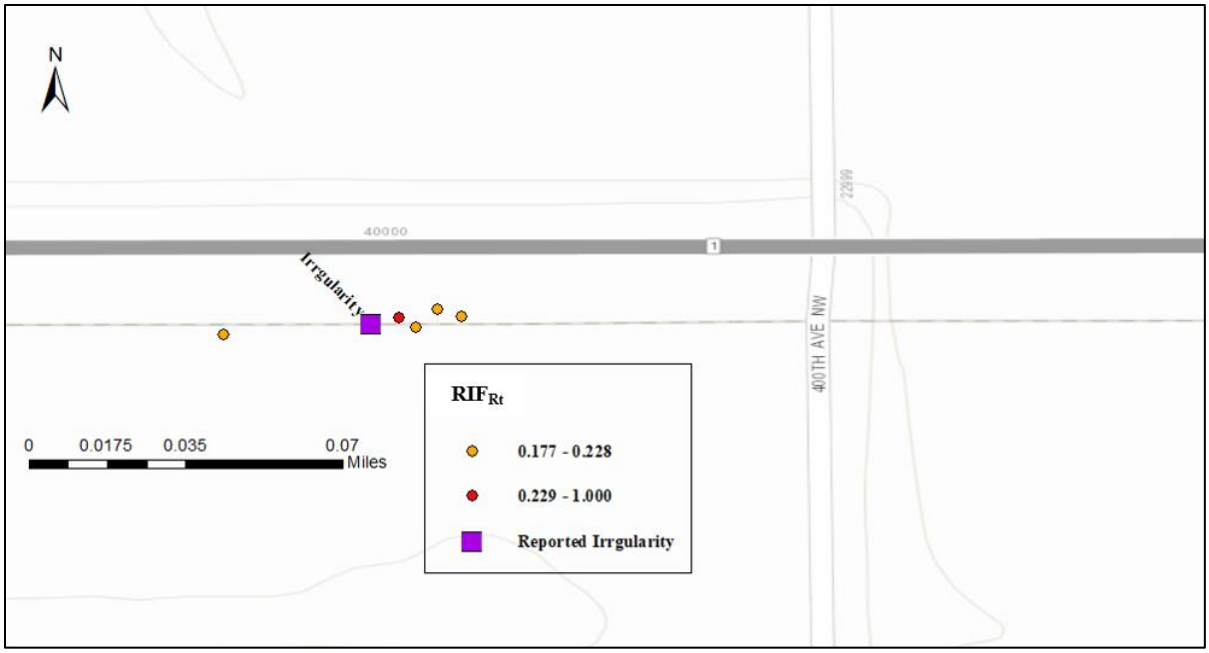


Figure 6.3. A map of the PIE cluster near the reported irregularity

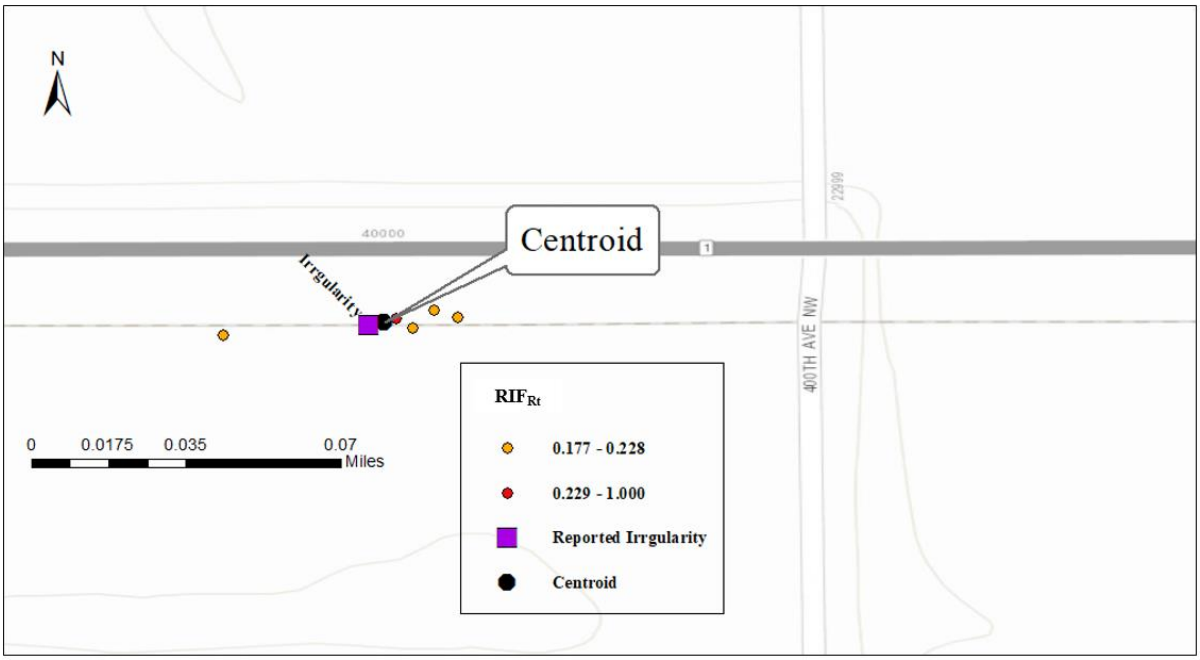


Figure 6.4. A map of the PIE cluster centroid relative to the actual irregularity

This research also validated the method for other Known anomaly called ground truth area (GTA). Below is the reference table 6.1.

