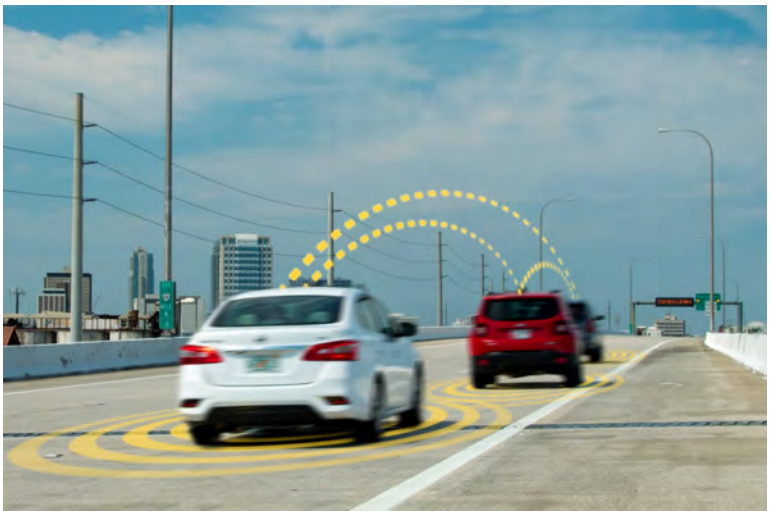
Connected Vehicle Pilot Deployment Program Performance Measurement and Evaluation – Tampa (THEA)

Phase 3 Evaluation Report

www.its.dot.gov/index.htm

Final Report – March 5, 2020
Publication Number: FHWA-JPO-20-829



Source: Tampa Hillsborough Expressway Authority, 2019

U.S. Department of Transportation

Produced by Tampa Hillsborough Expressway Authority (THEA) CV Pilot Team
U.S. Department of Transportation Intelligent Transportation Systems Joint Program Office (ITS JPO)

Notice

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-JPO-20-829	(Delete and insert information here or leave blank)	(Delete and insert information here or leave blank)

4. Title and Subtitle	5. Report Date
Connected Vehicle Pilot Deployment Program Performance Measurement and Evaluation– Tampa (THEA) CV Pilot Phase 3 Evaluation Report	March 5, 2021
	6. Performing Organization Code (Delete and insert information here or leave blank)
7. Author(s) Sisinnio Concas, Achilleas Kourtellis, Mohsen Kamrani, and Omkar Dokur Center for Urban Transportation Research, University of South Florida	**8. Performing Organization Report No.** (Delete and insert information here or leave blank)
9. Performing Organization Name and Address Center for Urban Transportation Research University of South Florida 4202 E. Fowler Ave., CUT 100 Tampa, FL 33620	**10. Work Unit No. (TRAIS)** (Delete and insert information here or leave blank)
	11. Contract or Grant No. (Delete and insert information here or leave blank)
12. Sponsoring Agency Name and Address ITS Joint Program Office 1200 New Jersey Ave., S.E., Washington, DC 20590	**13. Type of Report and Period Covered** Final Report
	14. Sponsoring Agency Code (Delete and insert information here or leave blank)

15. Supplementary Notes

Govind Vadakpat, Tampa CV site Agreement Officer's Representative and Sarah Tarpgaard, Agreement Officer

16. Abstract

This report presents the methodology and results of the performance measurement and evaluation of the Tampa Hillsborough Expressway Authority (THEA) Connected Vehicle Pilot Deployment. The study evaluated the deployment of CV technology in six use cases to address mobility and safety issues in specific areas of downtown Tampa, Florida. The study implemented a robust experimental design to enroll more than 1,000 private citizens to experience the functionality and impact of several vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) applications via a Human Machine Interface (HMI) installed in the vehicle's rearview mirror. The panel of participants was split into treatment (HMI enabled) and control (HMI disabled) groups using a randomized two-to-one matching (two treatment to one control) based on key socio-demographics. The study analyzed behavioral responses to the V2V and V2I applications using a before-after assessment and interrupted time-series analysis. The methodology allowed assessing the performance of each application with respect to each use case by estimating false positive, false negative, true positive, and true negative rates. Overall, findings provide evidence that the deployment contributed toward enhancing the mobility and safety of travelers and pedestrians. The analysis also shows heterogeneity in how the V2V and V2I applications performed in the six use cases under evaluation. This report details the analytical approach employed to evaluate the use cases and provides a discussion of limitations and main lessons learned from the deployment.

17. Keywords	18. Distribution Statement
Vehicle safety, mobility, connected vehicles, performance evaluation, V2V applications, V2I applications	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	214	

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

Acknowledgements

The Performance Measurement and Evaluation Support team would like to thank the entire THEA CV Pilot team for their contributions in successfully deploying the Pilot and their collaboration in the performance evaluation of this study. The authors thank USDOT for the timely, valuable, and constructive feedback offered by review teams. Finally, the CV THEA Pilot team would like to thank the Agreement Officer's Representative, Mr. Govind Vadakpat, for his leadership and continued support through the performance evaluation and measurement of the CV Pilot.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Acronyms

API	Application Programming Interface
App	Application
BSM	Basic Safety Message
CBD	Central Business District
CDF	Cumulative Distribution Function
ConOps	Concept of Operations
CoT	City of Tampa
COVID-19	Coronavirus Disease 2019
CU	Controller Unit
CUTR	Center for Urban Transportation Research, University of South Florida
CV	Connected Vehicle
DSRC	Dedicated Short-Range Communication
EEBL	Electronic Emergency Brake Light
ERDW	End of Ramp Deceleration Warning
FCW	Forward Collision Warning
FHWA	Federal Highway Administration
FN	False Negative
FP	False Positive
G	Gravity Unit
GPS	Global Positioning System
HART	Hillsborough Area Regional Transit
HMI	Human Machine Interface
HTTPS	Hypertext Transfer Protocol Secure
HV	Host Vehicle
Hz	Hertz
ID	Identifier
IE	Independent Evaluator
IMA	Intersection Movement Assist
I-SIG	Intelligent Traffic Signal System
ISM	Infrastructure Sensor Message
ITS	Intelligent Transportation System
JPO	Joint Program Office
JS	Java Script
JSON	Java Script Object Notation
LiDAR	Light Detection and Ranging
MAFB	MacDill Air Force Base
MAP	Map Data message

MMITSS	Multimodal Intelligent Traffic Signal System
MPH	Miles per Hour
MPS	Meters per Second
NPM	Node Package Manager
NTCIP	National Transportation Communications for Intelligent Transportation System Protocol
O&M	Operations and Maintenance
OBU	Onboard Unit
OSADP	Open Source Application Development Portal
OTA	Over the Air
PCE	Parameter Conformity Evaluation
PCW	Pedestrian Collision Warning
PDF	Probability Density Function
PED-X	Pedestrian Crossing
PID	Personal Information Device
PII	Personally Identifiable Information
PMESP	Performance Measurement and Evaluation Support Plan
PSM	Pedestrian Safety Message
PTP	Potentially True Positive
REL	Reversible Express Lanes
REST	Representational State Transfer
RSU	Roadside Unit
RV	Remote Vehicle
SAE	Society of Automotive Engineers
SD	Secure Digital
SDC	Secure Data Commons
SDD	System Design Document
SPaT	Signal Phase and Timing
SQL	Structured Query Language
SRM	Signal Request Message
SSL	Secure Sockets Layer
SSM	Signal Status Message
TECO	Tampa Electric Company
THEA	Tampa Hillsborough Expressway Authority
TIM	Traveler Information Message
TMC	Traffic Management Center
TN	True Negative
TP	True Positive
TSP	Transit Signal Priority
TTC	Time to Collision

TTI	Travel Time Index
UC	Use Case
USDOT	United States Department of Transportation
V2I	Vehicle-to-Infrastructure
V2V	Vehicle-to-Vehicle
VMT	Vehicle Miles Traveled
VTRFTV	Vehicle Turning Right in Front of Transit Vehicle
WSMP	WAVE Short Messaging Protocol
WWE	Wrong-Way Entry
XML	Extensible Markup Language

Executive Summary

The Tampa Hillsborough Expressway Authority (THEA) Connected Vehicle Pilot Deployment (CV Pilot) is one of the three CV projects selected by the Federal Highway Administration (FHWA) in September 2015 as part of a U.S. Department of Transportation (USDOT) Intelligent Transportation Systems Joint Program Office (ITS JPO) funded program. The Tampa CV Pilot identifies areas for improved traffic management in Tampa, Florida, that could be enhanced by the deployment of CV technology. The project considered several CV applications deployed across highway, transit, and pedestrian modes on a variety of facility and vehicle types in a connected urban environment.

This report presents the results of the implementation of the Performance Measurement and Evaluation Support Plan (PMESP) developed during Phase 1 (Concept Development) and Phase 2 (Design/Build/Test) of the deployment. Phase 3 consisted of maintaining and operating the system and implementing the PMESP. The PMESP adopted a use case approach with the goal of collecting and sharing data and evaluating the impact and contribution of vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) technology toward enhancing the mobility and safety of Tampa's travelers.

The study implemented a robust panel data experimental design to enroll more than 1,000 commuters to experience the functionality and impact of several V2V and V2I applications for a period of over 19 months. Participants' vehicles were equipped with aftermarket onboard units (OBUs) capable of deploying and delivering warnings via a Human Machine Interface (HMI) installed in the vehicle's rearview mirror. The panel was split into treatment (HMI enabled) and control (HMI disabled) groups using a randomized two-to-one matching (two treatment to one control) stratified by gender, age, income, and education.

The analysis relied on data collected from 49 roadside units (RSU) and travel logs stored by OBUs and transmitted over the air to secure storage locations for processing, analysis, and sharing with USDOT. The research team estimated more than 84,000 V2V and V2I interactions and identified 388 potentially true conflicts or situations where the CV applications correctly issued warnings to the participants.

To evaluate the safety benefits, the study analyzed the participants' behavioral responses to the deployment of the V2V and V2I applications. In the absence of in-vehicle video camera detection and recording, the research team developed data-driven algorithms to detect longitudinal and lateral reactions in response to evasive maneuvers. For each applicable use case, the analysis estimated false positive, false negative, true positive, and true negative rates by using algorithms developed to emulate the applications' logic to identify and quantify the number of V2V interactions and conflicts for each use case. The mobility impact evaluation relied on a before-after assessment using interrupted time-series regression analysis.

Findings provide evidence that the deployment contributed to enhancing the mobility and safety of travelers in the study area. Below is a summary of the main results.

Mobility Impact

The broadcasting of speed advisories via the speed harmonization End of Ramp Deceleration Warning (ERDW) application along Use Case 1 contributed to improvements compared with the baseline (before ERDW deployment) conditions.

The ERDW contributed to:

- 2.1 percent reduction in mean travel times
- 1.8 percent reduction in idle time or time spent traveling at less than one mile per hour.
- 1.8 percent reduction in queue length
- A travel time index (measured as peak hour travel time divided by off-peak travel time) reduction from 2.7 to 1.9.

It is important to note that the above findings can be considered short-term impacts due to the limited time that the improved ERDW application has been deployed (starting February 2020) and disruption in the data generation induced by travel behavior changes caused by the COVID-19 pandemic (starting March 20, 2020).

Use Case 4 (Transit Signal Priority) did not produce data conducive to the mobility evaluation because the Transit Signal Priority (TSP) application was not successfully deployed. Improvements to the TSP applications are currently being developed and generating test data.

Use Case 6 (Traffic Progression) also did not produce data conducive to the mobility evaluation. This is because the Intelligent Traffic Signal System (I-SIG) was not successfully deployed and did not generate the required data to conduct a before-after assessment.

Safety Impact

The safety evaluation uncovered heterogeneity in how the V2V and V2I applications contributed toward improved safety based on the use case being evaluated:

- Use Case 1 showed that the Forward Collision Warning (FCW) rate of conflicts per vehicle interaction did not change between before and after periods, with a rate of conflicts of 0.6 percent before and after FCW deployment. The analysis of the Electronic Emergency Brake Light (EEBL) warnings showed an increase in the rate of conflicts from 0.5 percent (before deployment) to 0.9 percent (after deployment). Furthermore the FCW application triggered nine True Positive (TP) warnings, in situations where there was a conflict as defined in this analysis. The EEBL triggered one TP warning. These show the potential of the applications to help drivers avoid conflicts.
- Use Case 2 focused on the safety improvements associated with the deployment of the Wrong-Way Entry (WWE) application. The analysis showed a contribution toward increased safety and uncovered that the application's performance varied according to the commuter peak travel flow:
    - In the PM peak period (3 p.m. to 12 a.m. weekdays), the application correctly warned drivers of entering the wrong way and identified 14 participants of 19 potentially true conflicts. The analysis found that the PM peak period was characterized by complex turning movements to

access the Selmon Expressway Reversible Express Lanes while driving back home. Global Positioning System (GPS) signal inaccuracy and the complexity of the warning delivery system contributed to a false positive rate of about 28 percent.

  - The AM period (6 a.m. to 10 a.m. weekdays) did not experience a single wrong-way occurrence during the entire deployment, but the application generated a high number of false positives. Most of these false warnings could be reduced by correcting the OBUs' wrong flagging of the Map Data message (MAP) identifying the revoked lanes and by improving the application's vehicle heading measurement while caught up in the morning queue. The false positive rate for the AM period was 2.8 percent.

- Use Case 3 (Pedestrian Conflicts) focused on improving pedestrian safety by deploying the Pedestrian Collision Warning (PCW) application. The application deployment required changing pedestrian detection technology during Phase 3. The improved PCW became operational in August 2020 and did not produce data conducive to evaluating the application's contribution to improved safety.

- Use Case 5 (Streetcar Conflicts) centered on the safety improvements on public transportation by deploying the Vehicle Turning Right in Front of Transit Vehicle (VTRFTV) application on the local streetcars. The analysis found that the VTRFTV deployed 61 warnings, of which 8 (13%) were classified as true positive during four unique events, but only 3 warnings (in one unique event) were shown to the participant driver due to the evaluation's experimental design. The conflict detection algorithm confirmed that the participant was not engaged in a conflict with the streetcar.

- Use Case 6 (Traffic Progression) deployed FCW, EEBL, and Intersection Movement Assist (IMA) applications to improve the safety of commuters traveling through a busy arterial characterized by several signalized intersections. The three safety applications issued 26 warnings classified as true positive, of which 8 were shown to drivers with the HMI enabled (i.e., treatment group). The conflict detection algorithm revealed that in two of these warnings the participants responded to the HMI message.

Participant Perceptions

The study implemented a series of surveys to collect socio-demographic, travel behavior, and specific feedback from the participant's use of the equipment and exposure to the applications:

- Overall, participants about were somewhat or very satisfied with the participation in the study (56%). This was not related directly to the applications but the overall experience, i.e. intake, installation, application experience, maintenance etc. Nineteen percent were indifferent and 25 percent somewhat dissatisfied. Overall satisfaction increased to 66 percent when considering the treatment group exposed to more interaction with the applications via the HMI.

- At the beginning of the study, most participants (66%) perceived safety as the greatest benefit of CV technology, followed by expectations about reduced congestion (56%) and a less stressful commute (54%). The perception about safety did not change as they participated in the study. Conversely, expectations about reduced congestion and a less stressful commute decreased to 33 and 30 percent, respectively. This could be due to factors influencing these perceptions, such as the localized

impact of the speed harmonization application (ERDW), lack of I-SIG implementation, and the currently low CV penetration rates.

- Before actively partaking in the study, about 46 percent of the participants expressed concerns about the impact of CV technology on their privacy. As they participated in the study, these concerns lessened as about 29 percent of the participants expressed some concerns.

- Concerns about the cost of CV technology increased during the study, with about 31 percent of the respondents being concerned in the interim and final phases of the study compared to 16.2 percent expressed at inception. These responses might have been somewhat affected by reliability issues associated with the aftermarket units requiring participants to be called back for repairs or OBU swaps.

Lessons Learned

The performance evaluation revealed some application-specific issues that can be resolved by improving the currently deployed OBU firmware with further research and development. Other factors affecting the application performance are more specific to the site deployment and the constraints associated with the OBU equipment: a highly dense urban environment posing challenges to V2I applications that are more dependent on accurate GPS and signal stability.

The safety evaluation revealed drawbacks in the development and implementation of V2V applications in the context of a research deployment, which are mostly due to the setup of the over the air (OTA) firmware updates and the applications' operational and functional parameters. These issues are related to the current landscape defining the aftermarket OBU industry. The OBU suppliers engaged in the deployment exhibited a high degree of variability in terms of research and development capabilities. This heterogeneity impacted the development, refinement, and level of maturity of some of the THEA CV Pilot applications (i.e., I-SIG, TSP, PCW). The lessons learned from these challenges documented in this study can serve to inform current and future deployments and to point to implementable solutions.

Not all the planned performance evaluation measures were actually implemented. As discussed in each use case's lessons learned of this report, technical issues related to maintaining and updating the OBU firmware, and technical difficulties in deploying some applications, affected the data generation and collection, thus preventing the measurement of some of the planned performance metrics.

In addition, the advent of the COVID-19 pandemic caused substantial changes in travel behavior with widespread effects across modes and to the study participants. This resulted in reducing the evaluation timeframe for some of the use cases and selected CV applications.

The experimental design was conceived with the goal of evaluating the performance of CV technology in the context of six use cases. This approach, while instrumental to the Pilot assessment, limited the analysis to considering only vehicles that traveled in a limited area, defining the geographic boundaries of each use case with a narrow focus on specific V2I and V2V applications.

Chapter 1. Introduction

The Tampa Hillsborough Expressway Authority (THEA) Connected Vehicle Pilot Deployment (CV Pilot) is one of the three CV projects selected by the Federal Highway Administration (FHWA) in September 2015 as part of a U.S. Department of Transportation (USDOT) funded program. The deployment consists of three phases: (1) Concept Development, (2) Design-Build-Test, and (3) Operations and Maintenance (O&M). The Pilot identified areas of traffic management in Tampa, Florida, that may be improved by the deployment of CV applications. The project team developed a system concept for deploying these CV applications and after approval by USDOT, designed, deployed, and currently operates the system.

The deployment includes several CV applications, deployed across highway, transit, and pedestrian modes of transportation on a variety of facility and vehicle types. This Pilot aims to create a connected urban environment to measure the effect and impact of CVs in Tampa's vibrant downtown.

During Phase 1 and Phase 2 of the deployment, THEA developed the Performance Measurement and Evaluation Support Plan (PMESP). The PMESP adopted a use case approach with the goal of collecting, measuring, sharing, and evaluating the impact and contribution of vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) technology toward enhancing the mobility and safety of Tampa's travelers.

1.1 Purpose of the Report

The purpose of this report is to summarize the findings of the implementation of the PMESP. The report details the analysis of the six use cases that were developed and defined in the PMSEP, the deployed data collection and experimental design, and the account of confounding factors and methods to control for them. The report also summarizes lessons learned for future pilot deployments.

1.2 Organization of the Report

The remainder of the report is organized into the following sections:

- **Chapter 2 – CV Pilot Performance Evaluation Objectives** defines the CV Pilot performance evaluation objectives organized in six use cases that are both quantitative and qualitative, depending on the case.
- **Chapter 3 – Experimental Design** discusses the experimental design that defined the participant recruitment campaign and the ensuing impact evaluation.
- **Chapter 4 – Data Collection and Sharing** discusses the processes to collect CV and non-CV data, the adopted archiving and sharing approaches with USDOT dedicated analytical and storage platforms.
- **Chapter 5 – Reporting to Stakeholders** presents the THEA CV Pilot Performance and Measurement Dashboard, an online platform to interactively track the CV Pilot progress and assess its performance.

- **Chapter 6 – System Impact Evaluation Methodology** describes the methodology used for system impact evaluation.
- **Chapter 7 – System Impact Evaluation Results** presents the results of the evaluation.
- **Chapter 8 – Conclusions** summarizes the findings and provides an overall discussion on the Pilot's lessons learned and contributions.

Chapter 2. CV Pilot Performance Evaluation Objectives

2.1 Introduction

The objective of the performance evaluation was to measure the impact of the THEA CV Pilot deployed V2V and V2I applications. The evaluation considered before-after changes and control-treatment differences in a series of established performance measures focused on mobility and safety in six uses cases, each developed to address a pre-existing safety or mobility concern affecting the study area.

2.2 Use Cases

The THEA CV Pilot developed six use cases (UCs) to describe the issues that the deployment sought to address (Table 2-1). The next section introduces each use case and the CV applications, along with maps showing deployment at specific routes and intersections.

Table 2-1. THEA CV Pilot Deployment Use Case Summary

Use Case	Condition	Location
UC1	Morning Backups	REL at E. Twiggs Street
UC2	Wrong-Way Entries	REL at E. Twiggs Street and Meridian Avenue
UC3	Pedestrian Conflicts	E. Twiggs Street at George E. Edgecombe Courthouse
UC4	Transit Signal Priority	Marion Street Transit Mall; Study area sections of Kennedy Boulevard and Jackson Street; Portions of Florida Avenue and Tampa Street
UC5	Streetcar Conflicts	Channelside Drive
UC6	Traffic Progression	Meridian Avenue and Florida Avenue

2.2.1 Use Case 1: Morning Backups

As drivers approach the end of the Selmon Expressway Reversible Express Lanes (REL), they enter a curve where the speed limit reduces from 70 miles per hour (MPH) to 40 mph. During morning rush hour, as vehicles exit the REL onto Meridian Avenue to make a right turn onto East Twiggs Street, the right turn lane backs up. An additional issue is that many of these drivers then want to make a right turn onto Nebraska Avenue, which is almost an immediate right turn after turning onto East Twiggs Street. The combination of these issues causes the queue to back up onto the REL. This backup causes exiting vehicles turning right to use the shoulder as part of the right turn lane. As drivers approach the REL exit, they may not be able to anticipate where the end of the queue is for the right-turn lane, potentially causing them to hard brake or attempt a rapid

lane change. Figure 2-1 shows the UC1 route. The goal of this use case was to deploy and measure the impact of the following applications:

- End of Ramp Deceleration Warning (ERDW)
- Electronic Emergency Brake Light Warning (EEBL)
- Forward Collision Warning (FCW)
- Intelligent Traffic Signal System (I-SIG).

As described in section 6.2.5 the I-SIG application was not successfully deployed.

Source: THEA, Global-5, 2017

Figure 2-1. Use Case 1 – Routes and Roadside Equipment

2.2.2 Use Case 2: Wrong-Way Entries

At the exit of the REL onto East Twiggs Street, there is a relatively easy opportunity for a driver to become confused and attempt to enter the REL going the wrong way. There are no gates or barriers at the westbound downtown terminus of the REL to prevent drivers from entering the REL going the wrong way. Drivers who are traveling on East Twiggs Street approaching the intersection where the REL ends and Meridian Street begins can mistakenly (or knowingly) enter the REL going the wrong way. Drivers approaching this intersection coming from downtown can inadvertently (or knowingly) make a left turn onto the REL exit. Conversely, drivers on East Twiggs Street approaching this intersection going toward downtown can inadvertently make a right turn onto the REL exit. Finally, drivers approaching the intersection on Meridian Avenue can potentially veer slightly to the left onto the REL exit. Each of these possibilities is a safety concern. Figure 2-2 shows the UC2 route. The goal of this use case was to deploy the Wrong-Way Entry (WWE) and I-SIG CV applications. As described in section 6.2.5, the I-SIG application was not successfully deployed.

Source: THEA, Global-5, 2017

Figure 2-2. Use Case 2 – Routes and Roadside Equipment

2.2.3 Use Case 3: Pedestrian Conflicts

At the George E. Edgecombe Hillsborough County Courthouse, there is one primary crosswalk for pedestrian access to the main parking garage. The crosswalk is marked and has a yellow flashing beacon to warn drivers that they are approaching a crosswalk. This crosswalk is the primary route for jurors, lawyers, and other people to get to and from the courthouse. During morning rush hour, there is significant pedestrian traffic as potential jurors unfamiliar with the area are attempting to arrive on time. This significant pedestrian traffic is compounded on Mondays and Tuesdays when new juror pools of up to 400 persons are required to report during rush hour. Lack of attention by drivers causes a safety concern for pedestrians trying to reach the courthouse. Some pedestrians elect to take a shortcut by crossing East Twiggs Street mid-block and outside the crosswalk. Figure 2-3 shows the UC3 location. Planned CV deployment at this location included the Pedestrian Collision Warning (PCW) application.

Source: THEA, Global-5, 2017

Figure 2-3. Use Case 3 – Location and Roadside Equipment

2.2.4 Use Case 4: Transit Signal Priority

Two express bus routes (24LX and 25LX) use the Selmon Expressway to connect the east and west sides of the metropolitan area and exit the Expressway to serve a stop in downtown. There are large residential communities in areas of Brandon, Riverview, and Fish Hawk to the east of downtown. Aside from the employment center associated with the central business district (CBD), MacDill Air Force Base (MAFB) is situated close to the western or southern terminus of the Selmon Expressway. CV technologies were deployed to attempt to create a "virtual transit connection" between the two portions of the expressway by providing more reliable transit mobility using Transit Signal Priority (TSP) as the express buses negotiate the surface streets of downtown in the morning and evening peak hours. Figure 2-4 shows the UC4 routes.

Source: THEA, Global-5, 2017

Figure 2-4. Use Case 4 – Routes and Roadside Equipment

Marion Street is a two-lane urban arterial road in the heart of the Tampa CBD that serves as a primary bus route, which terminates on the north end at the Marion Transit Center. During weekdays it is dedicated solely to transit vehicle use. Hillsborough Area Regional Transit (HART) operates several routes that converge onto Marion Street and head to Marion Street Transit Station. The third express route that was included in this use

case was Route 20X, which provides limited express bus service to and from the northern residential communities in New Tampa and Wesley Chapel through the CBD and south using the Selmon Expressway to and from MAFB. There are several morning drop-off points and evening pick-up points along the Marion Street transit facility.

As the express service departs the controlled-access highway system and enters the downtown grid, the buses experience congestion. When congestion occurs, the buses are unable to reach their stops promptly, causing them to fall behind schedule and compromising mobility. CV technology was used to address these mobility concerns. Buses and traffic signals communicated, and if a bus was behind schedule, the traffic signal system either gave the bus priority or flushed the queue, allowing the bus to reach its stop and assuming there were no other higher priorities. The buses on the routes described here benefitted from TSP while in the study area in this use case. Figure 2-5 shows the bus routes included in UC4.

Source: HART, CUTR, 2017

Figure 2-5. Use Case 4 – CV Pilot Transit Routes

Ten HART buses were equipped and assigned to the three express routes for the duration of the Pilot. CV applications planned for deployment of this use case included IMA, I-SIG, and Transit Signal Priority (TSP).

2.2.5 Use Case 5: Streetcar Conflicts

The Tampa Electric Company (TECO) Streetcar runs along Channelside Drive from the Amalie Arena area up Channelside Drive, north, and past the Selmon Expressway. The streetcar is a steel wheel on steel rail fixed-guideway system in a dedicated right-of-way. An overhead catenary powers it and the streetcar crosses intersections at grade. As a result, at various stops along the streetcar route, vehicles may have to turn right in front of a stopped or moving streetcar. As pedestrians disembark the streetcar and the streetcar prepares to depart, a vehicle may turn right in front of the streetcar. This situation occurs at signalized and non-signalized intersections, none of which have a right turn protected movement. CV technology was used to provide information to streetcar operators and drivers to improve safety around these locations. Figure 2-6 shows the UC5 route. The goal of this use case was to deploy the Vehicle Turning Right in Front of Transit Vehicle (VTRFTV) CV application.

