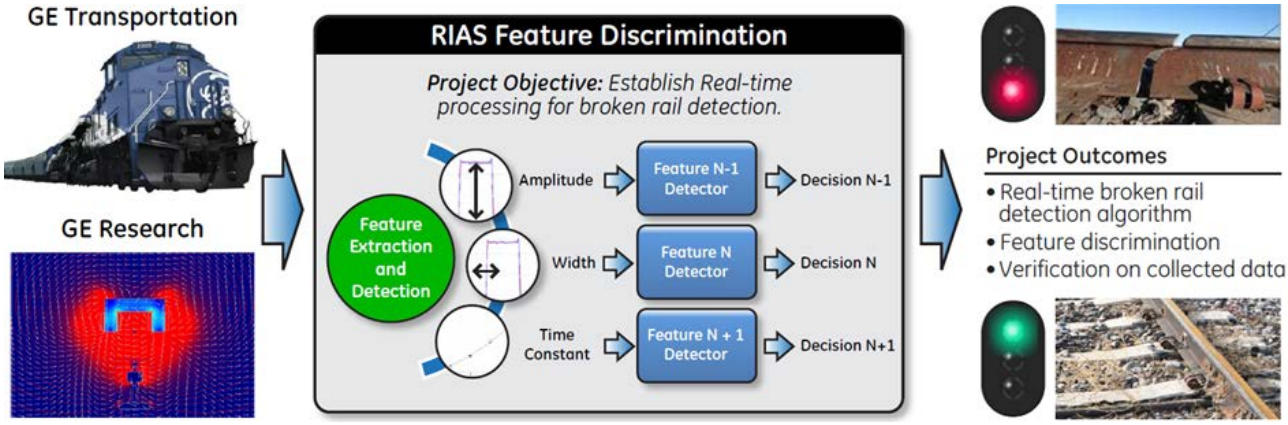
U.S. Department of Transportation

Federal Railroad Administration

Rail Integrity Alert System (RIAS) Feature Discrimination

Office of Research, Development, and Technology
Washington, DC 20590



.

DOT/FRA/ORD-16/33

Final Report
August 2016

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof. Any opinions, findings and conclusions, or recommendations expressed in this material do not necessarily reflect the views or policies of the United States Government, nor does mention of trade names, commercial products, or organizations imply endorsement by the United States Government. The United States Government assumes no liability for the content or use of the material contained in this document.

NOTICE

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the objective of this report.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 2016	Technical Report - September 2015

4. TITLE AND SUBTITLE Rail Integrity Alert System (RIAS) Feature Discrimination	5. FUNDING NUMBERS
6. AUTHOR(S) Y. Plotnikov, B. Matthews, G. Boverman, M. Nayeri, T. Frangieh	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) GE Global Research One Research Circle Niskayuna, NY 12309	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Department of Transportation Federal Railroad Administration Office of Railroad Policy and Development Office of Research and Development Washington, DC 20590	10. SPONSORING/MONITORING AGENCY REPORT NUMBER DOT/FRA/ORD-16/33
11. SUPPLEMENTARY NOTES COR: Jared Withers	
12a. DISTRIBUTION/AVAILABILITY STATEMENT This document is available to the public through the FRA Web site at http://www.fra.dot.gov.	12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This report describes GE Global Research's research, in partnership with GE Transportation, into developing and deploying algorithms for a locomotive-based inductive sensing system that has a very high probability of detecting broken rails with very few false-positives. These algorithms are described in detail in this report and their performance has been characterized by performing extensive field-testing on the data collected from a previously developed Rail Integrity Alert System (RIAS).

14. SUBJECT TERMS Broken rail, locomotive, inductive coils, probability of detection, probability of false alarm, support vector machine, receiver operating characteristics	15. NUMBER OF PAGES 69
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC

LENGTH (APPROXIMATE)

1 inch (in)	=	2.5 centimeters (cm)
1 foot (ft)	=	30 centimeters (cm)
1 yard (yd)	=	0.9 meter (m)
1 mile (mi)	=	1.6 kilometers (km)

AREA (APPROXIMATE)

1 square inch (sq in, in^2)	=	6.5 square centimeters (cm^2)
1 square foot (sq ft, ft^2)	=	0.09 square meter (m^2)
1 square yard (sq yd, yd^2)	=	0.8 square meter (m^2)
1 square mile (sq mi, mi^2)	=	2.6 square kilometers (km^2)
1 acre = 0.4 hectare (he)	=	4,000 square meters (m^2)

MASS - WEIGHT (APPROXIMATE)

1 ounce (oz)	=	28 grams (gm)
1 pound (lb)	=	0.45 kilogram (kg)
1 short ton = 2,000 pounds (lb)	=	0.9 tonne (t)

VOLUME (APPROXIMATE)

1 teaspoon (tsp)	=	5 milliliters (ml)
1 tablespoon (tbsp)	=	15 milliliters (ml)
1 fluid ounce (fl oz)	=	30 milliliters (ml)
1 cup (c)	=	0.24 liter (l)
1 pint (pt)	=	0.47 liter (l)
1 quart (qt)	=	0.96 liter (l)
1 gallon (gal)	=	3.8 liters (l)
1 cubic foot (cu ft, ft^3)	=	0.03 cubic meter (m^3)
1 cubic yard (cu yd, yd^3)	=	0.76 cubic meter (m^3)

TEMPERATURE (EXACT)

[(x-32)(5/9)] °F = y °C

METRIC TO ENGLISH

LENGTH (APPROXIMATE)

1 millimeter (mm)	=	0.04 inch (in)
1 centimeter (cm)	=	0.4 inch (in)
1 meter (m)	=	3.3 feet (ft)
1 meter (m)	=	1.1 yards (yd)
1 kilometer (km)	=	0.6 mile (mi)

AREA (APPROXIMATE)

1 square centimeter (cm^2)	=	0.16 square inch (sq in, in^2)
1 square meter (m^2)	=	1.2 square yards (sq yd, yd^2)
1 square kilometer (km^2)	=	0.4 square mile (sq mi, mi^2)
10,000 square meters (m^2)	=	1 hectare (ha) = 2.5 acres

MASS - WEIGHT (APPROXIMATE)

1 gram (gm)	=	0.036 ounce (oz)
1 kilogram (kg)	=	2.2 pounds (lb)
1 tonne (t)	=	1,000 kilograms (kg)
	=	1.1 short tons

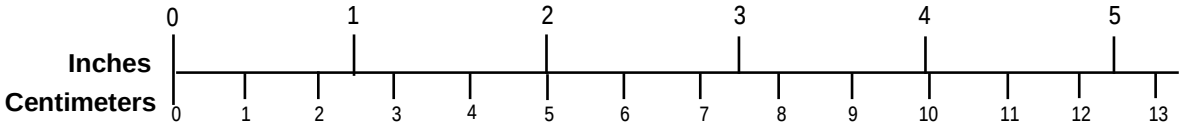
VOLUME (APPROXIMATE)

1 milliliter (ml)	=	0.03 fluid ounce (fl oz)
1 liter (l)	=	2.1 pints (pt)
1 liter (l)	=	1.06 quarts (qt)
1 liter (l)	=	0.26 gallon (gal)
1 cubic meter (m^3)	=	36 cubic feet (cu ft, ft^3)
1 cubic meter (m^3)	=	1.3 cubic yards (cu yd, yd^3)

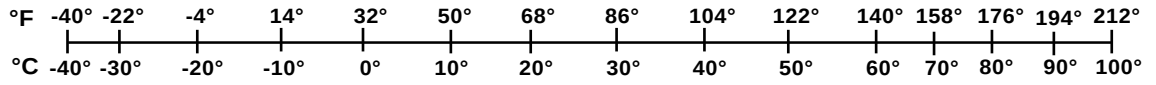
TEMPERATURE (EXACT)

[(9/5) y + 32] °C = x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSIO



For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 6/17/98

Acknowledgements

GE Transportation was responsible for establishing the Rail Integrity Alert System (RIAS) technical concept then developing and evaluating the system prototypes. The authors acknowledge the following individuals who organized the RIAS locomotive tests at GE, led the work on prototype development, and navigated and reviewed this program: Joseph Noffsinger, Ajith Kumar, and Jeffrey Fries of GE Transportation.

Contents

Acknowledgements iii

Illustrations v

Tables viii

Executive Summary 1

1. Introduction 2
   1.1 Background 2
   1.2 Objectives 3
   1.3 Overall Approach 3
   1.4 Scope 4
   1.5 Organization of the Report 4

2. RIAS Feasibility Study 5
   2.1 Laboratory Tests Summary 6
   2.2 Field Tests Summary 9
   2.3 Summary of Results 16

3. Establishing Identification of the Broken Rail Signal Produced by the RIAS 17
   3.1 Data collection system 17
   3.2 Data Format 19
   3.3 Railroad features covered by the collected data 20
   3.4 Structure and Parameters of the Designed Matched Filter 22
   3.5 Algorithm Performance, Test Results and Analysis 27
   3.6 Summary of the Results 32

4. Identifying Signal Patterns from the Track Features 33
   4.1 Locomotive Noise Study 33
   4.2 Signal Signatures of the Track Features 37
   4.3 Establishing and Testing a Track Feature Recognition Algorithm 39
   4.4 Summary of the Results 47

5. Algorithm Optimization 49
   5.1 Outline of Data Processing 49
   5.2 False alarm reduction and algorithm verification 50
   5.3 Signal Normalization 52
   5.4 Summary of the Results 56

6. Conclusion 57

7. References 58

Abbreviations and Acronyms 59

Illustrations

Figure 1. Timeline diagram for GE RIAS 4

Figure 2: Concept of the electromagnetic linkage to the rail. The green dashed line illustrates a connected axle-rail loop. 6

Figure 3: 1:15 scaled model of the electrical circuit of a locomotive 6

Figure 4: Voltage from the receive coil as a function of its position over the model 7

Figure 5: The 1:3 scaled down model with actual transmit and receive coils 8

Figure 6: Impedance plane scatter plot of the receive voltage from the 1:3 scaled model 9

Figure 7: The transmitting coil mounted between fuel tank and wheels. 10

Figure 8: Installation of a wire shunt across the insulated joint 10

Figure 9: Time plots of the RIAS signals: the insulated joint shunted three times 11

Figure 10: Joint GE Transportation and GE Research team in the cabin of a DC locomotive observing the RIAS performance 11

Figure 11: The RIAS data collected at 10 mph on the Erie test track 12

Figure 12: The RIAS data collected during locomotive motoring 12

Figure 13: RIAS signals at 5 mph linked to the rail features on the ground 13

Figure 14: High tonnage train at TTC, Pueblo, CO 14

Figure 15: The GE RIAS installed on an idling locomotive attached to the high tonnage train .. 14

Figure 16: Details and images of the three broken rail occurrences at the TTC tests 15

Figure 17: Comparison of the RIAS signal magnitudes recorded during the passing over break #1 (blue colors) and the signal recorded during the last pass before the break (red colors) 16

Figure 18: Working diagram and locations of the transmitting and receiving coils in the RIAS prototype. 17

Figure 19: RIAS prototype system diagram that was used for data collection in January - March of 2014 18

Figure 20: The RIAS prototype deployed on the NS inspection train 19

Figure 21: Locations of the railroad features on the HTL in TTC listed in Table 3 21

Figure 22: Block diagram of matched filter processing 23

Figure 23: Block diagram of direct demodulation processing. Shown for one excitation frequency; same structure for the second and subsequent excitation frequencies 24

Figure 24: Raw measurements made at Detector 1. (a) Noise-free data; (b) Noise-corrupted data 25

Figure 25: Quantitative comparison of frame size statistics. (a) Non-break mean vs. frame size. (b) Break mean vs. frame size 26

Figure 26: Comparison of class separability vs. frame size for various signal detection techniques 27

Figure 27: Matched filter amplitudes for 81 seconds of data including 3 insulated joints. Red, dashed lines are GPS-based occurrence times for insulated joints 28

Figure 28. Block diagram for detecting broken rails based on matched-filter amplitudes 29

Figure 29. Plot of "voltage drop" confidence score, i.e., median amplitude minus amplitude for IJs 1, 2, and 3, as listed in Table 2. Hand-labeled start and stop times of IJs in red, dashed rectangles 30

Figure 30. Receiver Operating Characteristic (ROC) curve for detecting the break state using the voltage drop score 31

Figure 31: RIAS signal after demodulation at two frequencies showing electromagnetic noise from the locomotive 34

Figure 32: Test data set for SNR-estimation example. (a) Matched-filter outputs. (b) Locomotive speed 35

Figure 33: Spectrum of test data set including excitation waveforms (a) and with components in the signal subspace removed (b) 35

Figure 34. Simulated noise based on autoregressive model (a) and SNR (b) 36

Figure 35. Data during dynamic braking DB8 37

Figure 36. Representative plots of matched filter amplitudes for an insulated joint (upper left, "Interlocking Signal") and 5 common track artifact types 39

Figure 37. System diagram for break detection based on thresholding (upper) and pattern recognition (lower) approaches 40

Figure 38. System signals from Break #1 41

Figure 39. Matched filter amplitude segment for a correctly detected insulated joint. The "Break State" is identified with a green rectangle. Signal mean and standard deviation before, during and after the break state are indicated with diamond markers and error bars 42

Figure 40. Similar to Figure 38, matched filter amplitudes, break state, and basic statistics of the signal are plotted for a false positive detection 43

Figure 41. First and second-order statistics for break state means as a feature vector for 4 manually labeled data classes from the CL1 data set 43

Figure 42. ROC curves for detecting IJs with SVM and GMM classifiers 46

Figure 43. ROC curves for detecting insulated single joints and joints with non-standard stagger widths using SVM and GMM classifiers 47

Figure 44. Block-diagram of procedure for classifier training and classification 49

Figure 45. SVM scores for a typical test data set. (a) True insulating joints (b) Noise, lubricators, etc 50

Figure 46. Minimum missed detection rate when the false-positive rate was zero for the seven train-test configurations 51

Figure 47. Closer look at the minimum miss rate for a false-positive rate of zero for the seven conditions (a) and the threshold achieving a minimal miss rate and zero false-positive rate (b) 52

Figure 48. Means with one-STDEV error bars for target and anti-target classes on CL1 and HTL data sets 53

Figure 49. An insulated joints from the CL1 database, unnormalized (upper panel) and after applying mean- and median-based normalization (middle and lower panels, respectively) 54

Figure 50. Means and one-STDEV errorbars for CL1 and HTL data sets after feature normalization with and norms 55

Tables

Table 1: Description of the Recorded Signals 19
Table 2: Railroad features on the TTC High Tonnage Loop 20
Table 3: Listing of manually labeled railroad artifacts for the CL1 and HTL databases 21
Table 4: Error trade-off operating points for break state detection with “voltage drop” score 32
Table 5. Breakdown of broad track feature types in the CL1 data set..... 38
Table 6. Training and test set breakdown for CL1 experiments..... 45
Table 7. “Area above the curve” for detecting non-standard IJs in ROC plots in Figure 42 47

Executive Summary

This report describes GE Global Research's research into the development and deployment of algorithms for a locomotive-based inductive sensing system capable of warning of broken rails with very high probability of detection and very few false-positives. The research into algorithms, which was conducted in partnership with GE Transportation and funded by the Federal Railroad Administration, is described in detail in this report, and the algorithms' performance are characterized by extensive field-testing on the data collected from a previously developed Rail Integrity Alert System (RIAS).

Broken rails, particularly in sections of running track not controlled by signals (i.e."dark territory"), represent a significant safety hazard and are the most common cause of reportable derailments. Unfortunately, dedicated inspection systems are expensive to operate with sufficient frequency and it would be cost-prohibitive to install fixed track circuits throughout all of the dark territory regions in the United States (and beyond).

Many methods for detecting broken rail have been proposed. However, to date, these methods cannot be widely adopted by the commercial rail industry because none of them have a sufficient level of detection accuracy. However, a novel on-locomotive electromagnetic monitoring system was introduced by GE and several prototypes were designed and field tested.

Previously in early 2014, GE Transportation used the resources of Norfolk Southern Corp. (NS) and Transportation Technology Center, Inc. (TTCI) to record system data from Class 1 railroads in Virginia and Georgia as well as from locomotives on a high tonnage loop in Colorado. The collected data, in the form of high-speed digitized system signals, were available for this project.

Lastly, the report describes how to maximize generalization performance by optimizing the hyper-parameters of the classifier and reduce false-alarms while maintaining a low rate of missed detections by selecting a particular point on the receiver operating characteristics (ROC) curve. For the collected field data, we were able to achieve a false-alarm rate of essentially zero with a missed detection probability of less than 5 percent.

GE plans pilot deployments of the broken rail detection system in 2016.

1. Introduction

The Rail Integrity Alert System (RIAS) performs continuous rail integrity testing from a mobile platform (i.e., a locomotive at revenue speeds). To evaluate the electrical conductivity of the rail, multi-frequency test signals are transmitted from the mobile platform into the rails via non-contact inductive coils; then they are received from the rail at another point and analyzed.

The RIAS does not require any wayside infrastructure for its normal operation. The system was designed to prevent derailments in most locomotives and rail lines, and mitigate derailment damage and casualties from through breaks in the rail caused by transverse defects or detail fractures. The system is able to test the rail section between the trucks (bogies) of locomotives and detect rail breaks deterministically, while not being subject to reliability issues such as ballast conductivity problems or shorted track circuits.

1.1 Background

In a recent study (Schafer, 2008), it was shown that broken rails caused 335 mainline derailments on Class 1 freight railroads from 2003-2006, with a direct economic impact of $176 million in equipment and track damage, and potentially larger impacts due to delays. Another study that examined the causes of derailments in detail found that, when severity was measured in terms of total derailed cars, broken rails were the leading cause of severe accidents (Dick, Barkan, Chapman, & Stehly, 2003). According to the FRA data (Railroad Safety Statistics), there were 213 derailments due to track defects in in 2013, with $56 million in reportable damages. It should be noted that, in many cases, derailment cost estimates reported within FRA guidelines may be significantly lower than the true economic impact.

GE has conducted several comprehensive studies on different broken rail detection methods, including impedance with an auxiliary conductor, transmitted acoustics, ultrasound and electromagnetic inspection systems, attached optical fiber, attached transmission line, video surveillance, and others (Weir, Hedeen, & Welles, 1996), (Bonanni, Van Stralen, Davenport, Barshinger, & Cheng, 2004). Some of these methods can be used from mobile platforms. However, practicality (i.e., cost, area coverage) and physical limitations (e.g. speed of inspection, mechanical coupling requirements, weather and day or night light conditions) prevent many investigated methods from being used in dark territory and these methods might interfere with existing wayside circuits.

If the broken rail detection systems were incorporated into the moving locomotives themselves, the solution would be of great value to the railroad industry. GE envisions a system that can detect broken rails in dark territory lacking both track circuits and train occupancy circuits, as well as regions of the track where signaling circuitry is present.

These investigations demonstrate that the RIAS technology is less vulnerable to weather conditions, does not require a mechanical coupling with rail, and it can perform detection tasks with locomotive and car fleets at maximum authorized freight speeds. The RIAS, equipped with real-time railroad feature detection and recognition, could be a critical component in safety improvements for the whole railroad industry.

1.2 Objectives

The project's objective was to develop an algorithm for detecting broken rails in real-time for on-locomotive alerts. Specifically, this FRA funded study focused on optimizing the break detection performance of the RIAS by maximizing the probability of detection while minimizing the false-alarm rate.

1.3 Overall Approach

The ability of the RIAS to detect broken rails has been established by a series of laboratory tests on scale models, which were followed by field tests using system prototypes developed by GE Transportation. The field collected data was used to establish the broken rail detection procedure, and the algorithm has been verified with separate and mixed data sets obtained from several field tests. The real-time implementation has been tested on the system prototype in laboratory conditions.

1.3.1 Technical Approach

As an RIAS prototype underwent preliminary testing, it became clear that the system must prevent false alarms by conclusively identifying certain track features that could otherwise be interpreted as a broken rail (certain switch layouts, insulated joints, etc.). While alarms could be suppressed using the GPS position of the feature and a database, analyzing the captured signals to identify the track feature conclusively would be the preferred approach, and the latter approach would provide maximum test coverage.

During this project, we used realistic rail data to design and implement machine learning algorithms to maximize the probability of detection and minimize the false-alarm rate. To demonstrate the ability of RIAS to perform broken rail detection, a laboratory evaluation of the physical principles was performed and the data collected from the railroad, including actual occurrences of rail breaks, was analyzed.

1.3.2 Modeling

In this project, scaled down electromagnetic models of the rails and locomotive frame (i.e., wheels and axles) were used to study non-contact signal transmitting and receiving, and establish limits of the electrical current injected by the system into the rail. Some artificially-generated signals were used to establish matching filter and investigate the performance of the broken rail detection algorithm in an early stage of this project.

1.3.3 Test

A large amount of previously collected RIAS data was available for this project. This unique data collection effort was coordinated and led by GE Transportation, in collaboration with Norfolk Southern Corp. (NS) and Transportation Technology Center, Inc. (TTCI). The data covered a range of different carrier platforms, such as: pulling and pushing AC and DC locomotives, an inspection train, and an idling locomotive attached to the high tonnage train. A new set of data that included extreme cases of locomotive noise interference became available to the project in November of 2014 as GE Transportation continued product development.

1.3.4 Analysis

Initially, matching filters were applied to the RIAS signals and it was determined that rail breaks can be reliably simulated by insulated joints in track. To improve detection reliability, a processing phase based on machine learning algorithms was evaluated. The performance of alternative approaches was compared by the means of analyzing Receiver Operating Characteristic (ROC) curves.

1.4 Scope

This project focused on demonstrating the feasibility of signal processing solutions for broken rail detection then developing and optimizing them. A through comprehensive analysis of the database of the recorded RIAS waveforms, which includes more than 120 GB of the digitized waveforms collected from several input channels at high resolutions and digitizing rates, was performed.

The processing algorithms will be implemented on an onboard processing platform capable of real-time decision making and alerts and the algorithms include detection and feature discrimination. The algorithm has been verified with the data collected from the GE RIAS.

The timeline diagram below illustrates the major development steps for the overall RIAS development program. The green color block illustrates the time frame and interactions of the FRA funded project with the overall technology development process.