Table 6.1. Known anomalies position assessment

Ground Truth Area (GTA)/ Known Anomaly	Mile Post	Lat	Lon	Distance Offset from GTA
Cross-level	MP 344.7	48.194843	-96.913964	3.9 meters (13.1 feet)
Rail Grade Crossing (RGC)	Near MP 344.7	48.194838	-96.912465	1.24 meters (4 feet)
Flacking, track Wiggle	MP 380.5	48.2391	-97.676611	3.14 meters (10 feet)

6.3.2. Position Estimation of Unknown Irregularity

The rail track irregularity not reported or observed by the railroad inspector is referred to as unknown irregularity. Early detection of these unknown irregularities is critical for maintenance and aids in timely exchange to avoid accidents. This case study uses equations 1 and 2 to calculate the RIF features. These RIF features follow the normal distribution (mentioned in chapter 4). Therefore, peak inertial events (PIEs) are visualized using a defined threshold and color-code scheme.

6.3.2.1. Position Localization Estimation

Localizing the position of unknown irregularity involves the extraction of traversal segments with frequently occurring PIEs irrespective of the direction of travel. The study extracted the traversals from the route Viking MP 320.2 and Radium MP 330.3.

This method required accumulating the most frequently occurring PIEs (red & orange) within the extracted traversal segments shows in figure 6.5. The method builds a PIE cluster across extracted traversal segments and then determines a centroid from the PIEs. The computed PIEs centroid is referred to as a Hot Spot area.

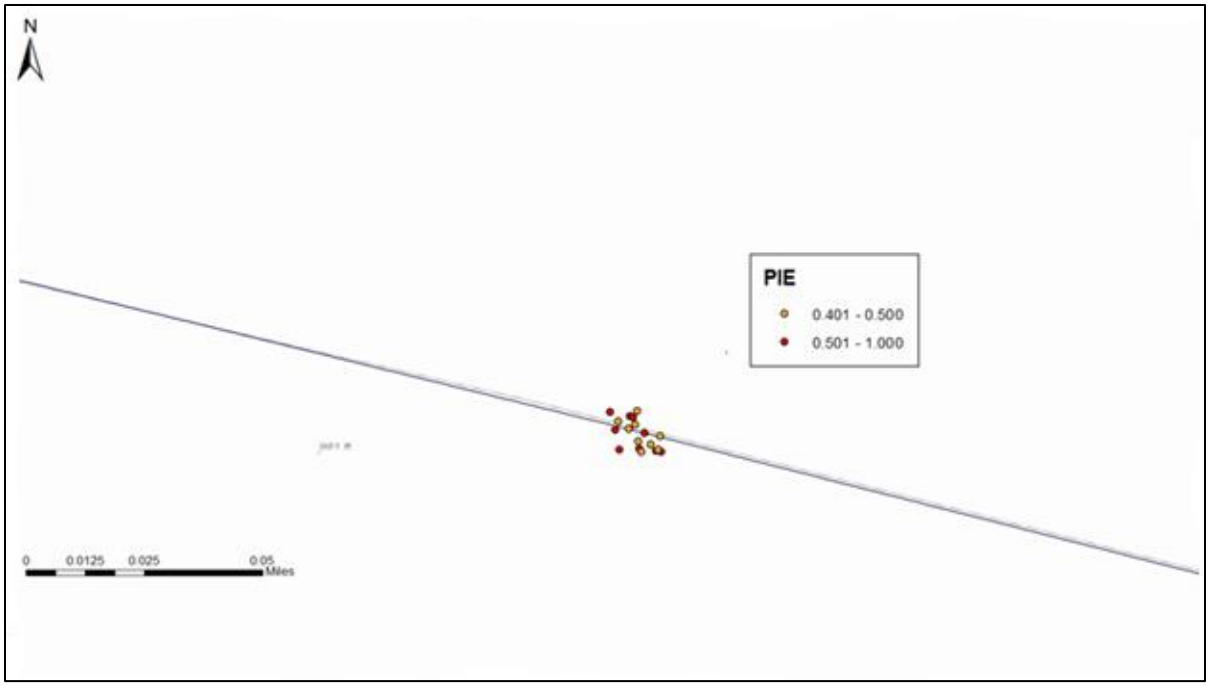


Figure 6.5. A map of most occurring and brightest PIEs cluster

6.3.2.2. Results

Figure 6.6 is a GIS map that shows the PIE cluster centroid. Figure 6.7 shows the hot spot area representing the PIEs cluster centroid position relative to the unknown irregularity. The hot spot area is at MP 325.614 which is 5.414 miles from the Viking 320.2 milepost. The position of the estimated hot spot area is validated and verified by the railroad inspector as show in figure 6.8. The result suggests that the estimated position of the unknown irregularity is within sight distance to locate the anomaly for follow-up inspections.