Source: THEA, Global-5, 2017

Figure 2-6. Use Case 5 – Routes and Roadside Equipment

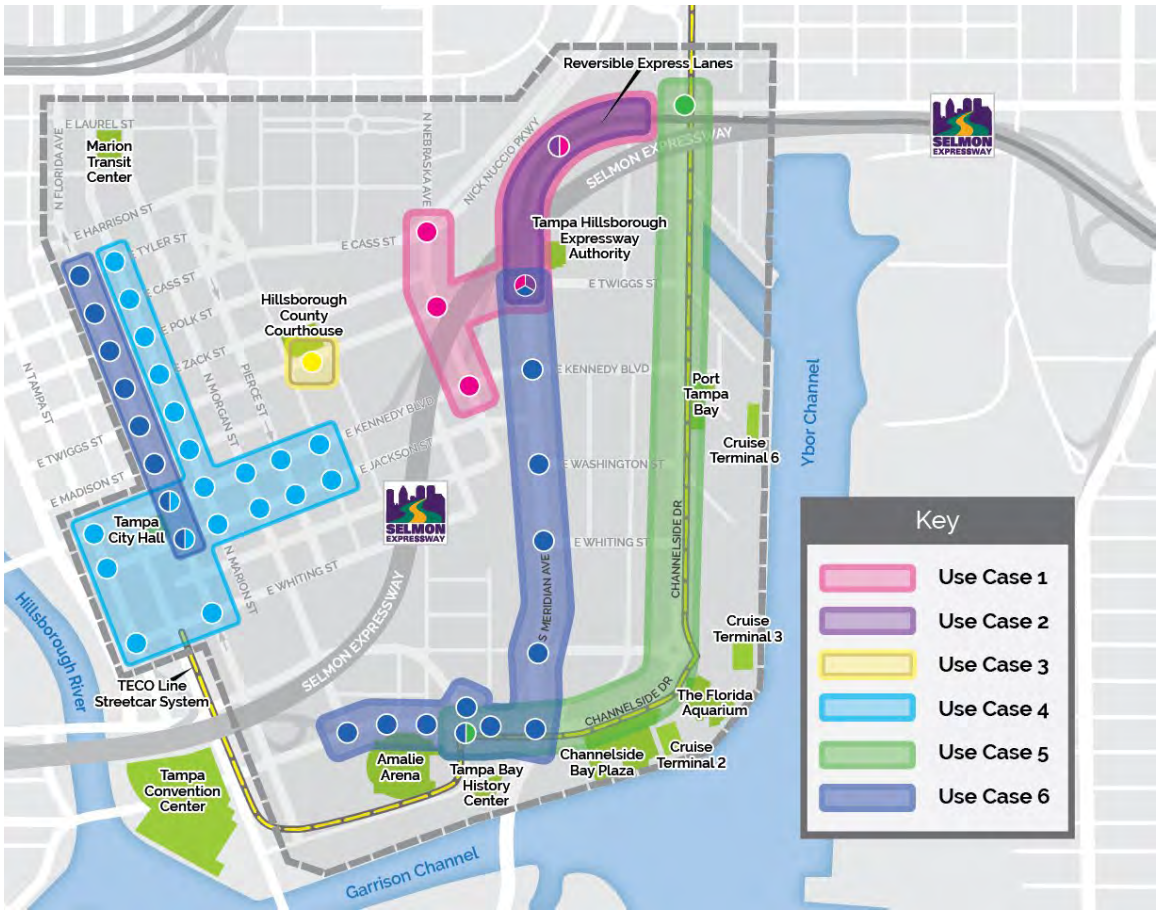
2.2.6 Use Case 6: Traffic Progression

Meridian Avenue has significant congestion and delay during morning peak hour periods. This congestion is due to many MAFB commuters exiting the Selmon Expressway at downtown and traveling through downtown arterial routes to reach the Base entrance. As some of these commuters use surface roads through downtown, they interact with other traffic and pedestrians, increasing the likelihood of conflicts. In addition to Meridian Avenue, Florida Avenue (sections within the study area) experiences similar issues for downtown commuters. I-SIG was the CV application primarily planned to be used in this use case, but as described in section 6.2.5, the I-SIG application was not successfully deployed during Phase 3. Figure 2-7 shows the two corridors for this use case.

Source: THEA, Global-5, 2017

Figure 2-7. Use Case 6 – Routes and Roadside Equipment

Figure 2-8 illustrates the combination of all the use cases in the study area.



Source: THEA, Global-5, 2017

Figure 2-8. THEA CV Pilot Deployment Locations

2.3 Planned and Actual Performance Measures

The PMESP identified a list of performance measures to assess the cases based on the four "pillars" of mobility, safety, environment, and agency efficiency. The performance measures were tied to the target values discussed in the Phase 2 approved Concept of Operations (ConOps) document [1]. As detailed in the approved ConOps, it is challenging to set performance targets. This is because the CV Pilot project is unprecedented. It is unlikely that the impact on safety would be directly measurable from the low numbers of crashes occurring over the course of the deployment and within the study limits.

Table 2-2 provides a snapshot of the performance measures originally addressed. As discussed later in Chapter 7, not all performance measures were evaluated, in particular the mobility measures associated with Use Case 4 (section 7.4) and Use Case 6 (section 7.6).

Table 2-2. Summary of Planned Performance Measures

Pillars	Performance Measures	UC1 Morning Backups	UC2 Wrong-Way Entries	UC3 Pedestrian Conflicts	UC4 Transit Signal Priority	UC5 Streetcar Conflicts	UC6 Traffic Progression
Mobility	Travel time	P/A	P	P			P/A
	Travel time reliability	P/A		P			P/A
	Queue length	P/A		P			P
	Vehicle delay	P	P	P			P
	Percent (%) arrival on green	P			P		P
	Bus travel time				P		
	Bus route travel-time reliability				P		
	Percent (%) arrival on schedule				P		
	Excess time spent in idle	P/A			P		P
Safety	Crash comparison	P/A	P/A	P/A		P/A	P/A
	Types of crashes	P/A	P/A			P/A	P/A
	Severity of crashes	P/A	P/A	P/A		P/A	P/A
	Type of conflicts	P/A	P/A	P/A		P/A	P/A
	Severity of conflicts	P	P	P		P	P
	Approaching vehicle speed	P		P			P
	Number of alerts from apps	P/A	P/A	P/A		P/A	P/A
Environmental	Emissions reductions in idle	P	P	P	P		P
	Emissions reductions in running	P	P	P	P		P
Agency Efficiency	Mobility improvements through the mobility pillar analysis	P/A	P/A		P		P
	Safety improvements through the safety pillar analysis	P/A	P/A	P/A		P/A	P/A
	Customer satisfaction through opinion survey and/or CV app feedback	P/A	P/A	P/A	P	P/A	P/A

P = Planned Performance Measures, A = Actual Performance Measures

2.3.1 Mobility

For mobility analysis the measures primarily used for evaluation were travel time, travel time reliability, and queue length. The vehicle delay was not used since the output from the Multimodal Intelligent Traffic Signal System (MMITSS) was not successful in providing these metrics. In addition, the percent arrival on green was not collected as the Centracs system used by the City of Tampa did not provide these data for the baseline period. For UC4, two issues prevented the use of the mobility measures for transit: (1) the TSP application was not successful during Phase 3, and (2) the equipped buses were not always running on the selected routes. Therefore, the measures associated with the bus mobility issues were not used (bus travel time, bus route travel-time reliability, percent arrival on schedule), but data to calculate them were collected.

2.3.2 Safety

All safety measures were collected and used except for the severity of conflicts and approaching vehicle speed. Through analysis of the types of conflicts, it became clear that it would not be ideal to assign severity for conflicts as the classification would be arbitrary. The identification of conflicts followed the parameters used by the vendors and the Society of Automotive Engineers (SAE) standards followed by the OBU vendors, and comparison of before-after periods required a fixed classification of conflicts. In addition, the team used two terms, *interactions* and *conflicts* as described in detail in section 6.4, to provide the type of severity for conflict scenarios.

2.3.3 Environmental

These measures were not estimated due to the lack of implementing the TSP (section 7.4) and I-SIG (section 7.6) applications.

2.3.4 Agency Efficiency

As outlined in the table, the measures for agency efficiency would come from the mobility and safety improvements as well as participant satisfaction via opinion surveys. All these measures were collected and used for analysis for the individual use cases.

Chapter 3. Experimental Design

The CV Pilot deployment provided a unique opportunity to implement an experimental design to optimize the level of control upon observed and unobserved confounding factors. Given the use case characteristics, the research team envisioned the implementation of specific experimental strategies to evaluate the impact of the deployed CV technologies. This chapter details the adopted designed experiments and the proposed methods to control for confounding factors.

3.1 Experimental Strategies

In an ideal context, confounding factors can be controlled for by conducting counterfactual analysis via random experimental design. Counterfactual modeling measures the potential outcome in the absence of an intervention, such as the implementation of the CV technologies. Empirically, there are different options for assessing the counterfactual, ranging from a simple before versus after comparison of outcomes to measuring responses in the context of a randomized experiment. The evaluation adopted two methods to control and minimize the impact of study area–specific and deployment-specific confounding factors:

1. Random Design
2. Before and After Comparison (Time Series Analysis).

3.1.1 Random Design

In a completely randomized design, study participants are randomly drawn from a representative sample and randomly assigned to a treatment group and a control group. The treatment group comprises those individuals who are assigned to the intervention (i.e., exposed to the CV applications via HMI enabled warnings) and the control group consists of those individuals who are excluded from the intervention (i.e., not exposed to the CV applications because the HMI is disabled). Random assignment to the treatment and control groups ensures the two groups are similar and have the same probability of being assigned to either one of the groups. Random assignment ensures that units assigned to the treatment and units assigned to the control are identical (over many iterations of the experiment).

3.1.2 Before and After Comparison (Time Series Analysis)

This method relies only on a comparison of time trends in performance measures (e.g., travel times, risk avoidance) without resorting to direct identification of treatment and control groups. The goal is to assess if the treatment (i.e., the CV technology deployment) has caused a change in pattern upon the baseline conditions using a pretest-posttest approach. The empirical analysis lends itself to comparing changes in the experimental subjects over time using an interrupted time series approach where the series is broken into intervals representing interventions [2, 3].

Whenever adopted, the approach utilized the performance measures data collected throughout the Pilot implementation, with a design strategy focused on the timing of the treatment. In this instance, the first period of data collection defined the baseline. In a subsequent period, the treatment (i.e., the CV technology) was applied as data collection and performance measurement continued. Study area–specific confounding factors were recorded concurrently to serve as explanatory variables in time series statistical analysis.

Table 3-1 summarizes the adopted experimental method(s) for each of the six use cases based on the actual participant recruitment considered and CV applications deployed.

Table 3-1. Adopted Experimental Design

Experimental Design	UC1 Morning Backups	UC2 Wrong-Way Entries	UC3 Pedestrian Conflicts	UC4 Transit Signal Priority	UC5 Streetcar Conflicts	UC6 Traffic Progression
Interrupted Time Series	✓	✓	✓	✓	✓	✓
Random Design	✓					✓

3.2 Participant Recruitment

THEA's CV Pilot Deployment team initiated a recruitment campaign to enroll participants into the study in early 2018. Participants were selected from the pool of existing THEA customers and recruitment efforts began early in Phase 2 and continued through the end of December 2018.

The Center for Urban Transportation Research (CUTR) team provided support throughout the participant recruitment campaign to implement the study's experimental design and to maintain the balance between group assignments (i.e., assignment of two treatment units for each control unit).

3.2.1 Experimental Group Assignment

Participants were assigned to either treatment or control using a randomized two-to-one matching (two treatment to one control) stratified by gender, age, income, and education. The research team processed 21 assignment requests throughout the installation campaign to accommodate the installation schedule. For each assignment, CUTR conducted pre-assignment and post-assignment checks to ensure balance. Table 3-2 through Table 3-5 report the sample distribution by gender, age, and income. The tests were based on the Pearson's and likelihood-ratio chi-square statistics, showing a balanced stratified sample split between treatment and control. The *frequency* represents the number of cases by cohort for each group. The *expected frequency* reports what would be the expected frequency to attain a perfect split for each cell. The *row percentage* measures the percent of cases that each cohort represents out of its row. Table 3-2, for example, shows that 266 participants were female and were assigned to the treatment, representing 42.8 percent of the treatment group. On the other hand, 174 participants were female and were assigned to the control, representing 44.5 percent of the control group.

Table 3-2. Assignment to Treatment and Control Stratified by Gender

Group	Female	Male	Did Not Answer	Total	Key
Treatment	266	344	11	621	Frequency
	270	337	14	621	Expected Frequency
	42.8	55.4	1.8	100	Row Percentage
Control	174	205	12	391	Frequency
	170	212	9	391	Expected Frequency
	44.5	52.4	3.1	100	Row Percentage
Total	440	549	23	1,012	Frequency
	440	549	23	1,012	Expected Frequency
	43.5	54.3	2.3	100	Row Percentage

Pearson Chi^2 (2) = 2.320; Pr = 0.313 | Likelihood-ratio Chi^2 (2) = 2.269; Pr = 0.322

Table 3-3. Assignment to Treatment and Control Stratified by Age

Group	25 and Under	26-35	36-45	46-55	56 and over	Did Not Answer	Total	Key
Treatment	20	128	161	184	114	14	621	Frequency
	22	131	152	188	112	15	621	Expected Freq.
	3.2	20.6	25.9	29.6	18.4	2.3	100	Row Percentage
Control	16	85	87	123	69	11	391	Frequency
	14	82	96	119	71	10	391	Expected Freq.
	4.1	21.7	22.3	31.5	17.7	2.8	100	Row Percentage
Total	36	213	248	307	183	25	1,012	Frequency
	36	213	248	307	183	25	1,012	Expected Freq.
	3.6	21.1	24.5	30.3	18.1	2.5	100	Row Percentage

Pearson Chi^2 (5) = 2.614; Pr = 0.759 | Likelihood-ratio Chi^2 (5) = 2.617; Pr = 0.759

Table 3-4. Assignment to Treatment and Control Stratified by Education

Group	High School	Assoc. Degree	Some College - No Degree	Bach. Degree	Graduate or Professional	Did Not Answer	Total	Key
Treatment	28	73	128	212	161	19	621	Frequency
	26	71	123	224	160	17	621	Expected Freq.
	4.5	11.8	20.6	34.1	25.9	3.1	100	Row Percentage
Control	15	43	72	153	100	8	391	Frequency
	17	45	77	141	101	10	391	Expected Freq.
	3.8	11.0	18.4	39.1	25.6	2.1	100	Row Percentage
Total	43	116	200	365	261	27	1012	Frequency
	43	116	200	365	261	27	1012	Expected Freq.
	4.3	11.5	19.8	36.1	25.8	2.7	100	Row Percentage

Pearson Chi^2 (5) = 3.555; Pr = 0.615 | Likelihood-ratio Chi^2 (5) = 3.582; Pr = 0.611

Table 3-5. Assignment to Treatment and Control Stratified by Income

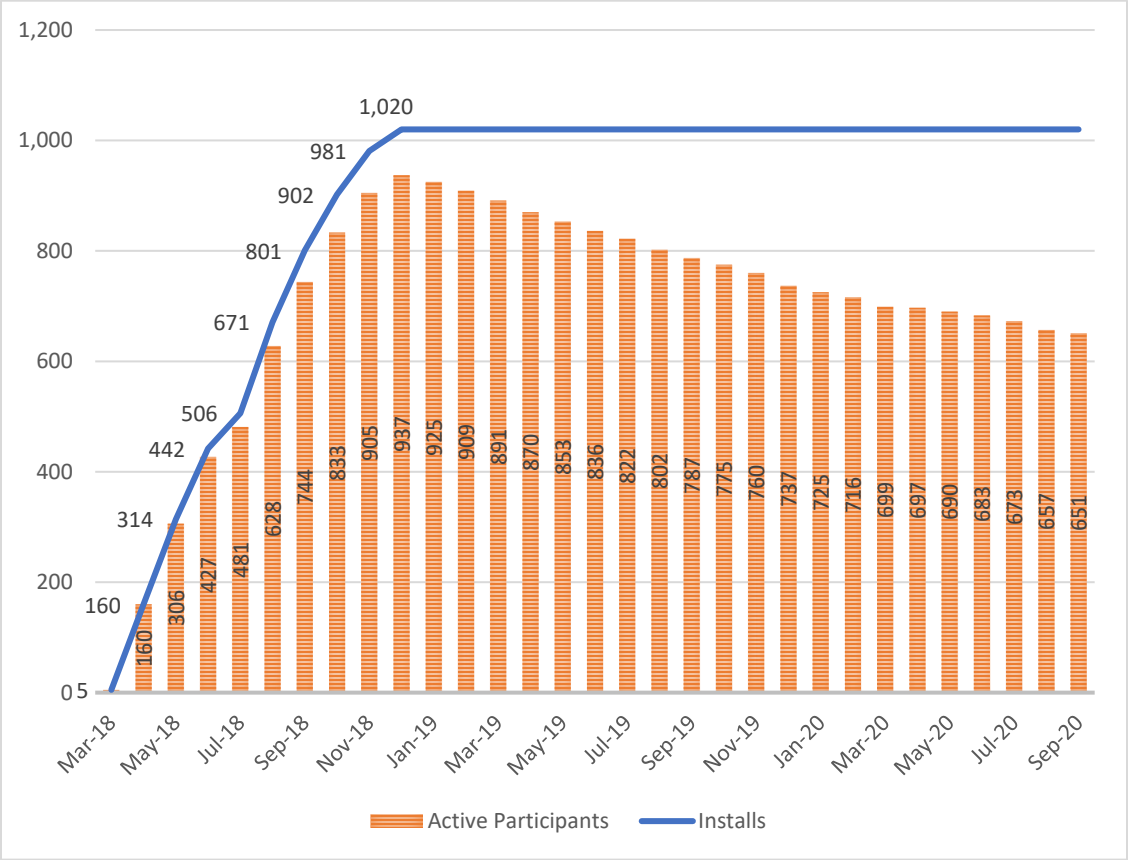
Group	Less than $25,000	$25,000 to $34,999	$35,000 to $49,999	$50,000 to $74,999	$75,000 to $99,999	$100,000 to $149,999	$150,000 or More	Did Not Answer	Total
Treatment	19	25	71	137	116	149	73	31	621
	20	30	74	133	112	149	73	31	621
	3.1	4.0	11.4	22.1	18.7	24.0	11.8	5.0	100
Control	13	24	50	79	67	93	46	19	391
	12	19	47	84	71	94	46	19	391
	3.32	6.14	12.79	20.2	17.14	23.79	11.76	4.86	100
Total	32	49	121	216	183	242	119	50	1,012
	32	49	121	216	183	242	119	50	1,012
	3.2	4.8	12.0	21.3	18.1	23.9	11.8	4.9	100

Pearson Chi^2 (5) = 3.349; Pr = 0.851 | Likelihood-ratio Chi^2 (5) = 3.296; Pr = 0.856

3.2.2 OBU Installation

Onboard unit (OBU) installation was performed at Hillsborough Community College by certified technicians and student trainees. A total of 1,020 private vehicles were equipped with aftermarket OBUs between March 2018 and December 2018. Figure 3-1 shows the number of installs and the number of active participants monthly (i.e., OBU generating at least one Basic Safety Message during the same month).

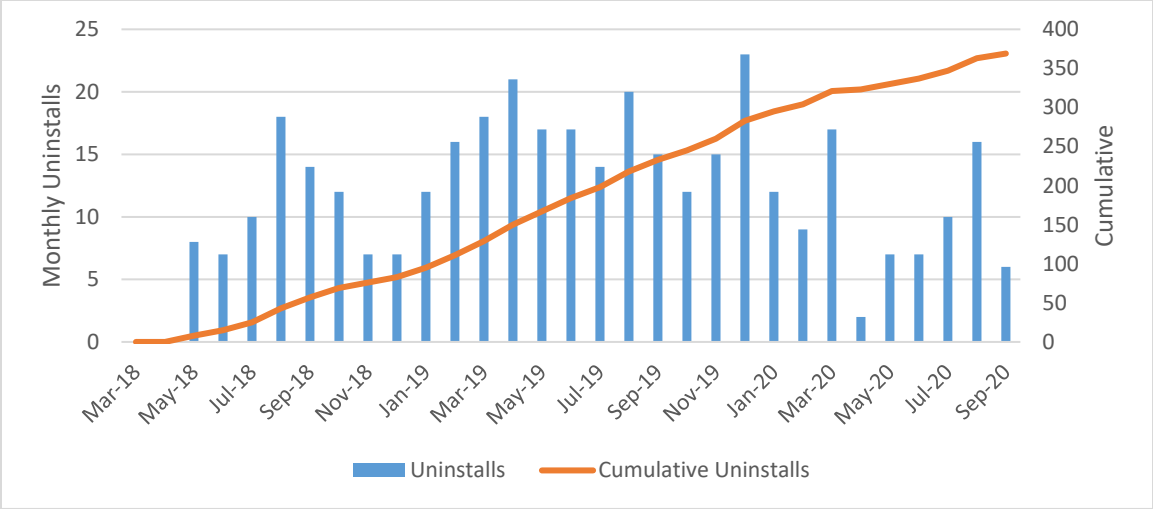
Over the course of the deployment, the total number of active participants reached 964. As installation progressed, some participants dropped out for various reasons, spanning from trading in their vehicles, to changing their commute patterns creating incompatibility with the participation incentive scheme, to relocating out of state. As of September 30, 2020, there were 651 active participants.



Source: CUTR, 2020

Figure 3-1. OBU Installation and Active Participants

Due to several reasons, participants chose to either uninstall the equipment, continue participation with a different vehicle, or drop out of the study. Figure 3-2 shows the number of uninstalls per month. Reasons for uninstalling included lack of incentive, moving out of the area, changing commute route, crashed vehicle, trade-in or sold vehicle, mechanical issues, and other personal reasons.



Source: CUTR, 2020

Figure 3-2. Number of OBU Uninstalls

3.3 Confounding Factors

The accuracy and effectiveness of performance measurement depend on the presence of concurrent confounding factors. Confounding factors are any events that might arise during the Pilot implementation that can be associated with having an apparent effect on some dependent variables of interest (i.e., performance measures). In a design experiment, confounding factors that are not accounted for during design could either understate or overstate the relevance of treatment effects upon treated units. In extreme cases, confounding factors can lead to spurious relationships between explanatory and dependent variables, with the variables having no direct causal connection, while it may be wrongly inferred that they do. Two types of confounding factors were likely to arise from the Pilot implementation:

- Study area–specific factors (e.g., climate, special events)
- Deployment-specific factors (e.g., participant-specific, technology-specific).

Factors that can be identified, recorded, and measured *a priori* (i.e., before Pilot implementation) are defined as *observed* factors. Factors that cannot be directly observed or measured are defined as *unobserved* factors. During performance measurement and assessment, observed factors can be accounted for by their proper inclusion as explanatory variables and modeling method, while unobserved factors can be accounted for by utilizing appropriate statistical techniques to reduce omitted-variable bias.

3.3.1 Study Area–Specific Factors

The research team sought to measure and record several time-variant factors, spanning from seasonal weather, planned events in the study area's main points of attraction, planned construction development, and seasonal cruise line tourism. These factors had the potential to generate confounding information across all use cases by influencing individual travel behavior.

3.3.1.1 Weather

Tampa is characterized by a subtropical climate with hot and humid conditions from mid-May through mid-October coinciding with the rainy season. Summertime weather is consistent from June through September and is characterized by mid-afternoon thunderstorms. These thunderstorms may last for only a few moments to several hours or even for an entire day. During the summer, average monthly rainfall increases to about 7.5 inches from the winter average of 2.5 inches. Localized weather conditions can have spatially heterogeneous effects on travel behavior. For example, thunderstorms can affect vehicle travel speed (e.g., traveling slower than usual), pedestrian trip patterns, and bus boarding differently at either the origin or end of a trip.

To control for weather-related factors at the aggregate level, a daily data log recording temperature, observed precipitation, and other weather-related occurrences was created using data from a third-party weather information service provider. Weather data were collected at 10-minute intervals describing humidity, visibility, and other conditions as detailed in Appendix C of the PMESP [4].

3.3.1.2 Special Events

The CV Pilot Study Area includes several site attractions, drawing visitors and residents to attend leisure or business events that can generate additional non-seasonal traffic with the potential to introduce confounding information throughout the Pilot. The research team collected road closure and event information from the City of Tampa (CoT) Traffic Management Center (TMC) that were regularly scheduled inside the CBD. The City has a public calendar for all events scheduled within it limits and a specific news feed for road closures for different reasons including construction, weather, and special events.

3.3.2 Deployment-Specific Factors

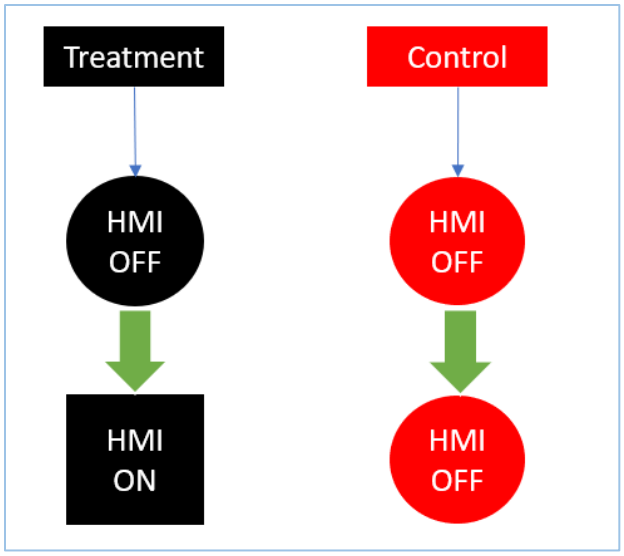
Deployment-specific confounding factors include all those factors or events that can be potentially triggered by the Pilot implementations. These include equipment malfunction instances as identified by the ConOps Failure/Anomaly/Exception Conditions and Safety Plan, and induced errors by linking data across platforms. Other confounding factors are likely to be introduced by participant identification and selection, their personal use of installed vehicle equipment, and improper use of downloaded applications.

During Phase 2, confounding factors were introduced at the installation point due to the manually intensive setup process at the participant appointment site. A few instances were recorded where participants' unique identifications were wrongly entered into the system, thus creating difficulties in ascertaining correct assignment to the experimental group.

3.3.2.1 Participant Group Assignment and OBU Firmware Updates

Over the course of Phase 3, the greatest source of deployment-specific confounding factors was the over the air (OTA) campaign to maintain and update onboard unit firmware and configuration files. One OTA campaign to update an OBU vendor's firmware mistakenly replaced the flag on participants who were set up with the goal of implementing the experimental group assignment to treatment (HMI enabled) and control (HMI disabled).

Figure 3-3 shows the initial group assignment setup of participants. The treatment group would initially have a silent period with HMI off and then HMI would turn on so they could see the warnings. The control group was set to have HMI off during the entire duration of the Pilot.



Source: CUTR, 2020

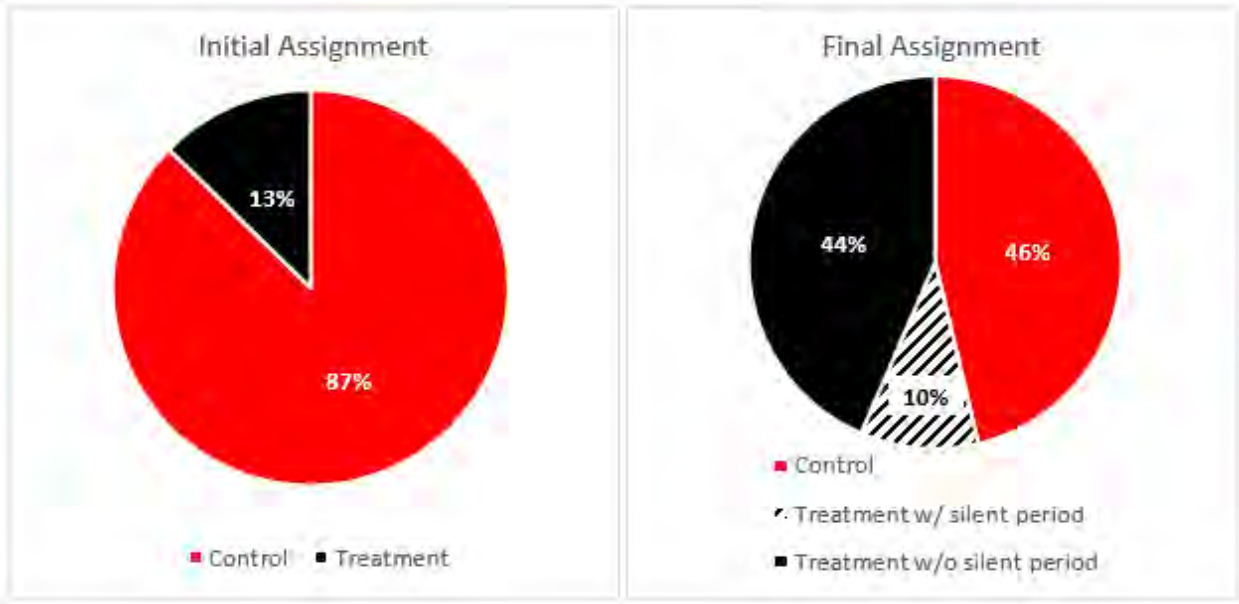
Figure 3-3. Initial Participant Assignment

Several challenges were presented in correctly assigning and keeping the two groups at installation and setup of the OBU or at maintenance, and at subsequent firmware updates of the OBUs. These issues caused some participants not to align with their original assignment. During the last quarter of 2019 and after receiving warnings from participants, the team analyzed the data to reclassify the groups. This reclassification was necessary to retain the experimental design integrity and to maximize the use of data from as many participants as possible.