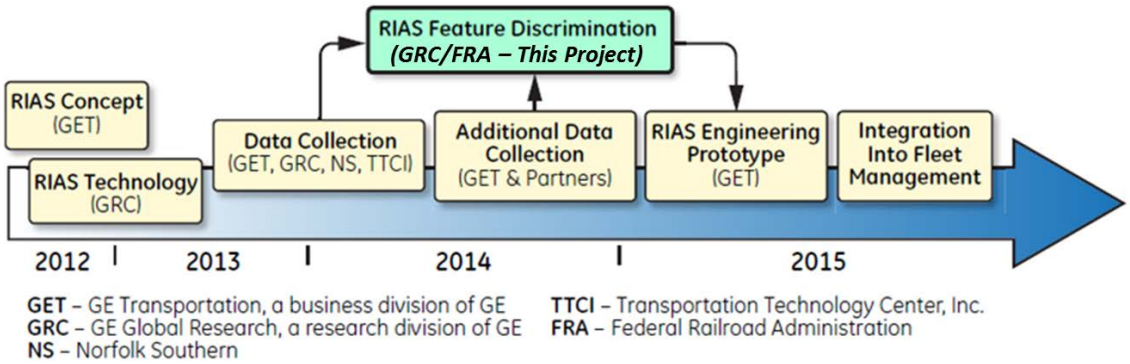


Figure 1. Timeline diagram for GE RIAS

1.5 Organization of the Report

The report includes a study of the physical principles used in the RIAS and whether the system is feasible in light of those principles. The report discusses a broken rail detection method and describes the algorithms that improved this method's reliability.

2. RIAS Feasibility Study

In this section, we demonstrate that the project's technical approach can detect a broken rail in real time from a locomotive frame. The RIAS should be able to detect broken rails using a sensor platform that is attached to the chassis of locomotives that are in the field, in operation, and moving at full speed.

These high level requirements were mandated by GE Transportation for the RIAS development and implementation

- Real-time broken rail detection
- No (a minimum number of) false alarm disruptions
- 180+ days between cleaning/calibration/service
- 79+ mph capable
- No fleet management impact
- Operational in both signaled territory and dark territory
- No wayside infrastructure required for detection
- Not dependent on real-time wireless connectivity
- Lower cost than track circuiting in dark territory
- Non-contact transmit-receive arrangement (>6" above the rail)
- Investigate additional functionality and use of the sensors

A RIAS concept has been developed that is based on electromagnetic principles. It establishes a constantly moving current loop from the car wheels. locomotive wheels, and axles (see schematic in Figure 2). In this approach, a transmitting (Tx) coil induces alternating current in the section of the track shown in green and a receiving (Rx) coil inductively measures this current. The loop between the two tracks is closed through the wheels and axles 2 and 3 of the train (or axles 3 and 4 in case of a 6-axle platform). Wheels and axles, under the pressure from the locomotive's weight, make good electrical contacts with the rail surface and as a result, several electrical contours are formed. The largest contour is made by the wheels and axles 2 and 3. A broken rail will interrupt the flow of current in the loop and a very small response will be measured by the receiving coil.

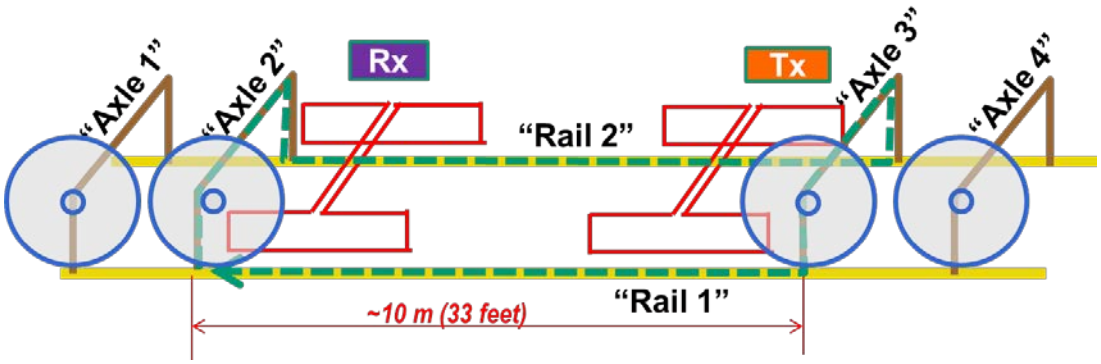


Figure 2. Concept of the electromagnetic linkage to the rail. The green dashed line illustrates a connected axle-rail loop.

An actual implementation will use multiple transmit and receive coils that operate on the same region of track using frequency-division multiplexing. In the RIAS prototype, the electromagnetic sensors continuously probe the audio-frequency (3-5 kHz) electrical properties of the track by inducing a relatively small (~100 mA) current in electrically connected loops of track.

2.1 Laboratory Tests Summary

A scaled down (1:15) model of the electrical circuit loops (using rails, wheels, and axles) was assembled and evaluated under laboratory conditions. It consisted of ½" steel bars that were connected together to imitate both rails and locomotive wheels. A picture of the laboratory setup is in Figure 3.

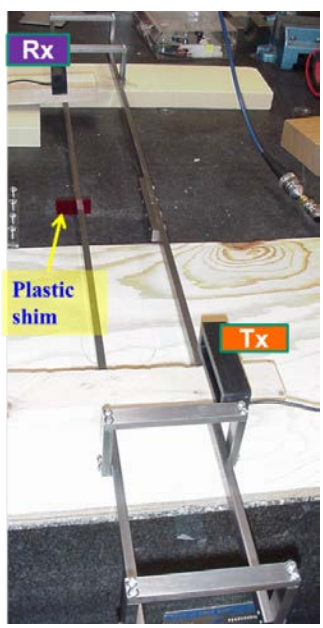


Figure 3. 1:15 scaled model of the electrical circuit of a locomotive

Two rectangular multi-turn coils injected an electrical current into the model to sense its strength. Cases of “good rail” and “broken rail” were investigated. To ensure a complete electrical disconnect in the middle portion, a plastic shim was inserted in the “break” point as shown in the figure above.

Coils of different shapes and sizes were tried on this model. Rectangular (80 mm x 30 mm) coils with the long side oriented parallel to the “rail” portion provided a signal capable of discriminating between a good and a broken connection.

To investigate the variability in the magnetic field around the physical model, a smaller (25 mm x 15 mm) coil with a ferrite C-core was used as a receiver. Two-dimensional maps (Figure 4) were constructed by mechanically moving the receiving coil over the model in 10 mm increments. The transmit coil (Tx) was excited by a 7 kHz alternating current and, to allow uninterrupted scan over the frame, it was at the same position under the bottom of the model frame.

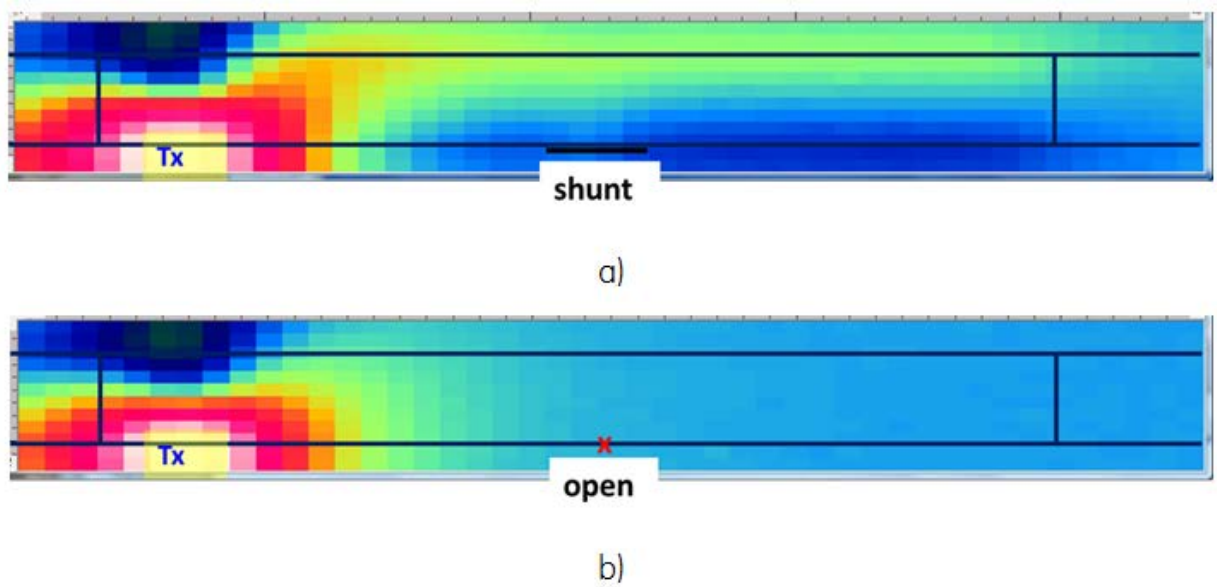


Figure 4. Voltage from the receive coil as a function of its position over the model

The receiving coil had the open ends of the C-core pointing downward or perpendicular to the scan increments (or the direction perpendicular to the map plane in Figure 4). As a first approximation, these maps can be seen as a density distribution of the vertical (z-component) of the magnetic flux generated by the rail-wheel-axle model. The highest field intensity exists around the transmit coil (left hand side of the maps) for both cases, and in the case of the open (broken) loop, the magnetic field is very uniform in areas away from the transmit coil.

We analyzed these charts to find the most sensitive areas that could be used to place transmit and receive devices on a real locomotive. In this scaled model, there was no preferred spot over the “rail” that produced a better break detection. Similar tests were conducted for variable locations of the transmit coil, and the results also suggested that there is no preferred position for the transmit coil.

The next phase of laboratory investigation was conducted on a more realistic track model for testing full scale components in laboratory conditions. This physical model simulates the rail-wheel-axle contour at the scale of 1:3, and it was assembled from steel pipes 0.75 inches in diameter. This model provides more realistic impedance parameters and possible trends of signal changes. Figure 5 shows this model together with full scale transmitting (at the front) and receiving (far right corner) coils. The middle section of one pipe had a "break" terminated with a bank of resistors. A laminated magnetic C-core with electrical coils was used for current injection into the model. A CAB receiver coil connected to an amplifier was used to monitor the level of the electrical current in the pipe loop. Both transmitting and receiving coils were positioned 6 inches above the pipe loop.

Figure 5. 1:3 scaled down model with actual transmit and receive coils

The resistance of the "break" was varied in the range of 0 to 5 Ohms, to a complete disconnect. Both magnitude and phase of the voltage response were measured in a frequency range from 1 to 10 kHz. Values of the receive voltage components as a function of the break resistance are plotted in Figure 6.

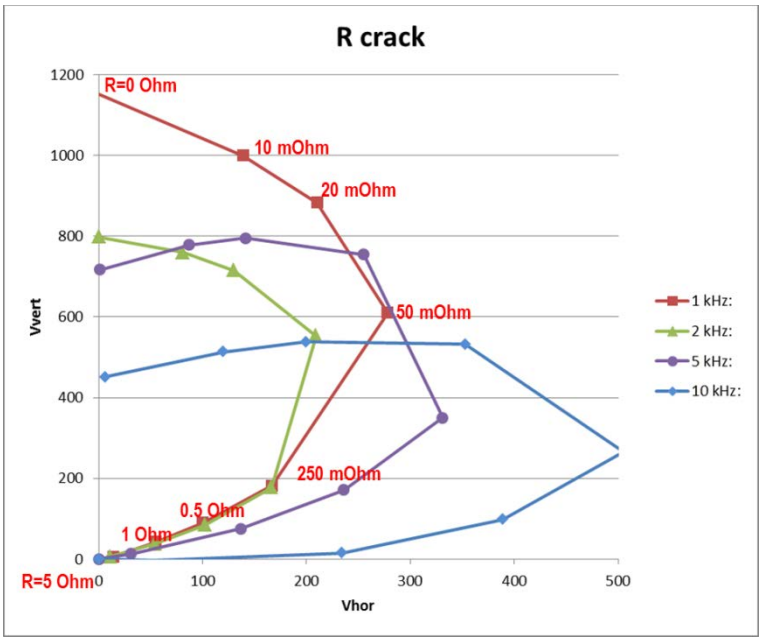


Figure 6. Impedance plane scatter plot of the receive voltage from the 1:3 scaled model

As can be seen from this chart, the largest voltage changes occurred at lower values of the break resistance (0 to 100 mOhm) and it is difficult to distinguish the signal changes if resistance is higher than 5 Ohms. The highest magnitude change was obtained at 1-2 kHz and signal was decreasing with higher frequencies. Also, the shape of these curves was different with the largest change in the real component measured at 10 kHz.

These were very encouraging results that provided solid support to the possibility that rail break detection can be achieved with a non-contact inductive linkage between the inspection system and a rail. A rail segment with the resistance on the order of 1 to 5 Ohms can be detected with this approach and interpreted as a break.

2.2 Field Tests Summary

The results of the laboratory tests on the small scaled model proved that the proposed technical approach was ready to become a full scale system that can detect electrical discontinuities in the rails. Though the laboratory tests may be able to test the physical principles of the electromagnetic break detection system, they are not be able to evaluate risks associated with the real scale tracks, real locomotives, outside weather, and other conditions that are difficult to adequately model in a lab.

To effectively progress with the risk mitigation plan, in 2013 a joint team of engineers from GE Transportation and GE Global Research developed and evaluated several prototypes of the RIAS. Details of the system prototype used during the field tests are given in Section 3.1.

During RIAS development, the team conducted an examination to determine if RAIS would fit into a typical locomotive. A preliminary investigation was conducted to determine possible locations for the transmitting and receiving hardware. The areas in front and behind the fuel tank were identified as possible locations. The transmitting and receiving coils and holding brackets were designed and built by the GE Transportation engineering team. The receiving and

transmitting coils were about 7” above the rails. The position of a transmitting coil on the right hand side of a locomotive is shown in 7.

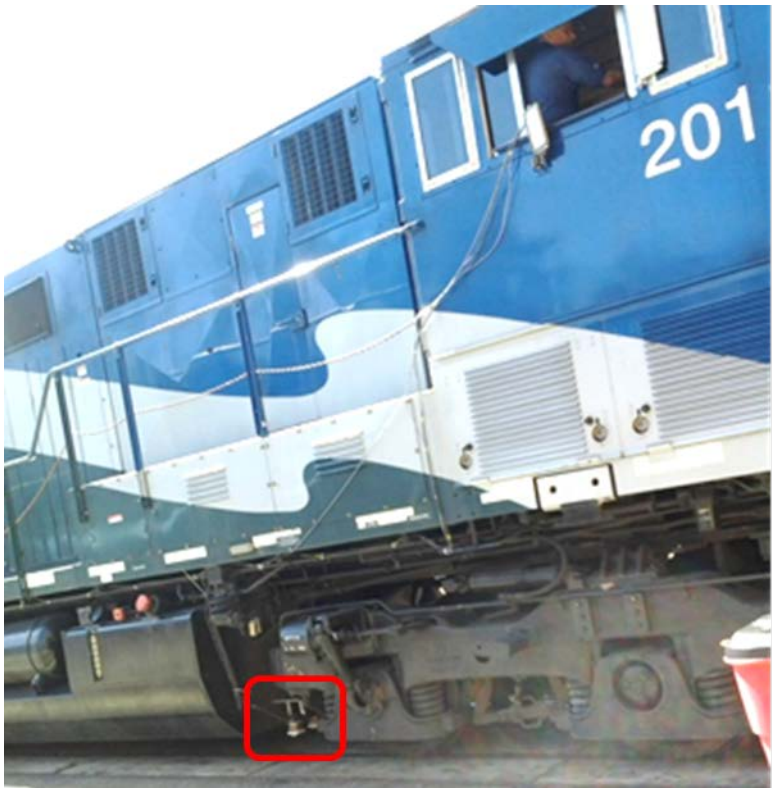


Figure 7. The transmitting coil mounted between fuel tank and wheels

Initial tests were conducted in static regime. To imitate a break in the rail, an insulated joint was used (Figure 8).

Figure 8. Installation of a wire shunt across the insulated joint

The joint was shunted with a cable, which represented a “normal” rail and the cable would be disconnected to simulate a “break”. The locomotive was placed such that the insulated joint was under the fuel tank, between axles 3 and 4. The shunting cable over the insulated joint was disconnected. Then, for a short period of time, the insulated joint was shunted three times. The RIAS signal increased compared to the open connection case as shown in Figure 9.

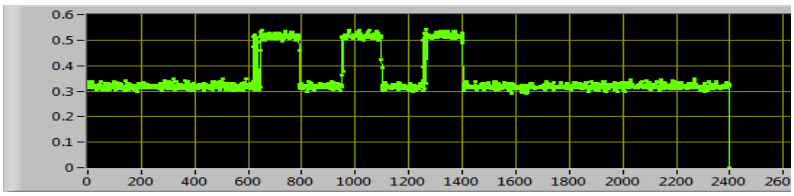


Figure 9. Time plots of the RIAS signals: the insulated joint shunted three times

The signals shown in Figure 9 clearly identify the area of connected and disconnected insulated joints. These results were obtained with the RIAS prototype installed on a locomotive with transmitter and receiver coils that were more than 6” above the rails.

Additional measurements were conducted with variable resistance across the insulated joint at different excitation frequencies. The results were similar to the behavior that was observed from the laboratory model (6).

Several adjustments were made to the RIAS prototype to make it functional, detect insulated joints, and record data for further analysis. In June and July of 2013, several test runs were conducted on the GE Transportation test track in Erie, PA, with DC and AC locomotives (Figure 10).

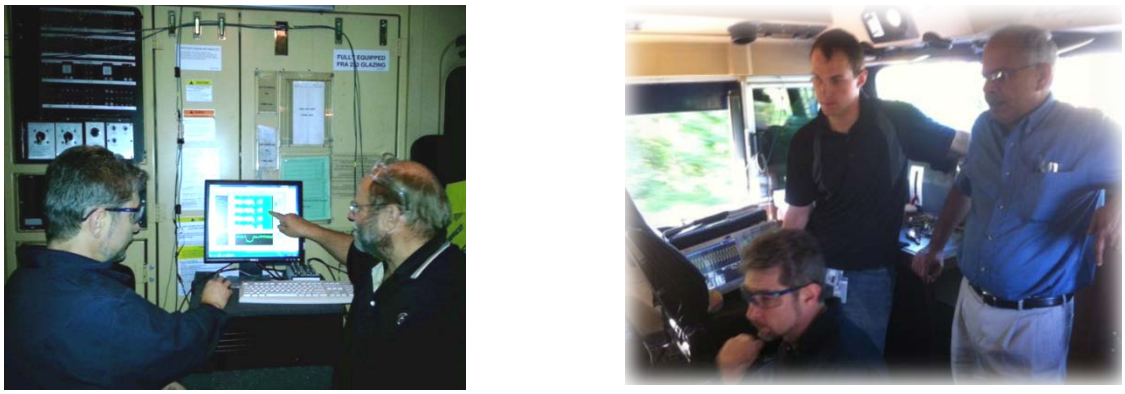


Figure 10. Joint GE Transportation and GE Research team in the cabin of a DC locomotive observing the RIAS performance

An example of RIAS data, which was collected while the locomotive traveled West at 10 mph, is shown in Figure 11. A shunt and an insulated joint can be distinguished in this chart.

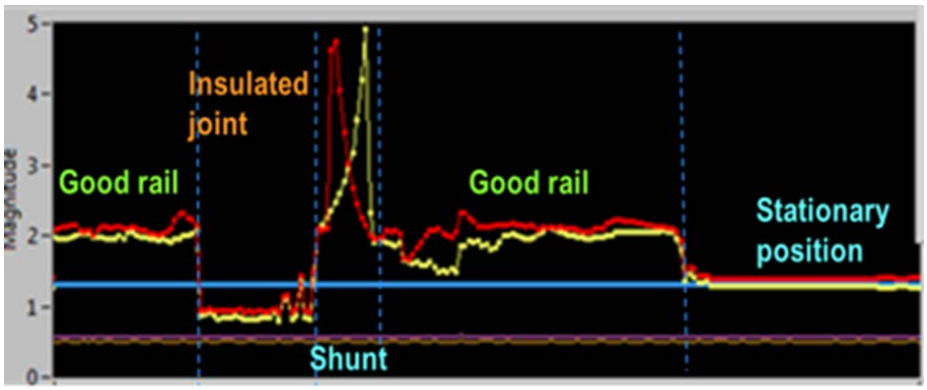


Figure 11. The RIAS data collected at 10 mph on the Erie test track

In this test, the locomotive was moving at low constant speed and the motoring noise was low. Another example of the RIAS signal is shown in Figure 12. This time the locomotive's motors were running on full power and produced significant electromagnetic noise, which was picked up by the RIAS receiving circuits. The full signal waveforms are plotted in vertical direction with 5 ms intervals in the horizontal direction. The signal change, which occurred while the locomotive was passing over an insulated joint, is still visible despite the strong presence of the background noise.

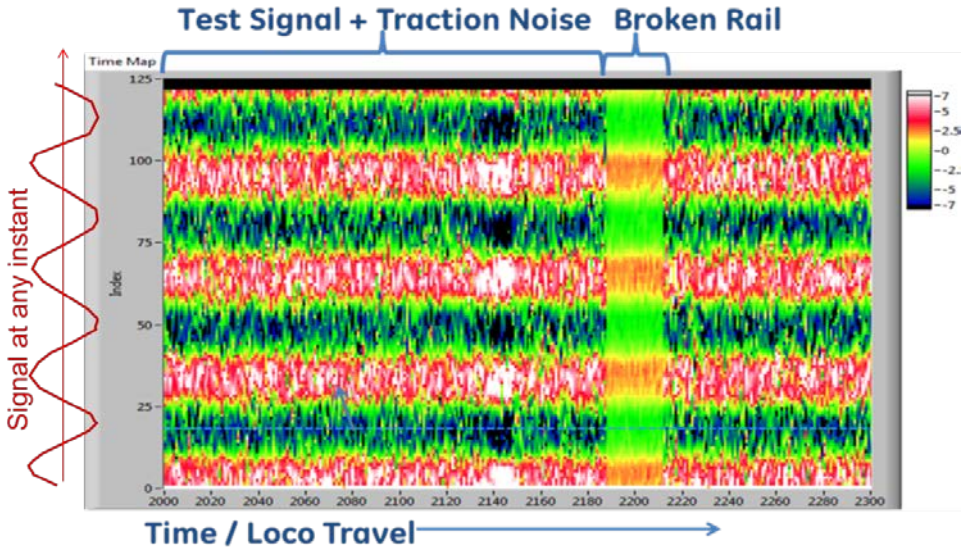


Figure 12. The RIAS data collected during locomotive motoring

A much cleaner time plot can be constructed by processing each waveform recorded by the RIAS. For example, by applying a sine transform at the same frequency as the transmit signal, a signal magnitude of this frequency component can be computed as a function of time.