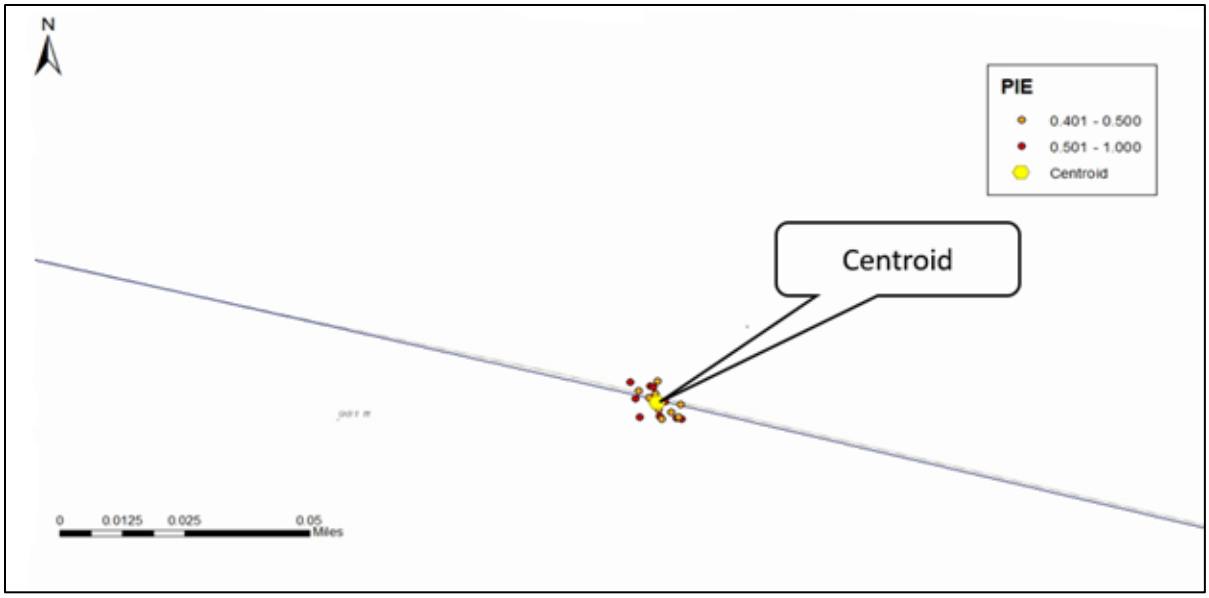


Figure 6.6. A map of PIEs cluster centroid

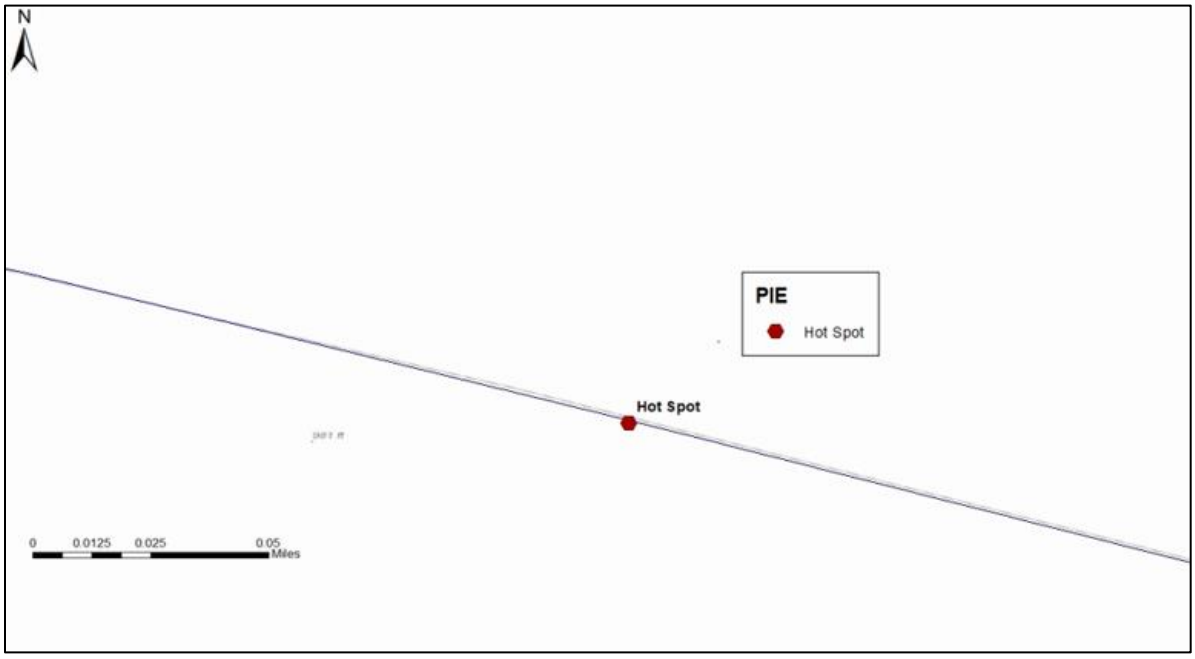


Figure 6.7. A map of Hot Spot representing unknown irregularity

Figure 6.8. A hot spot area position verified by railroad inspector at MP 325.614

6.4. Conclusion

This chapter introduced a signal processing and statistical method to estimate the position of peak inertial events from multiple traversals. The goal was to characterize its accuracy by comparing the estimated position of detected irregularities with the actual positions of irregularities that the railroad inspector observed. Another goal is to verify and validate the estimated position of unknown irregularities.