The audit showed combinations where participants in the control group had the HMI on, some participants in the treatment group had the HMI off, and some had their HMI switch on or off after an initial period. The team used the reclassification scheme to organize participants into three new groups:

1. Participants with their HMI off during the entire period were reclassified to the control group.
2. Participants with their HMI off initially and then HMI on (after an initial period) were reclassified to the treatment group with a silent period.
3. Participants with their HMI on during the entire period were reclassified to the treatment group without a silent period.

According to data received by participant vehicles, a total of 434 unique vehicles issued at least one warning during the CV Pilot deployment. Based on the initial assignment, 55 of those (13%) belonged to the treatment group and 379 (77%) belonged to the control group. When reclassified using the method described, 199 (46%) were classified as control group and 231 (43%) were classified as treatment group. Figure 3-4 shows the initial and final assignment classifications for participants. The team was then able to use all but four (1%) of the vehicles for analysis of the applications.



Source: CUTR, 2020

Figure 3-4. Group Assignment Reclassification

3.3.3 Experimental Design–Induced Confounding Factors

Participants in the CV Pilot deployment included drivers, pedestrians, and bus and streetcar operators. Although the primary objective of experimental design is to minimize the presence and influence of confounding factors, the experimental design approach under use case–specific constraints is likely to introduce some error in the form of:

- Participant self-selection
- Participant attrition
- Participant moral hazard.

3.3.3.1 Participant Self-selection

Participant recruitment identifies a treatment and control group following the suggested experimental design as discussed in section 3.1 of this document. The recruitment goal is to randomly select a pool of participants for treatment and control groups from a sample that is representative of a system's users. When sending out requests to participate in a study, some individuals will tend to self-select to either participate or exclude themselves due to their specific socio-economic, residential location, and travel behavior characteristics. Though the experimental design approach minimizes the difference between treatment and control units, self-selection would still be an issue as it also depends upon the adopted recruitment approach (e.g., phone, internet, snail mail, shopping center booth).

As discussed later in Chapter 7, results from the participant survey showed that a sizeable share of the study participants selected to enroll mainly due to the toll discount.

3.3.3.2 Participant Attrition

Once enrolled as participants, some individuals would likely exit the study due to triggering events, such as a change of job leading to a different commute pattern, vehicle replacement, lack of interest, or other similar factors. When measuring performance at the individual level, statistical methods (e.g., unbalanced panel data methods) were employed to reduce the impact of ensuing confounding factors.

As the toll incentive program approached its end, this led to an increased number of participants dropping from the study. While the Performance Measurement and Evaluation Support Plan called for a participant refreshment strategy, the CV Pilot funding constraints limited the participant pool replenishment.

3.3.3.3 Participant Moral Hazard

Other confounding factors were likely to arise due to participant moral hazards that might be induced by CV equipment or application. A moral hazard is a situation where an individual might undertake a riskier behavior knowing that it is protected against a risky situation.

Participant recruitment can reduce the impact of confounding factors due to moral hazard. In addition, selected participants were advised of the limits of the technology and were required to sign an Informed Consent Form to participate that explained those limits and their liability in using the app not as prescribed.

Chapter 4. Data Collection and Sharing

Throughout the deployment, the performance measurement and evaluation relied on data coming from three major sources: (1) CV Pilot generated data, (2) study participant surveys, and (3) third-party generated data (e.g., Bluetooth, weather, log events, emission database, and other traffic data). CV Pilot data were culled from all CV applications transmitting and receiving information. The performance evaluation team at CUTR established a dedicated server (i.e., CUTR server) to collect, process, and upload Personally Identifiable Information (PII) removed data to the ITS Data Hub and the Secure Data Commons (SDC). This chapter discusses the data collection and documents the procedures for data processing, archival, and sharing.

4.1 Data Sources

4.1.1 CV Data

The Tampa CV Pilot generated several datasets from the interaction between vehicles (via OBUs) and between vehicles and infrastructure (OBU/RSU interaction). Vehicles traveling or operating (i.e., public transportation and participant vehicles) generated data in the form of BSMs, which were collected by roadside units and transferred over the air to THEA's secured master server. At the same time, roadside units broadcast relevant information to onboard units.

4.1.1.1 RSU Data

Roadside units transmitted and/or collected the following data:

- BSMs from the participant and public transit vehicles (up to 10 Hz), also called "sniffed" BSMs or BSMs collected by a vehicle operating in range of an RSU.
- Signal Phase and Timing Message (SPaT) from RSUs (10 Hz)
- Map Data Message (MAP) from RSUs (1 Hz)
- Traveler Information Message (TIM) from RSUs at 1 Hz
- Signal Request Message (SRM) transmitted by OBUs within range of the Dedicated Short-Range Communication radio of an RSU.
- Signal Status Message (SSM) broadcast by RSUs for conveying back to OBUs the status of its SRM.
- Multimodal Intelligent Traffic Signal System (MMITSS), Java Script Object Notation (JSON) formatted Siemens-MMITSS calculated metrics
- Pedestrian Crossing (PED-X) – JSON-formatted structure with element "psm" containing the Pedestrian Safety Message (PSM) that triggered the collision alert as J2735 Message Frame.

4.1.1.2 OBU Data

Public transportation and participant vehicles recorded all received and transmitted data from interaction with nearby vehicles and roadside units in range via an OBU data log recording protocol. OBU data logs contain various data elements falling into one of the following categories:

- WAVE Short Messaging Protocol (WSMP) messages sent or received.
- Warnings issued to the driver.
- Internal system monitoring events (e.g., SD card space, security audits).

All the above data were uploaded to the SDC as raw data following the metadata structure detailed in Appendix C of the PMESP [4].

Driver warning event records are created whenever one of the applications triggers a warning. The OBU creates a unique warning ID used to identify multiple Warning Event Data records belonging to the same warning event. The OBU creates a set of Warning Event Data records per warning. Each record of the set represents a point in time before, during, and after the warning triggered. A Warning Event Data record always contains the host vehicle's (HV) Basic Safety Message at a given point in time within "hvBSM." Warnings that result from receiving a remote vehicle's (RV) Basic Safety Message populate the "rvBSM" field with the BSM of that vehicle. Before and after data records for the warning populate the "rvBSM" field with BSMs received from the same vehicle. The remote vehicle is identified by its Temporary ID contained within the BSM. Likewise, warnings that result from receiving a Pedestrian Safety Message populate the "vruPSM" field with PSMs from the vulnerable road user triggering the pedestrian collision warning. Due to their size and complexity (i.e., embedding several payloads), OBU data logs are Extensible Markup Language (XML) encoded and compressed as flat files for upload by the CUTR server to the SDC, along with a separate data dictionary.

4.1.2 Participant Surveys

In coordination with USDOT Independent Evaluators (IEs), CUTR designed a series of surveys to be administered throughout the deployment. The first survey (initial survey) was administered at the installation appointment to collect information about the participants' travel preferences to the study area, their knowledge about CV technologies, and to gauge the main reasons for joining the study. An exit survey was administered throughout Phase 3 to collect data as participants dropped from the study. The goal of this survey was to obtain information on the exit motivations and assess participant attrition.

Two separate surveys were administered to obtain information on the participant's experience with CV technologies and equipment, and for those exposed to warnings via the HMI, to obtain information on how they perceived and responded to the warnings. One survey was administered about midpoint of Phase 3 and one toward the end of Phase 3. These two survey instruments contained the same questions on CV application effectiveness to allow combining or comparing responses between waves.

4.1.3 Non-CV Data

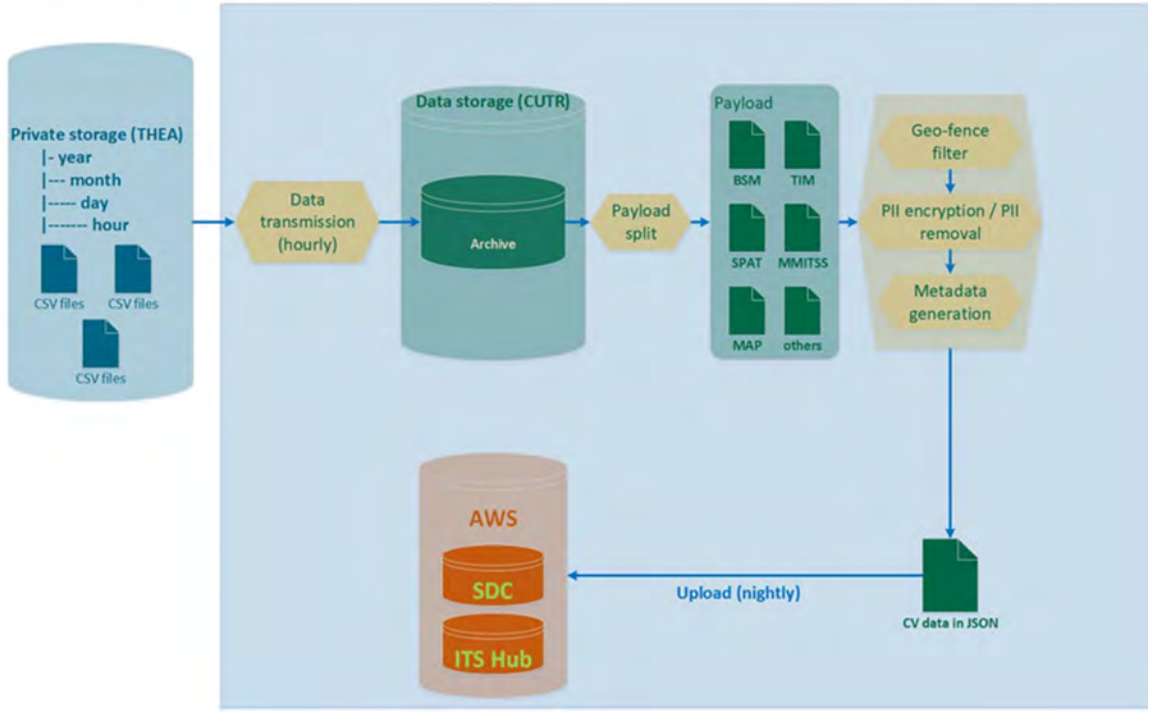
The performance evaluation and measurement support team collected data from other sources for the purposes of data integration and to measure observable confounding factors. The following datasets were collected throughout the deployment:

- Bluetooth reader data
- Transit GTFS (General Transit Feed Specification) data from Hart OneBusAway API (Application Programming Interface)
- Weather event data
- Crash data.

All the above data (except crash data) were uploaded to the SDC as raw data following the metadata structure detailed in Appendix C of the PMESP [4]. Even though crash data were collected and used in the analysis, the current agreement between the data provider and CUTR does not allow sharing this data outside those two parties.

4.2 Data Sharing

The CUTR server was configured and deployed to receive and archive all CV and non-CV data. CV data are stored as highly compressed flat files in THEA master server protected storage. As shown in Figure 4-1, the CV data are first split by payload (e.g., BSMs, TIMs, SPaTs), then subject to PII removal, and finally repackaged in a file format suitable for upload to the SDC and ITS Data Hub.



Source: CUTR, 2020

Figure 4-1. Data Sharing Flow from CUTR Server to USDOT

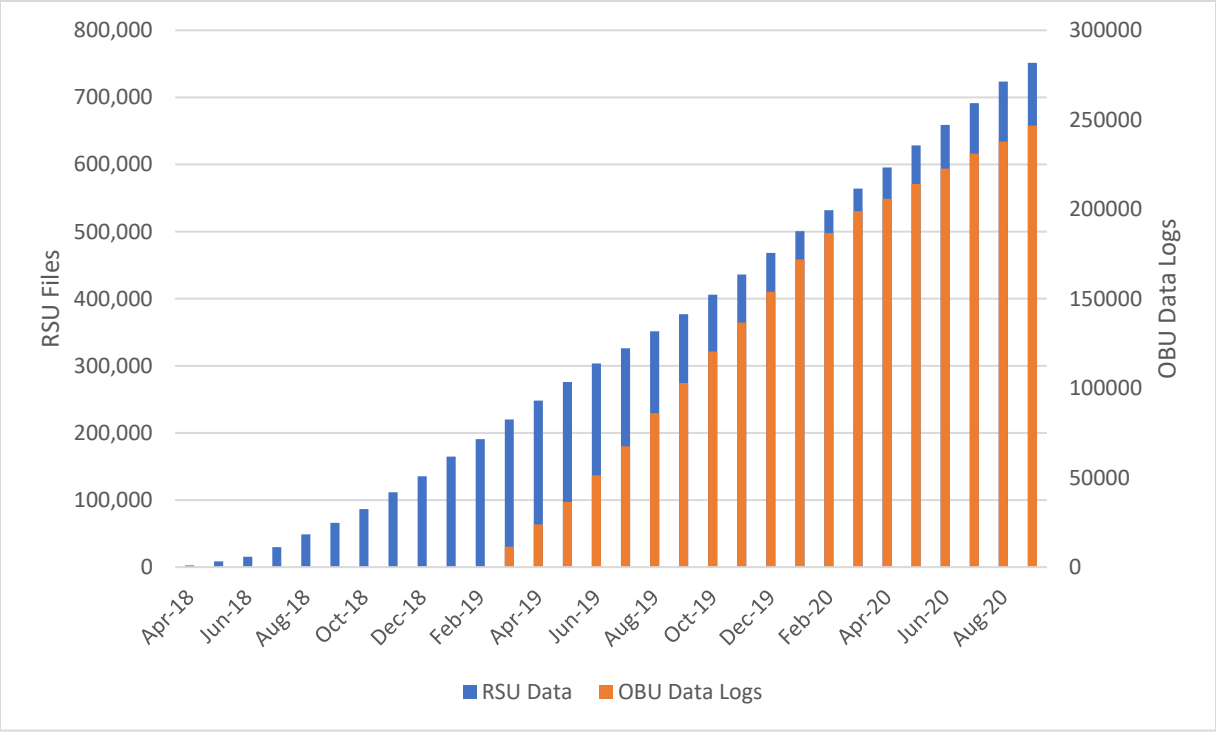
Of concern is any information contained in BSM data from the OBUs, RSUs/sensor data, and any other driving data that can be used to identify a vehicle. The contents in the raw data logs primarily contain BSMs generated from vehicles equipped with OBUs. The CUTR server performs PII removal as part of a nightly batch job before uploading data logs to the SDC. To ensure PII removal, the data logs are subject to cordon truncation to limit the data analysis to the geographic confines of the Tampa CV Pilot study area. This is achieved by establishing a geofence around the CV Pilot study area and by eliminating all records that place the vehicles outside the cordon. All remaining records are those collected within the study area. To conform to IE safety evaluation needs, the SDC data logs contain a new randomly generated identifier (ID). This ID remains constant over the study time frame to allow the IE conducting safety evaluation performance assessment. This field is removed from the data available in the ITS Data Hub.

The SDC developers designed and deployed a set of data governance procedures to manage access data generated and uploaded by the USDOT CV Pilots. The SDC developed and implemented a "New Dataset Provider Form," which was provided to the CV Pilot Team. The form allows the data provider specifying the information and access level for each data field.

In coordination with the SDC and ITS Data Hub developers, data transmission was carried in nightly batches. Each night at 1:15 a.m. EST, the CUTR server processed and uploaded CV data generated during the previous day. The upload schedule contained provisions to re-upload data that were not transmitted during the previous period due to technical issues encountered either by the CUTR server or the SDC and ITS Data Hub platforms.

The flow of data between servers followed security protocols established in the Data Management Plan and security protocols established by the SDC and ITS Data Hub developers.

Figure 4-2 reports the CV data collected and shared for the period beginning April 1, 2018, through September 30, 2020. The research team uploaded 751,595 highly compressed RSU data files and 246,740 OBU data logs. Note that the OBU data logs became available as Phase 3 of the CV Pilot began (March 2019), whereas the RSU data were being collected since Phase 2 of the deployment (April 2018).



Source: CUTR, 2020

Figure 4-2. CV Data Generation and Sharing

Chapter 5. Reporting to Stakeholders

All performance measures generated under the PMESP are shared with USDOT and only relevant CV Pilot stakeholders via an interactive dashboard available at www.cavdashboard.com. The public can access data collected in the Pilot via the ITS Data Hub available at https://www.its.dot.gov/data/.

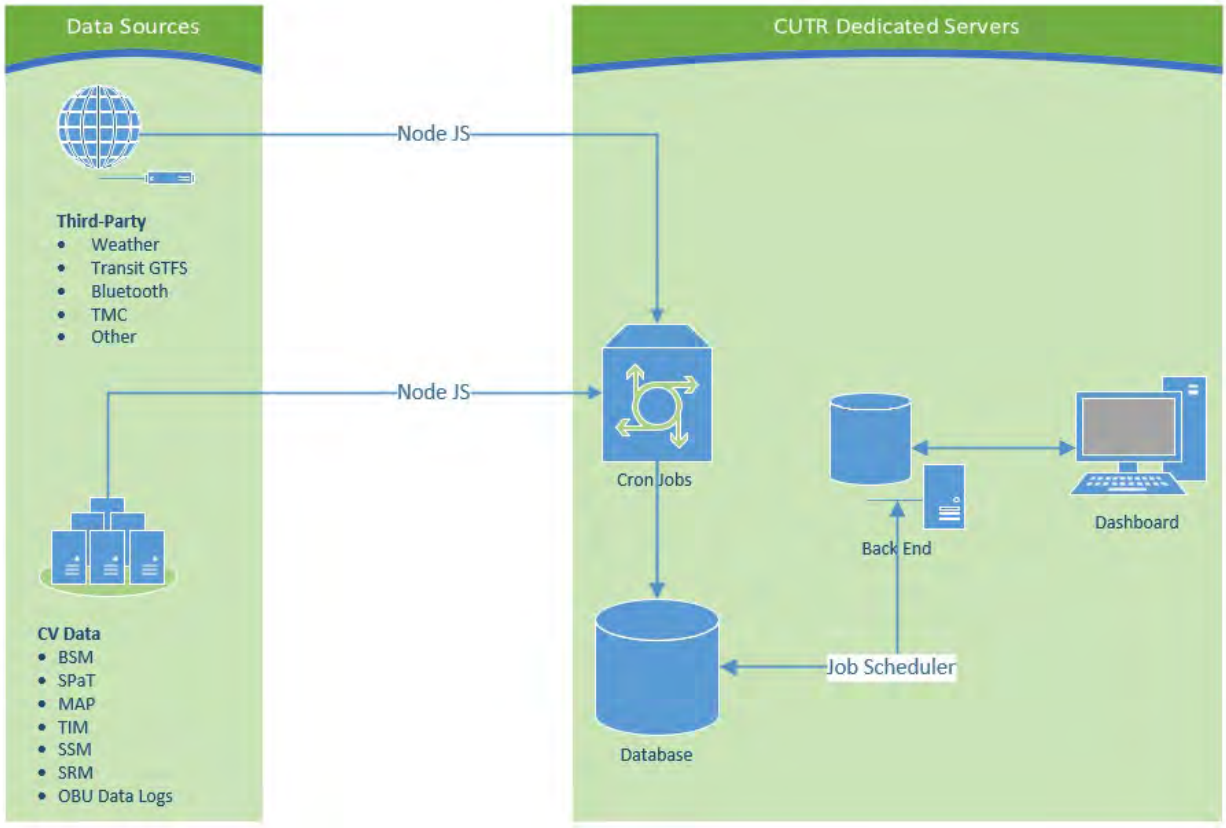
The Performance Measurement and Evaluation Dashboard (PMED) was developed initially to meet USDOT needs of performance reporting for several measures. The intent of the dashboard was to provide near real-time measures to the team and USDOT officials in a method that was developed with coordination with the USDOT.

5.1 Dashboard Flow

The portal relies on security features such as Hypertext Transfer Protocol Secure (HTTPS), data encryption using Advanced Encryption Standards (AES), and authorization tokens added at every layer of the entire architecture making it difficult to breach. The components of the dashboard and their integration are shown in Figure 5-1 and briefly explained below:

- **Dashboard Interface:** The dashboard is built using Hypertext Markup Language (HTML), Angular 7 JavaScript framework, NPM (Node Package Manager), RxJS (Reactive Extensions for JavaScript), TypeScript, and Leaflet (a map library). It is designed to support any form of device ranging from a smartphone to tablet or personal computer. The source code of the dashboard is completely uglified and minified so that no user can use the code to gain access. Also, the dashboard runs on HTTPS where it uses SSL (Secure Sockets Layer) to encrypt HTTP requests and responses. The flow of the dashboard is given in Figure 5-1.
- **Backend:** The data feeding into the dashboard are stored in a Structured Query Language (MySQL) database and served to the dashboard by a secure backend service in the form of a Representational State Transfer Application Programming Interface (REST API) built on Node.js and Express.js. These APIs are secured by unique access tokens that are generated by the backend when a user logs in and expire after three hours. This ensures that no user can access the data directly by calling a REST API. Further, the data being served to the dashboard are encrypted in the backend and decrypted back again when being received by the dashboard using AES, making the design more secure and reliable.
- **Job Scheduler:** The processed Pilot data from the performance evaluation team are synchronized to the MySQL database using scheduled jobs that run daily using job scheduler services. These services are built on Node.js and are an important part of the entire process.
- **Database:** A MySQL database is used to store the CV Pilot data to run the dashboard jobs.

F



Source: CUTR, 2020

Figure 5-1. THEA CV Pilot Dashboard Components

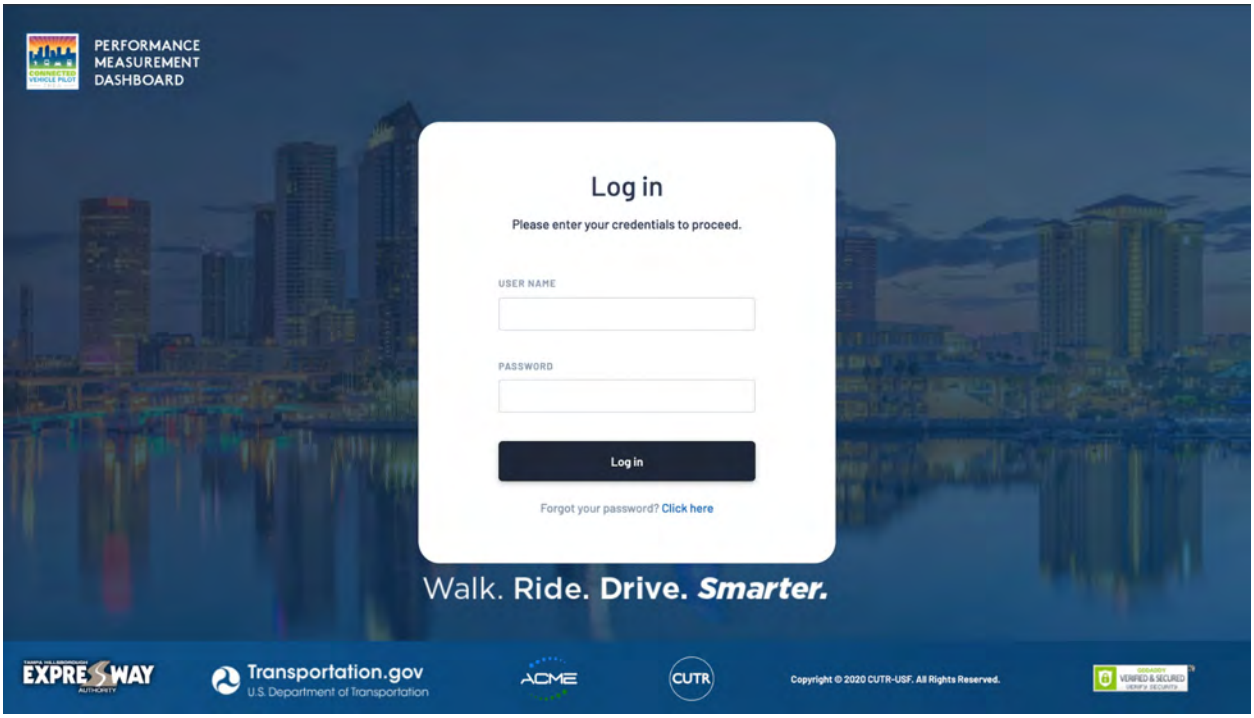
Daily, between 6 a.m. and 8 a.m., several job scheduler services begin processing various datasets of the Pilot, either in the form of disaggregate BSMs or aggregate performance measures computed by CUTR analysts. Specific jobs run to generate the warning profiles, which run throughout the day to check for new BSM data from the OBU data logs and to synchronize the existing BSM profiles around each event.

5.2 Landing Page

Figure 5-2 shows the login page. Upon entering credentials, an HTTPS request is sent to the backend API, which verifies the user's credentials, generates an authorization token, and sends the response back to the dashboard allowing the users to log in.

Once the user logs in, the token is saved in the browser session and is used to validate all future HTTPS requests from the dashboard to the backend. Each token is unique and expires once the user logs out or every three hours in case the user forgets to log out. This ensures that only authenticated users can use the dashboard.

After the user logs in, the dashboard shows different data layers based on the user type. The dashboard supports four different types of user levels: Administrator, Government Official, THEA Team, and USDOT Analyst.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 5-2. Dashboard Login Page

5.3 Performance Dashboard

This page provides a snapshot summarizing the continued monitoring of the overall THEA CV Pilot progress toward increased mobility and safety goals. Using near real-time dynamic data feeds from CV infrastructure, the dashboard displays key indicators reporting the operational health status of the deployed CV equipment, the amount of data generated by vehicles, and key mobility and safety performance measures related to the deployment of V2V and V2I applications. The elements on this page were developed in coordination with USDOT to present a high-level summary for all deployment aspects for each month during Phase 3 of the pilot.

This page can be accessed by administrators, USDOT government officials, and USDOT analysts. Figure 5-3 shows the landing page, which provides a high-level snapshot of the Pilot progress by way of performance gauges and is divided into three main types of indicators, with a map showing the aggregated warnings count for the current month. For each of the indicators there are record low, record high, and previous record high values.

Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 5-3. Performance Dashboard Page

5.3.1 Operational Indicators

The Operational Health Indicators provide information on the operational status of RSUs and vehicle OBUs, as well as the current participation rate by commuter vehicles. The operational health status is measured by three separate indicators monitoring equipment functionality on a 24/7 basis.

5.3.2 Application Activation Indicators

The Application Activation Indicators measure the amount of data generated by connected vehicles. There are three indicators measuring the volume of BSMs generated by vehicles and the number of warnings from the deployment of V2V and V2I safety and mobility applications.