Several runs over rail features were made at different speeds, different excitation frequencies, and in both forward and reverse locomotive movements to investigate the working range and possible shortcomings of the RIAS prototype. An example that features RIAS raw data and signal processed charts with the signals superimposed over the satellite images is presented in Figure 13. Insulated joints and shunts can be identified and their locations can be pinpointed on the map.

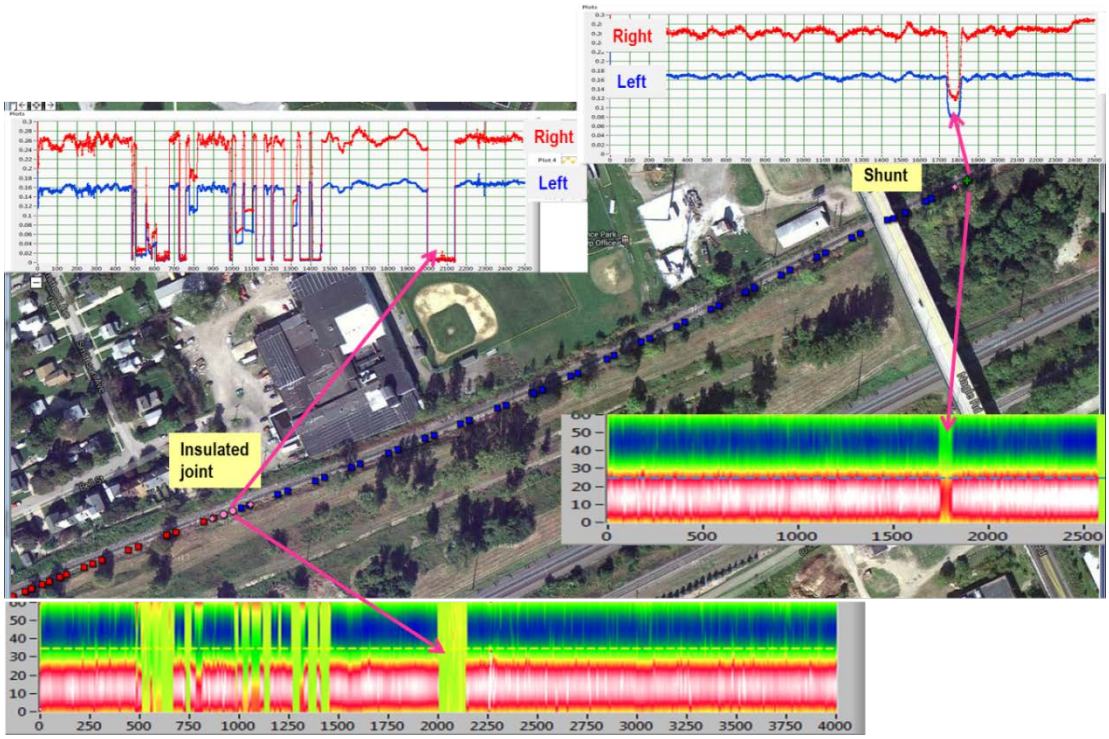


Figure 13. RIAS signals at 5 mph linked to the rail features on the ground

In response to the field tests in Erie, all the components of the RIAS prototype were modified and improved. Special attention was paid to the coil bobbins design and magnet core mount arrangement. The coils, cables, connectors were ruggedized to withstand outside weather conditions and survive the vibration of a moving train.

Two excitation frequencies were used, which allowed the team to perform separate analyses of the signals obtained from each transmitter (Tx1 and Tx2). Data from both receiving channels were recorded simultaneously. The two frequency components were mixed in the rail loop and ccould be extracted separately from the recorded data.

In January of 2014, the RIAS prototype was installed on a NS inspection train. The train was traveling during January and February through mostly signaled territory. There were no broken rails identified during that travel. While this provided us with a large amount of data suitable for digital signal processing and statistical analysis, the feasibility of detecting a real rail break had not yet been determined.

Later, in February and March of 2014, the RIAS prototype was tested at the Transportation Technology Center (TTC) in Pueblo, CO. The main goals for this test were to obtain signal signatures from natural rail breaks and validate the concept of broken rail detection by the RIAS. A high tonnage train (Figure 14, with the train seen on the back plane of the picture) autonomously running over the 2.7 miles high tonnage loop was used for the tests.

Figure 14. High tonnage train at TTC, Pueblo, CO

The RIAS prototype was installed on an idling locomotive (which provided electrical power for the System) that is shown in Figure 15. This locomotive was connected to the end of the high tonnage train, which was running over the high tonnage loop in continuous regime at the speed of 40 mph.

Figure 15. The GE RIAS installed on an idling locomotive attached to the high tonnage train

A total of about 800 passes were made over the loop. During this testing period three rail breaks occurred. All three broken rail signatures were collected by the system. Figure 16 has a summary of the breaks.

BREAK #1
Date: February 20, 2014
Time (Local): 00:48:55 (2:48:56 on IPC)
Type: Transverse Defect
Speed: 57.05 ft/sec
Calculated Window: 23.96 ft
Latitude: 38.453363
Longitude: -104.347807
Track: Bypass
File: DATA_02_19_14.bin-output-4.bin
Offset: 10449

BREAK #2
Date: February 20
Time (Local): 21:00:18 (23:00:38 on IPC)
Type: Gauge Corner Cracking and/or Transverse Defect
Speed: 58.78 ft/sec
Calculated Window: 22.83 ft
Latitude: 38.453287
Longitude: -104.348108
Track: Main
File: DATA_02_20_14.bin-output-6.bin
Offset: 3097

BREAK #3
Date: February 24
Time (Local): 23:34:28 (1:34:28 on IPC)
Type: Broken Weld
Speed: 58.74 ft/sec
Calculated Window: 24.55 ft
Latitude: 38.453167
Longitude: -104.335972
Track: Main
File: DATA_02_24_14.bin-output-9.bin
Offset: 4964

Figure 16. Details and images of the three broken rail occurrences during the TTC tests

Figure 17 illustrates the levels of the signal and background noise recorded during the TTC tests. The magnitudes of the RIAS signals from the two last loop runs are superimposed on the time plot in Figure 17. The first three voltage drops are due to existing rail features: insulated joints. These signal signatures are repeated every loop. The fourth clear signal drop (blue lines only) happened over the rail break during the last loop run (before the train was stopped). It is not seen at the previous pass (red lines).

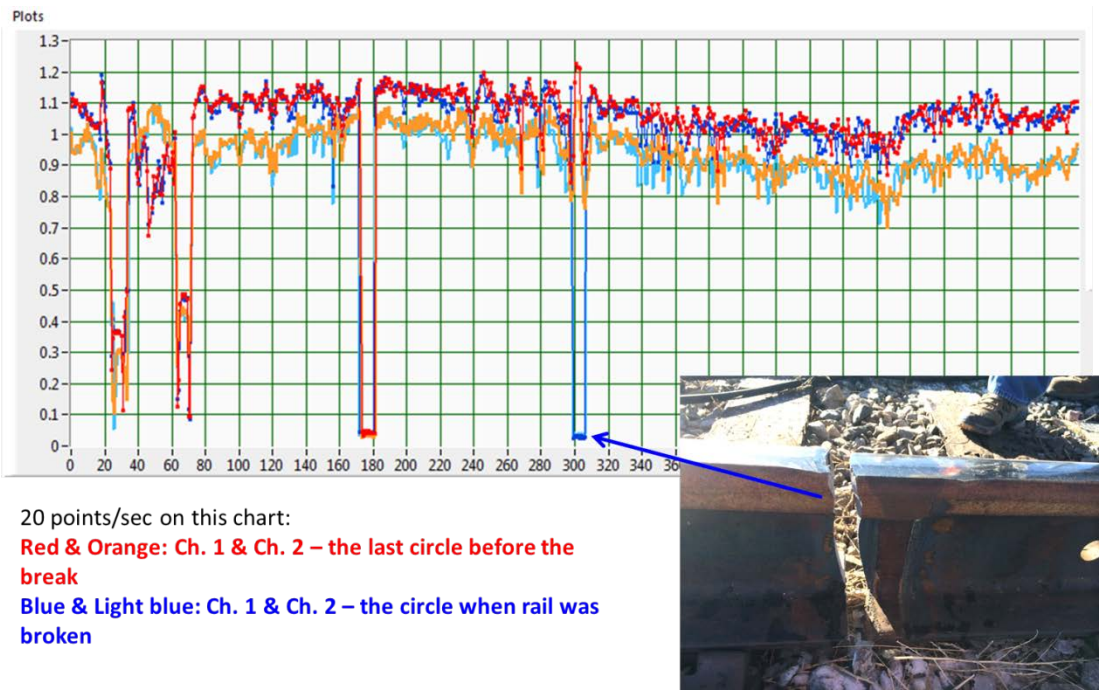


Figure 17. Comparison of the RIAS signal magnitudes recorded during the passing over break #1 (blue colors) and the signal recorded during the last pass before the break (red and orange colors)

Based on the results obtained during the RIAS field trials, we were able to address the most critical risks that were identified by laboratory tests. However, the team was concerned that RIAS might generate a high false alarm rate due to rail conditions and locomotive electromagnetic noise, and this aspect is analyzed in the following chapters of this report.

2.3 Summary of Results

- Laboratory tests demonstrated that detecting rail electrical disconnects with a electromagnetic non-contact system is feasible
- The system's physical principles were investigated on 1:15 and 1:3 scaled physical models of the rails-wheels-axles system.
- System functionality was tested with field trials with RIAS prototypes installed on real locomotives and conducted at different locations during different weather conditions
- The ability of RIAS to detect broken rails was proven with signals recorded from the high tonnage loop at TTC.

3. Establishing Identification of the Broken Rail Signal Produced by the RIAS

This section describes a broken rail detection algorithm and discusses the structure of the data collected by the RIAS prototype. Also, a system description for the RIAS prototype that has been used for data collection is used to summarize the railroad features that were passed by the inspection train during field tests. The signal processing approach and choice of filters applied to the data are described, as well as the matched filter used for data reduction. The established broken rail detection algorithm is evaluated on the collected data and results of evaluations are summarized.

3.1 Data Collection System

Section 2 of this report describes the laboratory and field tests that were used to demonstrate the feasibility of detecting a real rail break. Clear signal signatures were obtained while a locomotive with a RIAS prototype was passing over the breaks. Data collected during the field tests were available for post processing and were used in this project.

The mutual locations for the transmitting and receiving coils are schematically depicted in Figure 18. Each transmitter is driven by a single frequency harmonic signal. The two frequencies are injected into the same region of the track, mixed together, and then the multi-frequency response is picked up by the two receiving coils positioned at the opposite corners of the loop as indicated in Figure 18. This and similar configurations, details about the signal injection, and break detection procedures are covered by U.S. Patent No. 8,914,171.

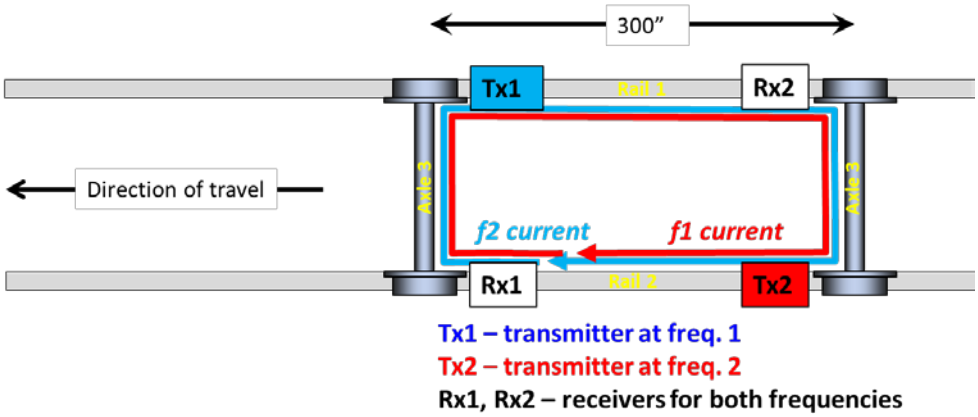


Figure 18: **Working diagram and locations of the transmitting and receiving coils in the RIAS prototype.**

On the tested RIAS prototype, two main tone frequencies of *f1*=3.83 and *f2*= 4.655 kHz were used to transmit signals into the rails. These frequencies were chosen from the frequency gaps currently unutilized by on-ground signaling circuits. This allows the system to avoid direct interference with existing railroad equipment.

After some adjustments and optimization, the system was assembled in early 2014 and configured as shown in a diagram in Figure 19.

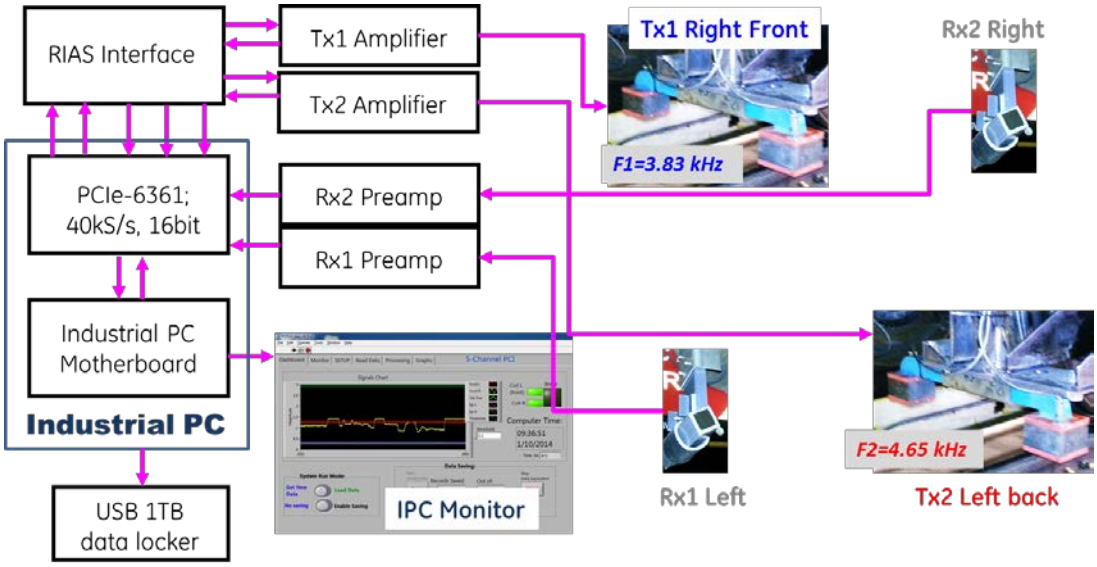


Figure 19: RIAS prototype system diagram that was used for data collection in January - March of 2014

An industrial personal computer (IPC) is used to coordinate the work of the RIAS prototype. It sets parameters for the transmit coils and synchronizes the data acquisition process. A waveform generator produces harmonic oscillations in the frequency range from 1 to 20 kHz, and a power amplifier is used to increase voltage and current of the transmit signal to a level that can excite the transmitting coils. The alternating current in the rails is monitored by two receiving coils located on the left and right sides of a locomotive. The receiving coils are connected to the preamplifiers, and the amplified signals are digitized in a PCI data acquisition card. The digitization process is synchronized with the waveform generator, which allows the magnitude and the phase of the receive signals to be computed. The signals are recorded to the hard drive of the IPC for further review and analysis.

When the system is installed, the coils are attached to the locomotive frame near the internal pair of the axles, while the system control and processing components are assembled in the cabin. The system control components as deployed on the NS inspection train are shown in Figure 20. The same components were mounted inside the locomotive cabin during the data collection at TTC.

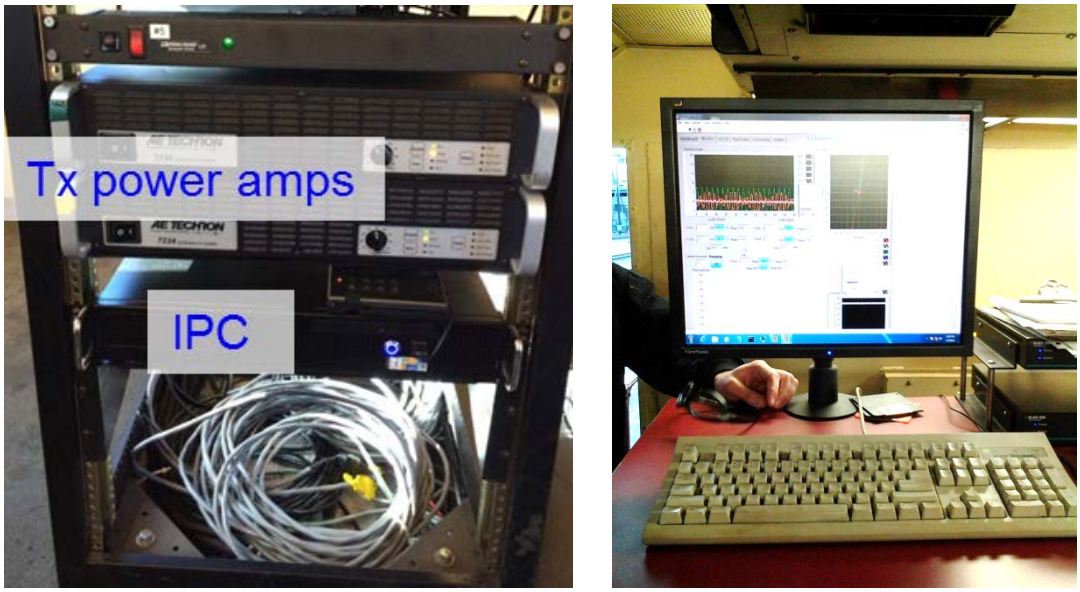


Figure 20: The RIAS prototype deployed on the NS inspection train

3.2 Data Format

Five RIAS signals were digitized with a sampling rate of 40 kS/s, which corresponds to 25 us intervals between digitized samples, which is five times higher than the Nyquist frequency. A multi-functional data acquisition board with 16 bit resolution from National Instruments is used in the RIAS prototype. The digitized signals and their functions are listed in Table 1.

The signal digitization is phase-locked and the sampling rate is ten times higher than the main tone frequencies. This allows full data post-processing and application of various approaches to extract useful information related to broken rail and track features and surrounding conditions.

Table 1. Description of the Recorded Signals

Signal	Function	Sampling parameters	Length
Channel 1	Voltage from Rx1	40 kS/s, 16 bit (Integer-16 format); Triggering rate: 2 Hz	19400 samples - corresponds to 485 ms
Channel 2	Voltage from Rx2		
Channel 3	Voltage from Tx1		
Channel 4	Current through Coil 1		
Channel 5	Current through Coil 2		

The digitized stream is translated into a single array of 16 bit integers that is stored on the hard drive. In the table above, each record requires 194 kB of storage: 19,400 samples x 5 channels x 2 bytes = 194 kB. The record's length is 485 ms but the triggering on the data acquisition board is done at 2 Hz rate or in 500 ms intervals. The graphical user interface was designed using the LabView environment

This protocol supports continuous real time data acquisition and recording on a hard drive for several hours without missing any records. A separate program was running on the computer during field tests simultaneously with the data acquisition in order to record RIAS coordinates from a GPS unit.

3.3 Railroad Features Covered by the Collected Data

GE collected several sets of data from motoring locomotives using the RIAS prototype hardware as described in Section 3.2. Several data sets were collected on the GE Transportation test track in Erie, PA. Approximately 1200 miles of track data were collected on a Class I railway and on the High Tonnage Loop (HTL) test track in early 2014. In this report, the railway data and the test track data are known as the "CL1" and "HTL" data sets. The team collected ancillary information on the location and condition of both the locomotive and tracks, the time of data collection, and track artifacts such as switches and railroad crossings during operation (i.e., automatically) and labeled manually, post hoc. The rail features that are physically present on the railroad at different sites are summarized below.

The East track of GE Transportation in Erie, PA, has several features including two insulated joints installed for testing, two hardwire shunts, and one road crossing. Data collected on the East track contain signals from RAIS runs but also include various degrees of locomotive noise.

Table 2 lists the railroad features present at TTC HTL. Signal responses to the six features of the loop were recorded multiple times, while only one occurrence for each of three breaks was recorded. The train was stopped immediately when the alarms from the wayside signaling network indicated that it passed a break and the rails were repaired.

Table 2: Railroad features on the TTC High Tonnage Loop

No	Feature	Details	Location	Latitude	Longitude
1	Insulated Joint	21" Stagger	Main & bypass	38.452918	-104.348927
2	Insulated Joint	Single	Main & bypass	38.452597	-104.349655
3	Insulated Joint	126" Stagger main; 57" stagger bypass	Main & bypass	38.452435,	-104.350022
4	Insulated Joint	41.5" stagger	Main	38.447487	-104.346108
5	Insulated Joint	20.5" stagger main, 41" stagger bypass	Main & bypass	38.453832	-104.338892
6	Insulated Joint	Single	Main & bypass	38.453893	-104.339280

7	Break 1	Single, (2/20 morning)	Bypass	38.453363	-104.347807
8	Break 2	Single, (2/20 evening)	Main	38.453287	-104.348108
9	Break 3	Single, (2/24)	Main	38.453167	-104.335972
	Total:	**6 Insulated Joints**	**3 breaks**		

A schematic diagram of the HTL is presented in Figure 21. The mutual positions of the rail features and breaks described in Table 2 are marked on this map.