The study developed a mathematical model to produce RIFs from the signals and defined a color-coding scheme to visualize PIEs using a geographical information system (GIS) platform. A PIE cluster accumulated from the multiple number of traversals provided an estimate

of the position of a measured irregularity. Two case studies are demonstrated to show the accuracy and reliability of the proposed system.

In first case study, the result is that the centroid is 13.1 feet (3.9 m) offset from the position of actual irregularity (GTA). Therefore, the estimated position of the irregularity is within a reasonable visual distance of the actual position of irregularity.

The second case study showed the estimated position of an unknown irregularity that was validated and verified by railroad inspector as a rail joint. The result shows that the hot spot area (PIEs cluster centroid) is 5.414 miles from the MP 320.2 milepost that can be seen during a follow-up manual inspection. Thus, the results will help to determine a model for quantifying the severity of track geometric irregularities such as profile, alignment, and warp.

7. CONCLUSION

A frequent and network-wide monitoring of railroad infrastructure is highly in demand to meet the emergent need of safety, reliability, and lower cost operations. This study investigated the track and railbed problems cause the derailments. It was determined that some existing practices requires line closure to perform inspections. However, some sophisticated technologies are more costly and complex to scale for more frequent and network-wide coverage. Thus, railroad companies think twice before investing and adopting the technology. Therefore, this research developed a generalized low-cost monitoring & evaluation system is called Railroad Autonomous Inspection Localization System “RAILS”.

The developed methods in this research will encourage transportation agencies to deploy and implement the technique to frequently scan the network for anomalies and repair them before they cause safety issues and vehicle damage.

7.1. Anomaly Position Detection and Localization Enhancement Conclusion

RAILS utilizes low-cost micro-electro-mechanical (MEM) accelerometer, gyroscope, and GPS receiver sensors embedded in the smartphone. However, it is found that their poor performance due to varied update rates and lower resolution is not suitable for high-accuracy and high-precision applications. In other words, it increases both false positive and false negative feature detection errors.

Therefore, this research embraces the ensemble averaging approach to enhance the signal quality by improving signal-to-noise ratio. The technique combines the signals from multiple traversals of a segment by ensemble averaging to reduce noise and enhance position resolution. This suggests that the signal quality improves continuously with the additionally available position repeatable signal stream. Subsequently, the finding that integrate distance interpolation,

signal alignment, and signal filtering to improve detection error and signal quality. Another finding is that an appropriate noise filtering is necessary to maximize the signal-to-noise ratio (SNR) of each signal prior to feature extraction. Thus, the developed methods that reduce detection error, and enhance the quality of the signals and features to get the ensemble averaging (EA) through heuristic alignment and signal filtering can be generalized for the practical applications produced sensor data from multiple traversals of railroad segments.

7.2. Position Accuracy Assessment Conclusion

The research evaluates the proposed approach that employed a Road Impact Factor (RIF) transform, which emphasizes features from the sensor signals in proportion to the amount of track geometry irregularity. The finding is that the estimated peak inertial events (PIEs) position is at a possible rail track irregularity. The study compared the estimated positions of detected irregularities with the actual positions of irregularities observed by the railroad inspector.

Another finding was the estimated position of an unknown irregularity in the route between mileposts MP 320.2 and MP 330.3. The railroad inspector verified the estimated position of the unknown irregularity at MP 325.614 as a rail joint. Thus, this finding validates and characterizes the accuracy of the method and its effectiveness.

7.3. Limitations

The following are the limitation that may impact the data and validation complications:

- *Inaccuracy in-field inspection due to human error*: The traversal volume plays a vital role in the precision of anomaly localization. In other words, signal quality improves with the continuous addition of data. However, data collection may be hampered due to the following:

- The railroad inspector forgets to plug in the phone charger during track inspection.
- The railroad inspector forgets to unplug the phone charger during non-inspection days. It will add stationary data that will limit the scope of detecting and localizing the anomaly with high accuracy and precision.

- *Impact on Ground Truth Data*: Generally, railroad inspectors work 8-10 hours shift to locate the anomaly in their assigned section of track. However, they frequently work beyond the scheduled workday, and sometimes on rest days [68]. The amount of territory an inspector covers depends on considerations such as the inspection method, the number of curves on the track, the type of rail (e.g., jointed or continuous welded), weather and track visibility, and time constraints. Time constraints contribute to time pressure that can hinder the inspectors' ability to detect defects effectively [68]. Therefore, a high-pressure work environment discouraged track inspectors from reporting issues or problems [69] and repairs.