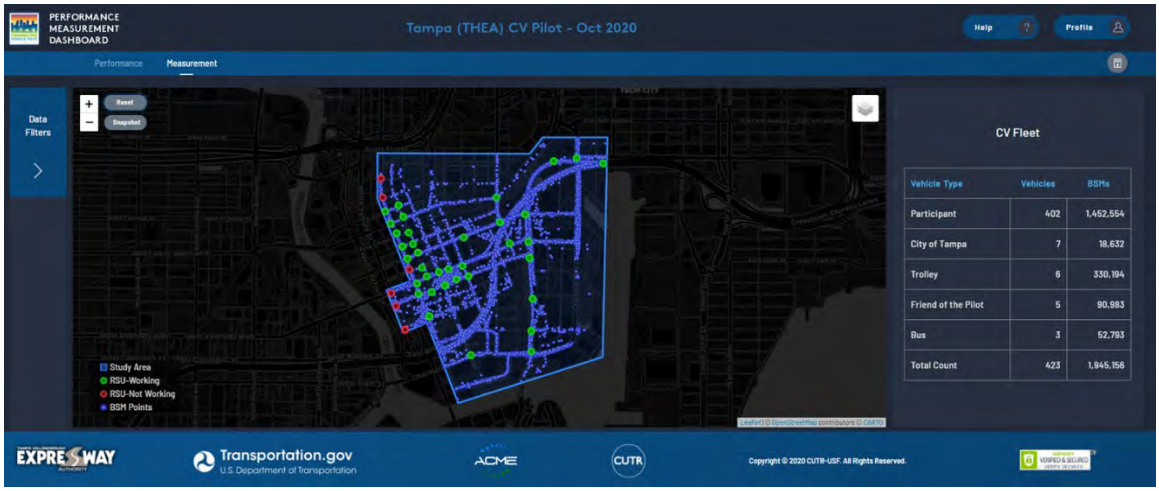
5.3.3 Performance Indicators

The Performance Indicators provide a summary of the performance measures established as part of USDOT's performance measurement and evaluation requirements of the THEA CV Pilot.

5.4 Measurement Dashboard

The Performance Measurement tab in dashboard can be viewed only by administrators, the THEA CV Pilot team, and USDOT analysts. This page provides more granular details about the CV fleet, the status of each RSU, BSM based performance measures, and a comprehensive assessment of the V2V and V2I warnings deployed since the Pilot's inception.

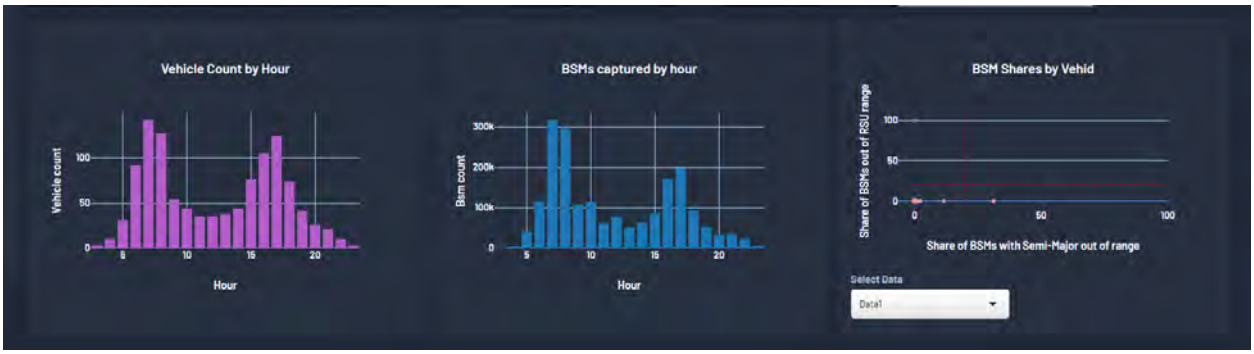
A series of data filers allows interacting with the database to focus on specific aspects of the Pilot performance, as shown in Figure 5-4 and Figure 5-5. Roadside units displayed in green and red indicate whether each RSU is operational or not. The BSMs are displayed in blue with frequency aggregation filters to quickly load the map. The CV fleet table shows the vehicle type, count of vehicles, and count of BSMs generated for each vehicle type during the selected day.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 5-4. Measurement Dashboard Page

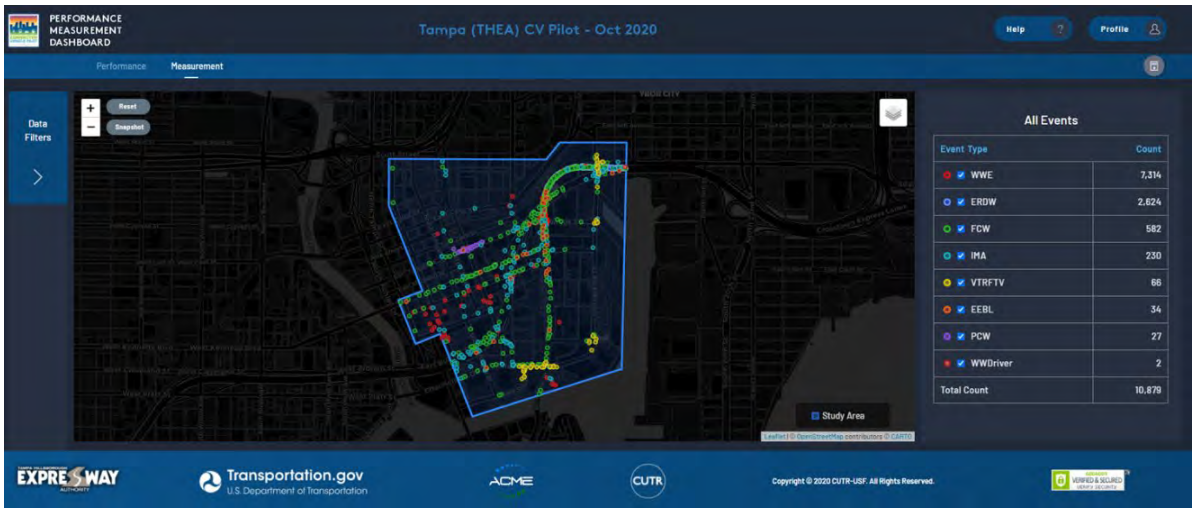
Figure 5-5 shows the histograms presenting vehicle counts and BSM counts by hour for the day chosen.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

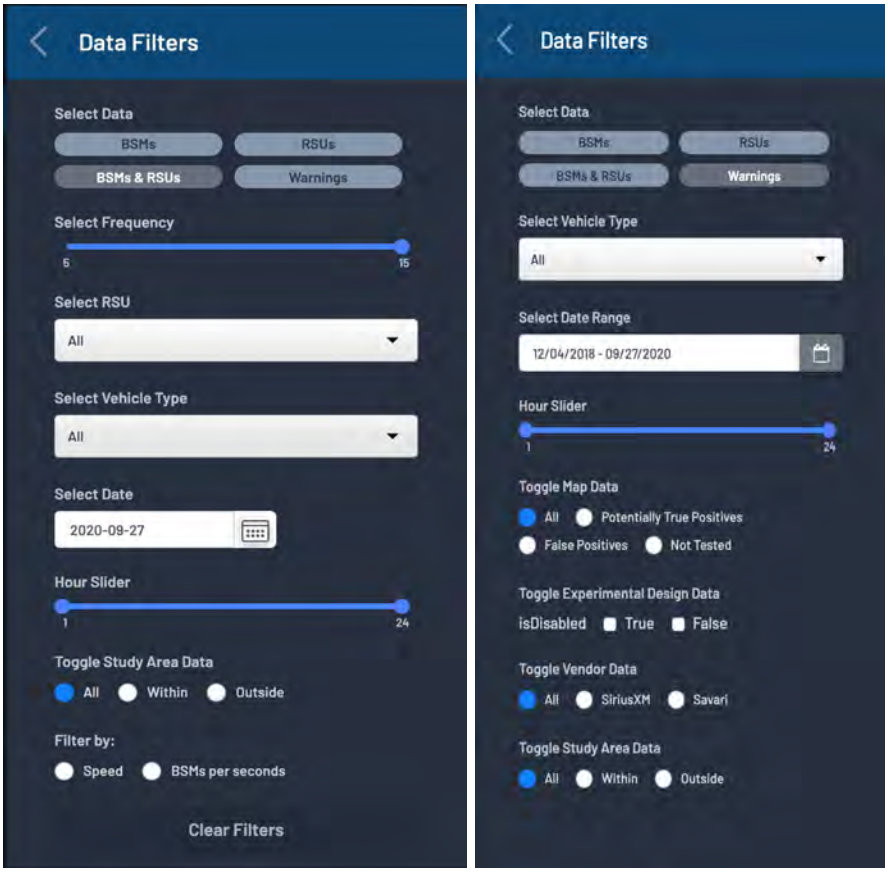
Figure 5-5. Data Analytics Example

Figure 5-6 shows collected warnings in the study area. The warning map displays all warnings generated throughout the study period and can be filtered further using the data filters sections (Figure 5-7). Along with the warnings on the map, a table containing the count of each warning type is also displayed. The data filters section is an important part of the dashboard, where the data can be filtered to meet user requirements. The filters change based on the type of data selected (BSMs or warnings).



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 5-6. Measurement Dashboard with CV Warnings



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

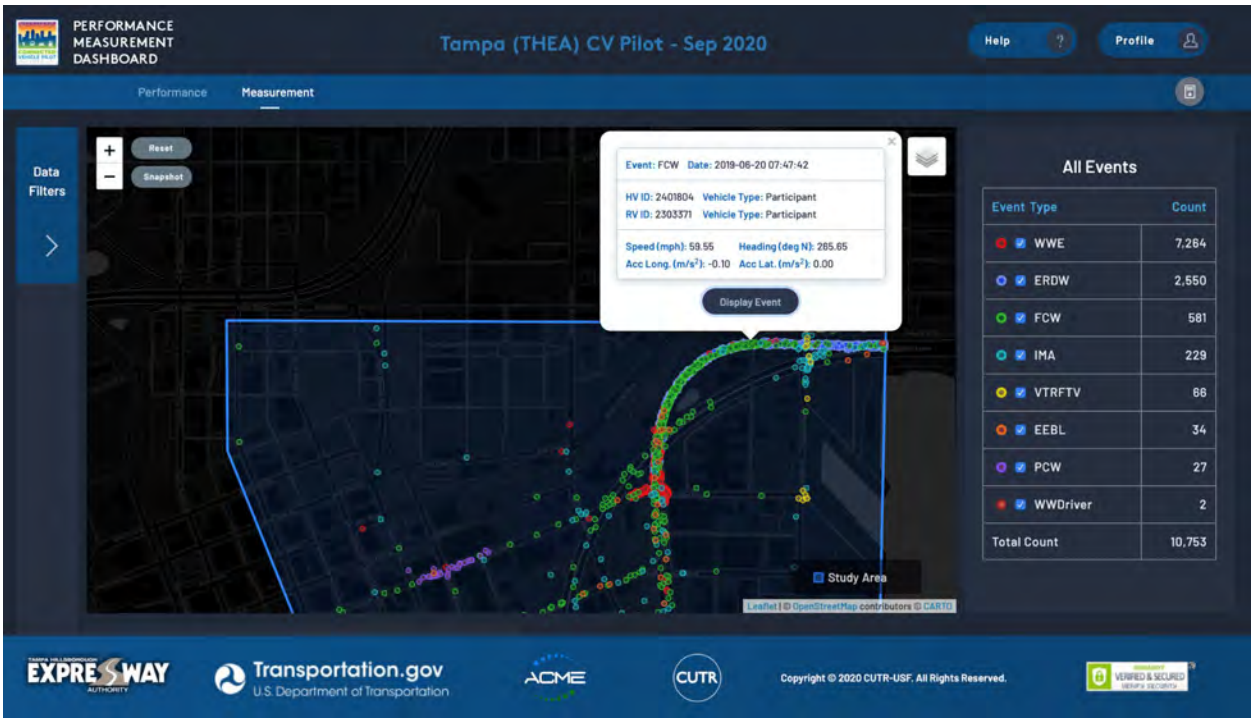
Figure 5-7. Dashboard Filters

5.5 Use Case Pages

Currently, six use case pages exist in the dashboard and each page focuses on the area describing the individual use case. The functionality is the same for each page. The pages follow the six use cases of the Pilot.

5.6 Warning Profiles

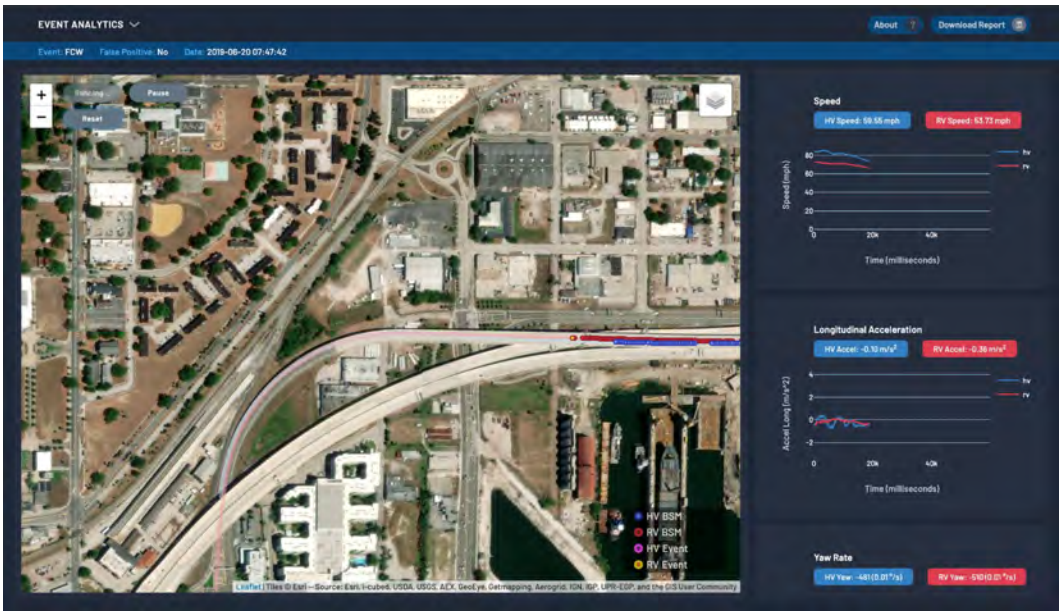
The last feature of the dashboard is the visual animation of every single event that triggers a warning in the study area. When a user selects a warning on the map, an informational pop-up appears with basic information about the event. The user may opt to display the event for further inspection. The warning profile is loaded in a new window by clicking on the Display Event button from the event pop-up as shown in Figure 5-8.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 5-8. Warning Event Profile Option

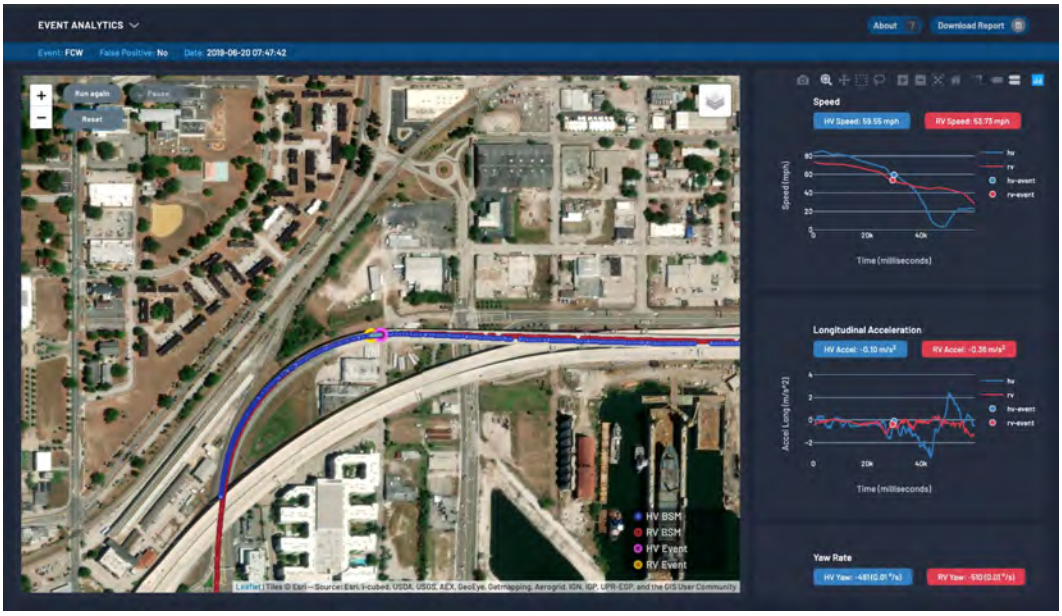
Figure 5-9 shows the warning profile page with a visual animation of the event through BSMs of the vehicles involved. Graphs showing speed, longitudinal acceleration, and yaw rate are also presented. The animation helps to recreate the event to better understand the conditions under which the warning was triggered and to analyze whether it was a true or false positive warning. The performance evaluation team developed a set of automated processes to conduct OBU Parameter Conformity Evaluation (PCE) as detailed in Chapter 7. PCE compares the actual values of the warning moments against the default OBU warning application deployment parameters. This is the first step leading to false positive identification.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

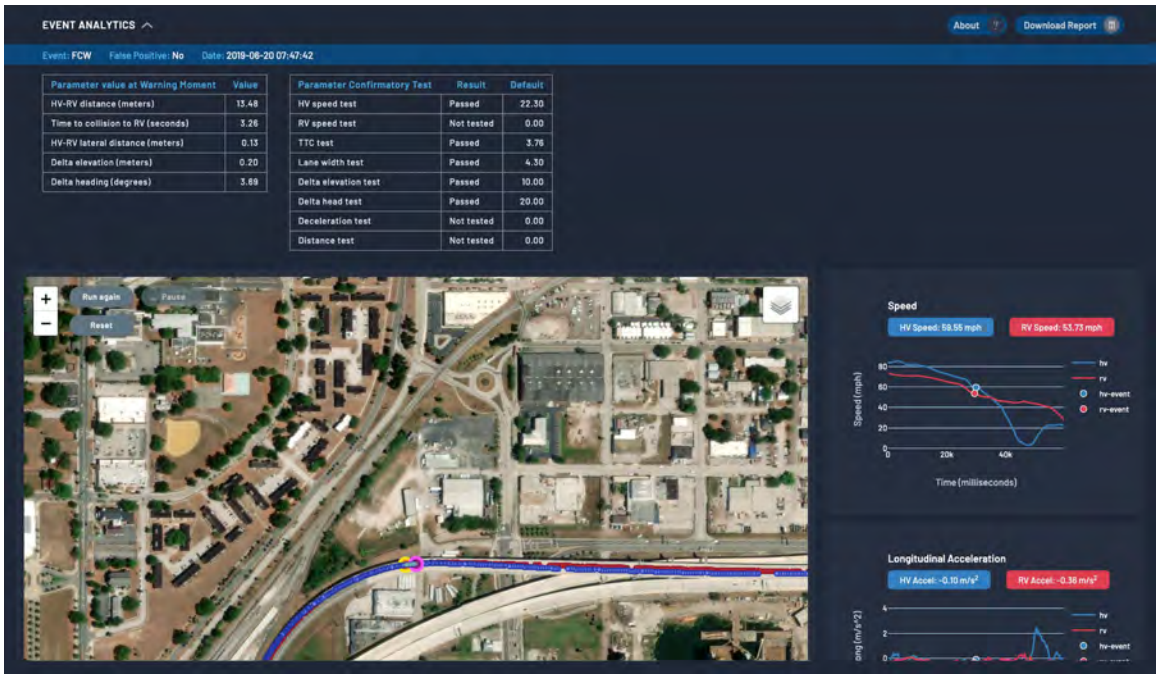
Figure 5-9. Warning Profile Animation

A complete warning profile and the event analytics are presented in Figure 5-10 and Figure 5-11, respectively.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 5-10. Complete Warning Profile



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 5-11. Warning Profile False Positive Analytics

Chapter 6. System Impact Evaluation Methodology

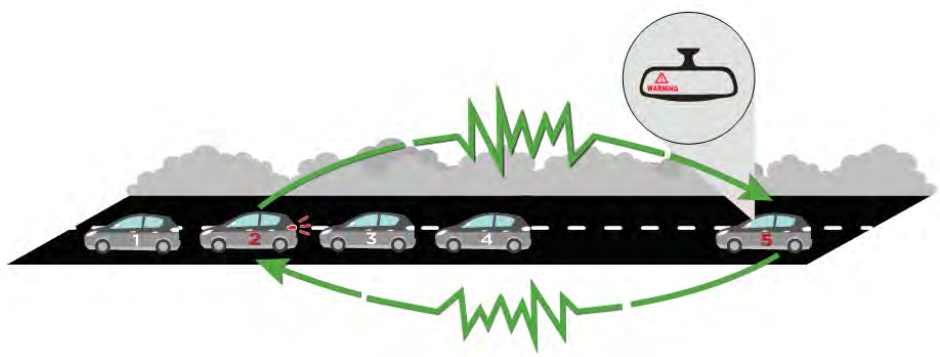
The methodology used for system evaluation first included evaluation of the V2V and V2I applications to determine if they met the requirements set by the project team. The applications were developed and deployed by aftermarket vendors using SAE standards [5, 6] and application specifications. Subsequently, the CUTR team evaluated if the applications met parameters set by the vendors. These are described in detail for each application in the following sections.

6.1 V2V Application Details

6.1.1 Electronic Emergency Brake Light

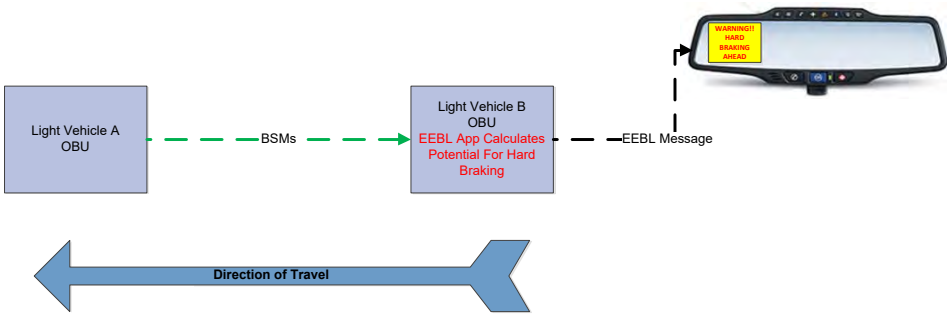
6.1.1.1 Functional Architecture and Parameters

The Electronic Emergency Brake Light (EEBL) application alerts the driver to hard braking in the traffic stream ahead. The alert is received from one or more vehicles in the same lane ahead, but not the immediate vehicle ahead. This provides the driver with additional time to perceive and assess situations developing ahead, as shown in Figure 6-1 and Figure 6-2 .



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-1. EEBL Functional Overview



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-2. EEBL Functional Flows

The reference parameters used by the vendors for operation of the EEBL application are shown in Table 6-1.

Table 6-1. EEBL Reference Parameter Values

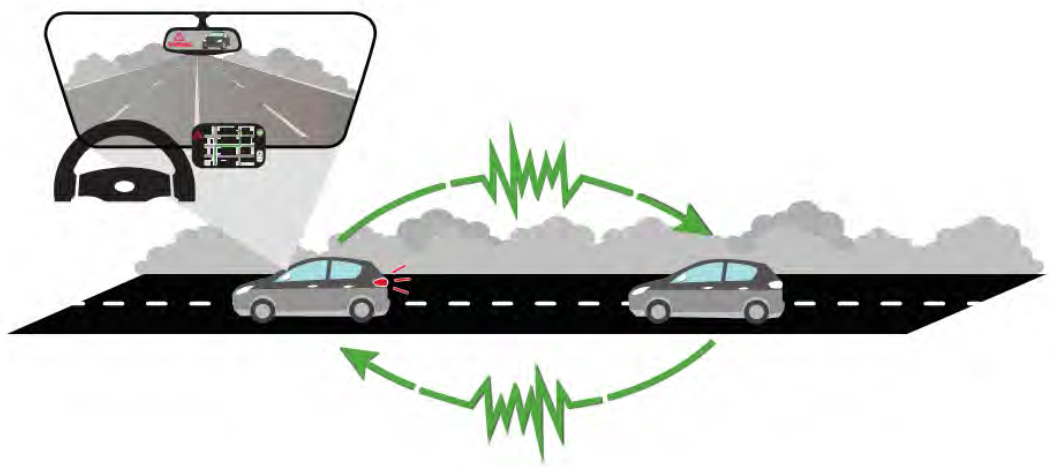
Parameter Type	Parameter Name (Unit)	Reference Value	Analytical Use
Operational	HV min. speed (kilometers per hour)	36	Interactions between vehicles
	RV min. speed (kilometers per hour)	36	
	HV-RV max. distance (meters)	100	
Configuration	HV-RV max. heading difference (degrees)	*	Potential conflicts between vehicles
	HV-RV max. elevation difference (meters)	5.4	
	HV-RV max. lane width (meters)	3	

** The vendors did not wish to publish this value.*

6.1.2 Forward Collision Warning

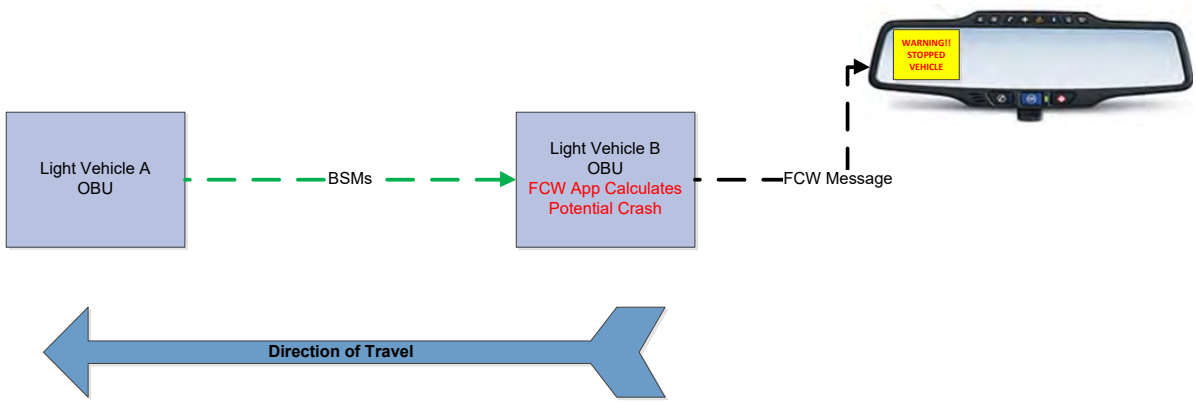
6.1.2.1 Functional Architecture and Parameters

The Forward Collision Warning (FCW) application alerts the driver to help avoid or mitigate the severity of crashes into the rear end of other vehicles on the road. FCW works lane by lane, responding to a direct and imminent threat ahead of the host vehicle. Anywhere two equipped vehicles interact, FCW deploys and provides a driver alert if the right conditions occur as follows: one vehicle is following the other and the lead vehicle brakes causing the closing distance to decrease (as calculated), which warrants an alert of a potential collision. Figure 6-3 and Figure 6-4 illustrate the FCW's functional architecture.



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-3. FCW Functional Overview



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-4. FCW Functional Flows

The applications use certain values as operational parameters that monitor the conditions between the host vehicle and any remote vehicles around it, and based on each scenario, assess the threat and issue warnings if necessary. For the FCW application, the reference parameters used by the vendors for the Tampa CV Pilot are shown in Table 6-2.