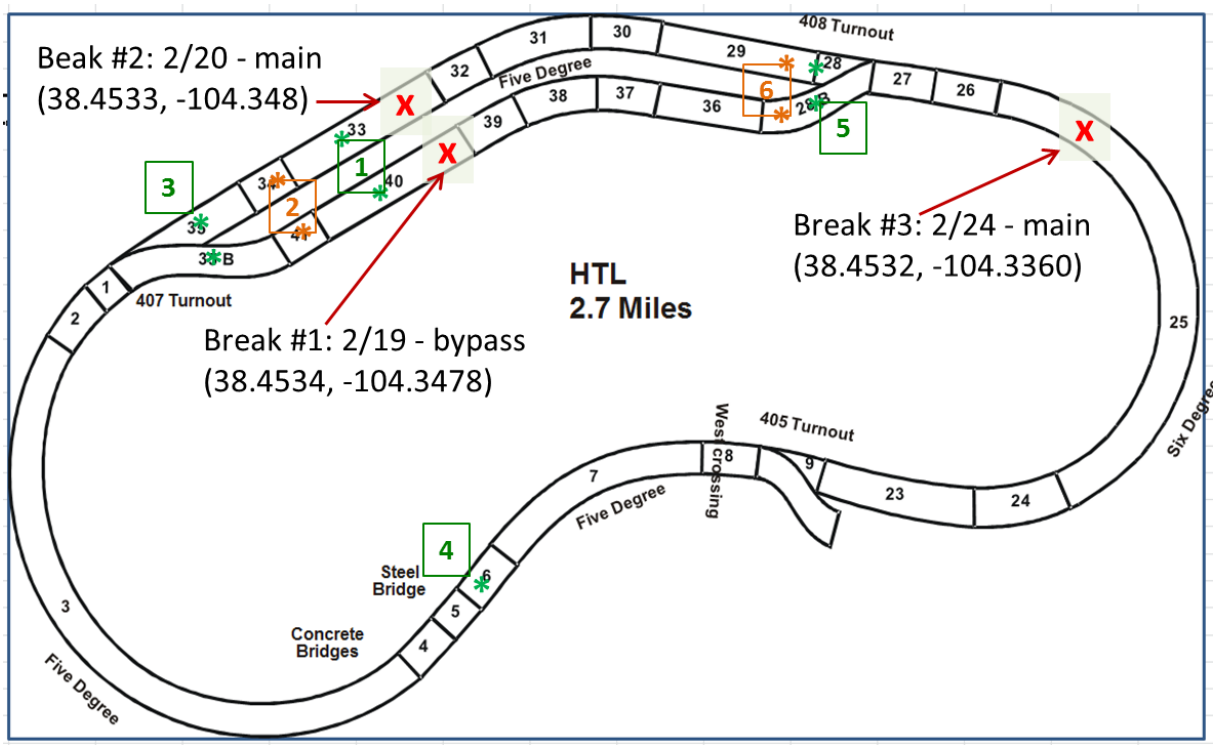


Figure 21: Locations of the railroad features on the HTL in TTC listed in Table 3

Detailed summaries of data collected in the CL1 and HTL data sets with a breakdown of the number of each track artifact type are given in Table 3. It should be noted that these artifacts were first identified by an automatic algorithm, and manually labeled after the data collection.

Table 3. Listing of manually labeled railroad artifacts for the CL1 and HTL databases

Data Set	Track Artifact Type	Count
CL1	AEI reader	14
	Diamond	16

	Intermediate Signal	222
	Interlocking Signal	538
	Joint (jointed rails and welded rails)	11
	Lubricator	137
	Single Crossover	52
	Repeater	5
	Shunt	31
	Switch	99
	Crossing	3
HTL	407FR	456
	407SW	418
	408FR	434
	408SW	434
	Break	3
	Bridge	453
	Feed	454
	Lubricator	146

3.4 Structure and Parameters of the Designed Matched Filter

3.4.1 Demodulation Algorithms

We investigated two broadly different methods to detect time-harmonic modulated signals: matched filtering and direct demodulation. The main difference between matched filtering and direct demodulation is that direct demodulation depends on the presence of a phase-synchronous reference signal, whereas with matched filtering, the phase of the excitation waveform is not assumed to be known a priori. Both algorithms begin by breaking up the raw data into smaller “chunks” or “frames”, since the raw data sampled at 40 kHz can be too large and unwieldy to work with. For each frame, we compute the amplitude and the phase of the received signal at each excitation frequency within the frame.

In matched filtering, the received signal was convolved with a finite-impulse response (FIR) filter that represented a truncated version of the excitation waveform. A window may be

employed to mitigate the frequency-domain effect of the time-domain truncation (as in the Fourier-domain, a rectangular window is represented by a sine function, with significant side lobes). It can be shown that matched-filtering with a time-reversed version of the excitation waveform maximizes the SNR of the FIR output assuming additive white Gaussian noise (AWGN). Other waveforms are optimal in the case of non-white and non-stationary noise, which we believe to be the case for realistic rail data and we plan to examine the design of optimal matched filters for this case in future work. The block diagram of the processing is shown in Figure 22.

The processing workflow is as follows:

1. The raw sampled data are broken up into discrete frames of some specified duration.
2. Within each frame, the raw data are convolved with matched filters representing time-reversed excitation waveforms, and are truncated in some way temporally.
3. The maximum response within the frame is computed and this response is reported as the amplitude of the frequency response for that frame, detector, and frequency.
4. Thus, for two detectors and two excitation frequencies, we report four amplitudes for each data frame.

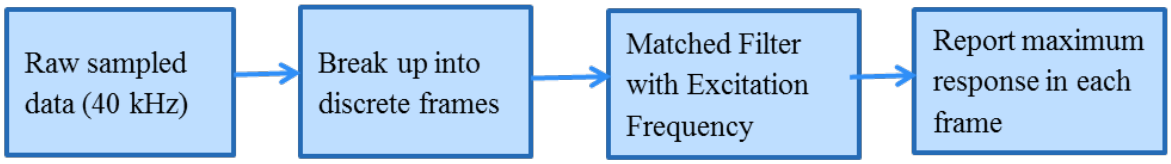


Figure 22. Block diagram of matched filter processing

Alternatively, in direct demodulation, we multiply the received signal in the time domain by phase-synchronous sine and cosine waveforms at each excitation frequency. As time-domain multiplication is equivalent to frequency-domain convolution, we have thus shifted the excitation waveform to baseband, with a low-pass filter needed to remove the aliased second-harmonic of the excitation waveform. Within each frame, we then integrate the multiplied, low-pass-filtered signals to compute the I and Q signals, the in-phase and quadrature components. The amplitude and phase of the received signal at each frequency can thus be computed by:

$$\text{Amplitude} = \sqrt{I^2 + Q^2}$$

$$\text{Phase} = \tan^{-1}\frac{Q}{I}$$

Figure 23 shows the block diagram for direct demodulation signal detection.

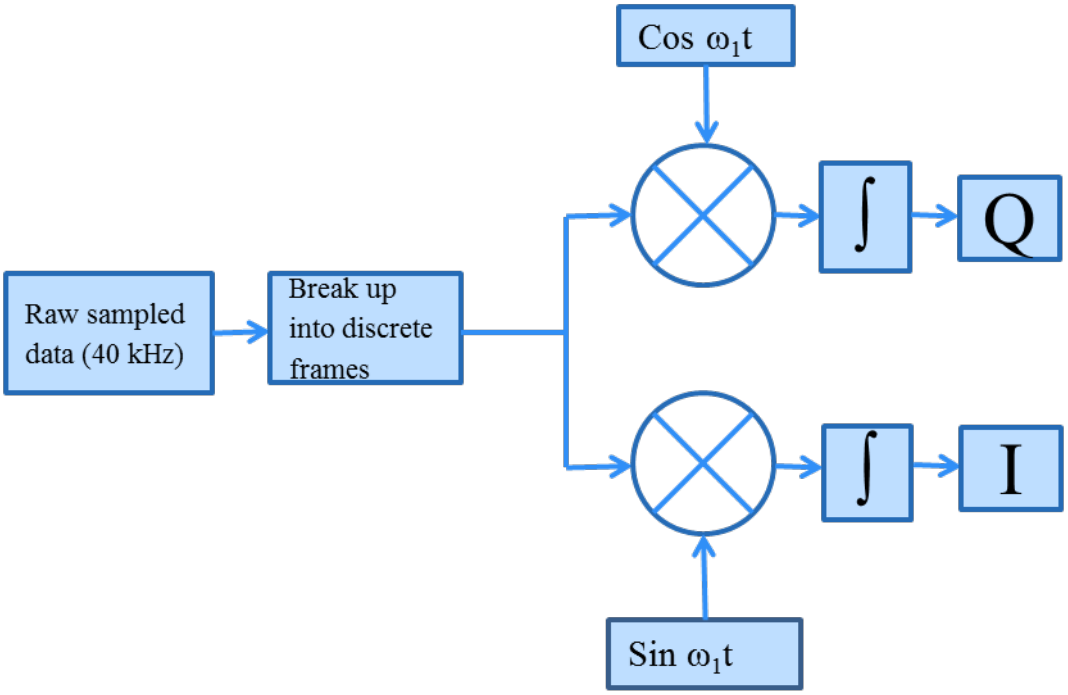


Figure 23. Block diagram of direct demodulation processing. Shown for one excitation frequency; same structure for the second and subsequent excitation frequencies

3.4.2 Simulation Study

Our simulation study demonstrates the feasibility of a RAIS-equipped train detecting a break at high speed and the measurements are corrupted by a very high level of white, Gaussian noise. This study assumed that:

- The train is moving at 70 mph
- The distance between front and rear axles is 25 feet
- Continuous data collection at 40 kHz
- The excitation waveforms are pure tones at 3.830 and 4.655 kHz
- There is an extremely high level of noise, such that the signal-to-noise ratio of the raw data was 0 dB

The excitation waveforms were assumed to be:

$$S_1(t) = \cos 2\pi f_1 t$$

$$S_2(t) = \cos 2\pi f_2 t$$

Then, at the detectors, the recorded waveforms are assumed to be:

$$R_1(t) = \mathrm{Re}\left[S_{11}e^{j2\pi f_1 t+\varphi_{11}} + S_{12}e^{j2\pi f_2 t+\varphi_{12}}\right] + N_1(t)$$

$$R_2(t) = \mathrm{Re}\left[S_{21}e^{j2\pi f_1 t+\varphi_{21}} + S_{22}e^{j2\pi f_2 t+\varphi_{22}}\right] + N_2(t)$$

Finally, the study assumed that there is no direct coupling between excitation and detection coils and thus that only white noise is detected during the break period.

$$R_{1,break}(t) = N_1(t)$$

$$R_{2,break}(t) = N_2(t)$$

The simulated noise-free and noise-corrupted raw sampled waveforms for Detector 1 are shown in 4. It is evident that the signal appears to be "buried" by the noise given the very low assumed SNR of the data-collection system.

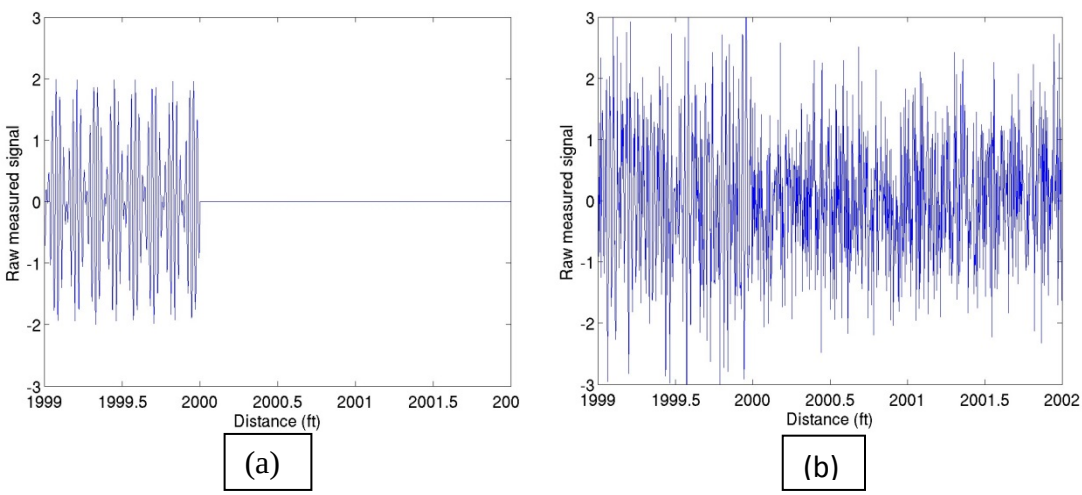


Figure 24. Raw measurements made at Detector 1. (a) Noise-free data; (b) Noise-corrupted data

3.4.3 Results for Matched-Filtering Processing

To simulate matched filtering, we vary the frame size and set the FIR filter to be 50 percent of the frame size, so the results for each data frame can be computed independently. We discard the response for the first half of the frame (where the processing would have required data from the previous frame).

In order to analyze the results quantitatively, the mean and standard deviation of the "break" and "non-break" detected signal amplitudes were computed as a function of frame size. The results are shown in Figure 25. As expected, the variance of "non-break" estimated amplitudes decreases monotonically as frame size increases. However, as frame size increases beyond 20 ms, the means and standard deviations of the estimated amplitudes during the break increase dramatically.

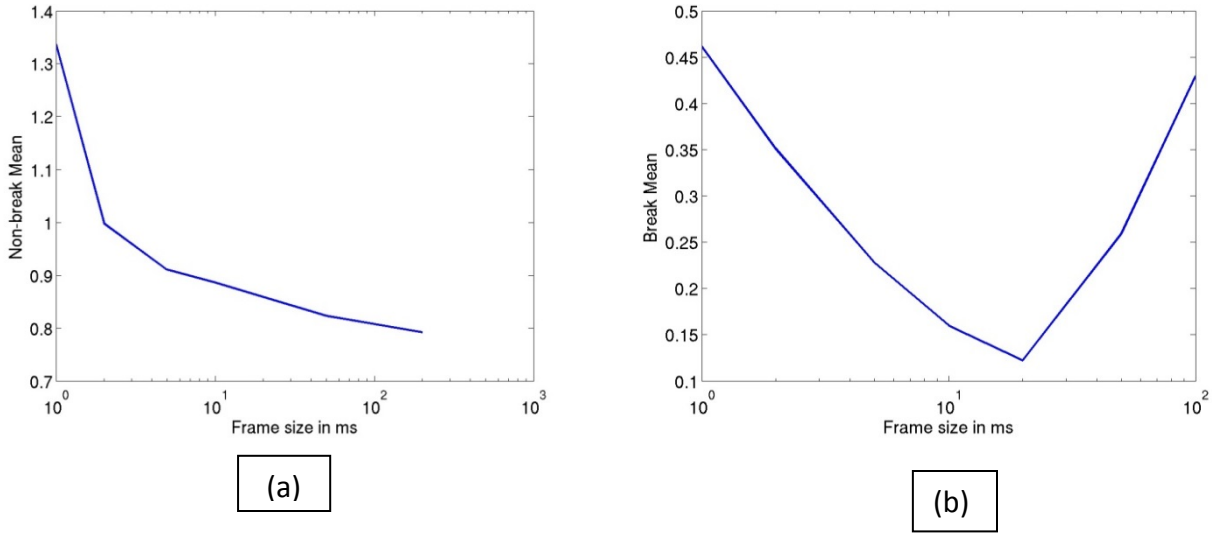


Figure 25. Quantitative comparison of frame size statistics. (a) Non-break mean vs. frame size. (b) Break mean vs. frame size

3.4.4 Direct Demodulation Processing

The same analysis is performed for estimation of the statistics vs. frame size for the case of direct demodulation, where phase-synchronous cosine and sine signals were assumed to be available. Unlike the case of matched filtering, there was no need to discard data at any point in the processing.

3.4.5 Comparison of Signal Detection Algorithms

In order to compare the two approaches, we introduce the following metric of class separability, defined intuitively as the difference between the means of two classes divided by the mean of the standard deviations. This definition allows us to compare detection approaches using a metric that is invariant to scaling factors:

$$Sep = \frac{\mu_1 - \mu_2}{0.5(\sigma_1 + \sigma_2)}$$

The comparison of the class separability vs. frame size for direct demodulation vs. matched filtering detection is shown in Figure 26. The improved separability of direct demodulation is due to the use of prior information giving the excitation signal phase. However, in the case of matched filtering, the maximum response is over the frame, while there may have been an anomalously high response due to noise alone. However, even in the presence of stationary white noise with high amplitude, the two classes are very separable with a 20 ms frame size using either detection approach.

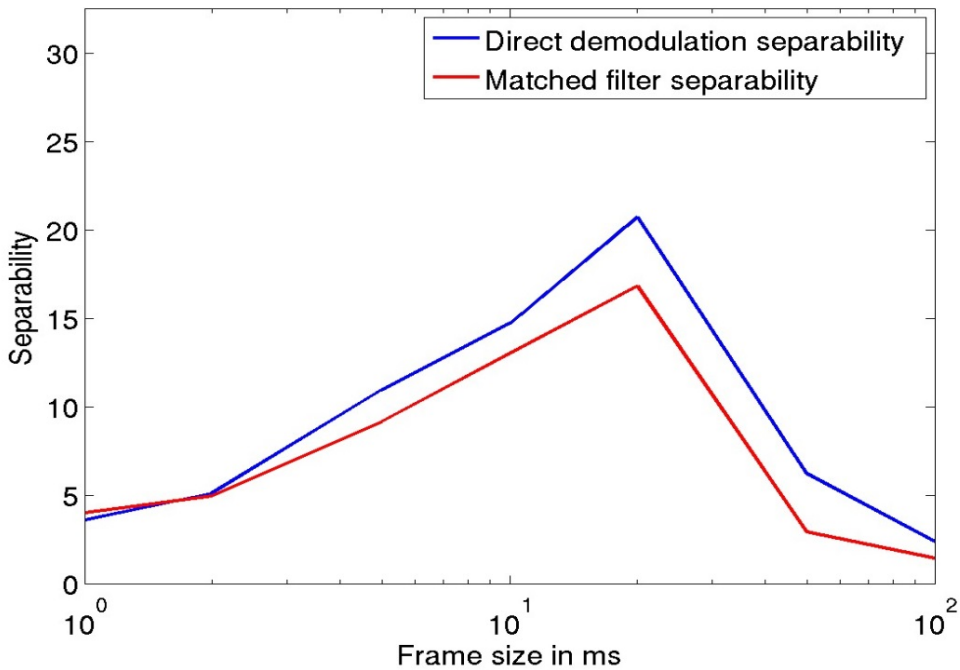


Figure 26. Comparison of class separability vs. frame size for various signal detection techniques

3.5 Algorithm Performance, Test Results and Analysis

3.5.1 Matched Filtering for Rail Data

In this section, the matched filtering method is applied to signals from the TTC HTL to compute estimates of time-varying excitation frequency amplitudes. Matched filtering methods, described in the previous section, were applied to voltage signals measured at Receiver 1 and Receiver 2 at two distinct frequencies, and , resulting in 4 matched filter amplitude signals. Frequencies 3830 Hz and 4655 Hz were chosen to match the known excitation frequencies.

A characteristic drop in amplitude is observed at the crossing of each insulated joint, similar to the simulation study in Section 3.4. The “Receiver 1, Frequency 2” and “Receiver 2, Frequency 1,” signals exhibit significantly lower voltages during the break than the remaining signals.

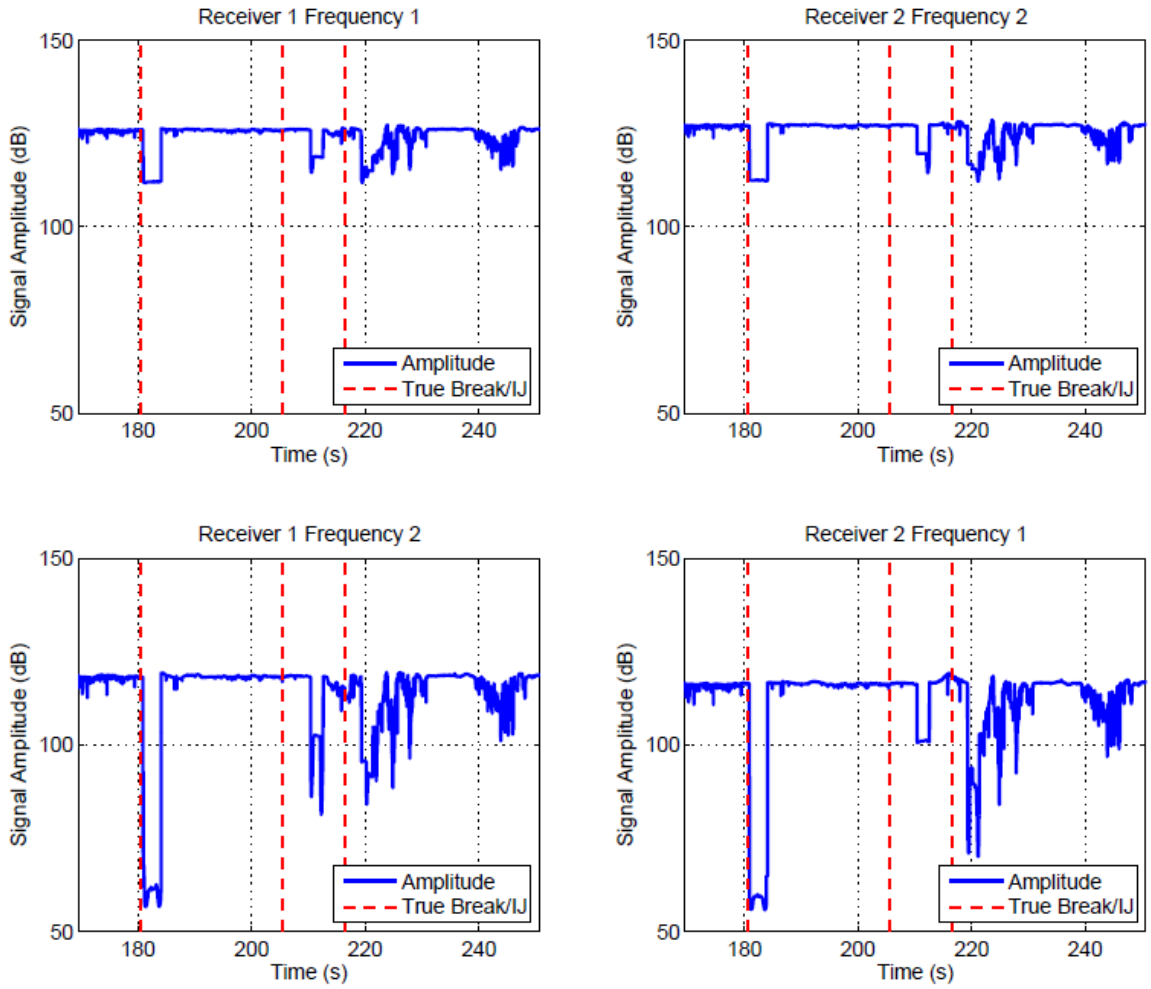


Figure 27: Matched filter amplitudes for 81 seconds of data including 3 insulated joints. Red, dashed lines are GPS-based occurrence times for insulated joints

3.5.2 Detecting the Break State

If a break or insulated joint is crossed, a significant drop in voltage for the matched filter amplitude signal occurs. This drop in voltage could be used to detect broken rails. While a broken rail detection system's ultimate goal is to determine whether a broken rail occurred and record the time at which it occurred, detecting the change in amplitude caused by a broken rail (or the "break state") is, itself, is a useful result, and it can predict the accuracy of a broken rail detection system. Results obtained from detecting the break state can be used to predict trade-offs between miss and false alarm errors for detecting broken rails. In this section, we will develop methods of detecting the Break State and evaluate the accuracy of these methods.