  Moreover, many inspectors felt that their territory was too large to inspect in the time available. Hence it limits their ability to complete inspections [69]. Consequently, it may impact the identification of the ground truth area (GRA).
- *Weather Condition*: Weather conditions such as severe winter storms sometimes disrupt the entire transportation system [70]. The railroad operations degrade in such conditions due to lower visibility, icing, snowdrifts, and cold temperatures. Therefore, it affects the data collection pattern and the number of traversals. Also, extreme cold weather leads to phone discharge that may disrupt the data collection.

Additionally, uneven or rapid temperature variations could increase the track maintenance requirements [70]. For example, a track exposed to extreme heat poses misalignment or warp. In winter, the freeze-thaw cycle adversely degrades the track surface. Thus, it may worsen known track defects or raise new problems.

7.4. Future Work

The promising results from this research suggest that future work on this study should be continued. The successful implementation of the low-cost onboard sensing monitoring technology in this research shows the significant potential for economic benefits and accident prevention due to the track and roadbed problems.

The finding from this research has some extracted features that labeled as joint track roughness and warp. These features have been verified by the local railroad inspector. Subsequently, these extracted features from the composite signals will then be used to develop the machine learning models that can be trained to classify the types of detected railway track irregularities. The machine learning models involves supervised and semi-supervised learning that requires labeled features. These labeled features have two values type and intensity that will be the input for the machine learning models. Additionally, all these research algorithms and models including machine learning process will annotate with maps to make the system work autonomously in future.

Therefore, this capability enhanced the follow-up inspection and repairs that will help railroad industry to prioritize inspection resources and equipment to categorize the issues and make informed decisions. Moreover, such solution could integrate with the PTC system as another sensor that strength the communication network and cloud-based system.

8. REFERENCES

[1] L. Janušová and S. Čičmancová, "Improving safety of transportation by using intelligent transport systems," *Procedia Engineering*, vol. 134, pp. 14-22, 2016.

[2] X. Zhang, L. Jia, X. Wei and N. Ru, "Railway track condition monitoring based on acceleration measurements," in *In The 27th Chinese Control and Decision Conference*, Qingdao, China, 2015.

[3] R. S. Barbosa, "New method for railway track quality identification through the safety dynamic performance of instrumented railway vehicle," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 38, no. 8, pp. 2265-2275, 2016.

[4] P. Xu, Q. Sun, R. Liu, R. R. Souleyrette and F. Wang, "Optimizing the alignment of inspection data from track geometry cars," *Computer-Aided Civil and Infrastructure Engineering*, vol. 30, no. 1, pp. 19-35, 2015.

[5] Boardman, Joseph H., "FRA ’ s Current Safety Regulations and Rulemaking," Federal Railroad Administration, U.S. Department of Transportation, 27 June 2006. [Online]. Available: https://www.transportation.gov/testimony/fras-current-safety-regulations-and-rulemaking-proceedings. [Accessed 01 07 2022].

[6] X. Liu, M. R. Saat and C. P. Barkan, "Analysis of causes of major train derailment and their effect on accident rates," *Transportation Research Record*, vol. 2289, no. 1, pp. 154-163, 2012.

[7] P. Lu, R. Bridgelall, D. Tolliver, L. Chia and B. Bhardwaj, "Intelligent Transportation systems approach to railroad infrastructure performance evaluation: track surface abnormality identification with smartphone-based App," Mountain-Plains Consortium, Fargo, 2019.

[8] E. M. Vinberg, M. Martin, A. H. Firdaus, Y. Tang and A. Qazizadeh, "Railway applications of condition monitoring," KTH Royal Institute of Technology, Stockholm, Sweden, 2018.

[9] Z. Zhang, X. Liu and K. Holt, "Positive Train Control (PTC) for railway safety in the United States: Policy developments and critical issues," *Utilities Policy*, vol. 51, pp. 33-40, 2018.

[10] J. C. Peters and J. Frittelli, "Positive Train Control (PTC): overview and policy issues," Congressional Research Service, Washington, DC United States, 2018.

[11] G. W. Flintsch, B. Ferne, B. Diefenderfer, S. Katicha, J. Bryce and S. Nell, "Evaluation of traffic-speed deflectometers," *Transportation research record,* vol. 2304, no. 1, pp. 37-46, 2012.

[12] R. Bridgelall, "Connected Vehicle Approach for Pavement Roughness Evaluation," *Journal of Infrastructure Systems* , vol. 20, no. 1, pp. 1-6, 2014.

[13] A. Malekjafarian, E. J. O'Brien and D. Cantero, "Railway track monitoring using drive-by measurements," in *In The Fifteenth East Asia-Pacific Conference on Structural Engineering and Construction*, Xi'an, China, 2017.

[14] L. Wang, P. D. Groves and M. K. Ziebart, "GNSS shadow matching: Improving urban positioning accuracy using a 3D city model with optimized visibility scoring scheme," *NAVIGATION: Journal of the Institute of Navigation,* vol. 60, no. 3, pp. 195-207, 2013.

[15] GPS Product Team, Federal Aviation Administration, "Global Positioning System (GPS) Standard Positioning Service (SPS) Performance Analysis," Washington, DC, USA, 2014.