Table 6-2. FCW Reference Parameter Values

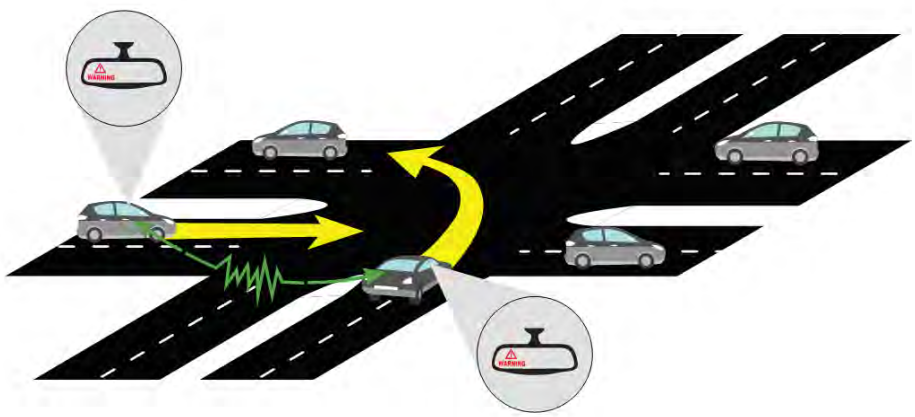
Parameter Type	Parameter Name (Unit)	Reference Value	Analytical Use
Operational	HV min. speed (kilometers per hour)	32	Interactions between vehicles
	HV-RV max. distance (meters)	100	
Configuration	HV-RV max. time to collision (seconds)	4	Potential conflicts between vehicles
	HV-RV max. heading difference (degrees)	*	
	HV-RV max. elevation difference (meters)	5.4	
	HV-RV max. lane width (meters)	3.0	

** The vendors did not wish to publish this value.*

6.1.3 Intersection Movement Assist

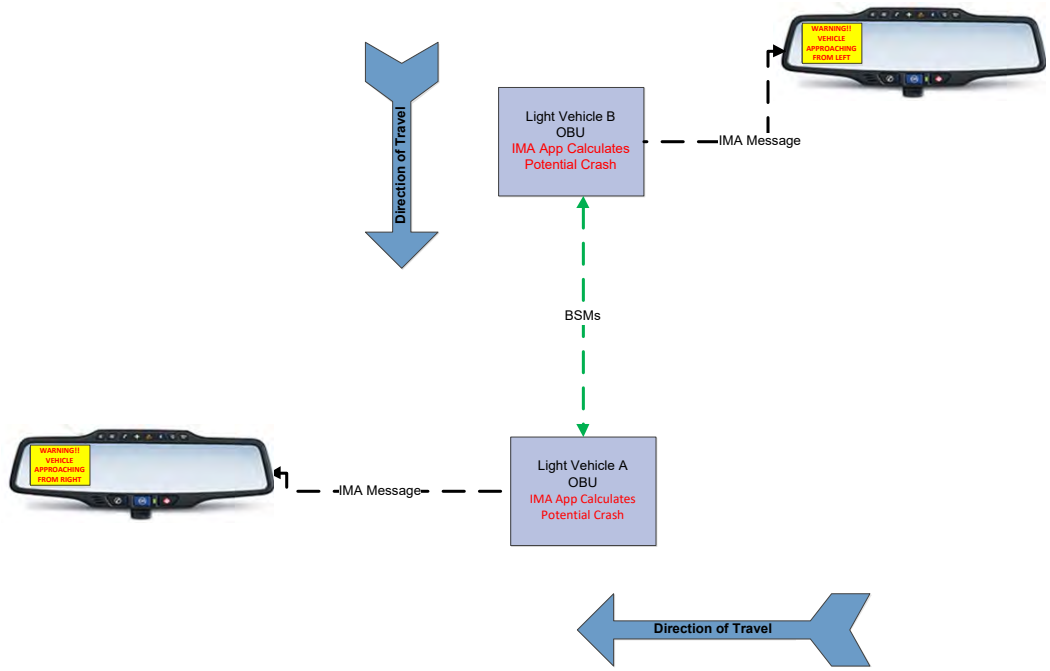
6.1.3.1 Functional Architecture and Parameters

The Intersection Movement Assist (IMA) application warns the driver of a potential collision when two or more vehicles are approaching one another by using the relative position, speed, and heading of those vehicles. The IMA application receives BSMs from approaching vehicles adjacent to the vehicle equipped with IMA. If IMA determines there is a high probability of a collision, the application warns the driver. Figure 6-5 and Figure 6-6 illustrate the functional flow architecture of the IMA application.



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-5. IMA Functional Overview



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-6. IMA Functional Flows

The parameters used by the vendors and CUTR to evaluate the operational status of the IMA application are shown in Table 6-3.

Table 6-3. IMA Reference Parameter Values

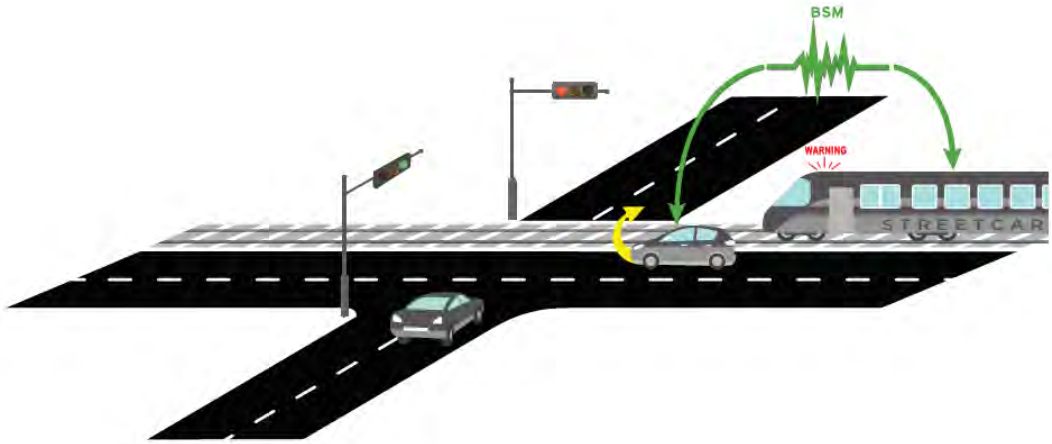
Parameter Type	Parameter Name (Unit)	Reference Value	Analytical Use
Operational	HV-RV max. lateral distance (meters)	70	Interactions between vehicles
	HV-RV max. longitudinal distance (meters)	86	
	HV-RV max. elevation difference (meters)	5.4	
	HV-RV min. heading difference (degrees)	65	
	HV-RV max. heading difference (degrees)	115	
Configuration	HV min. speed (kilometers per hour)	3.6	Potential conflicts between vehicles
	RV min. speed (kilometers per hour)	36	

6.1.4 Vehicle Turning Right in Front of Transit Vehicle

6.1.4.1 Functional Architecture and Parameters

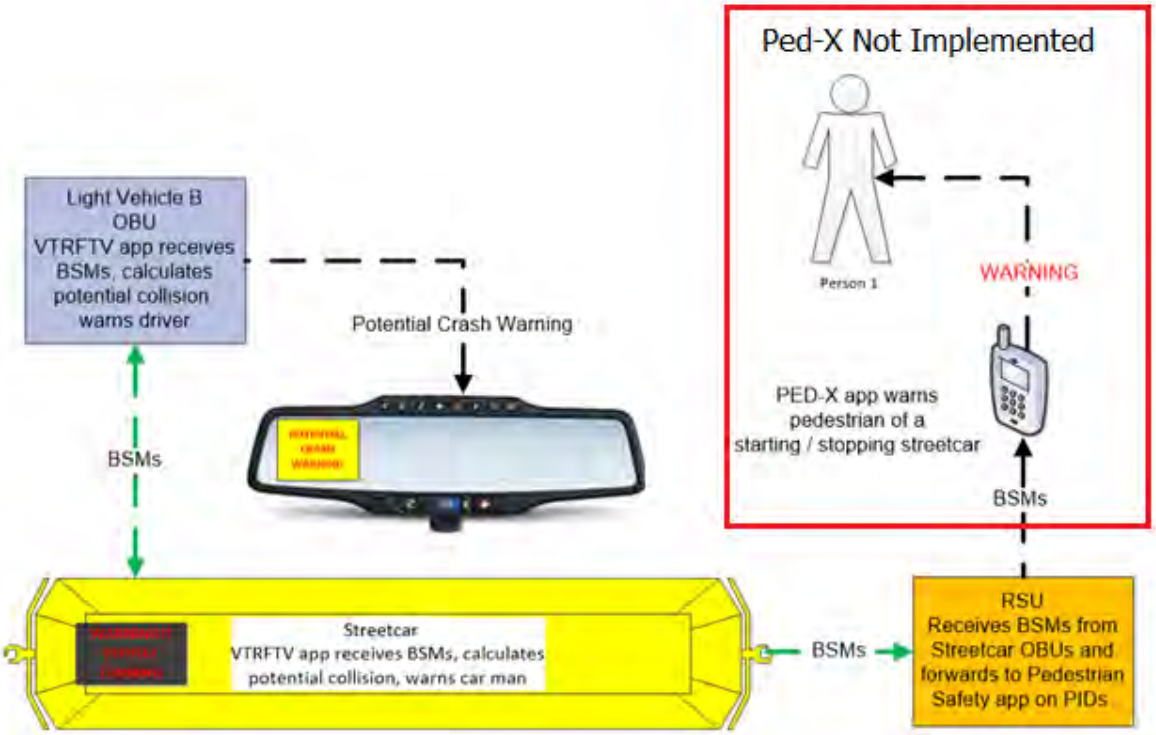
The Vehicle Turning Right in Front of Transit Vehicle (VTRFTV) application warns the streetcar operator of a vehicle turning right at an intersection the streetcar is approaching. Using BSMs that are being sent and received, the application determines the vehicles that are on a potential collision trajectory. Equipped vehicles receive a similar warning that they are on a collision course with a streetcar. Figure 6-7 shows the application's

functional overview and Figure 6-8 shows the functional flows. The pedestrian application component included in Figure 6-8 was not implemented due to the Global Positioning System (GPS) inaccuracy of Personal Information Devices (PIDs).



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-7. VTRFTV Functional Overview



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-8. VTRFTV Functional Flows

The reference parameters used by the vendors for the operation of the VTRFTV application are shown in Table 6-4.

Table 6-4. VTRFTV Reference Parameter Values

Parameter Type	Parameter Name (Unit)	Reference Value	Analytical Use
Operational	Streetcar–Participant Car max. heading difference (degrees)	*	Interactions between vehicles
	Streetcar–Participant Car elevation difference (meters)	5.4	
	Streetcar–Participant Car max. distance (meters)	100	
Configuration	Streetcar TTC (seconds)	4.2	Potential conflicts between vehicles
	Streetcar min. speed (kilometers per hour)	3	
	Participant Car min. speed (kilometers per hour)	10	
	Streetcar–Participant Car lane width (meters)	6	

** The vendors did not wish to publish this value.*

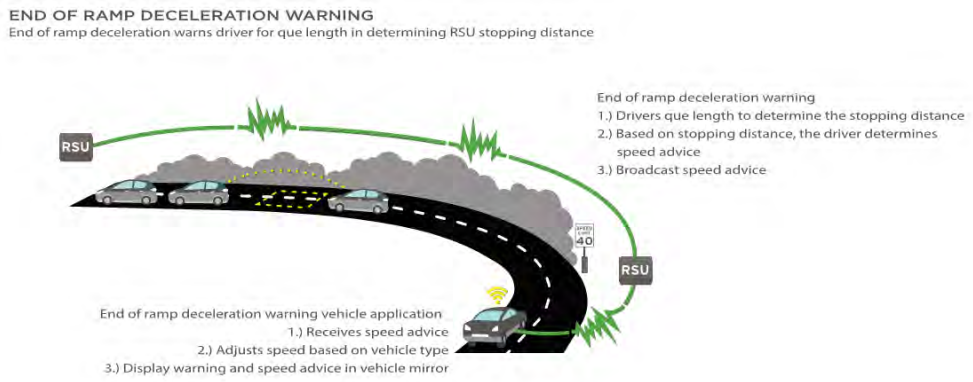
6.2 V2I Application Details

6.2.1 End of Ramp Deceleration Warning

6.2.1.1 Functional Architecture

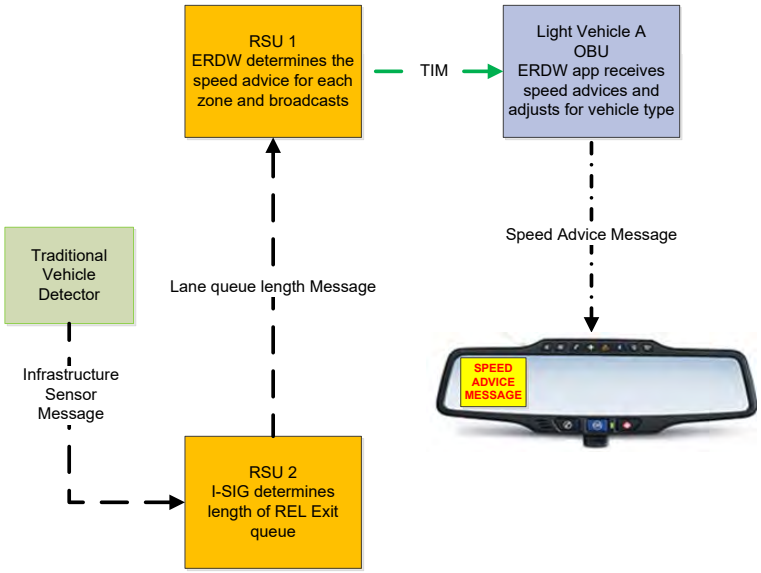
According to the original THEA CV Pilot functional architecture, the End of Ramp Deceleration Warning (ERDW) application computes a geo location of stopped traffic/vehicle queue based on the longest lane queue length computed by the I-SIG application. Overlapping I-SIG applications at Twiggs and Meridian estimate the queue length from the end of the REL. An Infrastructure Sensor Message (ISM) is generated when a vehicle passes a traditional vehicle detector (e.g., radar or video) and is provided to I-SIG only. I-SIG uses the ISM to enhance its queue length estimation. The REL is divided into one or more speed zones extending from Twiggs to the Selmon's main lanes. Based on the end of queue location, the RSU sends a Traveler Information Message with the recommended speed for each zone based on the safe stopping distance noted in the *Official Florida Driver License Handbook*. As the driver approaches the end of queue, the recommended speed TIM drops to within the safe stopping distance or the posted speed, whichever is lower for that zone.

There is a complementary OBU application that receives the recommended safe speeds as TIMs. The OBU application adjusts the recommended safe speed based on vehicle type and sends a message to the HMI for display to the driver. For example, a loaded heavy truck would adjust the TIM speed to a lower speed for that stopping distance and vehicle weight for each zone. Figure 6-9 and Figure 6-10 illustrate the functional flow architecture for the ERDW application.



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-9. ERDW Functional Overview



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-10. ERDW Functional Flows

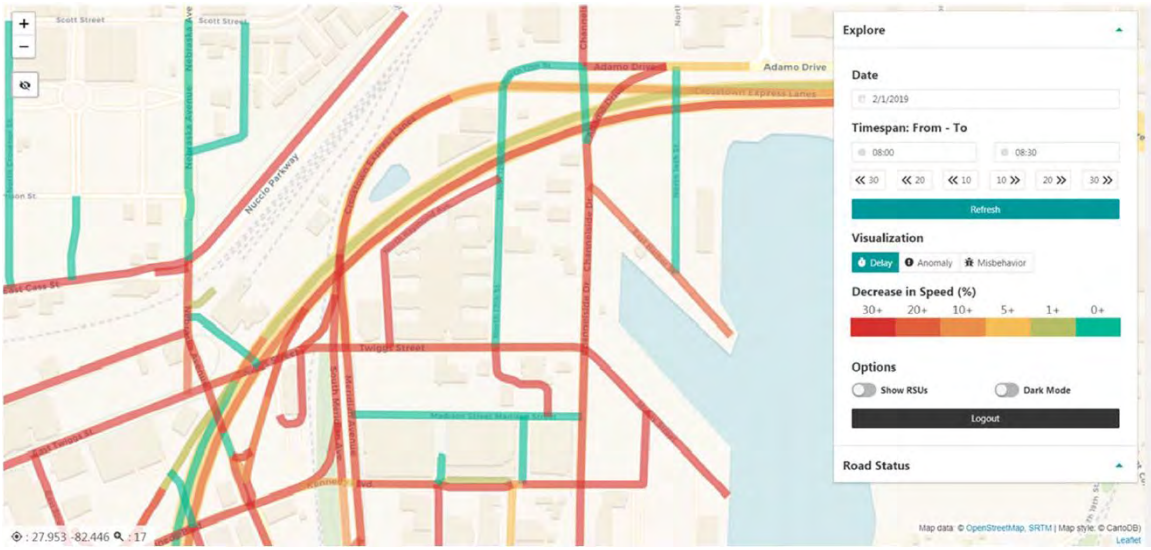
6.2.1.1.1 Adopted Resolution to Queue Length Estimation

The I-SIG implementation for Use Case 1 was designed to feed information to the Multimodal Intelligent Traffic Signal System (MMITSS) application to estimate queue length (along with other traffic delay measures) so that advisory speeds would be sent to participants via the ERDW application as they approached the REL exit. Subsequent performance tests revealed that the MMITSS application was not correctly estimating queue length [7].

During Phase 3 of the deployment, Siemens Mobility provided the following queue length estimation solution, which was implemented in February 2020.

MMITSS queue length estimation is replaced using queue length measurements derived from the participants' BSMs. Data collection on the REL to date indicates that one or more participant vehicles are typically traveling inbound on the REL during morning backup.

Figure 6-11 shows BSMs analyzed between 8:00 and 8:30 a.m. to display the percent decrease in speed of participant vehicles, color coded by location. The UC1 requirement to measure queues per lane was deferred as cars in lanes with longer queues unpredictably moved to other lanes with shorter queues. Under the new approach, the queue length estimate fed to the ERDW application is the longest queue measured across all REL lanes.



Source: Siemens. 2020

Figure 6-11. Speed Decrease Derived from BSMs

BSM data are aggregated on each segment of the road. Since the REL's queue stretches over multiple road segments, the multiple segments must be broken down further to detect queue length at a finer granularity. Figure 6-12 shows the segmentation, where the geometric representation is converted to a sequence of points that are distanced approximately 8 to 11 meters from each other, with their geocoded coordinates associated with each index of the sequence. The geometric representation is used to create a vector/array of length equal to the number of points, with each index being associated with the location.

Source: Siemens, 2020

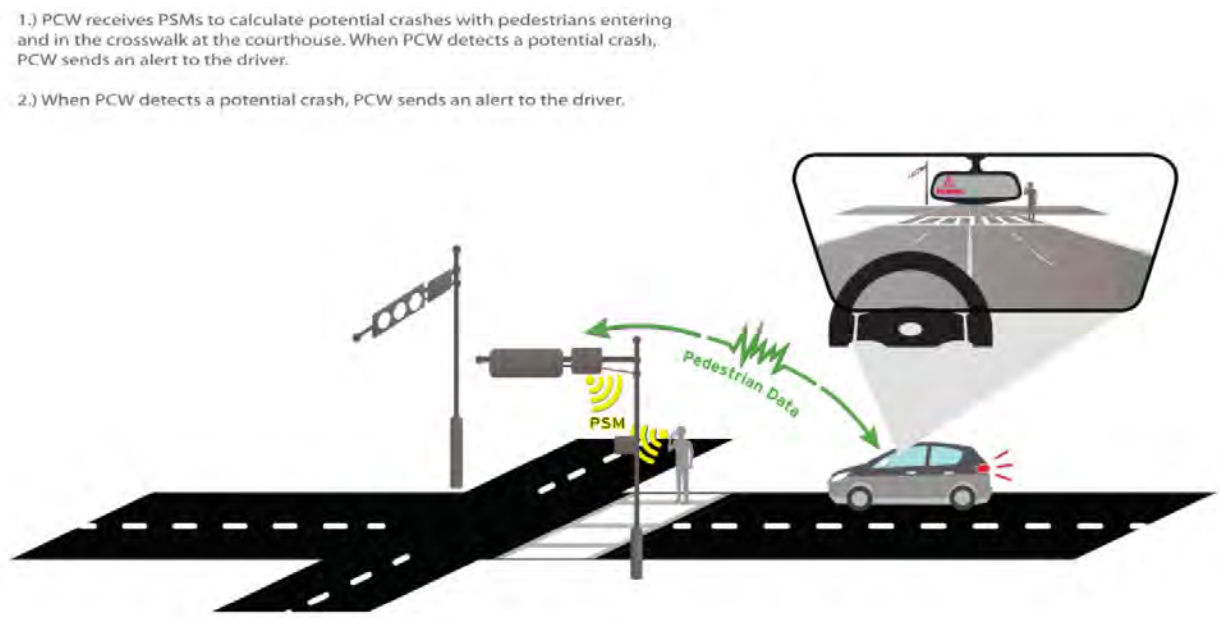
Figure 6-12. REL Index Locations

When a connected vehicle sends a BSM, the streaming data are processed to update the speed information at the location of the vehicle. The value is stored at the index associated with that location. The vector/array is updated if the vehicle's speed is lower than the speed observed earlier at the same location within the same minute. A queue detection algorithm then parses the vector/array and extracts the queue information that is needed to determine the tail end of the vehicle queue. These queue length estimates are then fed back to the RSU to broadcast the Traveler Information Messages indicating the speed advice for each zone approaching the end of queue.

6.2.2 Pedestrian Collision Warning

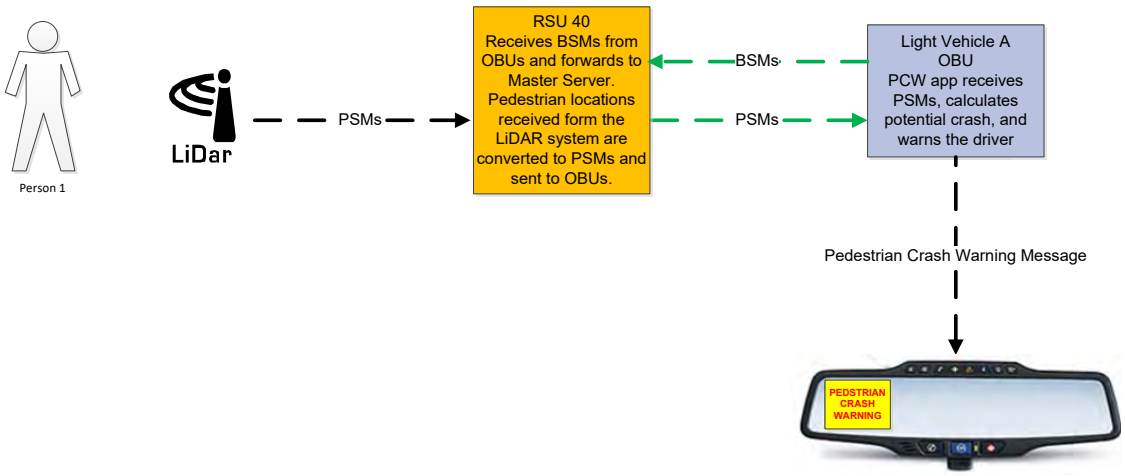
6.2.2.1 Functional Architecture and Parameters

The PCW (Pedestrian Collision Warning) application was designed to work at the midblock crosswalk on East Twiggs Street at the Hillsborough County Courthouse to improve pedestrian safety. Initially, two Light Detection and Ranging (LiDAR) sensors installed at the crosswalk located pedestrians in the area, translated the information to PSMs, and then broadcast them over Dedicated Short-Range Communication (DSRC) to the OBUs. OBU equipped vehicles using the PCW application warned drivers who were on a collision course with a pedestrian in the crosswalk. Figure 6-13 shows the application's functional overview and Figure 6-14 shows the functional flows that originally included a LiDAR. However, LiDAR was replaced with thermal imaging sensors because upon deployment of the LiDAR system, the research team concluded that the operational reliability of the LiDAR sensors was not adequate to support UC3 goals within the project's time frame.



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-13. PCW Functional Overview



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

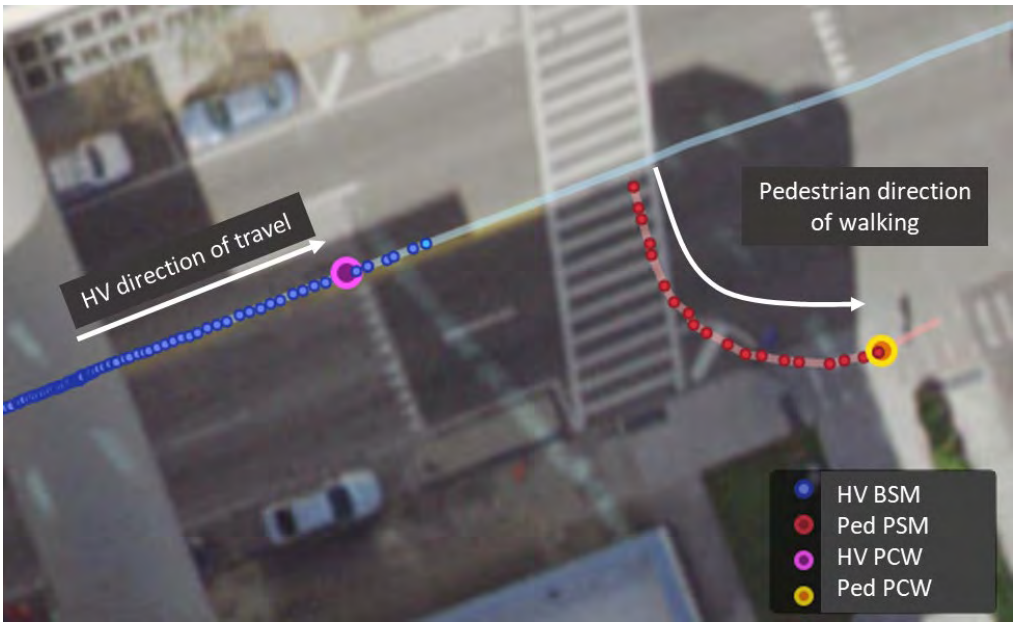
Figure 6-14. PCW Functional Flows

6.2.2.1.1 Issues with LiDAR Sensors and Adopted Resolution

As per the System Architecture Document [8], the system to detect and generate PSMs for Use Case 3 included two LiDAR sensors that would detect, identify and track pedestrians on the sidewalk and marked crosswalk, generate PSMs, and broadcast via two roadside units, one at each side of the crosswalk. The two units would serve as redundancy in case an obstruction (e.g., a box truck) blocked the view of one of the sensors. The system was first tested per the Operational Readiness Plan on April 24, 2018, on a closed course mockup crosswalk at Hillsborough Community College, which served as a testing site. Upon completion of the utility power installation, the LiDAR sensors, RSUs, and PED-X application were installed at

the courthouse and became operational, along with two surveillance cameras mounted on the same poles to provide a remote view of the UC3 area.

After becoming operational at the site, several issues surfaced. The LiDAR system seemed to not track pedestrians until they were located halfway through the crosswalk. The two LiDAR sensors had to work harmoniously to detect pedestrians crossing the road from one side and finishing on the other side. One of the LiDAR sensors failed and therefore could not track pedestrians until they were halfway through the crosswalk. Figure 6-15 shows an example of a pedestrian trajectory tracked for only half of the crosswalk.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 6-15. LiDAR Failure to Track Pedestrian on Crosswalk

After this issue was identified, the sensor vendor replaced the LiDAR sensor with a different model, but the old sensor software driver was incompatible with the output format of the new sensor, requiring further development and regression testing.

Another issue with the LiDAR sensors was that they could not provide tracking of one pedestrian with the same ID. This was one of the requirements of the Performance Measurement and Evaluation. CUTR conducted tests in September 2019 to determine if the pedestrians were assigned IDs correctly (one ID for a pedestrian crossing the crosswalk). The tests showed that in a matter of minutes, and with only one pedestrian crossing the crosswalk, the system would generate thousands of PSMs with several hundred IDs. Figure 6-16 shows PSMs generated on September 19, 2019, between 8:41 p.m. and 8:59 p.m. Between 8:41 p.m. and 8:48 p.m., a CUTR researcher crossed the crosswalk a few times and was the only pedestrian present.

Between 8:48 p.m. and 8:59 p.m., one pedestrian and one bicyclist passed the area on the sidewalk (did not cross). As shown in Eventually, it was concluded that the operational reliability of the LiDAR sensors was not adequate to support the research goals within the limited time remaining in Phase 3 of the project; thus, the LiDAR sensors were replaced with video and thermal imaging sensors [9].

during this time when no pedestrians crossed the crosswalk, the LiDAR system reported 142 pedestrian IDs and 2,362 Pedestrian Safety Messages. This test showed two failures of the LiDAR system: first, the system reported several different IDs for only one pedestrian and second, the LiDAR system generated PSMs when no pedestrians were present.