A block diagram illustrates the system for broken rail detection in Figure 28. Matched filter amplitude values are grouped into short blocks, and a confidence score (i.e., some value reasonably commensurate with the posterior probability of the break state) is computed for each block. Finally, the broken rail detection system can allow thresholds to be set for the confidence

score value as well as the estimated duration of the break state, so a final decision as to whether a break occurred, and when.

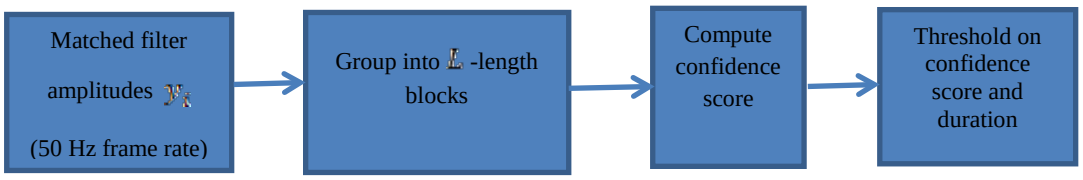


Figure 28. Block diagram for detecting broken rails based on matched-filter amplitudes

Grouping the matched filter amplitude values into blocks of length L, a confidence score can be computed that indicates whether each block is in the break state or not. The confidence score should model a significant drop in amplitude. A simple method for computing a confidence score based on a drop in voltage is illustrated in Figure 29 for a segment of data recorded on February 18th. The "voltage drop" score, $Vdrop_j^k$, for the block of data containing amplitude observations y_j to y_k , is computed as the median signal amplitude value, minus the mean amplitude for block.

$$Vdrop_j^k = median - \frac{1}{k-j+1}\sum_{i=j}^{k} y_i$$

"Voltage drop" score values are computed for blocks of 10 matched filter amplitude values, with a spacing of 200 ms between blocks. In Figure 29, confidence scores for all 4 signals are plotted in red, along with labels for 3 occurrences of the break state that corresponds to IJs 1, 2, and 3. The confidence score in Figure 28 is generally elevated in the break state. Otherwise, the score is low, although several spurious peaks occur following IJ #3.

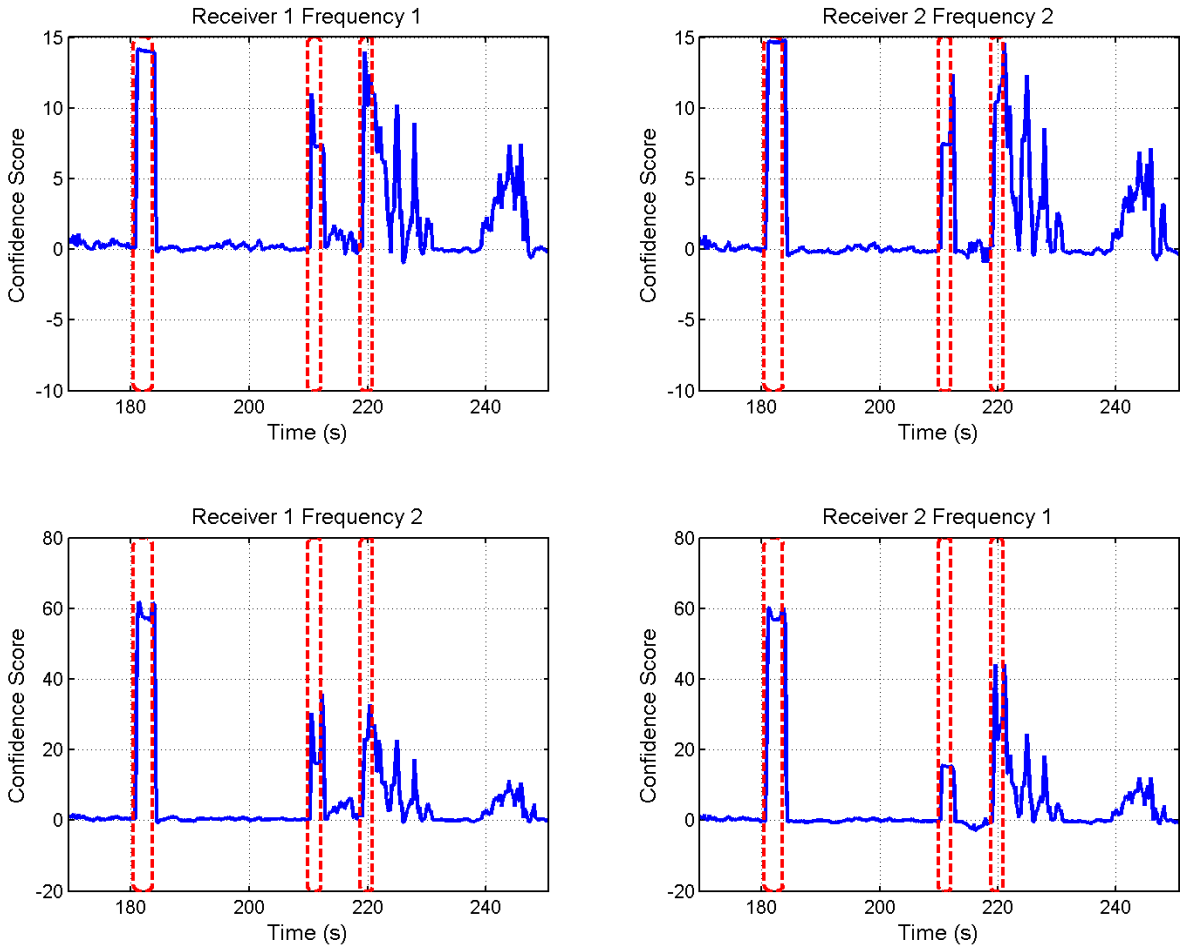


Figure 29. Plot of "voltage drop" confidence score, i.e., median amplitude minus amplitude for IJs 1, 2, and 3, as listed in Table 2. Hand-labeled start and stop times of IJs in red, dashed rectangles

3.5.3 Evaluating Break State Detection

Once a complete set of labels for the break state is available, we can evaluate the detection accuracy of any confidence score using the receiver operating characteristic (ROC) curve. This curve illustrates the trade-off between true-negative or "miss" errors and false-positive or "false alarm" errors, and it is a plot of the probability of missed detections vs. the probability of false alarm errors as the threshold is varied across its entire range. A plot of the ROC curve in Figure 30 covers the voltage drop confidence score for all 4 signals. These curves were computed with a short segment (comprising one track loop) of hand-labelled data that was recorded on 2/18.

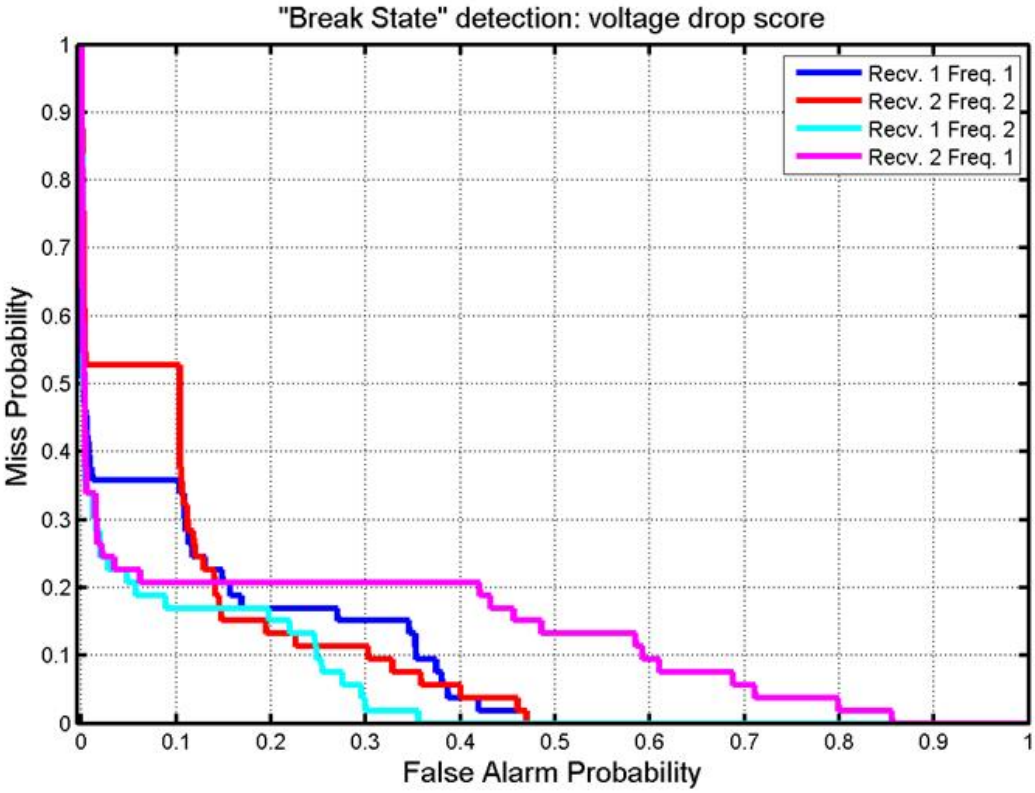


Figure 30. Receiver Operating Characteristic (ROC) curve for detecting the break state using the voltage drop score

A trade-off between false alarm and miss errors can be performed with respect to the thresholds for all the signals. In Table 4, threshold values were selected to achieve two detection operating points, as "equal error rate" and "low false alarm rate." "Equal error rate" operating points were selected to achieve approximately equal false alarm and miss probabilities, while "low false alarm rate" thresholds were selected to achieve a false alarm probability equal to approximately 0.01. While error rates are nearly equal for all of the four signals, the "Recv. 1, Freq. 2" and "Recv. 2, Freq. 1" signals achieve significantly lower miss probabilities at the "Low false alarm rate" operating point.

Table 4: Error trade-off operating points for break state detection with "voltage drop" score

	Equal error rate		Low false alarm rate		
	Threshold	**EER**	**Threshold**	**FA prob.**	**Miss prob.**
Recv. 1, Freq. 1	0.3974	0.1822	7.276	0.0081	0.4151
Recv. 2, Freq. 2	0.3067	0.1698	7.538	0.0134	0.5283
Recv. 1, Freq. 2	0.5959	0.1710	8.054	0.0112	0.2642
Recv. 2, Freq. 1	0.4096	0.2136	9.024	0.0096	0.2642

3.6 Summary of the Results

- The data collection system used on the RIAS prototype is described and an example of component arrangements during field tests is given. The location of the transmitting and receiving coils are described and diagrammatically illustrated
- The system prototype packaged signals into 16-bit integer format continuous stream, which allowed the team to save 8-12 hours of raw data on a computer hard drive
- The motoring noise coming from the locomotive motors was considered as white noise for simplicity
- The railroad features that were detected during various field tests are organized into tables
- There were two approaches to detecting time-harmonic modulated signals: matched filtering and direct demodulation. Optimization of the processing parameters for both approaches was conducted. The separability of direct demodulation was improved due to the use of prior information about the excitation signal phase
- An investigation of methods of detecting broken rails based on computed matched filter amplitudes was conducted
- We have investigated both "break state" and broken rail detection on real data that was collected from the TTC test track loop, using the matched filtering approach. We computed a very simple confidence score for the break state that was based on the difference between the estimated signal amplitude and its median value
- Using a small, hand-labeled portion of the data, we selected a threshold and applied it to the task of broken rail detection based on the confidence score. We computed ROC curves to analyze the trade-off between miss and false alarm errors in detecting the break state

4. Identifying Signal Patterns from the Track Features

This section establishes a procedure for discriminating railroad features that are detected by the RIAS. A track feature recognition algorithm is described and evaluated, and this section also includes RIAS electrical signatures that were taken from a range of railroad features .. This procedure uses thresholding as a preprocessing step, followed by a more sophisticated detection algorithm based on signal processing, statistical pattern recognition, or machine learning. Examples in which one class of the railroad feature is discriminated against others are provided and discussed. In addition, a procedure and the data collected with exaggerated noise produced by the traction motors are included. Preliminary analysis shows that this noise is in-band with the RIAS transmitting frequencies and is not removed by the matched filtering. An approach for the locomotive noise reduction is presented.

4.1 Locomotive Noise Study

When data was collected at NS and TTC, the locomotives with RIAS prototypes installed were idling to provide the electrical power for the system. As a result, the collected data have a relatively small noise contribution from the locomotive themselves. In contrast, an upcoming implementation of the RIAS is designed to be installed on a pulling/pushing diesel-electromotor locomotive that can produce excessive electromagnetic noise.

In July 2014, a Technical Advisory Panel from GE Transportation reviewed the field test conditions and proposed a set of additional tests with an AC locomotive that would include the severe noise conditions. Our initial processing demonstrated that the noise exceeds the levels that we observed previously and interferes with our developed processing algorithm. This chapter is devoted to our attempt to assess and mitigate the worst case scenario locomotive noise.

4.1.1 Locomotive Noise Examples

In November of 2014 a new data set was collected from the test track adjacent to the GE Transportation plant in Erie, PA. Four locomotives were used to simulate some cases of motoring noise. During the tests, the head unit was applying traction while the last unit was applying brakes to exaggerate the noise produced by the traction motors by simulating a train load.

The test conditions included:

- Traction motors #1, #2, #5, #6 cut in and out
- Idle motors; notch 1, 6, 8
- Dynamic braking 1 and 8; with/without simulated ground fault
- Variable (oscillating speed) 0 to 10 mph; 0 to 25 mph; 25 to 20 mph

A brief analysis of the data revealed that an excessive amount of noise is received by the RIAS prototype. One of the highest noise levels that was recorded during dynamic braking was captured in Figure 31. The plots shown in this chart are the outputs of the direct demodulation applied to the raw data as described in Section 3.4. The lower frequency (3.83 kHz) signals from

the two receiving coils are shown in shades of blue color. The plots in red shades are the signals demodulated at 4.65 kHz.

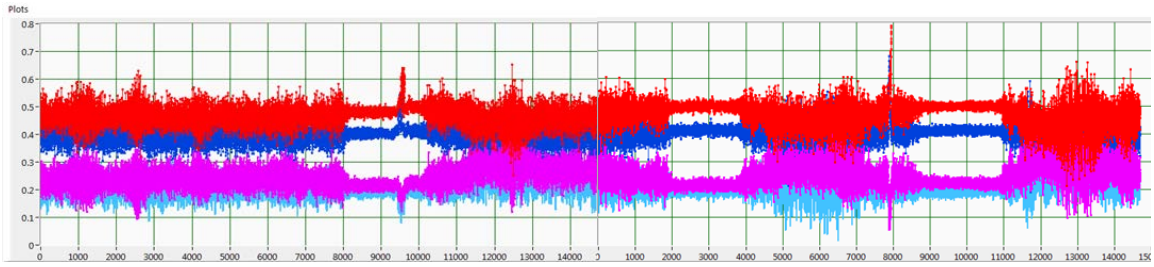


Figure 31: RIAS signal after demodulation at two frequencies showing electromagnetic noise from the locomotive

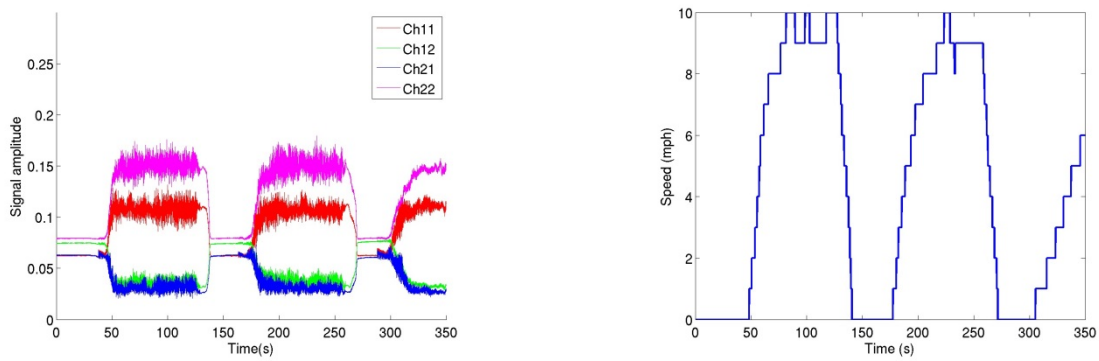
The plots above indicate that a significant amount of noise is still present, even after the signals processing, and that noise might affect the feature detection and discrimination. The locomotive noise covered a large range of frequencies that included in-band interference with the transmitting RIAS frequencies, and that interference is not removed by the filtering performed during signal demodulation. There was very complex, highly spectrally "colored" noise in data sets with simulated ground faults.

4.1.2 Estimation of Signal-To-Noise Ratio

The signal-to-noise ratio (SNR) is important for establishing an upper bound on system performance. Assuming Gaussian but not necessarily white, noise, one can show that the optimal detection scheme utilizes the Likelihood Ratio Test, where we compare the null hypothesis of a signal being present at the excitation frequency to the hypothesis that there is no connection between the transmitter and receiver. In reality, in addition to the coupling of the coils through the track, there is a certain degree of direct electromagnetic coupling, and thus the formulation of the problem as a hypothesis testing problem is not entirely accurate. However, it is a reasonable approximation of real-world performance that we are likely to encounter.

Ideally, in order to estimate the SNR we would estimate the power in the signal using a matched filter or similar approach and then "listen" to the noise without the signal present. As the logistics of this approach would be difficult, our strategy is to instead remove components of the measured data in the signal subspace (with dimension 2 for each frequency) and then to interpolate the expected noise power in the signal subspace using a parametric autoregressive all-pole model. The optimal model order is selected using a variant of the well-known Akaike Information Criterion (Akaike, 1974).

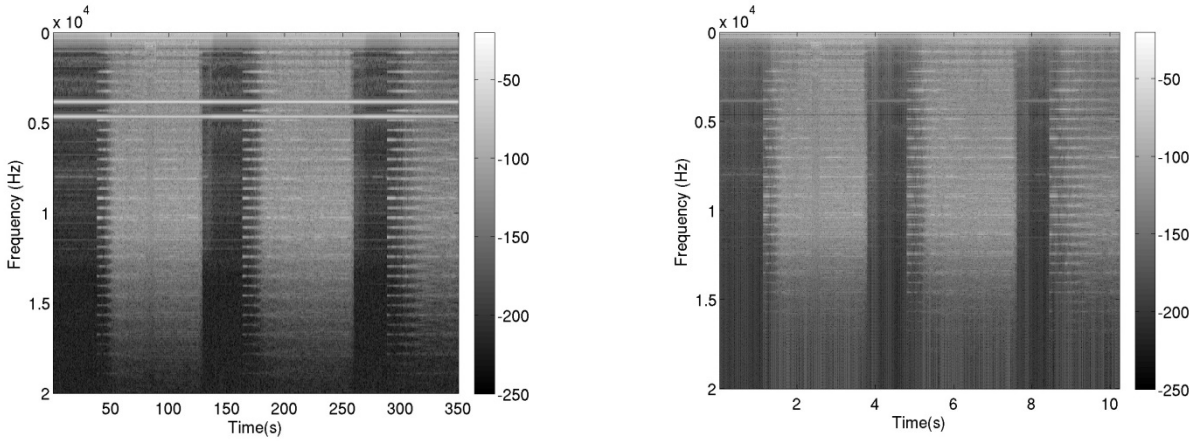
In order to test the performance of our methods, we focused on a particular test run where noise was maximized through a simulated ground fault. The matched-filter output and locomotive speed for this data set are shown in Figure 32. As the train shifts from idling to motion, there was a change in the noise level and the signal level, with two of the signals increasing in amplitude and two decreasing in amplitude.



(a) (b)

Figure 32: Test data set for SNR-estimation example. (a) Matched-filter outputs. (b) Locomotive speed

In order to estimate the properties of the noise, we first removed the components of the recorded data waveforms in the "signal subspace", which was composed of sine and cosine waveforms at the two excitation waveforms. This procedure is illustrated in Figure 33, where the excitation signals were removed (both excitation frequencies are seen only in the left spectrum map) by projection onto the orthogonal complement of the signal subspace.



(a) (b)

Figure 33: Spectrum of test data set including excitation waveforms (a) and with components in the signal subspace removed (b)

Since the noise could not be measured without the signal present, we assumed that the noise spectrum is spectrally continuous and then we interpolated the spectrum that should be seen due to noise only in the signal sub-band. This is accomplished using a parametric autoregressive model, but other non-parametric approaches are also possible (for example, smoothing the spectrum with a Gaussian kernel.)