[16] C. Q. Gómez, M. A. Villegas, F. P. García and D. J. Pedregal, "Big data and web intelligence for condition monitoring: A case study on wind turbines," in *Handbook of research on trends and future directions in big data and web intelligence*, IGI global, 2015, pp. 149-163.

[17] A. O. Eggen, O. Rommetveit, A. Reitlo and E. O. Midtbø, "Handbook on condition monitoring of wind turbines," in *European Wind Energy Conference*, Marseille, France, 2009.

[18] S. Bagavathiappan, B. B. Lahiri, T. Saravanan, J. Philip and T. Jayakumar, "Infrared thermography for condition monitoring–A review," *Infrared Physics & Technology* , vol. 60, pp. 35-55, 2013.

[19] M. J. Neale and B. Woodley, "Condition Monitoring Methods and Economics," Bruel & Kjaer, 1978.

[20] M. Cocconcelli, L. Capelli, J. C. C. Molano and D. Borghi, "Development of a methodology for condition-based maintenance in a large-scale application field," *Machines,* vol. 6, no. 2, p. 17, 2018.

[21] V. J. Hodge, S. O'Keefe, M. Weeks and A. Moulds, "Wireless sensor networks for condition monitoring in the railway industry: A survey," *IEEE Trans. Intell. Transp. Syst.,* vol. 16, no. 3, p. 1088–1106, 2015.

[22] H. Tsunashima, "Railway Condition Monitoring, Present and Future," in *The Proceedings of the Transportation and Logistics Conference*, 2017.

[23] R. W. Ngigi, C. Pislaru, A. Ball and F. Gu, "Modern techniques for condition monitoring of railway vehicle dynamics," *Journal of physics: conference series,* vol. 364, no. 1, pp. 18-20, 2012.

[24] Office of Railroad Safety, "Track inspector rail defect reference manual," Federal Railroad Administration, 2015.

[25] Federal Railroad Administration, "Research Result- RR00-06," December 2000. [Online]. Available: https://railroads.dot.gov/sites/fra.dot.gov/files/fra_net/2159/rr00_06.pdf. [Accessed 18 July 2017].

[26] USDA Agricultral Marketing Service, "Grain Transportation Report," August 2018. [Online]. Available: https://www.ams.usda.gov/services/transportation-analysis/gtr. [Accessed 14 December 2019].

[27] J. Sadeghi, "Development of railway. track geometry indexes based on statistical distribution of geometry data.," *Journal of Transportation Engineering,* vol. 136, no. 8, pp. 693-700, 2010.

[28] The Canadian National Railway Company, "Industry Track Inspection," 19 September 2018. [Online]. Available: https://www.cn.ca/-/media/Files/Customer-Centre/Track-Specifications/industry-inspection-outline-en.pdf. [Accessed 18 July 2019].

[29] R. Bridgelall, P. Lu, D. Tolliver, N. Dhingra and B. Bhardwaj, "MPC-21-446/; Benefit Cost Analysis of Railroad Track Monitoring Using Sensors Onboard Revenue Service Trains," 12 December 2021. [Online]. Available: https://rosap.ntl.bts.gov/view/dot/61622. [Accessed 18 January 2022].

[30] H. Berger, "Nondestructive Testing of Railroad Rail," *Transportation Rsearch Record,* vol. 77, pp. 22-26, 1980.

[31] Q. Li, Z. Zhong, Z. Liang and Y. Liang, "Rail inspection meets big data: Methods and trends.," in *In 2015 18th International Conference on Network-Based Information Systems,* IEEE, 2015.

[32] S. T. Vipparthy, "Inspection of defects in rails using ultrasonic probe.," *The e-Journal of Nondestructive Testing,* vol. 18, no. 10, 2013.

[33] P. Lu, R. Bridgelall and R. Tolliver, "Rolling-stock Automatic In-situ Line Deterioration and Operating Condition Sensing," in *In 92nd Annual Meeting of the Transportation Research Board*, Washington, 2013.

[34] J. W. Park, T. G. Lee, I. C. Back, S. J. Park, J. M. Seo, W. J. Choi and S. G. Kwon, "Rail Surface Defect Detection and Analysis Using Multi-Channel Eddy," *Rail Surface Defect*

Detection and Analysis Using Multi-Channel Eddy Current Method Based Algorithm for Defect Evaluation., vol. 40, no. 3, pp. 1-12, 2021.

[35] R. H. Priewald, C. Magele, P. D. Ledger, N. R. Pearson and J. S. Mason, ""Fast magnetic flux leakage signal inversion for the reconstruction of arbitrary defect profiles in steel using finite elements.," *IEEE Transactions on Magnetics,* vol. 49, no. 1, pp. 506-516, 2012.

[36] L. Peng, S. Zheng, P. Li, Y. Wang and Q. Zhong, "A Comprehensive Detection System for Track Geometry Using Fused Vision and Inertia," *IEEE Transactions on Instrumentation and Measurement,* vol. 70, pp. 1-15, 2020.

[37] T. Predrag, S. Stanislav and M. Dick, "Analysis of vehicle/track interaction measurement data using the V/TI Monitor system.," *Građevinar,* vol. 70, no. 02, pp. 105-119, 2018.