Eventually, it was concluded that the operational reliability of the LiDAR sensors was not adequate to support the research goals within the limited time remaining in Phase 3 of the project; thus, the LiDAR sensors were replaced with video and thermal imaging sensors [9].

Source: CUTR, 2020

Figure 6-16. PSM Test Results

Table 6-5. PSM Test Results

Minute	Unique IDs	Total PSMs
48	9	203
49	13	202
50	12	126
51	13	168
52	10	60
53	7	61
54	11	151
55	9	244
56	19	349
57	9	230
58	18	289
59	12	279
Total	***142***	***2,362***

6.2.2.1.2 Adopted Resolution

To achieve the research goals within the time frame, replacing the LiDAR sensor with video and thermal imaging sensors was recommended by Siemens Mobility Inc. The thermal imaging sensor detects heat radiated as infrared light, such as pedestrians, cyclists, and internal combustion vehicles. Thermal imaging is more costly but detects heat sources in total darkness and challenging situations such as fog or dust. The video imaging sensor detects light reflected from objects, such as pedestrians, cyclists, internal combustion vehicles, and inanimate objects. A video imaging sensor is less costly but requires sunlight or street lighting to illuminate the objects. Both detector types were recommended for testing and deployment in parallel for risk mitigation and to inform future deployments. The change in equipment is document by Siemens Mobility Inc. The new system was installed in May 2020 and subsequent testing using test vehicles was conducted in June, July, and August 2020. On August 5, 2020, the system began full operation and deployment to participants.

Table 6-6 reports the application parameters used to perform Parameter Conformity Evaluation and true conflict identification.

Table 6-6. PCW Reference Parameter Values

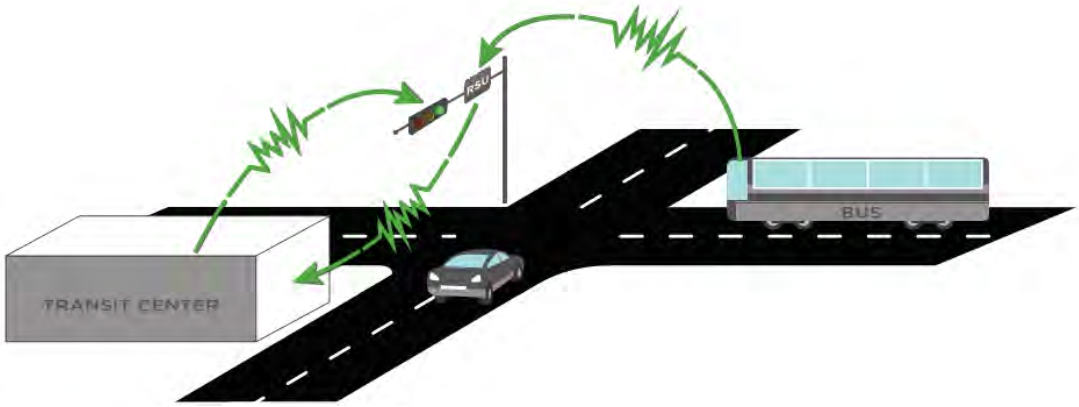
Parameter Type	Parameter Name (Unit)	Reference Value	Analytical Use
Configuration	HV min. speed (kilometers per hour)	8	Potential conflicts between vehicle-pedestrian
	HV-Pedestrian max. distance (meters)	70	
	HV-Pedestrian TTC (seconds)	6	

6.2.3 Transit Signal Priority

6.2.3.1 Functional Architecture

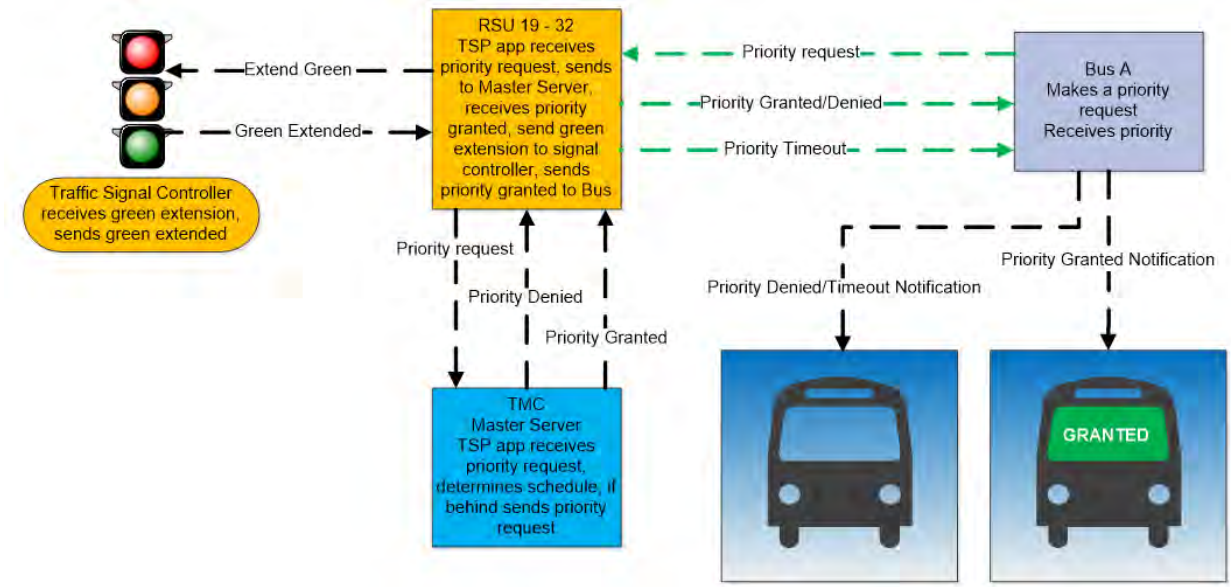
According to the THEA CV Pilot System Design Document (SDD), the Transit Signal Priority app provides signal priority to transit at intersections along arterial corridors only if the bus is behind schedule. TSP is part of MMITSS.

If the bus is behind schedule, priority will be granted for the bus. The OBU sends a Signal Request Message to the RSU. The RSU forwards that to the Transit Server at the Traffic Management Center. The Transit Server determines if the bus is behind schedule. If the bus is behind schedule, the SRM is returned from the Transit Server to the RSU. If the signal is green in the bus's travel direction, the RSU selects the controller phase via National Transportation Communications for Intelligent Transportation System Protocol (NTCIP) objects to extend the green, allowing the bus to proceed through the intersection. If the signal is yellow or red in the bus's travel direction, the RSU requests the shortest cycle via NTCIP objects to provide a green to the bus as quickly as possible. At the same time, the RSU sends the Signal Status Message to the approaching bus to inform the driver of priority received. Figure 6-17 and Figure 6-18 show the functional overview and flows of the application.



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-17. TSP Functional Overview



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2020

Figure 6-18. TSP Functional Flows

Transit Signal Priority is part of MMITSS, which is available on the Open Source Application Development Portal (OSADP). As part of this application suite, TSP must be used in conjunction with I-SIG. TSP provides signal priority to transit at intersections and along arterial corridors. The OBU sends a Signal Request Message to the RSU. The RSU forwards that to the Transit Server (housed on the Master Server) at the Traffic Management Center. The Transit Server determines if the bus is behind schedule. If the bus is behind schedule, the SRM is returned from the Transit Server to the RSU. The RSU determines priority for all SRMs received from all approaching vehicles, and then selects the controller phase via NTCIP objects to extend the green, allowing the bus to proceed through the intersection. At the same time, the RSU sends the Signal Status Message to approaching vehicles to inform them of which have received priority to extend the green and which have been denied priority [8].

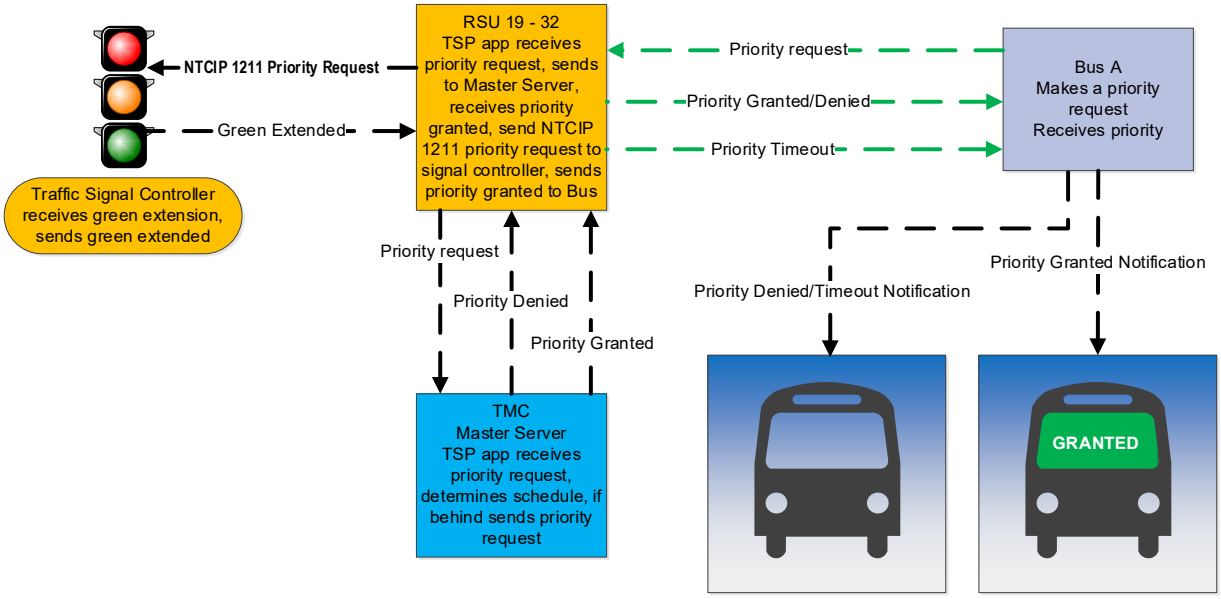
6.2.3.1.1 Issues with TSP Deployment

Transit Signal Priority was deployed and tested as follows:

1. MMITSS was ported to run on the Siemens RSU and uploaded to the OSADP as planned.
2. From Central, Controller Units (CUs) were commanded to run normal coordinated signal plans.
3. CUs were tested and verified to run MMITSS signal plans.
4. Research goals of the THEA Pilot required operational TSP for data collection.
5. TSP was part of MMITSS.
6. MMITSS version available on OSADP did not support coordination, each CU ran free.
7. City of Tampa would not approve use of MMITSS signal plans on coordinated corridors.
8. Research goals required TSP to operate on UC4 coordinated corridors.
9. No TSP data was being collected due to inability of MMITSS to operate coordinated.

6.2.3.1.2 Adopted Resolution

In order to achieve the research goals, replacing MMITSS and NTCIP 1202 standard messages with NTCIP 1211 standard messages was proposed as shown in Figure 6-19.



Source: Transit Signal Priority Operational Plan Description, Siemens, 2020

Figure 6-19. Proposed TSP Operation

The Signal Request Messages and Signal Status Messages between the RSU and OBU of the bus are maintained and unaffected in the proposed resolution as follows:

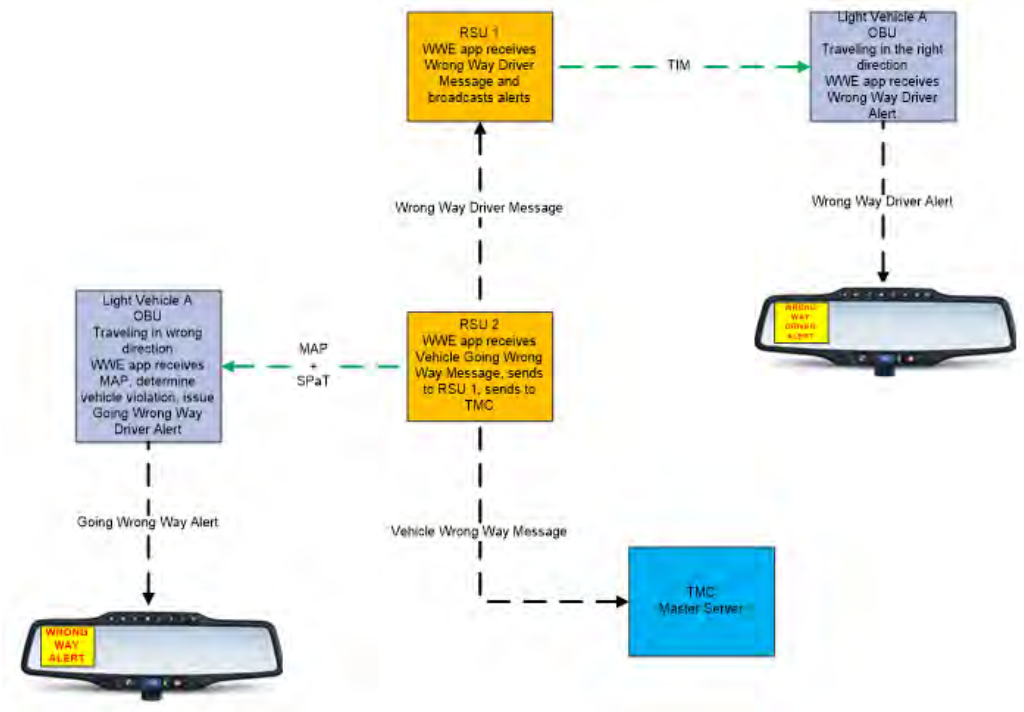
1. From Central, Controller Units (CUs) are commanded to run normal coordinated signal plans.
2. MMITSS remains disabled, all CUs run coordinated.
3. SRMs received from buses are converted to NTCIP 1211 priority requests to the CU.
4. NTCIP 1211 priority status responses from the CU are converted to SSMs to the buses.

This proposed resolution had not been thoroughly tested as of the date of this report.

6.2.4 Wrong-Way Entry

6.2.4.1 Functional Architecture

According to the THEA CV Pilot SDD, the Wrong-Way Entry (WWE) application warns OBU equipped vehicles approaching a RSU equipped intersection when the vehicles are not traveling in the allowed direction. At the East Twiggs Street and North Meridian Avenue intersection, the RSU broadcasts Map Data (MAP) and SPaT messages [10]. Figure 6-20 shows the functional flow architecture for the WWE application.



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

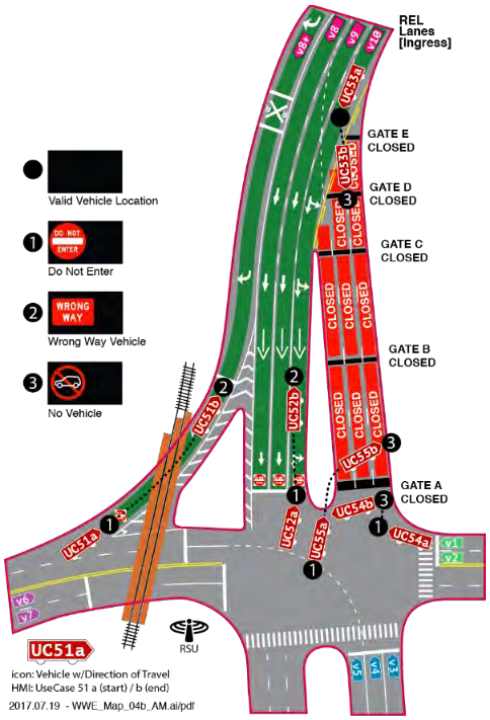
Figure 6-20. WWE Functional Flow

The WWE application has multiple warning levels that are all recorded with the same warning type (WWE) in the OBU data logs:

1. **DOT NOT ENTER** warning if determined that the vehicle is advancing to enter the REL going the wrong way.
2. **WRONG-WAY** warning if determined that the same vehicle has continued up the REL the wrong way.
3. **NO TRAVEL LANE** warning if the vehicle enters the outbound or inbound closed section of the REL.
4. **WRONG-WAY VEHICLE** warning to the legal inbound driver after the wrong-way violation occurs. This feature produces a different warning titled “Wrong-Way Driver” and is not part of this assessment.

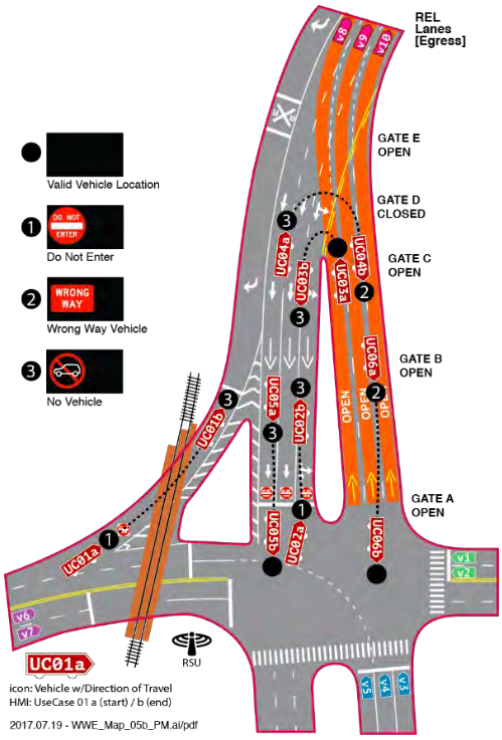
The analysis in this report refers to the first three types of WWE warnings. The three warnings are not identified separately in the warning data. They are all recorded with the generic WWE warning type. To determine the first warning, or pre-warning, the OBU analyzes the trajectory, speed, and allowed movements of vehicles on the Selmon Expressway’s morning (AM) and afternoon/evening (PM) REL operations, as shown in Figure 6-21 and Figure 6-22.[1]

[1] The figures represent the original REL turning movements defining the WWE zones. Changes to MAP configurations were subsequently made to reduce the number of AM false positives.



Source: Vehicle Systems – Onboard Unit (OBU) HMI Specification, 2018

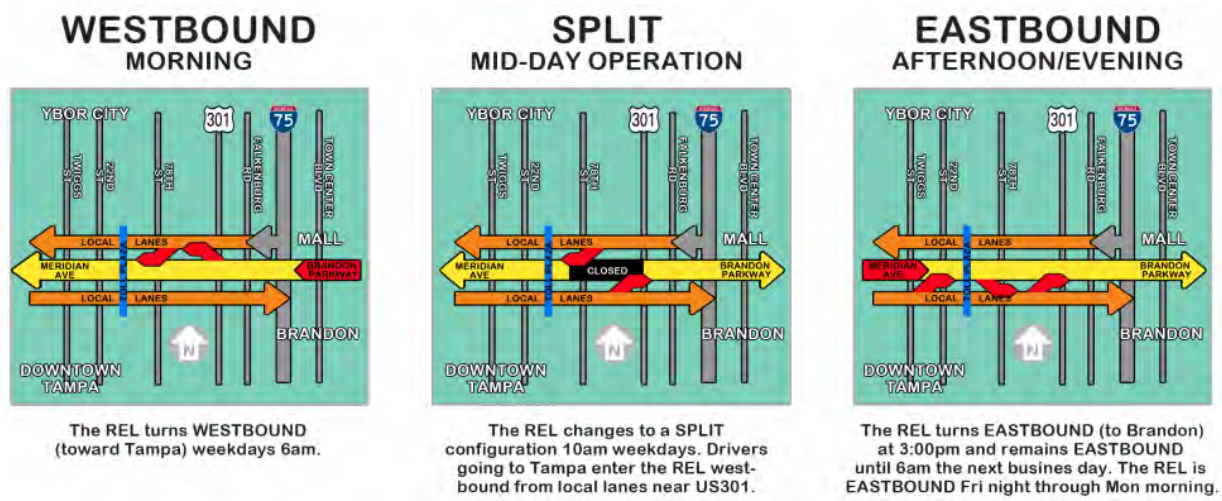
Figure 6-21. AM REL Movements



Source: Vehicle Systems – Onboard Unit (OBU) HMI Specification, 2018

Figure 6-22. PM REL Movements

According to the PMESP, the WWE application as described in Use Case 2 of the CV Pilot is intended to warn drivers entering the Reversible Express Lanes (REL) the wrong way [4]. The REL is a three-lane bidirectional expressway that provides a direct connection between Brandon and downtown Tampa, allowing for express travel of people in cars and buses. The schedule for the REL permits travel into downtown Tampa (westbound) in the morning between 6 a.m. and 10 a.m., split operation mid-day between 10 a.m. and 1 p.m., and travel toward Brandon (eastbound) in the afternoon/evening between 3 p.m. and 6 a.m. (Figure 6-23).



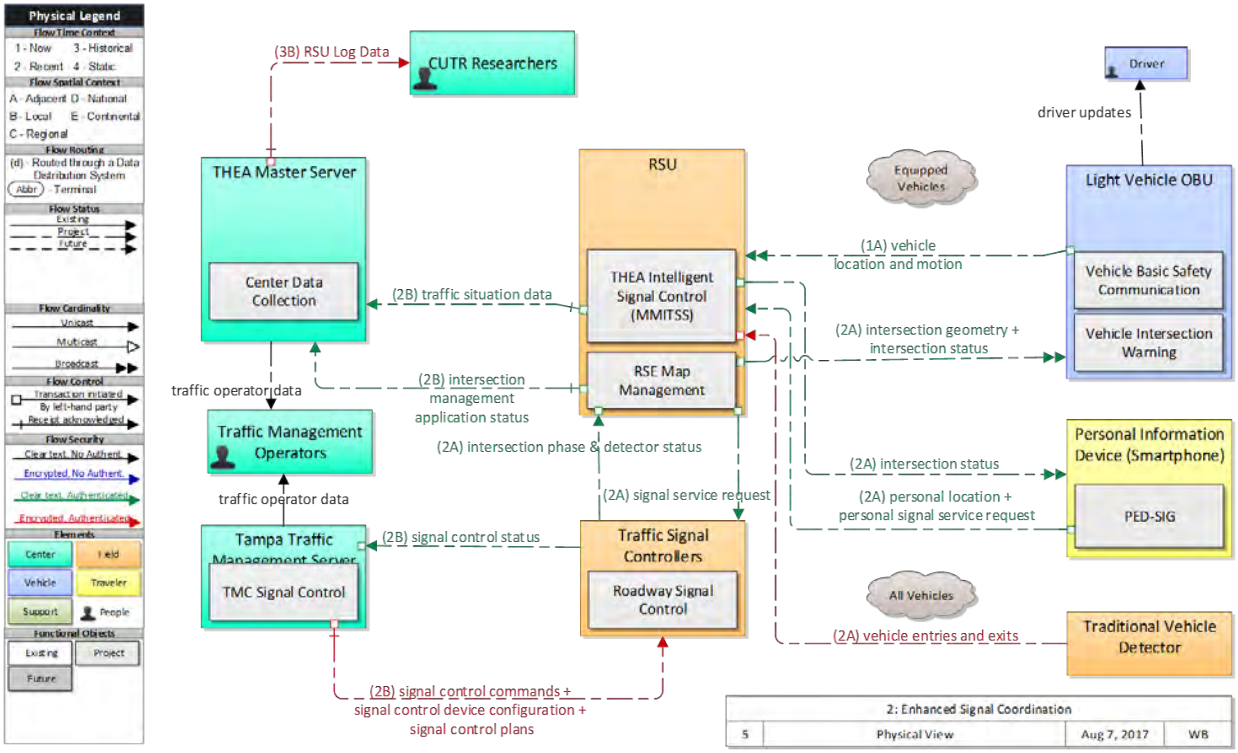
Source: THEA, 2020

Figure 6-23. REL Operational Directions

6.2.5 I-SIG

6.2.5.1 Functional Architecture

According to the original THEA CV Pilot functional architecture, the Intelligent Traffic Signal System (I-SIG) implementation for Use Case 6 was designed to feed information to the MMITSS application to estimate queue length, along with other traffic delay measures, to improve traffic progression in the relevant roadway sections. Figure 6-24 illustrates the functional architecture for Use Case 6.



Source: System Architecture Document, Publication FHWA-JPO-17-459, 2018

Figure 6-24. Use Case 6 Traffic Progression Physical Architecture

6.2.5.2 Deployment of I-SIG

During Phase 3, performance tests revealed that the MMITSS application was not successful in correctly estimating queue length [7]. In addition, I-SIG deployment at the signalized intersection did not occur due to integration issues between MMITSS, I-SIG, and the signal controllers. A series of tests were conducted during the first half of 2020 to explore what would allow the signal controllers to communicate with the I-SIG without relying on key input measures from MMITSS. Researchers used test vehicles (i.e., Friends of the Pilot) driving at selected intersections. The data generated from these tests were uploaded to the Secure Data Commons (SDC). No additional data were available from the participants, thus preventing the before-after assessment originally planned for this use case.

6.3 Mobility Analysis Methodology

The mobility impact assessment relied on a comparison of time trends in the mobility performance measures before and after the End of Ramp Deceleration Warning's deployment.[2] The goal was to assess if the treatment (i.e., the implementation of the ERDW) caused a change in the following mobility performance measures upon baseline conditions:

- Travel time (average travel times at five-minute intervals)
- Travel time reliability (the 95th percentile travel times)
- Queue length (maximum queue length in meters).

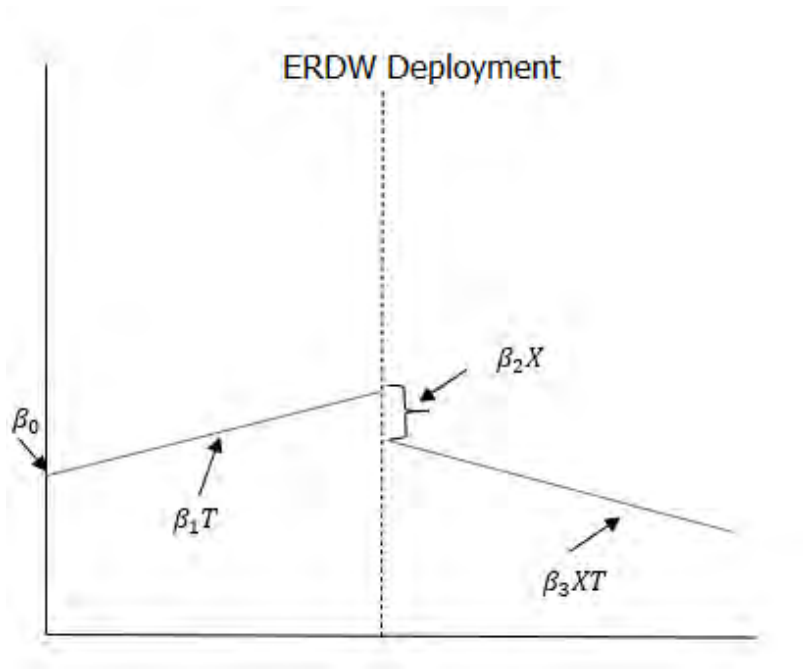
To test the impact of ERDW deployment, the analysis adopted an interrupted time-series approach where the outcome variables (i.e., the mobility measures) were observed before and after application deployment [11, 12]. The goal was to test the hypothesis that the introduction of the ERDW affected the level and trend of the outcome variables of interest. Empirically, the assessment was carried out via regression modeling using the following equation:

$$Y_t = \beta_0 + \beta_1 T_t + \beta_2 X_t + \beta_3 X_t T_t + \beta_4 Z_t + \epsilon_t \quad (1)$$

where Y_t is the outcome variable (e.g., mean travel time) observed daily over the analysis period, X_t is a dummy indicator representing the ERDW intervention (pre-intervention period is 0, otherwise 1), T_t is time measuring days over the analysis period, $X_t T_t$ is an interaction term, and Z_t is a vector of controls for confounding factors (e.g., weather).

Figure 6-25 provides a visual depiction of the model and its interpretation. The parameter β_0 represents the predicted baseline value of the impact measure (mean travel time) at the beginning of the analysis period. The parameter β_1 is the slope or trend of the impact measure before ERDW deployment. The parameter β_2 represents the change in the impact measure immediately following ERDW deployment and β_3 is the difference between the pre-intervention and post-intervention slopes of the outcome (mean travel time). The parameter β_4 evaluates the impact of confounding factors. The goal was to estimate a statistically significant parameter β_2 to indicate an immediate impact or in β_3 to indicate an effect over time, in this case a reduction upon the initial baseline parameters (lower travel times or reduced queue length).