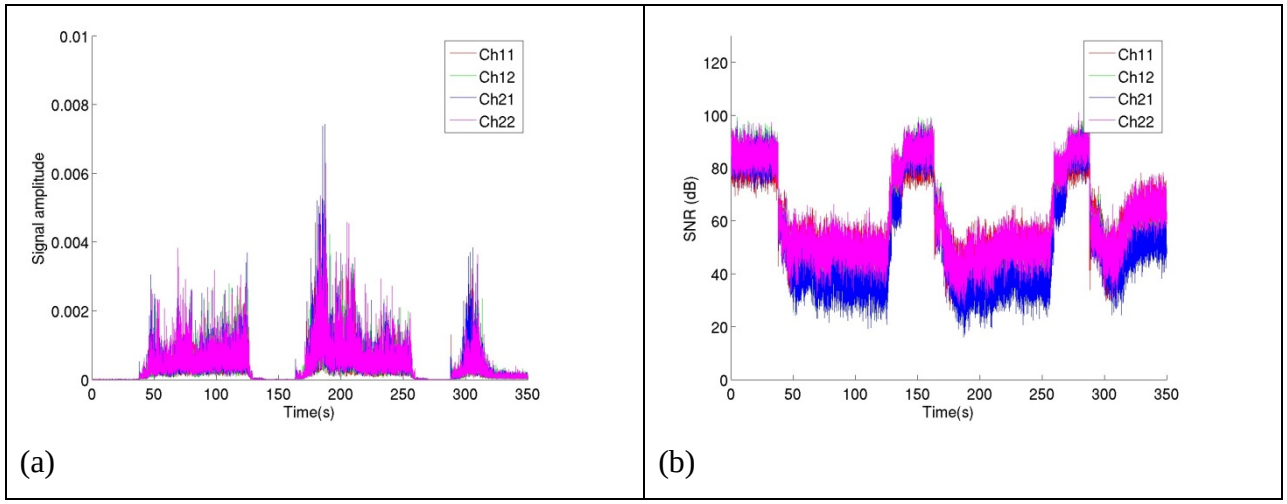
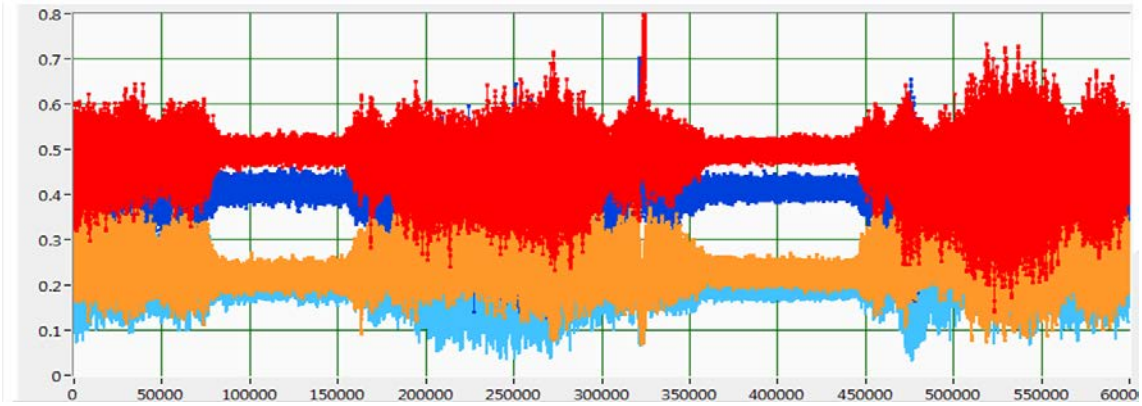


Figure 34. Simulated noise based on autoregressive model (a) and SNR (b)

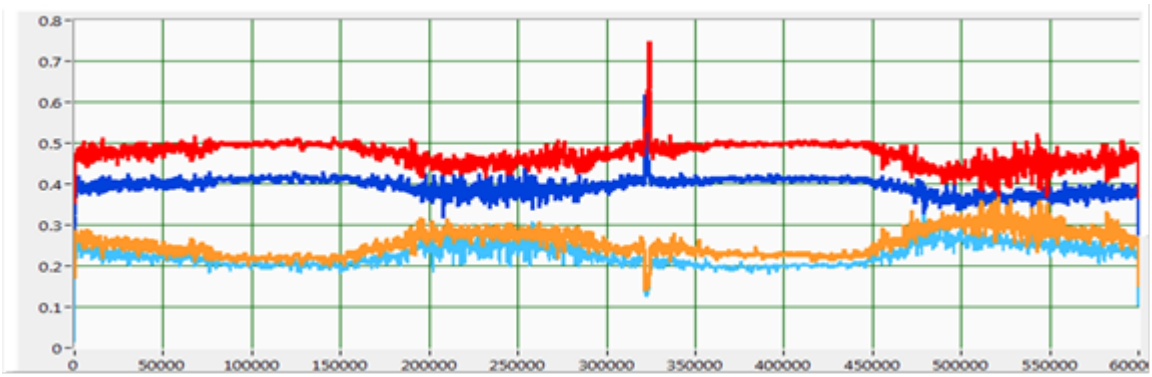
Finally, in order to estimate the SNR, we generated simulated noise by filtering white Gaussian noise with variance (as estimated by the autoregressive model of the specified order) and then ran this simulated data through the matched filter. Then the SNR can be generated by computing 20 log10(S/N), where S is the actual measured signal and N is the result of matched-filtering of the simulated noise data. The results of this procedure are shown in Figure 34, where the SNR for all channels was best during the idle condition and it approached a minimum of approximately 30 dB during motion.

4.1.3 Locomotive Noise Reduction Approaches

Since the receiving signals still have significant presence in the locomotive electromagnetic noise after the application of the match filter or direct demodulation, other noise suppression techniques have been investigated. An example of processed data during DB8 is presented in Figure 35(a). A short processing time of 3.2 ms (128 points of the digitized raw signal) yields high resolution signal profiles that depict railroad feature signatures in great detail. However, the amount of signal variations due to the noise remains high. For example, a hardwire shunt that is present in the time window shown in this plot is hidden under the noise.



a) Signals from two receivers after passing through the matched filters



b) The same signals after application of a median filter

Figure 35. Data during dynamic braking DB8

Increasing the amount of low-pass filter that is applied to the signal—which is obtained after the matched filter is employed or a longer processing time is used—would remove a significant portion of the random noise. A downside of this approach is that a signal which represents a particular rail feature can be corrupted. Particularly, the sharp changes of the signal (as in a case of a broken rail) can be smoothed, which in turn might cause a miss for the feature defection algorithm.

Alternatively, a median filter can be used with a better effect: while it reduces the noise, the sharp changing signal features remain intact if the processing window remains relatively short. In order to reduce the noise from the signals presented in Figure 35 (a), a median filter with 480 points sliding window was applied. The resulting signal (Figure 35 (b)) has suppressed noise and reveals the hardwire shunt signature.

4.2 Signal Signatures of the Track Features

The RIAS detection system is designed to detect rail breaks during locomotive operation, and when it passes over a break, the system provides a consistent, expected response in dual-tone signals induced in the rail. However, many types of common track artifacts in railroads are also capable of causing a response in the RIAS signal. Artifacts that affect the wheel-to-rail shunt quality (e.g., rusty rails or wayside lubricators) or affect the impedance of the loop

between axles (e.g., unbonded rail joints) can produce a response in the RIAS signal. Many of these track artifacts generate responses in the RIAS excitation signal that could potentially be interpreted as a break. A reliable system for break detection should reject all potential false positives and, when possible, identify the track feature which caused the false positive. Below, we discuss broad classes of track features that can cause false alarms for a break detection system.

As discussed in Section 3.1.2, GE collected approximately 1200 miles of track data on a Class I railway (also known as the "CL1" data set). This data set included information about the location and type of many common track features encountered along the railway. The track features collected in the CL1 data set can be organized into several broad types, including insulated joints (IJs), automatic equipment identification (AEI) reader signals, unbonded joints, poor shunting, and crossings (XING). Table 5 lists all of these features and includes counts of each feature type; it also subdivides insulated joints and poor shunting artifacts according to ancillary information about their location and type. Of the 168 poor shunting artifacts, for example, 137 occur at wayside lubricator stations. Insulated joint pairs were also divided according to whether the distance between the staggered joint pair was of standard length (i.e., between 32 and 56 inches).

Table 5. Breakdown of broad track feature types in the CL1 data set

Feature Type	Count	Location (Count)
Insulated Joints	909	Diamond (16) IntSig (222) IxlSig (538) Joint (1) OS (52) Repeater (5) Switch (99) Yard (10)
AEI reader	14	
Joint	10	
Poor Shunting	168	Wayside Lubricator (137) Poor Shunting (31)
Railroad Crossing (XING)	9	

Figure 36 provides the representative plots of matched filter amplitudes for the track features in Table 5. Each of these track feature types affects the rail impedance in such a way that creates a temporary drop in the estimated matched filter amplitude. As a result, all of these track feature types are potential sources of false alarms for a broken rail detection system.

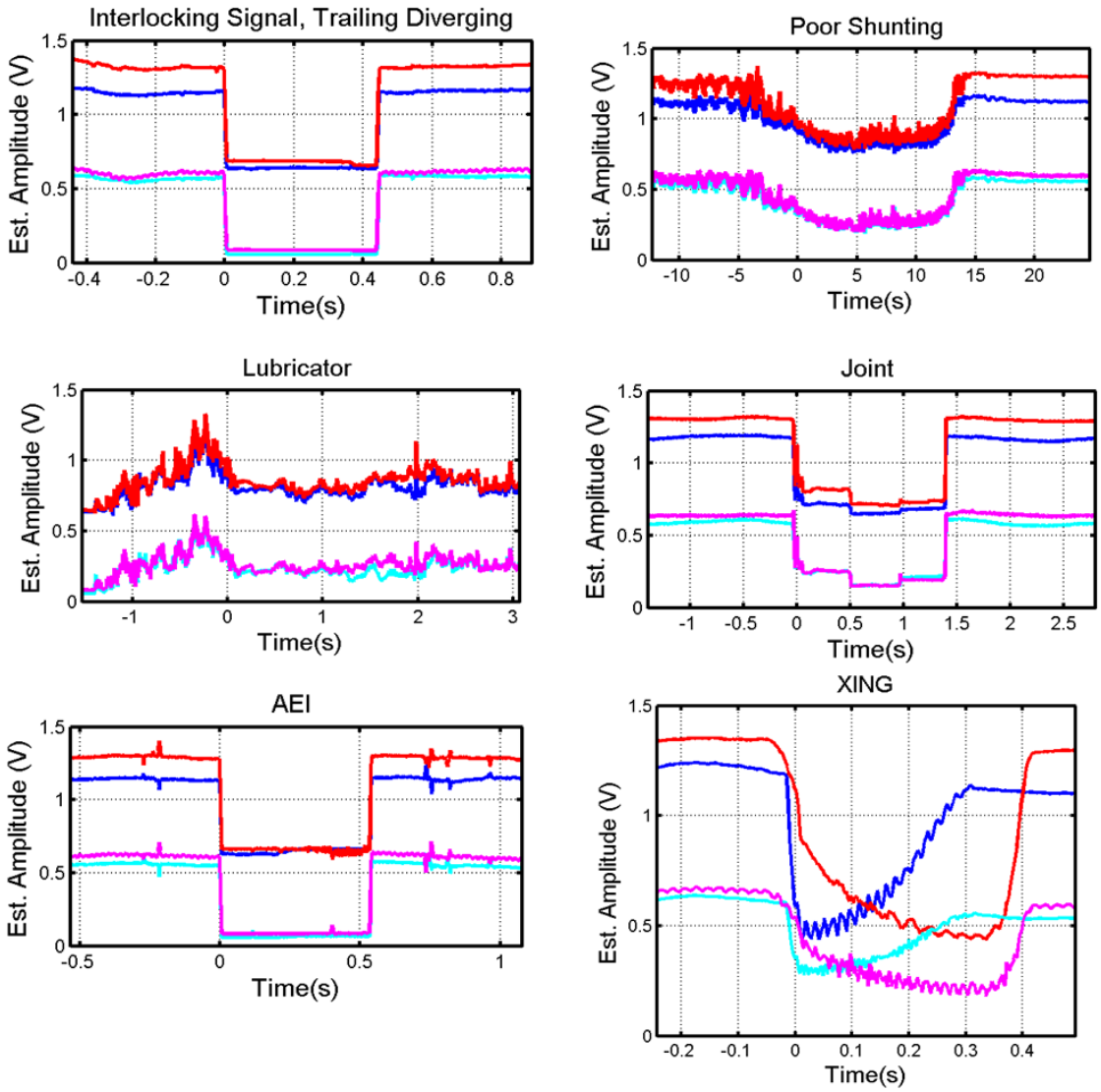


Figure 36. Representative plots of matched filter amplitudes for an insulated joint (upper left, "Interlocking Signal") and 5 common track artifact types

4.3 Establishing and Testing a Track Feature Recognition Algorithm

Earlier we evaluated an approach that uses thresholding on the confidence score and checks if the width of the break state is inside of an acceptable range. While broken rails can be identified by thresholding on voltage drop, trough duration, or other easily measurable signal characteristics, the results in Section 3 show that this approach is prone to high true-negative (miss) and false-positive (false alarm) error rates.

As this project progressed, we proposed a more effective approach that detects broken rails and significantly reduces false positive identifications. In this approach, thresholding is performed as

preprocessing step, then a more sophisticated detection method is used that is based on signal processing, statistical pattern recognition or machine learning.

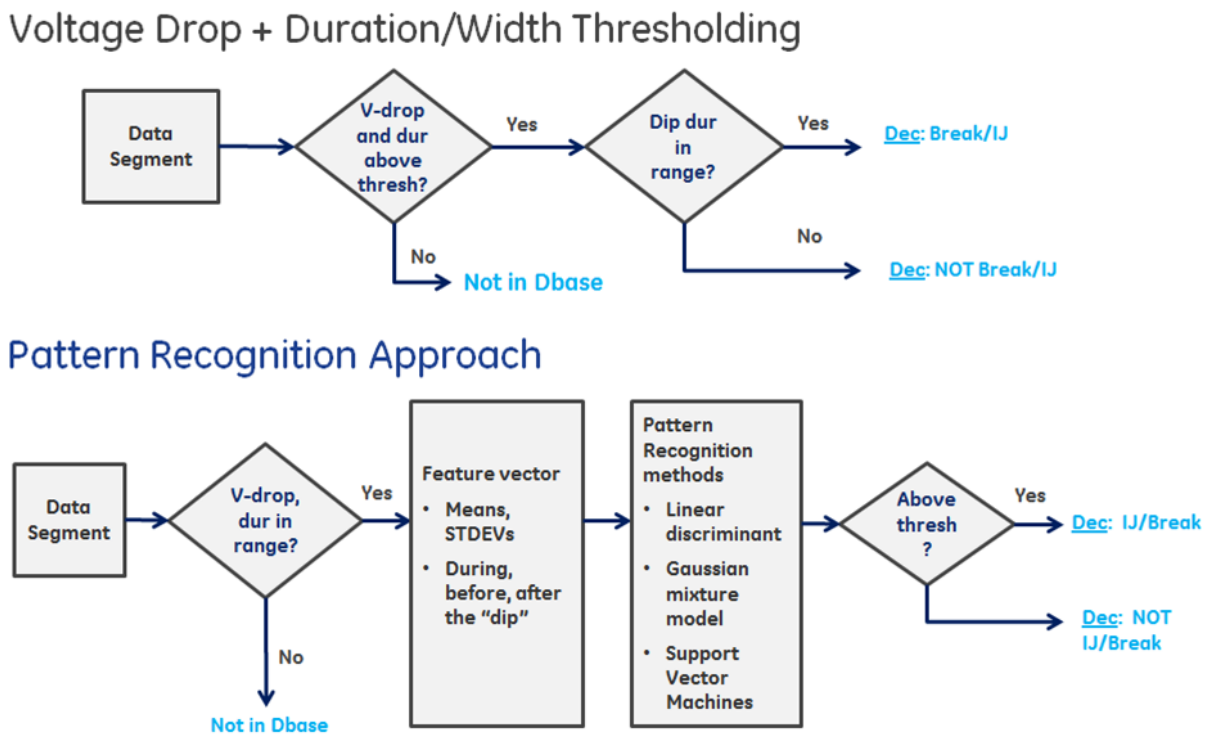


Figure 37. System diagram for break detection based on thresholding (upper) and pattern recognition (lower) approaches

A system diagram illustrating the improved method (also known as the "Pattern Recognition Approach") is provided in Figure 37. After a first-pass detection based on voltage drop or duration is performed, a vector of representative measures or statistics of the signal (i.e. a feature vector) is computed from a segment of the RIAS signal that is in the vicinity of the first-pass detection time. Supervised statistical pattern recognition and machine learning approaches, such as Gaussian mixture models (GMMs) and support vector machines (SVMs), are used to classify the set of feature vectors as broken rails and, potentially, other class types as well.

4.3.1 Feature Vector Representations

As seen above, the formulation of a representative feature vector, i.e., a multivariate representation of the signal segment first identified by thresholding, plays a critical role in the broken rail detection approach proposed in this chapter. The length vector is typically composed of representative measures or statistics of the signal, or the results of multivariate signal transformations.

A feature vector, which represents track events detected by thresholding, should effectively characterize broken rails and insulated joints to distinguish them from other railroad features such as shunts, railroad crossings, wayside lubricators, etc. A plot of the 4 matched filtered amplitude signals for Break #1 of the TTC test loop is shown in Figure 38. As discussed in the Section 3.5.2 the break and insulated joint track features are characterized by a sudden, sustained

drop in the estimated matched filter amplitude voltage (which we call the “break state”). The drop in amplitude is observed on all 4 channels, and it can be detected by thresholding on a confidence score based on the estimated amplitude.

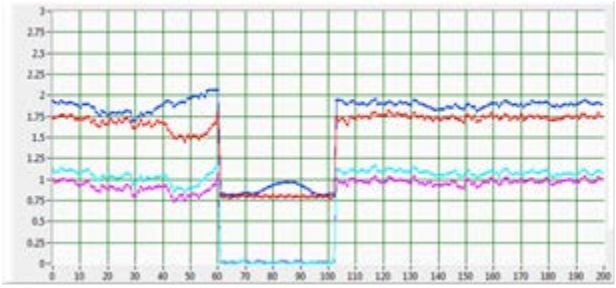


Figure 38. System signals from Break #1

In this section we investigate computing basic summary statistics (such as means and standard deviations) of amplitude values in the break state and also immediately before and after the break state. In all cases, we form all collected measurements and statistics into a single feature vector for each putative railroad feature, first detected by thresholding.

Figure 39 contains a plot of estimated matched filter amplitudes for an insulated joint recorded on 3/3/2014. The insulated joint event in Figure 39 was first detected according to the methods described in Section 3.5. Specifically, a confidence score for the Break State was computed for the “Receiver 1, Frequency 1” channel, and start and stop times for the Break State were estimated based on thresholds listed in Table 4 in Section 3.5.3, under the “Low False Alarm Rate” column heading. The detected break state is indicated in Figure 38 with a green rectangle.

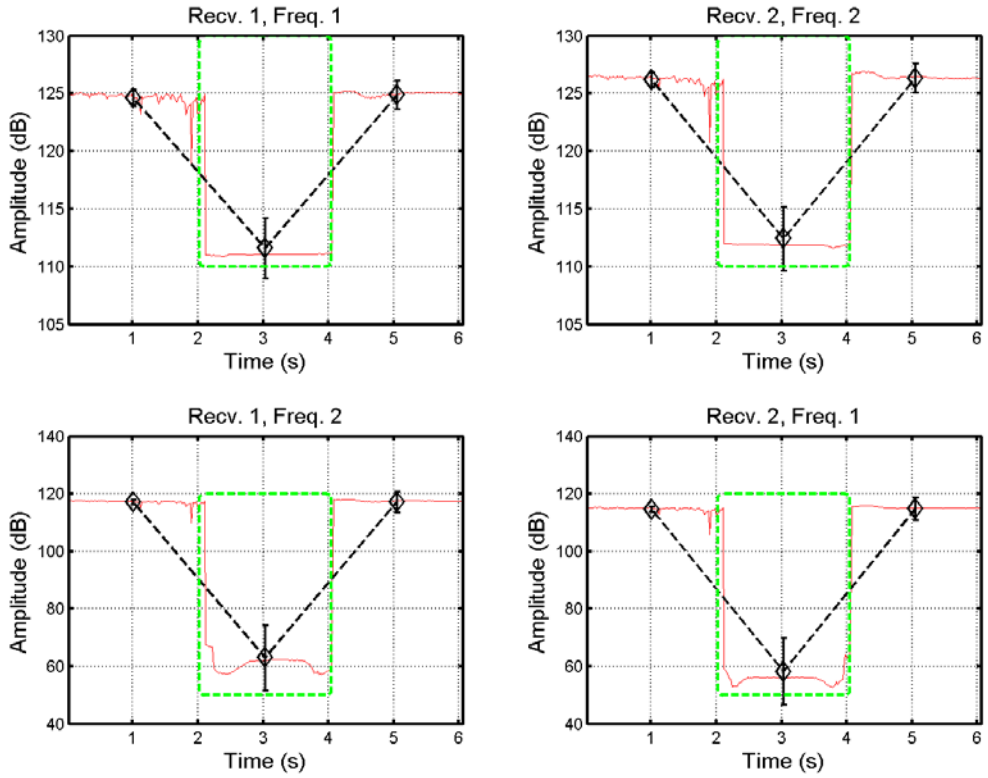


Figure 39. Matched filter amplitude segment for a correctly detected insulated joint. The "Break State" is identified with a green rectangle. Signal mean and standard deviation before, during and after the break state are indicated with diamond markers and error bars

The mean and standard deviation of the matched filter amplitude signal in the break state are also plotted in Figure 39, and it includes black markers and error bars for each channel. Basic statistics are plotted for signal segments immediately before and after the break state as well. A similar plot of matched filter amplitudes is given in Figure 40 for a false positive detection. The likely cause of the artifact in Figure 40 is poor shunting due to either a rusted rail or wayside track lubricators. Comparing basic statistics plotted in Figure 39 and Figure 40, the break state means for Receiver 1, Frequency 2 and Receiver 2, Frequency 1, are significantly greater for the false positive detection (85.3 and 82.3, respectively), than for the insulated joint (63.0 and 58.2, respectively).