[38] H. Tsunashima, H. Mori, M. Ogino and A. Asano, "Development of Track Condition Monitoring System Using On-board Sensing Device.," *Railway Research: Selected Topics on Development, Safety and Technology,* vol. 145, 2015.

[39] H. Mori, H. Tsunashima, T. Kojima, A. Matsumoto and T. Mizuma, "Condition monitoring of railway track using in-service vehicle.," *Journal of Mechanical Systems for Transportation and Logistics,* vol. 3, no. 1, pp. 154-165, 2010.

[40] M. Bocciolone, A. Caprioli, A. Cigada and A. Collina, "A measurement system for quick rail inspection and effective track maintenance strategy.," *Mechanical Systems and Signal Processing,* vol. 21, no. 3, pp. 1242-1254, 2007.

[41] X.-f. Wu, C. Chen, J.-j. Bu and G. Chen, "Sensor network architecture for intelligent high-speed train on-board monitoring.," *Journal of Zhejiang University-SCIENCE A,* vol. 12, no. 12, pp. 921-925, 2011.

[42] P. Weston, C. Roberts, Y. Graeme and E. Stewart, "Perspectives on railway track geometry condition monitoring from in-service railway vehicles.," *Vehicle System Dynamics,* vol. 53, no. 7, pp. 1063-1091, 2015.

[43] S.-S. Kim, C. Park, Y.-G. Kim and C. Park, "Parameter characteristics of rail inspection measurement system of HSR-350x.," *Journal of mechanical science and technology,* vol. 23, no. 4, pp. 1019-1022, 2009.

[44] Y. M. K. Santur and E. Akın, "Learning based experimental approach for condition monitoring using laser cameras in railway tracks.," *International Journal of Applied Mathematics Electronics and Computers,* Vols. Special Issue-1, pp. 1-5, 2016.

[45] C. Alippi, E. Casagrande, F. Scotti and V. Piuri, "Composite real-time image processing for railways track profile measurement.," *IEEE Transactions on instrumentation and measurement,* vol. 49, no. 3, pp. 559-564, 2000.

[46] Y. M. Karakose, O. Yamanand, K. Murat and E. Akin, "A new approach for condition monitoring and detection of rail components and rail track in railway.," *nternational Journal of Computational Intelligence Systems,* vol. 11, no. 1, pp. 830-845, 2018.

[47] R. Bridgelall, L. A. Chia, B. Bhardwaj, P. Lu, D. Tolliver and N. Dhingra, "Enhancement of signals from connected vehicles to detect roadway and railway anomalies.," *Measurement Science and Technology,* vol. 31, no. 3, p. 035105, 2019.

[48] C. Li, L. Shihui, C. Cole and M. Spiryagin, "An overview: modern techniques for railway vehicle on-board health monitoring systems.," *Vehicle system dynamics,* vol. 55, no. 7, pp. 1045-1070, 2017.

[49] N. Madenas and A. Davies, "Integrating PLM Systems with Maaintenance Information Across the Supply Chain," in *In Proceedings of the 12h International Conference on Manufacturing Research (ICMR2014),* 2014.

[50] Z. Kais, M. Sallak, W. Schon, S. Rangra and R. Sacile, "A Uml Approach For Modeling And Verifying Of Railway Signalling Systems Specifications," in *In MOSIM 2014, 10ème Conférence Francophone de Modélisation, Optimisation et Simulation,* 2014.

[51] P. Lu, R. Bridgelall, D. Tolliver, L. Chia and B. Bhardwaj, "MPC 19-384: Intelligent Transportation Systems Approach to Railroad Infrastructure Performance Evaluation: Track Surface Abnormality Identification with Smartphone-Based App," Mountain-Plains Consortium, Fargo, 2019.

[52] MonoDAQ, "$2 smartphone accelerometer vs. high-end IEPE vibration measurement," 15 July 2019. [Online]. Available: https://www.monodaq.com/applications/2-smartphone-accelerometer-vs-high-end-iepe-vibration-measurement/. [Accessed 18 May 2020].

[53] M. Liu, "A Study of Mobile Sensing Using Smartphones," *International Journal of Distributed Sensor Networks,* vol. 9, no. 3, p. 272916, 2013.

[54] M. Masoud, Y. Jaradat, A. Manasrah and I. Jannoud, "Sensors of smart devices in the internet of everything (IoE) era: big opportunities and massive doubts.," *Journal of Sensors 201,* 2019.

[55] K. A. Nguyen, Y. Wang, G. Li, Z. Luo and C. Watkins, "Realtime tracking of passengers on the london underground transport by matching smartphone accelerometer footprints.," *Sensors,* vol. 19, no. 19, p. 4184, 2019.

[56] O. Azeroual, "Data wrangling in database systems: purging of dirty data.," *Data,* vol. 5, no. 2, p. 50, 2020.

[57] T. Stobierski, "Data Wrangling: What it is & Why it's Important," Harvard Business School Online, 19 January 2021. [Online]. Available: https://online.hbs.edu/blog/post/data-wrangling. [Accessed 04 December 2021].

[58] Euro NCAP, "EUROPEAN NEW CAR ASSESSMENT PROGRAMME," Euro NCAP, 2012. [Online]. Available: https://www.euroncap.com/en. [Accessed 01 07 2022].