[2] The same methodology would apply to evaluate other V2I mobility-focused applications, such as TSP and I-SIG once they are deployed and generate data for the *after* period.



Source: CUTR, 2020

Figure 6-25. Visual Depiction of Interrupted Time-Series ERDW Assessment

6.4 Safety Analysis Methodology

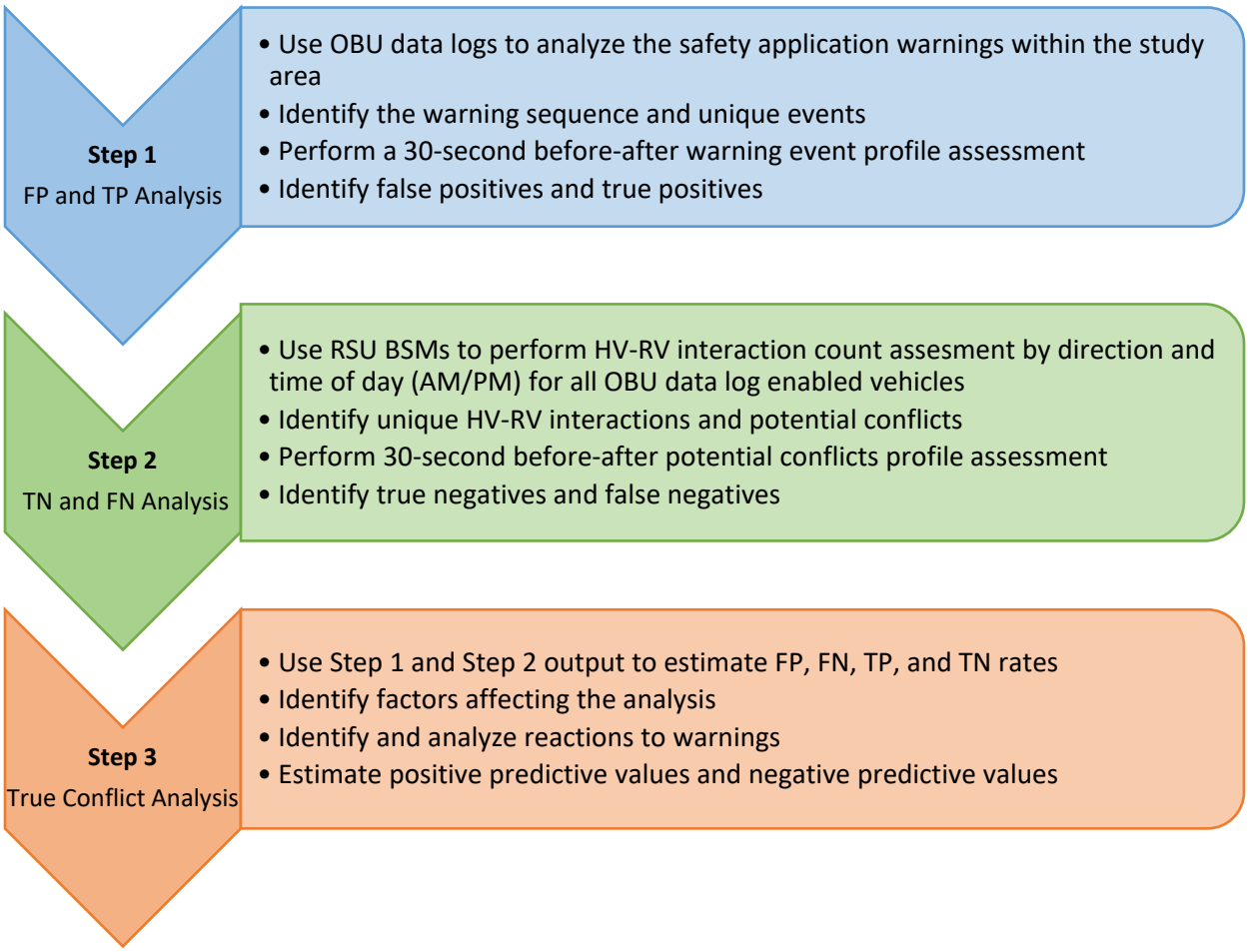
The goal of the safety analysis was to verify that warnings were issued when there was a conflict. Several confounding factors might lead V2V applications to either deploy in the absence of a conflict or fail to deploy when a conflict warrants it. The identification of these instances is paramount to assessing the applications' effectiveness. This analysis adopted the following terms:

1. **True Positive (TP) –** An instance of a warning issued when there is a conflict.
2. **True Negative (TN) –** An instance of no warning issued when there is not a conflict (i.e., normal conditions)
3. **False Positive (FP) –** An instance of a warning issued when there is not a conflict.
4. **False Negative (FN) –** An instance of a warning not issued when there is a conflict.

In this analysis, the term *instance* represents an interaction between a host vehicle (HV) and a remote vehicle (RV) while traveling within the study area.[3] Study area here refers to the geographic boundaries defining each use case. The term *interaction* defines a condition where the HV and RV interact and the V2V applications become operational and begin evaluating the requirements to issue a warning. The term *conflict* defines a

[3] According to SAE J2735, a host vehicle (HV) "is the equipped vehicle about which a given use case can be constructed." The remote vehicle (RV) is an "equipped vehicle (or vehicles) which play supporting roles in the use case by interacting with the HV in some way." [9]

condition under which all the application parameters are met to issue a warning. Figure 6-26 outlines the steps for conducting the safety assessment.



Source: CUTR, 2020

Figure 6-26. Steps in Methodology for Safety Analysis

6.4.1 Step One – False Positive and True Positive Analysis

Step one consists of a detailed analysis of the onboard unit data logs to count the total number of warning events collected over the analysis time frame. The OBU data logs contain all the information leading to the identification and analysis of warnings. Whenever the OBU deploys an application, the OBU data log timestamps and records the event via an anchor Basic Safety Message, or HV BSM. Automatically, a 30-second profile before-after the HV BSM timestamp (for a total of 60 seconds) is recorded at a frequency of 10 Hz. The analysis includes only data from vehicles capable of issuing the warnings, recording the events in their OBU data logs, and transmitting data OTA.

For each recorded warning, warning moment–specific parameters such as HV speed, HV-RV distance, elevation, lane width, and time to collision (TTC) are calculated and compared with reference values used by OBU vendors to set up the applications (operational and configuration parameters listed in the application

details). This automated process labels warnings as FP with parameter values that fail to meet the application reference values. Warnings meeting the application reference value thresholds are visually inspected by dynamically mapping their BSM profiles to classify them as FP or TP.

6.4.2 Step Two – False Negative and True Negative Analysis

While FP and TP events can be identified using the method described in step one, measurement of FP and FN rates involves the quantification of all TN and FN instances. This requires analyzing all movements made by each participant vehicle over the analysis time frame and identifying all interactions with nearby vehicles in situations where the V2V application is expected to be both operational and ready to deploy the warning as required. The process to estimate FN and TN events consists of the following steps:

1. For the analysis time frame, get daily RSU BSMs, apply the UC geofencing polygon, and filter to retain the required time period BSMs. Using RSU BSMs compensates for the likelihood of getting incomplete BSM profiles by solely relying on sent BSMs recorded in OBU data logs.
2. Remove duplicate BSMs using vehicle ID, timestamp, and BSM message count sequence.
3. Reduce the frequency from 10 Hz to 1 Hz by keeping the first BSM generated in each second.
4. Apply hour, minute filters, and determine the number of unique vehicles that generated BSMs.
5. Given the number of vehicles (*n*), determine the number of all possible V2V interactions using the following equation:

$$no.of\ possible\ V2V\ interactions = C(n,2) = \frac{n!}{2!(n-2)!} \quad (2)$$

6. Apply the OBU application reference parameter values to determine the number of interactions. These parameters are the same ones used in the FP assessment of step one.
7. Generate a per-vehicle count table of time spent, number of interactions based on operational parameter values, and number of conflicts based on configuration parameter values.
8. Aggregate the results of step 7 to estimate the number of interactions and number of conflicts over the analysis time frame.
9. Compare conflicts with the TP and FP events identified using the OBU data logs warning analysis to estimate the number of FN events.
10. Estimate the number of TN as *TN = Total Interaction – TP – FP – FN.*

6.4.3 Step Three – True Conflict Analysis

In step three, the outputs from steps one and two are used to perform the conflict analysis to produce all relevant performance rates and proceed to analyze participant reactions to warnings as detailed in the next section. The following rates are used to determine the effectiveness of an application.

The false positive rate is computed as

$$FP\ rate = \frac{FP}{FP+TN} \times 100 \quad (3)$$

where *FP* is the number of false positives and *TN* is the number of true negatives. The sum of *(FP + TN)* is equal to the total number of negatives (*N*).

The false negative rate is computed as

$$FN\ rate = \frac{FN}{FN+TP} \times 100 \tag{4}$$

where *FN* is the number of false negatives. The sum of *(FN + TP)* is equal to the total number of positives (*P*).

The true positive rate is computed as

$$TP\ rate = \frac{TP}{TP+FN} \times 100 \tag{5}$$

where *TP* is the number of true positives. The sum of *(TP + FN)* is equal to the total number of positives *(P).*

Finally, the true negative rate is computed as

$$TN\ rate = \frac{TN}{TN+FP} \times 100 \tag{6}$$

where *TN* is the number of true negatives. The sum of *(TN + FP)* is equal to the total number of negatives *(N)*.

An important note is that these rates (FP, FN, TP, TN) are not the same as the ratio/percent/share of FP, FN, TP, TN warnings to all warnings received. For example, if an application triggered 100 warnings and out of those 30 are FP and 70 TP, the ratio/percent/share of FPs is 30/100, or 33 percent. However, the *FP rate* as defined above requires the calculation of TN, which are instances of no warning issued when there was not a conflict, such as normal driving situations where the application correctly did not trigger warnings. Then the FP rate can be calculated using the formula provided above. Similarly, for the *TP rate*, the count of missed alarms, or FN, is necessary.

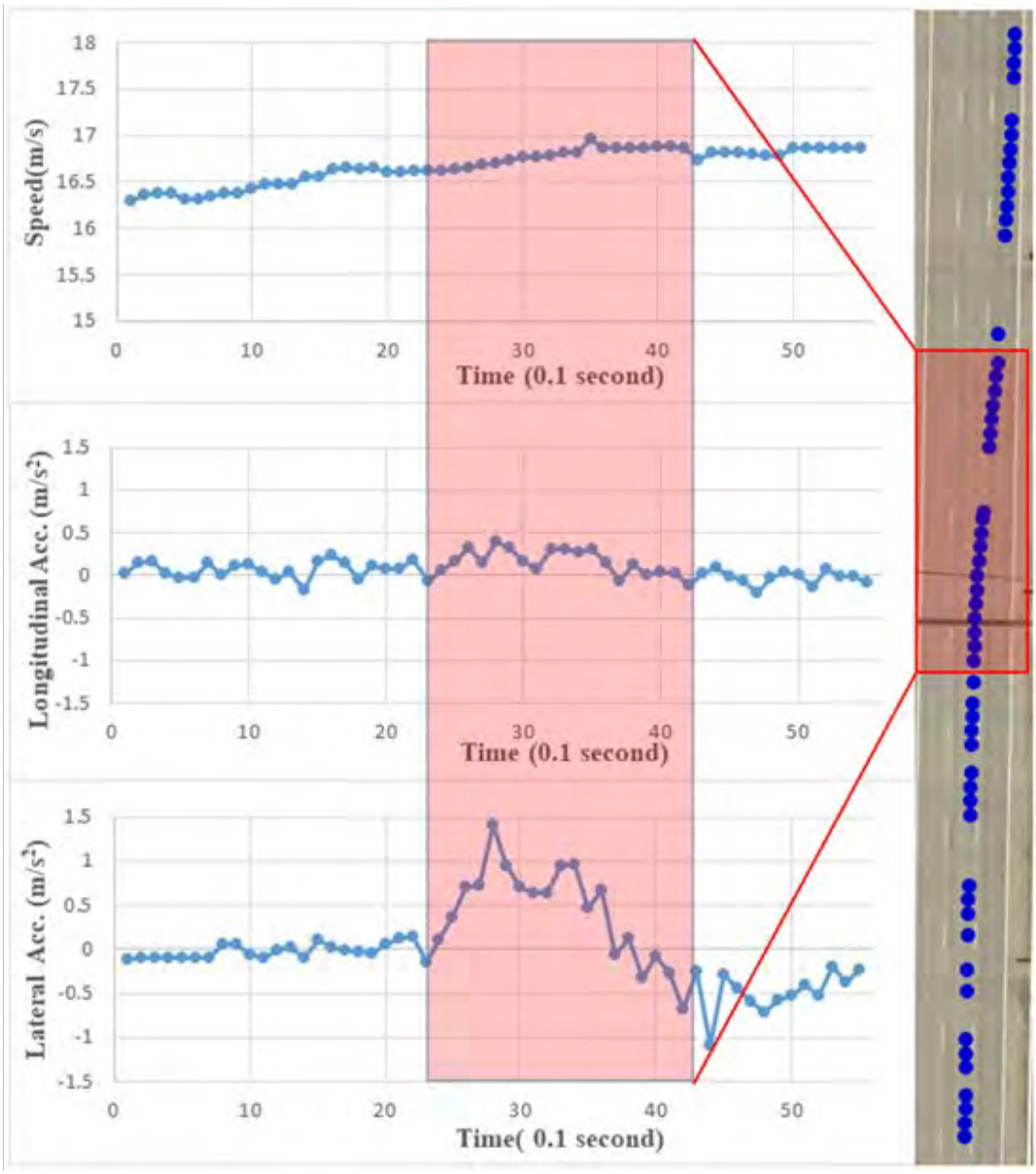
6.4.4 Participant Reaction to Warnings

The analysis of reactions to warnings is based on the monitoring and measurement of participant vehicles' instantaneous longitudinal and lateral movements, which are reflected in the vehicles' longitudinal and lateral acceleration values.

In the absence of in-vehicle video camera detection and recording, the research team modified and adapted a previously developed data-driven lane change detection algorithm to spot the location of road hazard debris using BSM Part I data [13]. The algorithm can detect longitudinal and lateral reactions in response to evasive

maneuvers. Before applying it to the participants' V2V warning assessment, researchers calibrated the algorithm using data generated by the THEA CV Pilot test vehicles.

Figure 6-27 presents a snapshot of a test vehicle trajectory and speed at 0.1-second intervals, showing longitudinal and lateral acceleration while performing a lane change maneuver. In this instance, the test vehicle was engaging in a maneuver to avoid a test cone used to represent road debris on the REL. The figure highlights the lane change moments in the red-shaded area to indicate a noticeable change in lateral acceleration and no considerable change in longitudinal acceleration. In fact, the speed graph indicates that the driver was increasing speed while changing lanes.

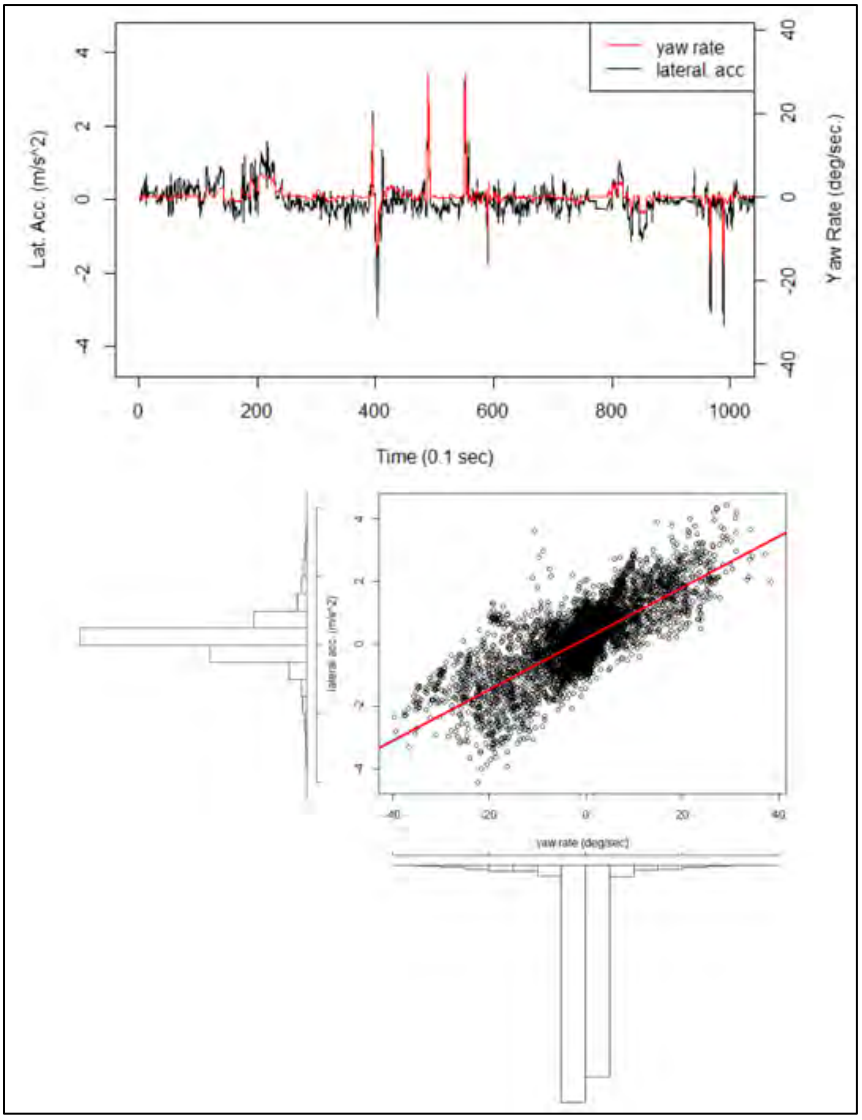


Source: CUTR, 2020

Figure 6-27. Vehicle Speed and Acceleration Profiles during a Lane Change

Drivers react to a collision threat by engaging their brakes, changing lanes, or a combination of both. Using OBU data logs, braking reactions can be identified through longitudinal acceleration of host vehicle sent BSMs around the anchor point of a warning event. Similarly, a lane change or swerving can be discerned from lateral acceleration values.

The research team implemented the algorithm on the entire OBU data logs warning event database and RSU BSM databases with the goal of identifying and assessing conflicts. The analysis revealed data gaps in the lateral acceleration values for vehicles equipped with one of the aftermarket OBUs. Therefore, yaw rate was substituted to identify lane change maneuvers. Using data collected from test vehicles performing evasive maneuvers, the team tuned the algorithm to use yaw rate instead of lateral acceleration to detect lane changes. Figure 6-28 shows that yaw rate is highly correlated with lateral acceleration and thus can be used to identify lateral movements of vehicles.

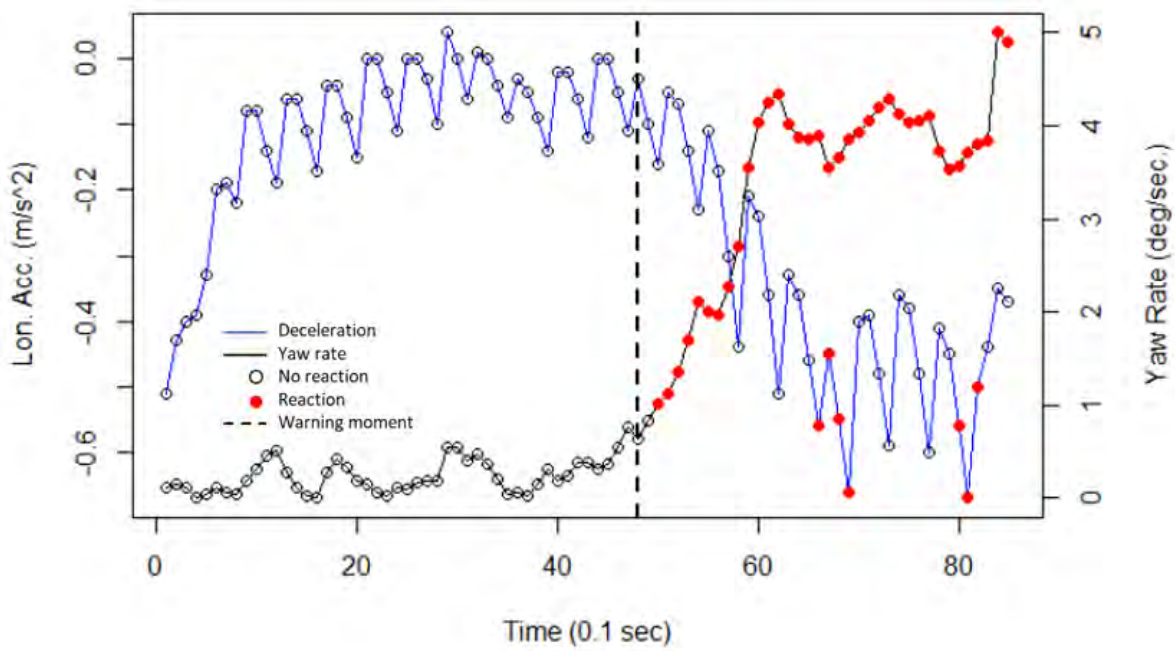


Source: CUTR, 2020

Figure 6-28. Lateral Acceleration and Yaw Rate from BSM Data

To analyze the participants' behavioral responses to warnings generated by V2V applications, the conflict detection algorithm was run by building 5- and 10-second profiles before-after the warning moment identified as true positive.

Figure 6-29 shows the algorithm post-processing output. In this example, the driver decelerated after the warning moment with deceleration values beyond -0.5 meters per second squared (mps^2), which was set as the threshold for this purpose. In addition, the driver exhibited no reaction before the moment of warning. Therefore, in this instance, it is reasonable to assume that the driver reacted to the warning and/or a traffic situation requiring the driver to slow down. The -0.5 mps^2 value was selected based on the test vehicle BSM profile data during the process of fine-tuning the algorithm.



Source: CUTR, 2020

Figure 6-29. Example of Expected Driver Reaction to Warning

Chapter 7. System Impact Evaluation Results

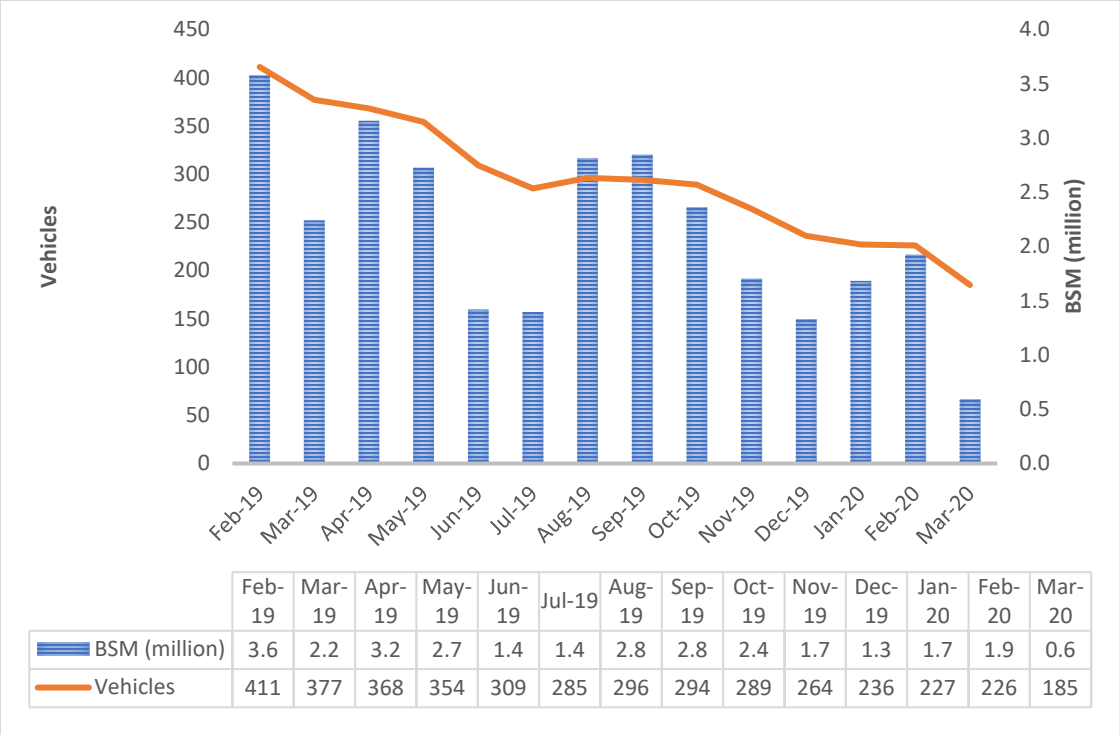
7.1 Use Case 1: Morning Backups

Use Case 1 analyzed the impact of V2V and V2I deployment on participants during their morning commute, between 6:00 and 10:59 a.m. on weekdays. The morning rush hour and backup occurs during this time frame when the REL operational direction is westbound from Brandon toward downtown Tampa. The analysis excludes national holidays because the REL operates eastbound on a 24-hour basis during those days.

7.1.1 Analysis Dataset

7.1.1.1 Mobility Dataset

The dataset consists of the measures above computed at the vehicle level using the first and last BSM as each vehicle traveled through the REL study area. The *before* period began February 4, 2019 (Monday), and ended January 31, 2020 (Friday). On March 20, 2020, THEA set the REL operational direction to eastbound on a 24-hour basis (leaving downtown Tampa toward Brandon) in response to the COVID-19 pandemic. The *after* period began February 3, 2020 (Monday) and ended March 20, 2020 (Friday). The dataset consists of 29.8 million RSU BSMs collected from 587 unique participant vehicles traveling through the UC1 study impact area. Figure 7-1 shows the monthly distribution of the data sample over the analysis period.



	Feb-19	Mar-19	Apr-19	May-19	Jun-19	Jul-19	Aug-19	Sep-19	Oct-19	Nov-19	Dec-19	Jan-20	Feb-20	Mar-20
BSM (million)	3.6	2.2	3.2	2.7	1.4	1.4	2.8	2.8	2.4	1.7	1.3	1.7	1.9	0.6
Vehicles	411	377	368	354	309	285	296	294	289	264	236	227	226	185

Source: CUTR, 2020

Figure 7-1. Mobility Evaluation Analysis Dataset

7.1.1.2 Safety Dataset

The safety analysis consists of OBU data log and RSU BSM data collected between August 1, 2019, and March 20, 2020. The applications were using different (looser) parameters prior to August 2019 and were therefore not included in the analysis as they issued warnings under different conditions.

7.1.2 Mobility Impact

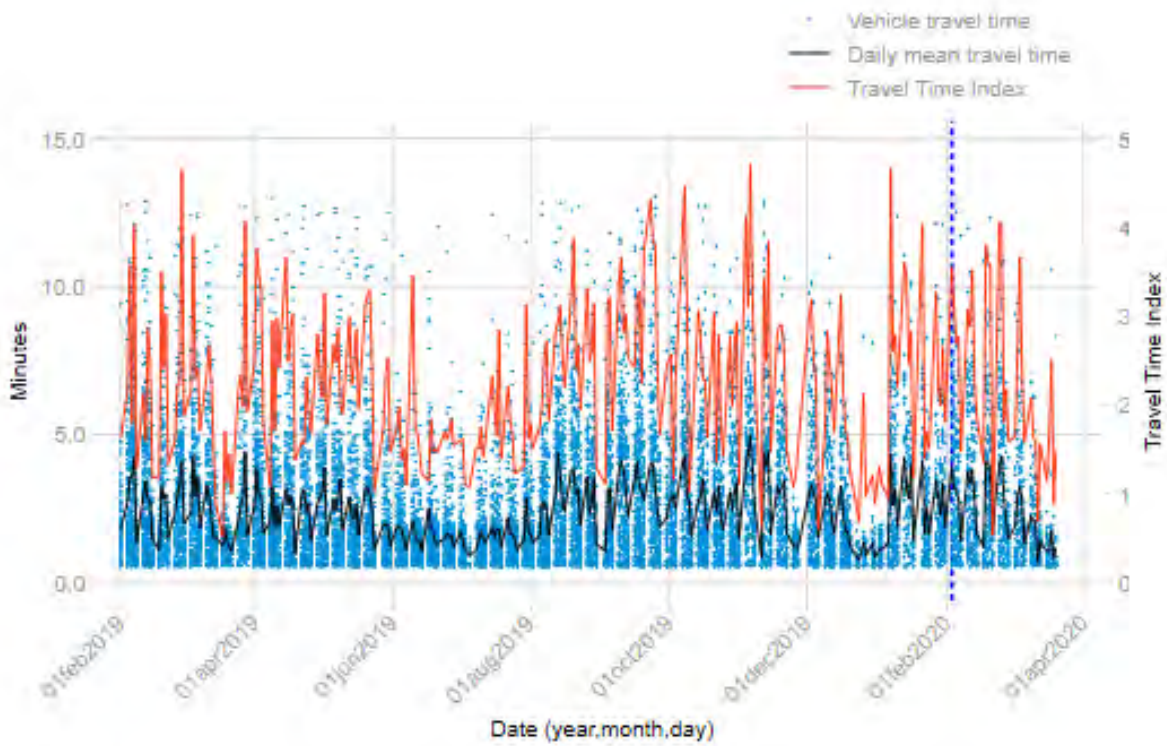
To determine the impact of CV application deployment on mobility, the analysis focused on the following measures and their comparison on a before and after basis:

- Travel time (average travel times at five-minute intervals)
- Travel time reliability (by comparing the 95th percentile travel times)
- Idle time (time spent traveling at a speed of less than one mile per hour)
- Queue length (maximum queue length in meters, hourly).