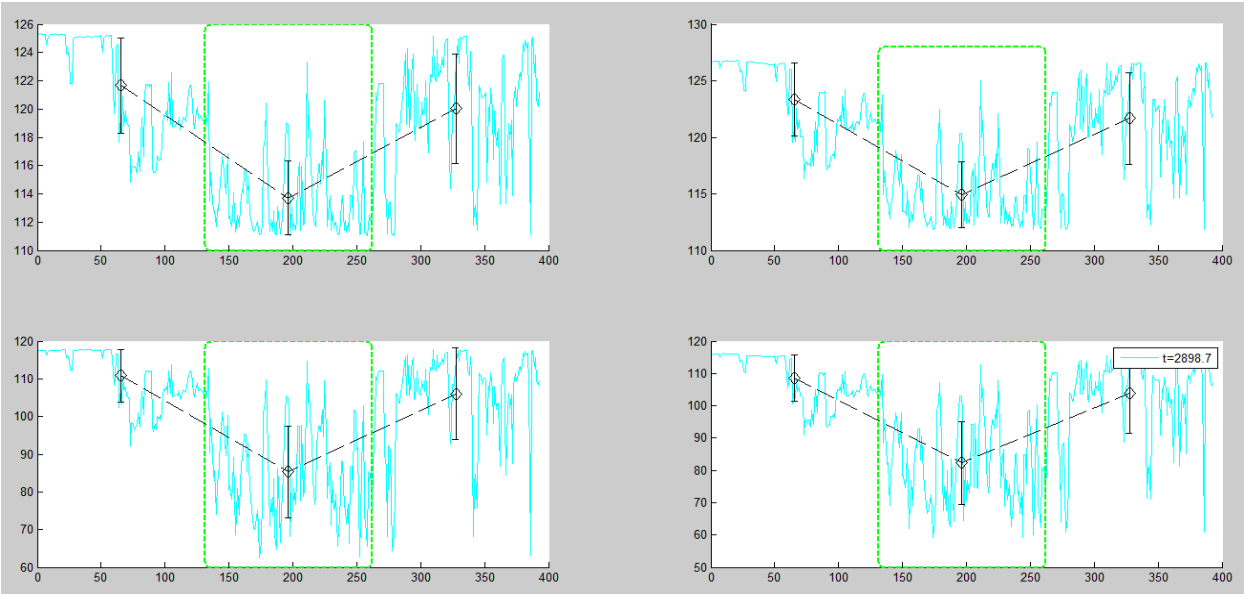


Figure 40. Similar to Figure 38, matched filter amplitudes, break state, and basic statistics of the signal are plotted for a false positive detection

Plots of means and standard deviations of break state computed for CL1 data set for 4 feature types are given in Figure 41.

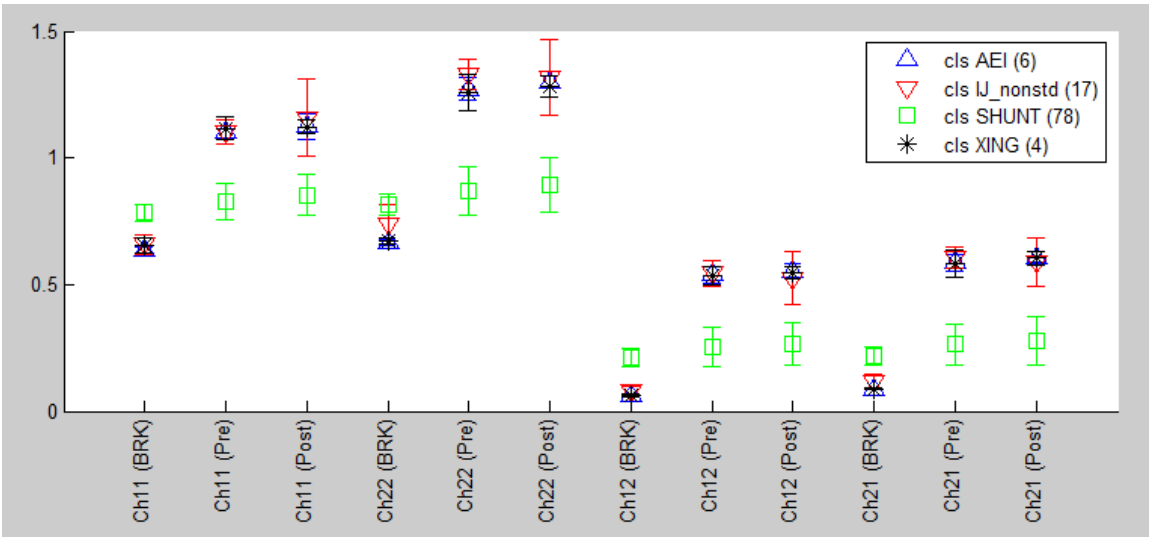


Figure 41. First and second-order statistics for break state means as a feature vector for 4 manually labeled data classes from the CL1 data set

While the single (non-paired) insulated joint (labeled “IJ_nonstd”), AEI, and crossing (labeled “XING”) classes overlap significantly, the class of poor shunting artifacts (labeled “SHUNT”) is well separated from the others. The means of the 4 matched filter amplitude signals before, during and after the break state comprise 12 quantities which can be used to compose a feature vector for pattern recognition.

4.3.2 Support Vector Machine Classifier

SVM classifiers are part of a family of pattern classification methods often referred to as *kernel machines*. Kernel machine methods involve a two-parameter function $K(\boldsymbol{x},\boldsymbol{y})$, where $\boldsymbol{x}$ and $\boldsymbol{y}$ are vector quantities, called a “kernel function,” which produces a 1-dimension result. If the form of the kernel function is chosen correctly, then its result is equivalent to the inner product of $\boldsymbol{\phi}(\boldsymbol{x})$ and $\boldsymbol{\phi}(\boldsymbol{y})$, where $\boldsymbol{\phi}(\boldsymbol{x})$ is one-parameter function that maps an input vector $\boldsymbol{x}$, in what is referred to as the “input space” to a higher dimensional space, referred to as the “feature space.” An example of a kernel function called the “radial basis function” kernel is given below

$$K(\boldsymbol{x}_1,\boldsymbol{x}_2) = e^{-\|\boldsymbol{x}_1-\boldsymbol{x}_2\|^2/\sigma^2}$$

where the parameter σ is often called the “radius” and is typically optimized by experiment. The SVM model itself consists of a vector of weights $\boldsymbol{\alpha}$, whose length, N, is equal to the number of data points in the training set. The training procedure for the SVM algorithm is to find the set of weights $\boldsymbol{\alpha}$ such that

$$\hat{\alpha} = \underset{\alpha}{\arg\max}\ -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j K(\boldsymbol{x}_i,\boldsymbol{x}_j)$$

where $\boldsymbol{y}$ is a vector of labels, also of length N, for the training set, such that $y_i = +1$ for any training vector $\boldsymbol{x}_i$ belonging to the target class (i.e., rail breaks and IJs, for our problem) and $y_i = -1$ for any training vector $\boldsymbol{x}_i$ belonging to the anti-target class (i.e., artifacts due to poor wheel-to-rail shunting). A quadratic programming optimization technique is typically used to optimize this expression and train the SVM model. Finally, when the SVM model is trained, a continuous-valued, 1-dimensional confidence score $z(\boldsymbol{x})$ is computed for an unseen vector $\boldsymbol{x}$ (i.e., a vector not used in the training set) according to the relation below:

$$z(\boldsymbol{x}) = \sum_{i=1}^{N}\alpha_i y_i K(\boldsymbol{x}_i,\boldsymbol{x})$$

where each $\boldsymbol{x}_i$ is a feature vector in the training set.

A decision as to whether a broken rail has been detected in the test vector $\boldsymbol{x}$ is determined by setting a threshold on the confidence score $z(\boldsymbol{x})$. The trade-off between false-positive and true-negative error rates for Stage 2, can be evaluated with the ROC curve.

A comprehensive tutorial on support vector machine classifiers can be found in [1].

4.3.3 Experiments with Classifiers

We compared two well-known pattern recognition methods that can detect broken rails: GMMs and SVMs. The contents of the labeled CL1 data set, as described earlier, contains the 1110 track features listed in Table 5. Originally, these features were identified using a threshold-based detection method on the RIAS signal by experts from GE Transportation and GE Global Research, and the data set is composed largely of insulated joints and other track features which are most likely to cause false-positive errors for broken rail detection algorithms. We evaluated our classification methods with these previously identified data points.

Feature detection experiments were done by randomly selecting training and test sets consisting of approximately equal counts of each broad feature type. The training and test set breakdown for experiments in this section is given in Table 6. The training and test sets consist of 461 and 466 insulated joints as well as 17 and 18 non-standard insulated joints, respectively. These are single (non-paired) or non-standard stagger width joints that closely resemble real broken rails.

Table 6. Training and test set breakdown for CL1 experiments

Feature Type	Training Set	Test Set
Insulated Joints	461 (17 non-standard IJs)	466 (18 non-standard IJs)
Auto Equipment ID	6	6
Unbonded Joints	5	5
Poor Shunting	79	79
Crossings	4	5
Total	**555**	**561**

SVM and GMM classifiers were trained and evaluated on the CL1 data set as specified in Table 6. In all GMM experiments in this section, we use models with $K = 4$ Gaussian mixture components. All experiments used a 24-dimensional feature vector that consists of means and standard deviations of signal segments that were before, during, and after the break state, for all 4 channels.

Detection results are evaluated with ROC curves, illustrating the trade-off between the miss (or true-negative) error probability and the false alarm (or false-positive) error probability. ROC plots for detecting insulated joints using SVM and GMM classifiers are given in Figure 42. Results are plotted for detecting insulated joints against the other four classes together (i.e., auto equipment IDs, unbonded joints, poor shunting, and crossings) as well as individually. The "area above the curve," (AAC) a scalar metric for evaluating ROC plots, equal to the fraction of the plot area above an ROC curve[1]. SVM and GMM classifiers achieve detection performance of

[1] Depending on the plotting convention, either "area above the curve" (AAC) or "area under the curve" can be used to evaluate performance in an ROC curve.

93.9 percent and 93.1 percent of area above the curve for detecting insulated joints against all other classes together.

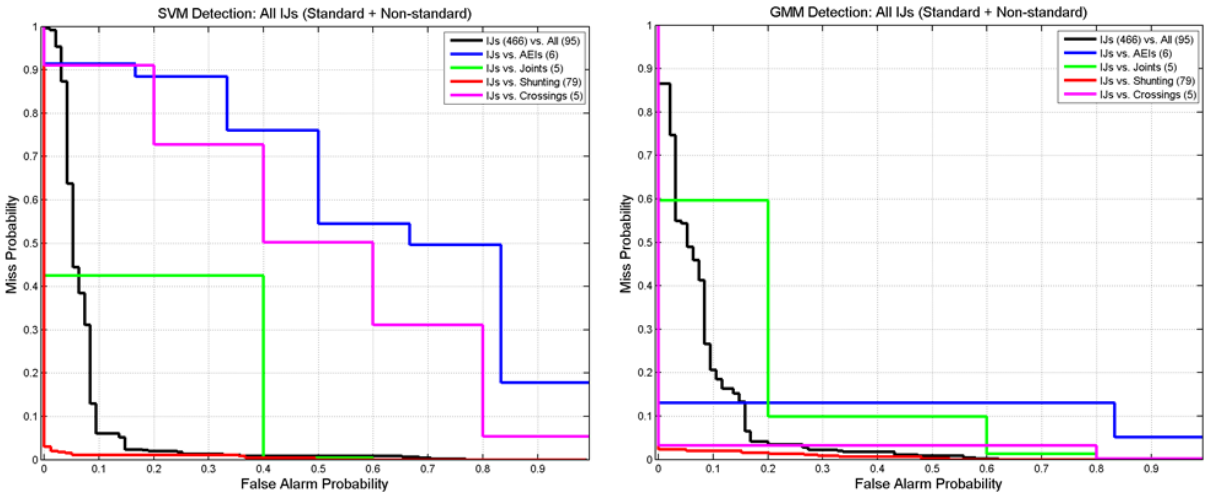


Figure 42. ROC curves for detecting IJs with SVM and GMM classifiers

ROC plots are also given in Figure 41 for discriminating insulated joints against all other classes, individually. These results are intended to demonstrate which track feature types are most easily confused with insulated joints or broken rails. For both SVM and GMM classifiers, detection performance is significantly improved when insulated joints are evaluated against poor shunting track artifacts only. SVM and GMM classifiers achieve 99.6 percent and 99.4 percent of the area above the curve for these experiments. The ability to detect insulated joints against auto equipment identification, unbonded joints, and crossings is significantly worse. As illustrated in Section 4.3.1, all of these classes exhibit very similar responses to insulated joints in the RIAS signal, and are not well-separated from insulated joints with respect to the input feature vector. Furthermore, these three classes each have very few examples in the test set. It should be noted that the ability to discriminate insulated joints from poor shunting features is critical since GPS location data are often unavailable for these track in artifacts in both signal territory and dark territory.

Figure 43 displays the performance for detecting non-standard insulated joints against all other feature type in ROC plots, and Table 7 lists the fraction of area above the curve. Performance for detecting non-standard IJs against all other classes together is comparable to results for the full set of IJs. For only detecting non-standard IJs against poor shunting features, both SVM and GMM classifiers obtain perfect detection performance (i.e., 100 percent of area above the curve).

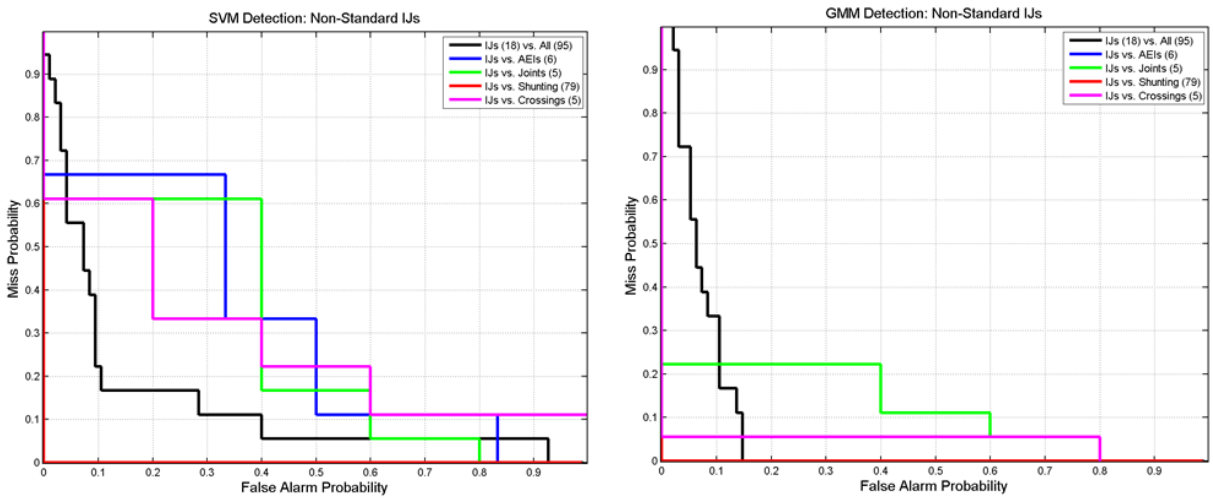


Figure 43. ROC curves for detecting insulated single joints and joints with non-standard stagger widths using SVM and GMM classifiers

Table 7. "Area above the curve" for detecting non-standard IJs in ROC plots in Figure 43

Detection Task	Area above curve	
	SVM	GMM
Non-Standard IJs (18) vs. All (95)	0.949	0.928
CL1-IJs vs. AEIs (6)	0.472	0.500
CL1-IJs vs. Joints (5)	0.711	0.844
CL1-IJs vs. Shunting (79)	1.000	1.000
CL1-IJs vs. Crossings (5)	0.722	0.578

4.4 Summary of the Results

- The procedure for data collection with exaggerated noise produced by the traction motors is described. Preliminary analysis shows that this noise is in-band with the RIAS transmitting frequencies and is not removed by the matched filtering (or direct demodulation)
- The proposed approach of including a median filter into the signal processing procedure provides visible noise reduction while preserving patterns associated with the railroad features
- A summary of various track features is provided with examples of electrical signatures as received and processed by the RIAS

- A track feature recognition algorithm was established and evaluated. The approach uses thresholding as a preprocessing step, followed by a more sophisticated detection algorithm based on signal processing, statistical pattern recognition or machine learning
- Example of discrimination one class of the railroad feature against others are provided and discussed
- The new approach described in this Section is more effective for detecting broken rails, while significantly reduces false positive identifications

5. Algorithm Optimization

In this section, we discuss methods for characterizing the false-alarm rate and reducing it. Our goal was to minimize the false-positive rate (ideally to have a false-positive rate of zero on our test data set) while achieving a missed detection rate of at most 5 percent.

To reduce the false-alarm rate, we collected data from multiple sources, specifically from the HTL at the TTC test facility and from the Class 1 railroad, then trained and tested the classifier on ensembles of data from the two sources. While the performance of a classifier trained on data from one source and tested using data from another railroad is very poor, a training procedure that is enriched with data from both sources improves performance considerably. This is analogous to what has been done in the speech recognition community where there has been effort over time to collect recordings of multiple speakers, both male and female, with various regional accents. We have also examined the performance of alternative approaches to signal normalization.

5.1 Outline of Data Processing

Figure 44 outlines the steps involved in the data processing for broken-rail detection. The first step is "front-end" processing, where the raw data, sampled at a rate of 40 kHz in our prototype system, are filtered at the two excitation frequencies.

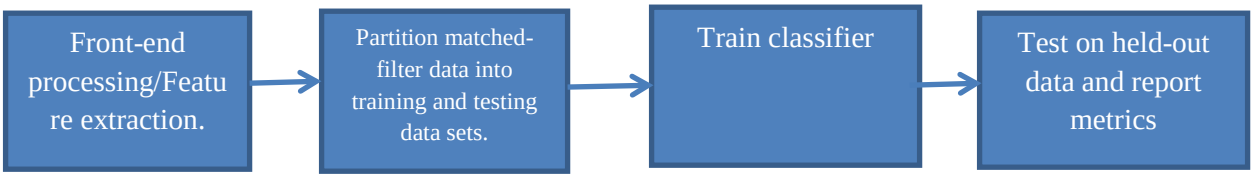


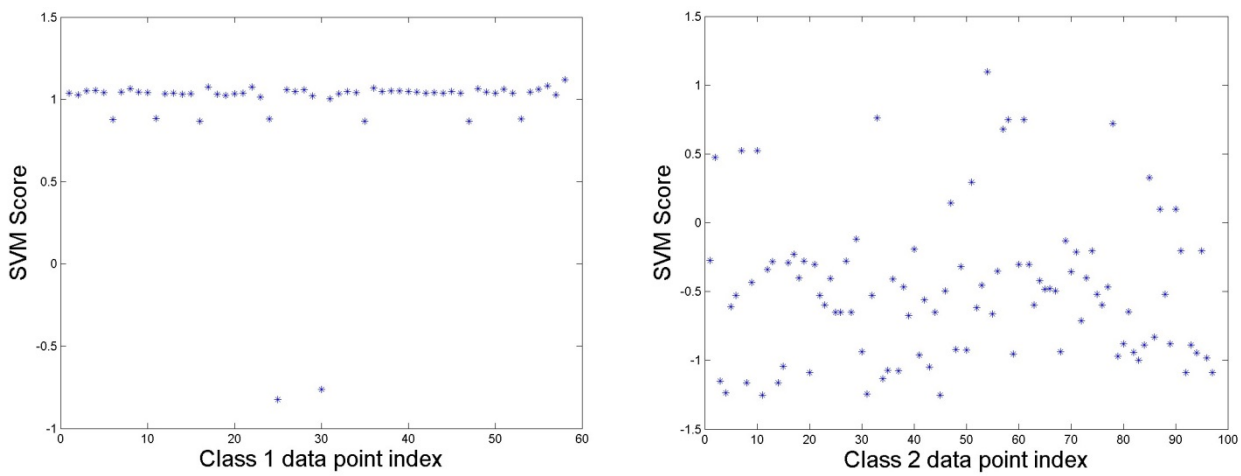
Figure 44. Block-diagram of procedure for classifier training and classification

The goal of the classification procedure is to separate physical track artifacts such as insulating joints and breaks from the variations in amplitude due to other causes. The classification procedures that we are utilizing are within the "supervised classification" paradigm. In this approach, each detected dip is manually assigned a class label based on its recorded GPS location, correlating to known track features such as insulating joints, signals, crossings, lubricators, among others, though the vast majority of the detected track artifacts were insulating joints. If a dip was detected in the absence of a nearby known track artifact, it was assumed to be a false-positive detection.

In order to evaluate the generalization performance for unseen data, we partitioned the labeled data samples into training and test sets, set the parameters of a classifier using the training data, and validated its performance with the held-out test data that were not seen in the training phase. In reality, the data were typically partitioned further on in the training procedure, since many classifiers had associated hyper-parameters related to their complexity and ability to generalize, and a cross-validation procedure is often beneficial when the values of these hyper-parameters are set. For example, in most of the work described in this report, we show results for a SVM classifier, which has associated with it parameters such as the number of support vectors to

employ, the “width” or variance of each radial basis function, and the regularization parameter. In the training procedure, the training data are portioned into random folds where we optimize the hyper-parameters by maximizing their performance on a cross-validated classification task.

In the final outcome of the training procedure, each feature vector is assigned a score and the classifier parameters are set to maximize the ability to distinguish between the two classes with a simple threshold. An example illustrating the SVM scores generated for instances of the two classes is shown in Figure 45, where the SVM scores for Class 1, comprising the true insulating joints, is shown in Figure 45 (a), whereas the scores for the “anti-target “ class comprising examples of noise, track lubricators, and other causes of false-positive dip detections are illustrated in Figure 45 (b). It is striking that, to paraphrase Leo Tolstoy, all insulating joints are alike, but each source of noise is noisy in its own way. Almost all of the true insulating joints are given a score very close to 1 with a number of outliers which could, for example, correspond to labeling errors while there is a very large spread to the scores for the noise examples, with only a few of these examples being assigned scores near or above a value of 1.



(a) (b)

Figure 45. SVM scores for a typical test data set. (a) True insulating joints (b) Noise, lubricators, etc

Therefore a threshold of just below 1.0 would detect almost all of the true insulating joints, with two unavoidable missed detections, while there is one false-positive that would appear to be very difficult or impossible to eliminate without missing many of the true detections.