[59] Federal Aviation Administration, "Acceleration in Aviation: G-Force," 14 December 2021. [Online]. Available: http://www.faa.gov/pilots/safety/pilotsafetybrochures/. [Accessed 24 February 2022].

[60] R. Bridgelall, Y. Huang, Z. Zhang and F. Deng, "Precision enhancement of pavement roughness localization with connected vehicles.," *Measurement Science and Technology,* vol. 2, no. 025012, p. 27, 2016.

[61] B. W. Yap and C. H. Sim, "Comparisons of various types of normality tests.," *Journal of Statistical Computation and Simulation,* vol. 81, no. 12, pp. 2141-2155, 2011.

[62] F. Ali, P. Rawat and S. Malvia, "Comparative analysis and survey of LMS and RLS adaptive algorithms.," *International Journal of Computer Applications,* vol. 161, no. 3, pp. 26-29, 2017.

[63] E. A. Oni, I. D. Olatunde, K. O. Babatunde and D. O. Okpafi, "Improvement of Audio Signal Quality Using Adaptive Filtering and It's Performance Advantage over Non - Adaptive (Linear) Filtering," *International Journal of Electrical and Electronic Science,* vol. 5, no. 2, pp. 39-45, 2018.

[64] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series.," *Mathematics of computation,* vol. 19, no. 90, pp. 297-301, 1965.

[65] S. W. Smith, "Statistics, Probability and Noise," in *The Scientist and Engineer's Guide to Digital Signal Processing,* San Diego, CA: California Technical Publishing, 1997, pp. 11-24.

[66] E. Punskaya, "Design of FIR Filters," 29 January 2010. [Online]. Available: https://www.vyssotski.ch/BasicsOfInstrumentation/SpikeSorting/Design_of_FIR_Filters.pdf. [Accessed 23 May 2019].

[67] I. . I.-Y. Choi, U. Ju-Hwan, J. S. Lee and H.-H. Choi, "The influence of track irregularities on the running behavior of high-speed trains.," *Proceedings of the Institution of*

Mechanical Engineers, Part F: Journal of rail and rapid transit, vol. 227, no. 1, pp. 94-102, 2013.

[68] L. Al-Nazer, T. Raslear, R. Wilson, J. Kidd and Q. N. America, "Quantification of the Sensitivity of Two Prevalent Track Inspection Systems.," No. DOT/FRA/ORD-17/03. United States. Federal Railroad Administration. Office of Research, Development, and Technology, 01 May 2017. [Online]. Available: https://rosap.ntl.bts.gov/view/dot/37340. [Accessed 12 March 2020].

[69] M. France, G. Melnik, H. Safar and J. Multer, "Reducing Hazards Associated with Visual and Automation-Aided Track Inspections.," No. DOT/FRA/ORD-21/18. United States. Federal Railroad Administration. Office of Railroad Policy and Development, 01 May 2021. [Online]. Available: https://rosap.ntl.bts.gov/view/dot/55878. [Accessed 11 June 2022].

[70] M. A. Rossetti, "Analysis of weather events on US Railroads.," 13 Jauary 2007. [Online]. Available: https://rosap.ntl.bts.gov/view/dot/9745. [Accessed 20 March 2021].

APPENDIX. ROAD IMPACT FACTOR (RIF) CALCULATION

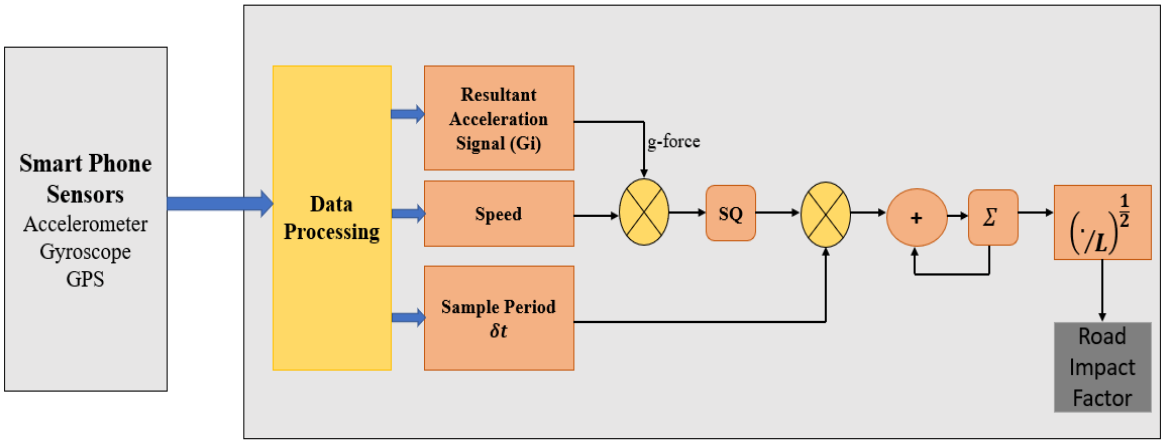


Figure A1. Road Impact Factor (RIF) computation flow chart