5.2 False alarm reduction and algorithm verification

Our main approach to reducing the false-positive detection rate has been to enrich the training and testing data with examples from each of the two railroads from which we have collected data. Specifically, we have examined the following training and testing conditions:

1. Train using HTL data, test on CL1 data
2. Train using CL1 data, test on HTL data
3. Train on both databases, test on both databases
4. Train on both databases, test on HTL
5. Train on both databases, test on CL1
6. Train using CL1 data, test using CL1 data
7. Train using HTL data, test using HTL data

For each of these cases, we have randomly drawn ensembles of training and test data which were necessarily overlapping due to the small size of the data sets after dip-detection. In order to assess the performance of our classifiers trained with the random ensemble data, we have computed the missed detection rate and false-positive rate on the testing data for a finely sampled set of thresholds and have reported the minimum missed detection rate in the region where the false-positive rate was identically zero. The results for this minimum miss rate analysis for the seven listed cases are shown in Figure 46.

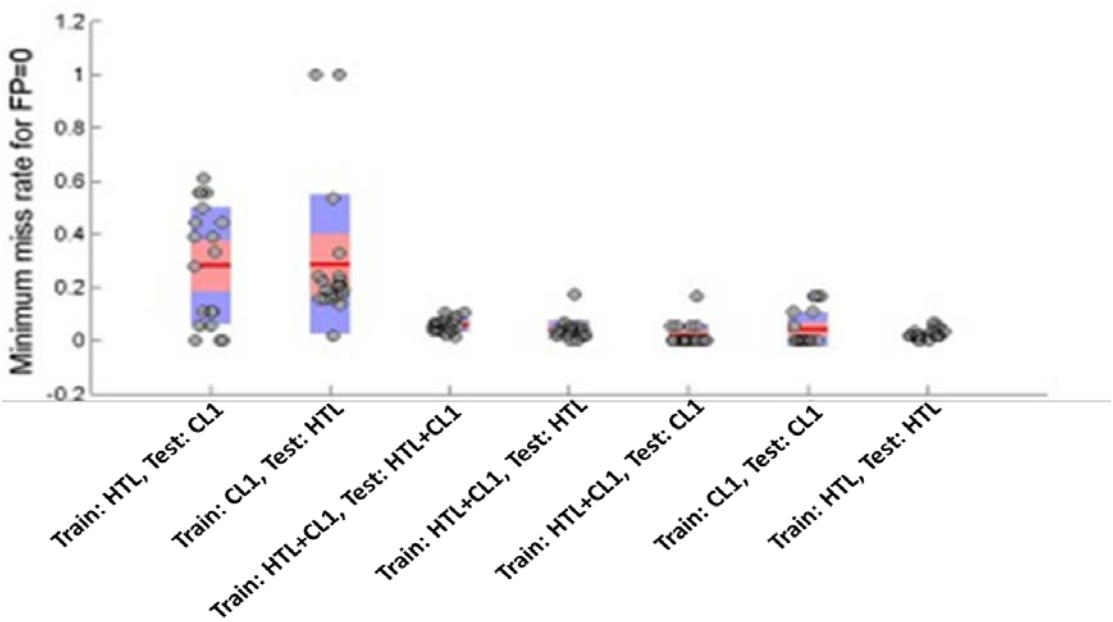


Figure 46. Minimum missed detection rate when the false-positive rate was zero for the seven train-test configurations

Each small dot represents the performance for a particular ensemble. We see that the performance for mismatched cases where we train the classifier using data from one railroad and test on data from the other railroad, not particularly surprisingly, is quite poor. Conversely, we see that much better accuracy can be achieved when we combine data from the two sources. A closer examination of the statistical performance of each classification condition is shown in Figure 48 (a) and the threshold needed to achieve a minimal miss rate for a zero false-positive rate is shown in Figure 47 (b) for each random ensemble.

(a)

(b)

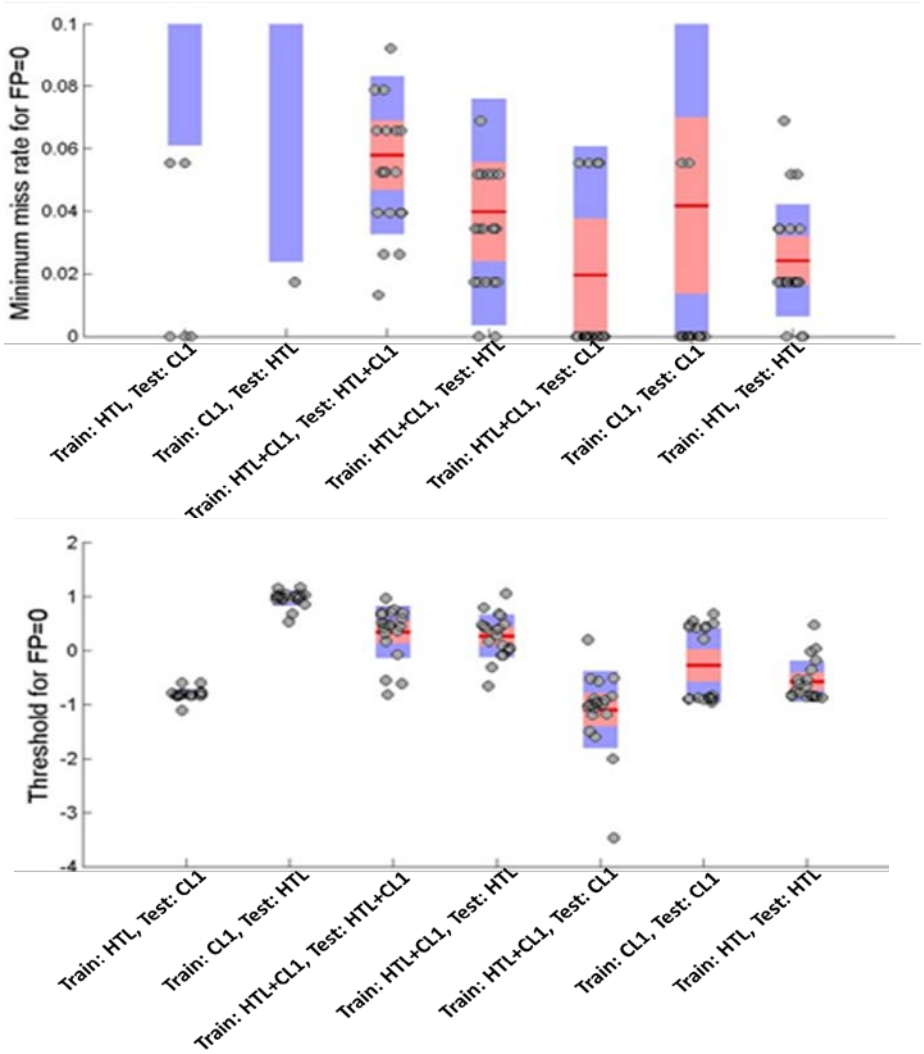


Figure 47. Closer look at the minimum miss rate for a false-positive rate of zero for the seven conditions (a) and the threshold achieving a minimal miss rate and zero false-positive rate (b)

As indicated by the example above, a threshold of 1.0 will conservatively optimize our detection criteria for most of the cases of interest. In some cases, when training occurs on both of the databases and testing only occurs on the CL1 data, a lower threshold appears to be possible. This particular case is very important for the verification as it includes training on a large range of feature signatures and a test on the actual railroad. The mean value for the missed rates in this case is at 2 percent (red line in the chart) while overall variations are below 7 percent.

5.3 Signal Normalization

The RIAS detection system is designed to be implemented and deployed in a variety of locomotive types, different geographic locations, a wide range of sensor design specifications,

and other environmental and operating conditions. Physical differences between locomotive installations can lead to variations in the signal levels that are observed at the receiver. While these effects can be mitigated with careful calibration at setup time, it is still difficult to ensure consistent signal levels across different conditions.

Our method for detecting broken rails involves computing a vector of representative measures or statistics of the signal, referred to as a feature vector, in the vicinity of a p rail break. Variation in feature vector values, due to changes in observed signal levels at the receiver in changing conditions, can negatively impact the performance of pattern recognition methods for broken rail detection, especially when new, unseen data are mismatched with previously trained models. Plots of mean feature values with one-STDEV error bars are given in Figure 48 for the CL1 and HTL databases, for the 12-dimensional feature vector method described in Section 4.3.1.

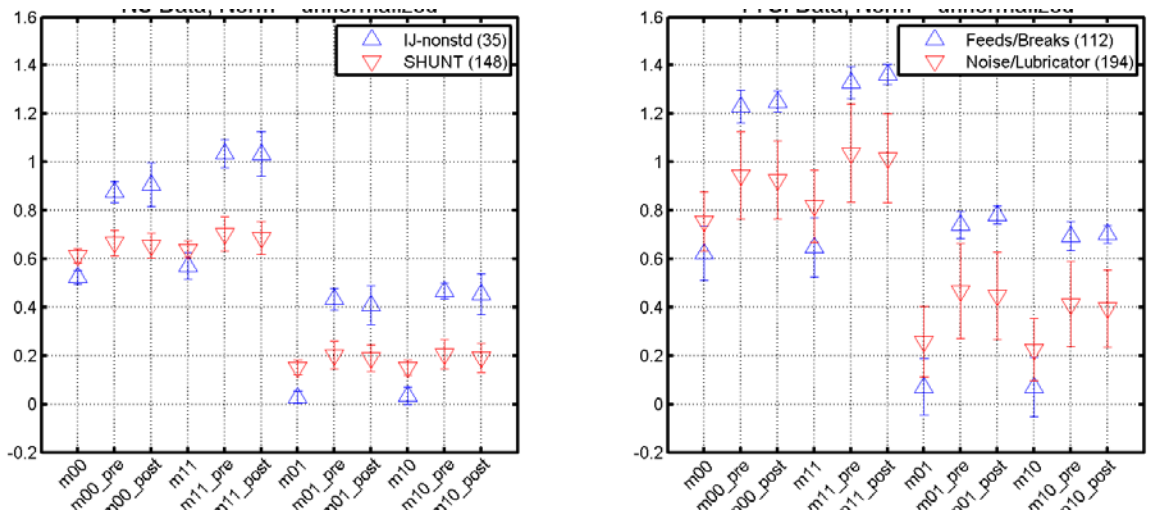


Figure 48. Means with one-STDEV error bars for target and anti-target classes on CL1 and HTL data sets

Mean feature vector values for the HTL database are consistently and significantly greater than corresponding values for CL1 for most feature dimensions. The feature dimensions labeled "m00_pre" and "m00_post", for example, have mean values for the target class (i.e., the "IJ-nonstd" class for CL1 data and the "Feeds/Breaks" class for HTL data) of 0.88 and 0.91, respectively, while the corresponding mean values for the HTL database are significantly higher at 1.23 and 1.21.

5.3.1 Normalizing by the Long-Term Median

Plots of the four estimated tone amplitude signals for one insulated joint taken from the CL1 databases is given in the upper panel of Figure 48.

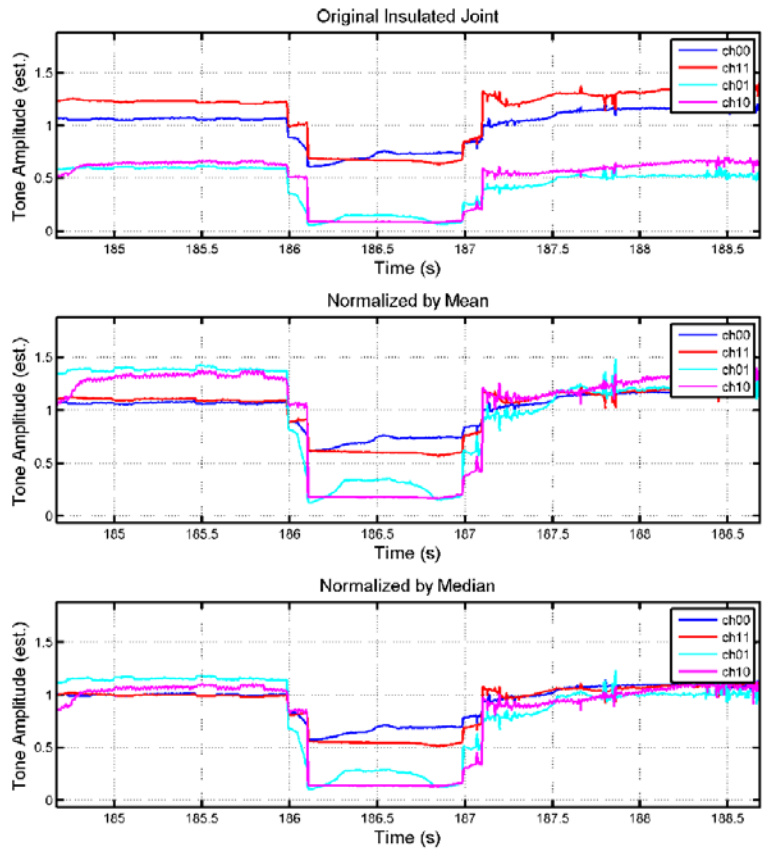


Figure 49. An insulated joints from the CL1 database, non-normalized (upper panel) and after applying mean- and median-based normalization (middle and lower panels, respectively)

An effective way to normalize tone amplitude signals is to apply a scaling such that the nominal amplitude (i.e., the signal amplitude in the absence of response-inducing track artifacts such as broken rails, insulated joints, rail shunts, poor wheel-to-rail shunting, etc.) is unity for each of the 4 signals. To achieve this, we estimate the nominal signal level by computing either the mean or median of each signal over a long window, and scaling the signal by its inverse. Plots of the tone amplitude estimates after scaling by the mean and median are given in the middle and lower panels, respectively, of Figure 49.

5.3.2 Scaling Feature Vectors to Unit Length/Norm

While median and mean normalization are effective, computing the median and mean values require that a long window be maintained, and real-time implementations find this difficult since extra storage space and computation are required. A common preprocessing step in multivariate pattern recognition approaches is to normalize each feature vector independently according to its length or $l_p$norm in feature space.

For an n-length vector $\boldsymbol{x}$, its $l_p$norm $\|\boldsymbol{x}\|_p$ is defined as

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$

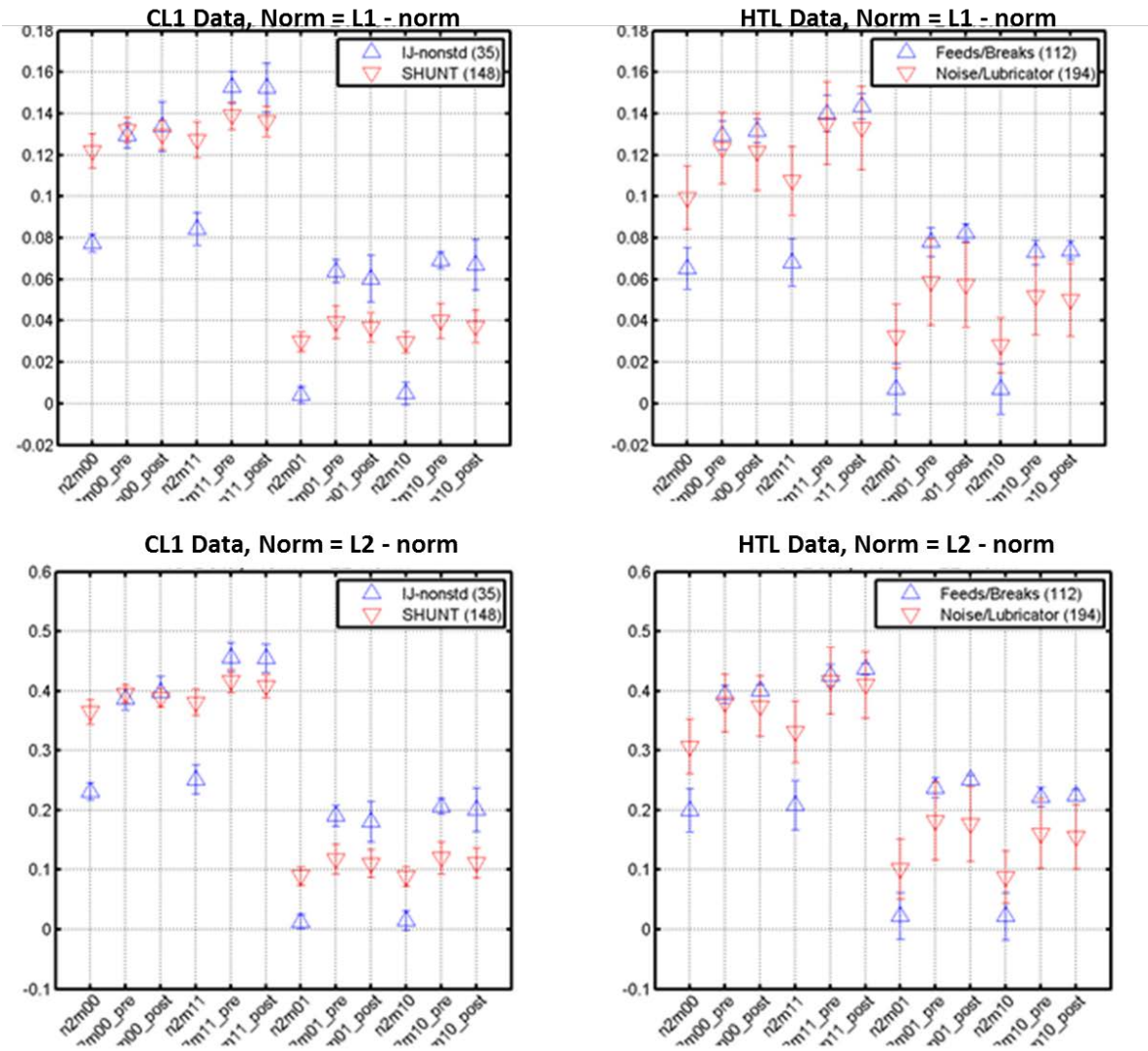


Figure 50. Means and one-STDEV errorbars for CL1 and HTL data sets after feature normalization with and norms

In this section, we analyze the l_1 and l_2 norms, which are commonly used for normalization in pattern recognition approaches. Note that $\|\boldsymbol{x}\|_1$ and $\|\boldsymbol{x}\|_2$ are equal to the sum and root-mean-square, respectively, of the vector $\boldsymbol{x}$. Mean and one-STDEV error bar plots for the CL1 and HTL databases are given in Figure 50.

As with median and mean normalization, normalizing feature vectors to unit length results in a similar range of feature values for the CL1 and HTL data sets.

5.4 Summary of the Results

- The overall signal processing procedure, which was verified on the field test data that was collected at multiple sites, was described in this section
- The performance of SVM as the feature recognition algorithm of choice, and its success rate for detecting broken rail on the field trial data, was discussed
- The training procedure for SVM classifier is enriched by including data from CL1 and HTL sites. A much better accuracy of feature discrimination is achieved when we combine data from the two sources. For example, we achieved a false alarm rate of zero and a miss rate of about 2-6 percent on the combined HTL/CL1 data sets
- To further reduce the apparent condition mismatch between these two data sets, we attempted to improve detection performance across mismatched conditions by investigating two methods for normalizing data: normalizing by a mean or median and normalizing each feature vector by its norms
- A normalization based on the l_p norm was introduced and successfully applied to reduce the inter-variations of matched filter responses across different data sets

6. Conclusion

GE has been developing RIAS, a non-contact on-locomotive system that detects and monitors rail artifacts, particularly broken rails. Several prototype RIAS systems were built, installed on locomotives, and tested. RIAS not only stored information about the voltages and currents measured on the inductive coils, but it also stored data that was related to train speed and position from GPS receivers, and this data made it possible to correlate received signals with known track artifacts, such as insulating joints, crossings, and signaling equipment.

This report presented a systematic analysis of the RIAS, beginning with scale models tested in the laboratory with known impedances and ending with full-scale tests involving an inspection train from NS and test trains at the TTC test facility in Colorado.

While the scale model and initial locomotive tests showed the intuitive feasibility of the approach, this project focused on developing a machine-learning-based method for reliably detecting broken rails with low probability of error. This approach is based on an initial filtering of the data to obtain a single time-trace for each receiver-frequency pair, followed by a hard-thresholding of the signals to identify candidate break events. The final step of the algorithm was to train an SVM classifier on labeled data given a known track database and optimizing the hyperparameters of the classifier (specifically, the Gaussian kernel width, the regularization parameter, and the threshold) to minimize the cross-validated error rate.

The performance of the classifier was improved by combining data sets from different locomotives and different signal normalization approaches. For the collected field data, the team was able to achieve a false-alarm rate of essentially zero with a missed detection probability of less than 5 percent, whereas the purely threshold-based system, without the additional support vector machine, could only achieve a missed detection rate of approximately 26 percent.

Finally, the team has tested real-time signal generation, filtering, and data storage on a prototype LabView-based system. The developed broken rail detection approach will be further verified on larger sets of data to be collected on the Class 1 railroads as GE Transportation progresses with product validation.

7. References

1. Bidigare, P., & Nayeri, M. (2003). RF Tags: Radar as a Communications Channel. ASAP. Lexington, MA.

2. Dick, C. T., Barkan, C. P., Chapman, E. R., & Stehly, M. P. (2003). Multivariate statistical model for predicting. Transportation Research Record: Journal of the Transportation Research Board, 1825(1), 48-55.

3. Railroad Safety Statistics. (n.d.). Federal Railroad Administration Office of Safety Analysis Website. Retrieved from http://safetydata.fra.dot.gov/OfficeofSafety

4. Schafer, D. H. (2008). Effect of train length on railroad. Master's Thesis, University of Illinois at Urbana-Champaign.

5. Akaike, Hirotugu. (1974). A new look at the statistical model identification. Automatic Control, IEEE Transactions on 19.6, 716-723.

Abbreviations and Acronyms

AEI	Automatic Equipment Identification
DAC	Digital-to-analog converter
FIR	Finite-Impulse Response
GMM	Gaussian Mixture Models
GPS	Global Positioning System
HTL	High Tonnage Loop
IPC	Industrial personal computer
NS	Norfolk Southern Corp.
RIAS	Rail Integrity Alert System
ROC	Receiver Operating Characteristic
SVM	Support Vector Machine
TTC	Transportation Technology Center

.