The approach to impact assessment relies on interrupted time-series analysis to compare the effect of the ERDW deployment on the above mobility measures. Travel time, idle time, and travel time reliability measures were obtained using the RSU BSMs from participant vehicles over the analysis period.

Figure 7-2 illustrates a scatterplot of participant vehicle travel times over the analysis period, with the black and red lines showing the daily mean travel time and travel time index. The travel time index (TTI) is measured as peak hour travel time divided by off-peak travel time. For reference, the figure shows a vertical dotted blue line

representing the beginning period of ERDW deployment (February 3, 2020). The recurring spikes in the graph show intra-weekday variability during the analysis period.



Source: CUTR, 2020

Figure 7-2. Travel Times Before-After ERDW Deployment

Table 7-1 reports the sample descriptive statistics of the peak hour travel time obtained from each vehicle traveling in the study area over the entire evaluation time frame.

Table 7-1. Sample Descriptive Statistics

Travel Time (Minutes) - Peak					
ERDW Deployment	**Observations**	**Mean**	**Max**	**St. Dev**	**95th Percentile**
Before	17,613	2.5	13.0	2.1	6.7
After	1,484	2.7	12.5	2.2	7.0
Overall Sample	**19,097**	**2.5**	**13.0**	**2.1**	**6.8**

Travel Time (Minutes) – Off-Peak					
ERDW Deployment	**Observations**	**Mean**	**Max**	**St. Dev**	**95th Percentile**
Before	6,752	1.0	3.7	0.5	2.0
After	637	1.1	3.7	0.6	2.1
Overall Sample	**7,389**	**1.0**	**3.7**	**0.6**	**2.0**

Travel Time Index					
ERDW Deployment	**Observations**	**Mean**	**Max**	**St. Dev**	**95th Percentile**
Before	17,600	2.4	17.8	2.0	6.3
After	1,484	2.5	14.9	2.0	6.5
Overall Sample	**19,084**	**2.4**	**17.8**	**2.0**	**6.3**

Idle Time (Minutes)					
ERDW Deployment	**Observations**	**Mean**	**Max**	**St. Dev**	**95th Percentile**
Before	18,457	1.1	3.5	0.5	2.0
After	1,578	1.2	2.6	0.5	2.1
Overall Sample	**20,035**	**1.1**	**3.5**	**0.5**	**3.2**

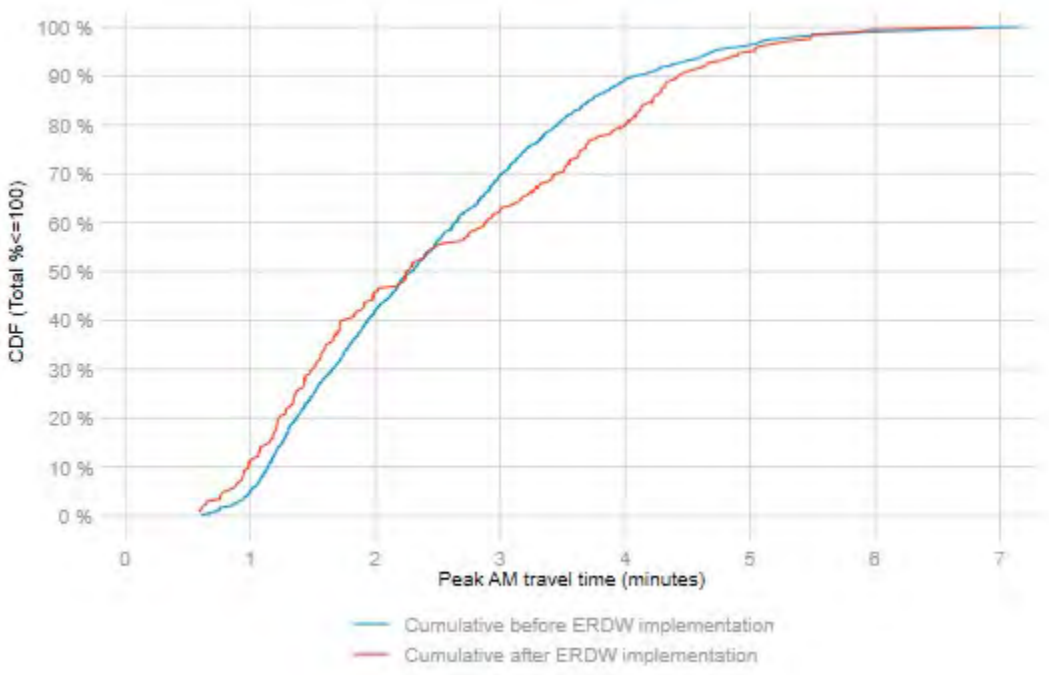
7.1.2.1 Interrupted Times Series Analysis Results

The final dataset used in the regression models consists of daily mean travel times and travel time index data culled at 10-minute intervals during peak morning travel. The 10-minute interval binning allows merging the travel time dataset with the weather data (only available at 10-minute intervals) to check the impact of adverse weather events not captured by the time trends. This interval binning captures variability in travel time and travel time index data during the evaluation time frame conducive to inference modeling.

7.1.2.1.1 Impact on Average Travel Time

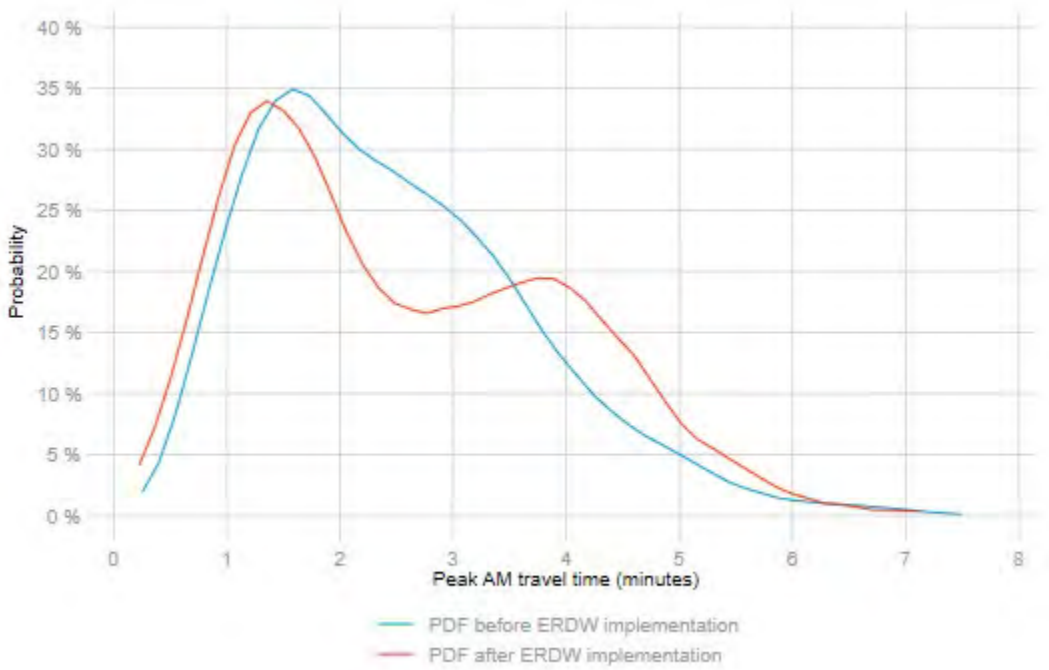
Figure 7-3 plots and compares the cumulative distribution functions (CDFs) of travel times for before and after ERDW application deployment. The CDF measures the probability a vehicle making a trip lasts a given number of minutes, as measured by the horizontal axis of the figure. The CDF shows differences between the two distributions for trips lasting up to 2.5 minutes and for trips taking longer than 2.5 minutes.

Figure 7-4 shows the probability density functions (PDFs) using kernel density smoothing and highlights differences between the before and after ERDW deployment periods. This is further confirmed by conducting a metric based test of equality of density based on the metric entropy of Maasoumi and Racine, which rejects the null of equality at the 0.01 percent [14].



Source: CUTR, 2020

Figure 7-3. Before-After ERDW Deployment Cumulative Travel Time Distribution Functions



Source: CUTR, 2020

Figure 7-4. Before-After ERDW Deployment Travel Time Probability Density Functions

Table 7-2 reports the regression analysis results, with three model specifications that progressively account for the inclusion of controls for weather (Model 2) and the onset of the COVID-19 pandemic (Model 3). The impact of weather is accounted for by using 10-minute interval measurements on the presence and intensity of rain.[4] The impact of the onset of COVID-19 is measured by a dummy indicator set equal to one beginning March 13, 2020. This is because travel demand on the REL began declining on this day (Figure 7-2), while the REL operational profile was changed to accommodate eastbound travel only on March 20, 2020.

Table 7-2. ERDW Deployment Impact on Travel Time – Estimation Results

Variable	Regression Models		
	Model 1	Model 2	Model 3
Time Trend (*T*)	0.00134**	0.00115**	0.00114**
	(3.04)	(2.53)	(2.51)
Treatment (*X*)	1.062***	1.074***	0.935***
	(6.14)	(6.21)	(5.09)
Post-intervention (*XT*)	-0.0677***	-0.0671***	-0.0529***
	(-9.43)	(-9.28)	(-5.48)
Tuesday†	0.224**	0.216**	0.213**
	(2.25)	(2.16)	(2.14)
Wednesday	0.0745	0.0803	0.0771
	(0.80)	(0.86)	(0.83)
Thursday	0.0561	0.0564	0.0532
	(0.60)	(0.60)	(0.57)
Friday	-0.974***	-0.984***	-0.982***
	(-12.07)	(-12.11)	(-12.09)
Rain		0.737**	0.782**
		(2.04)	(2.16)
Covid-19 Effect			-0.566**
			(-2.53)
Constant Term	2.421***	1.832***	1.797***
	(28.97)	(6.07)	(5.93)
Sample Size	1,602	1,602	1,602
Adjusted R-square	0.173	0.175	0.177

*Standard errors in parenthesis: * p<0.05, **p<0.01, ***p<0.001*

†Weekday baseline is Monday

[4] As detailed in the PMESP, weather data came from World Weather Online, which provides national weather broadcast services and downloadable data via its DarkSky automated protocol interface. The dataset includes several weather measurements. One alternative model specification considered a combined index derived by conducting a principal component analysis on all weather measures to produce a weather index. The inclusion of this variable did not provide additional explanatory power to the model over the use of the dummy indicator for rain.

The preferred model is Model 3. The results show that at the beginning of the analysis period, the estimated average travel time (10-minute frequency) is about 1.8 minutes from the REL posted speed sign of 40 mph to Twiggs Street and Meridian Avenue (constant term 1.797). As indicated by the statistically significant parameter associated with the time trend variable (*T*), each day the average travel time increases by a fraction of minute (0.00114).

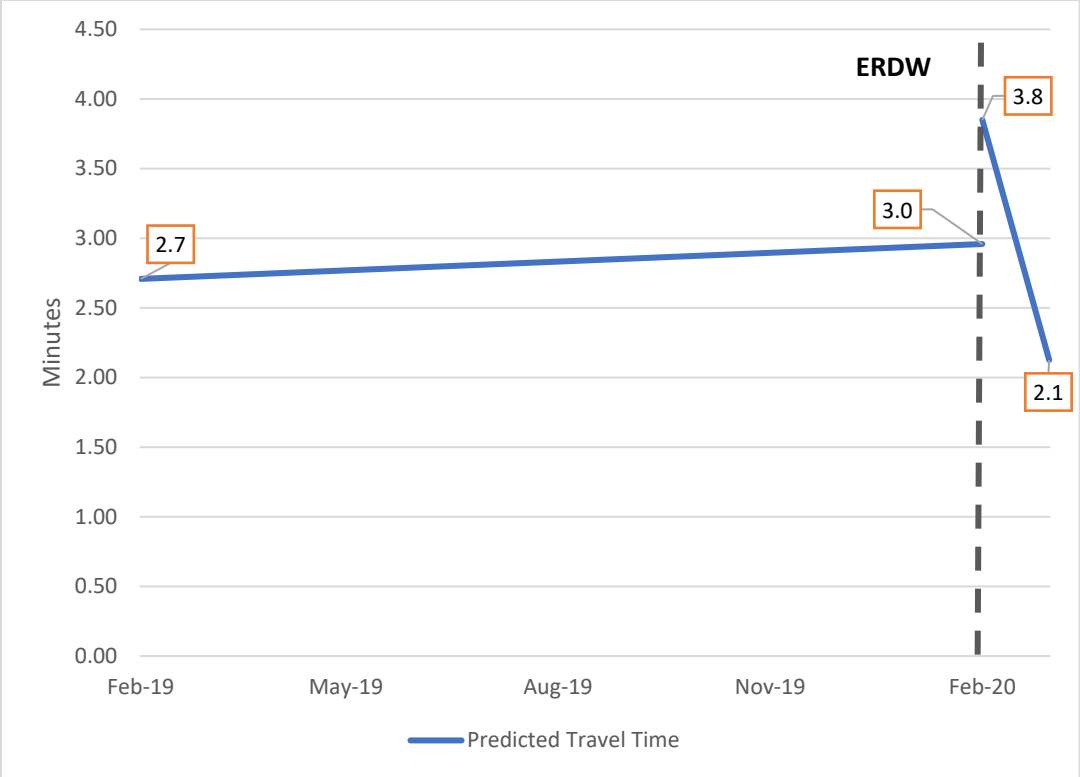
The relevant parameters associated with deployment of the ERDW (treatment and post-intervention) are all statistically significant. At the beginning of the deployment of the ERDW application on February 3, 2020 (treatment variable), the average travel time shows an upward increase as indicated by the parameter associated with the treatment variable (0.93) of about a minute with respect to the baseline. On the other hand, the days following the deployment of the ERDW are associated with declining travel times, as indicated by the negative value associated with the post-intervention variable (*XT*). The controls for weekday travel (baseline Monday), weather, and the COVID 19 pandemic are all statistically significant. The presence of rain increases travel time by about almost a minute, while the onset of the pandemic is associated with a decrease in travel time as indicated by the negative sign of the estimated parameter. This is substantiated by the observed reduction in the number of participants vehicles beginning on March 13, 2020, as highlighted by Figure 7-1.

To relate the impact of the ERDW application on average travel times, Figure 7-5 reports the adjusted predictions of the ERDW deployment's effect on travel times at specific values of the sample. The marginal effects are reported assuming no adverse weather conditions (i.e., no rain), limiting the impact before the onset of the pandemic (i.e., COVID-19 = 0), and with Tuesday as the representative weekday.

At the beginning of the analysis period of February 2019, the estimated mean travel time (10-minute frequency) through the UC1 study impact area is about 2.7 minutes. After the deployment of the ERDW application in March 2020, the average travel time is about 2.1 minutes.

The estimation of Model 3 using a natural log transformation of the dependent variable (not reported here) reveals that the ERDW contributes to a 2.1 percent reduction in mean travel times with respect to the baseline (pre-intervention period).[5]

[5] The natural log transformation of the dependent variable produces an estimate of the post-intervention parameter of -0.021. Applying the proportional change transformation [(exp(-0.021) -1] x 100, the estimated proportional change is -2.07 percent.



Source: CUTR, 2020

Figure 7-5. Adjusted Predictions of ERDW Impact on Travel Times

7.1.2.1.2 Impact on Travel Time Reliability

The approach used to estimate the impact on average travel time is applied to estimate the impact on travel time reliability. Reliability is measured using the travel time index (TTI), which is computed as the peak morning travel time divided by the free flow off-peak travel time. For example, a TTI value of 1.5 means that a commuter traveling through the use case would need to budget 1.5 times the amount of time needed during off-peak travel. Because of how the TTI is derived, the CDF and PDF do not change compared to those reported for the peak travel time (Figure 7-3 and Figure 7-4) and are thus not reported here.

The interrupted time series regression models follow the same specification applied to estimating the impact on average travel times. The results show that at the beginning of the analysis period, the estimated travel time index is about 1.8, meaning it takes on average about twice as much time with respect to off-peak conditions to travel through UC1. As indicated by the statistically significant parameter associated with the time trend variable (*T*), each day the 95th percentile travel time reliability shows a small but steady decrease (-0.0493). The relevant parameters associated with the deployment of the ERDW (treatment and post-intervention) are all statistically significant. At the beginning of the deployment of the ERDW application on February 3, 2020 (treatment variable), travel time reliability shows an upward increase as indicated by the parameter associated with the treatment variable. This could be due to localized traffic conditions specific to that day (first Monday of the month), which adds more instability to travel times.

Table 7-3 reports the results, with Model 3 as the preferred model. While the impact of weather (i.e., rain) is statistically significant, the early onset of the pandemic is not statistically significant.

Table 7-3. ERDW Deployment Impact on Travel Time Reliability – Estimation Results

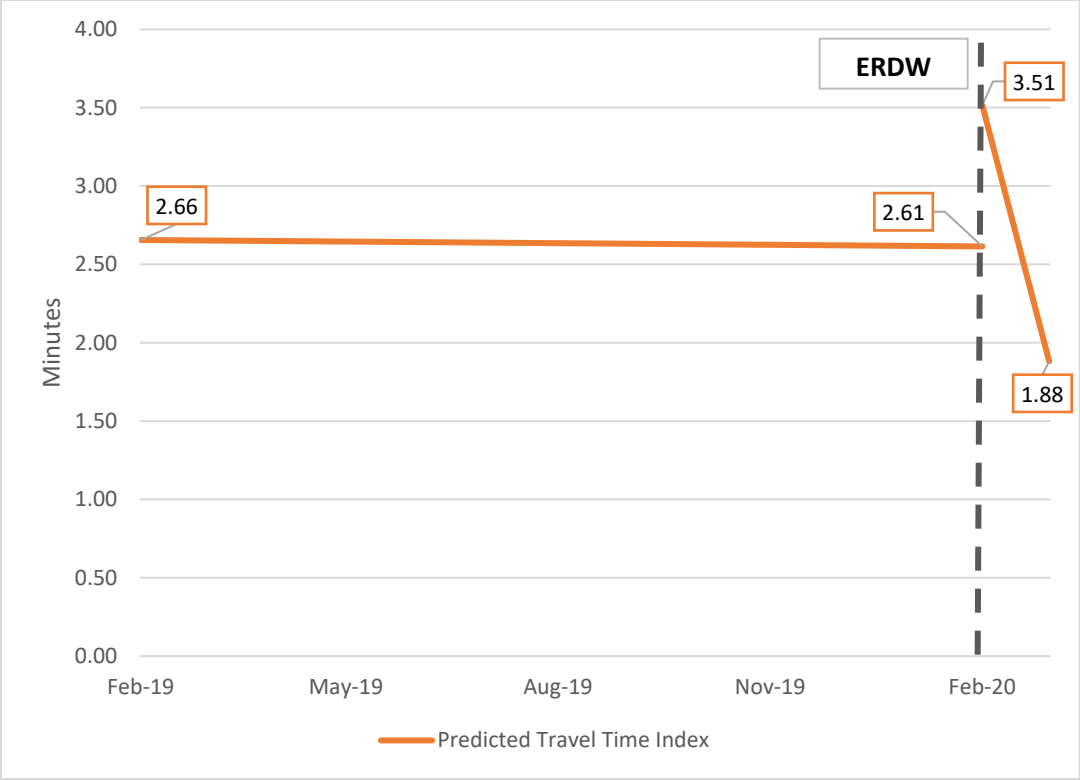
Variable	Model 1	Model 2	Model 3
Time Trend (*T*)	0.0000325	-0.000161	-0.000162
	(0.08)	(-0.39)	(-0.39)
Treatment (*X*)	0.953***	0.969***	0.949***
	(5.87)	(5.98)	(5.58)
Post-intervention (*XT*)	-0.0520***	-0.0513***	-0.0493***
	(-7.22)	(-7.11)	(-5.46)
Tuesday†	0.135	0.123	0.123
	(1.49)	(1.35)	(1.34)
Wednesday	-0.00640	-0.00339	-0.00391
	(-0.07)	(-0.04)	(-0.04)
Thursday	-0.0386	-0.0465	-0.0470
	(-0.46)	(-0.55)	(-0.55)
Friday	-0.860***	-0.870***	-0.870***
	(-11.10)	(-11.15)	(-11.16)
Rain		0.841**	0.846**
		(2.72)	(2.73)
Covid-19 Effect			-0.0800
			(-0.34)
Constant Term	2.501***	1.825***	1.821***
	(32.01)	(7.11)	(7.09)
Sample Size	1602	1602	1602
Adjusted R-square	0.138	0.141	0.141

*Standard errors in parenthesis: * p<0.05, **p<0.01, ***p<0.001*

†*Weekday baseline is Monday*

On the other hand, the days following ERDW deployment are associated with declining travel times, as indicated by the negative value associated with the post-intervention variable (*XT*).

Figure 7-6 reports the adjusted predictions of the effect of the ERDW deployment on travel times at specific values of the sample. The marginal effects are reported assuming no adverse weather conditions (i.e., no rain), limiting the impact before the onset of the pandemic, and Tuesday as the representative weekday. At the beginning of the analysis period (February 2019), the estimated TTI is about 2.7. After the deployment of the ERDW application in March 2020, the estimated TTI is about 1.9.

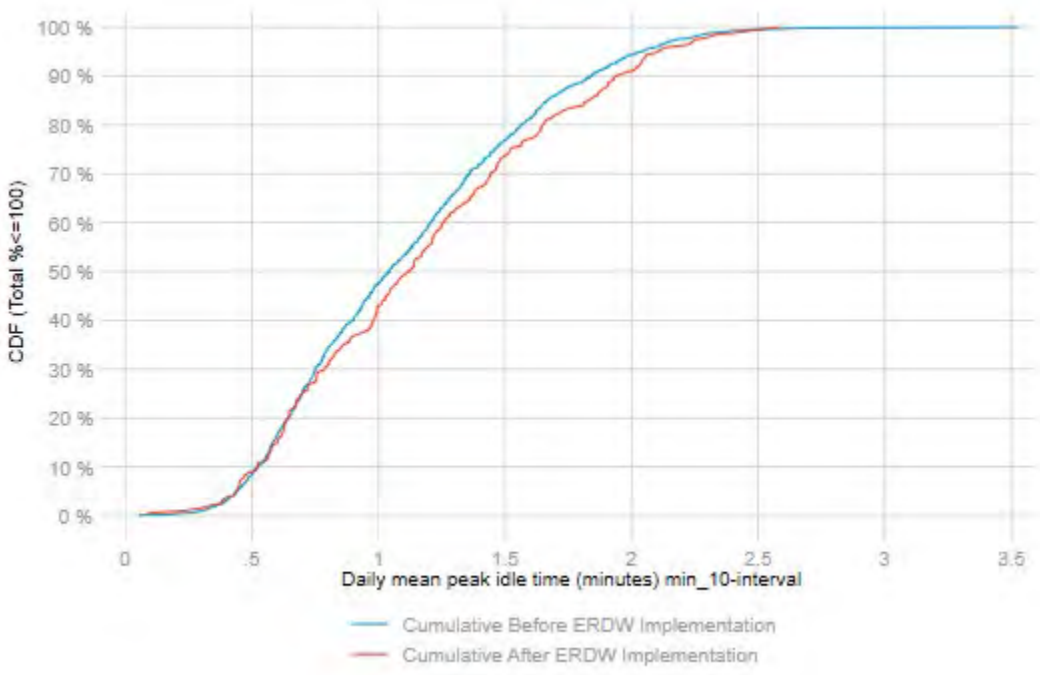


Source: CUTR, 2020

Figure 7-6. Adjusted Predictions of ERDW Impact on Travel Time Index

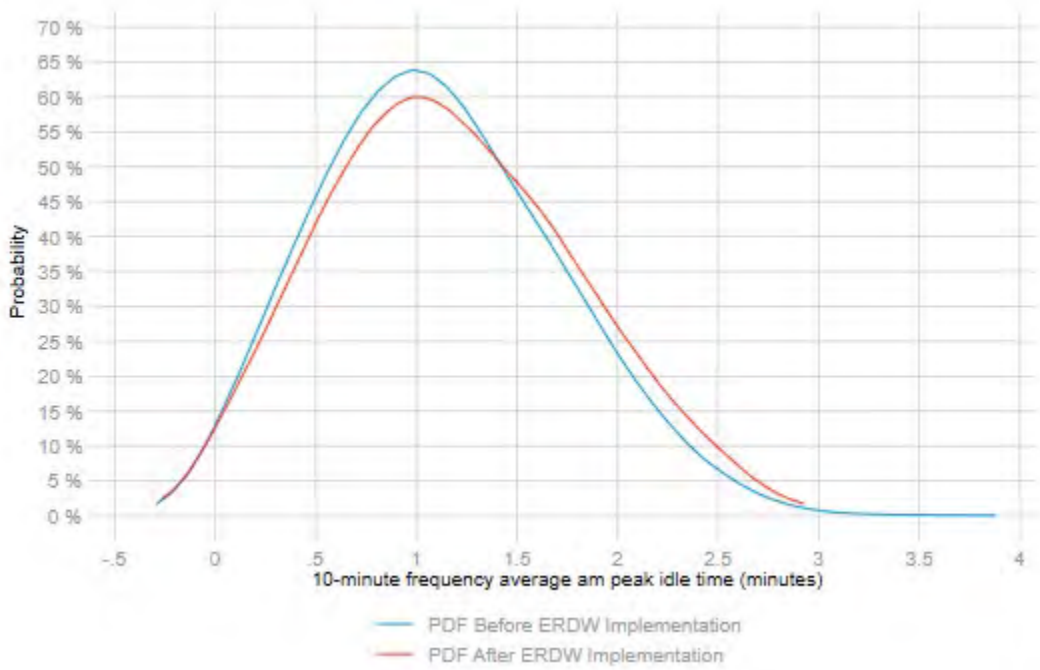
7.1.2.1.3 Impact on Idle Time

The approach applied earlier is extended to estimate the impact of the ERDW on time spent on idle by the participants. Idle time is measured as the amount of time spent in UC1 traveling at a speed of less than one mile per hour. Figure 7-7 and Figure 7-8 show the CDF and PDF of time spent in idle and compare the before and after periods.



Source: CUTR, 2020

Figure 7-7. Before-After ERDW Deployment Idle Time Cumulative Distribution Functions



Source: CUTR, 2020

Figure 7-8. Before-After ERDW Deployment Idle Time Probability Density Functions

Table 7-4 reports the results of the regression analysis with Model 3 as the preferred model. The estimated parameters are statistically significant. When looking at intra-day variability, Friday has a significant impact in reducing time spent on idle compared to the baseline day (i.e., Monday). The adverse impact of rain is reflected by the positive and statistically significant parameter (0.384).

Table 7-4. ERDW Deployment Impact on Idle Time – Estimation Results

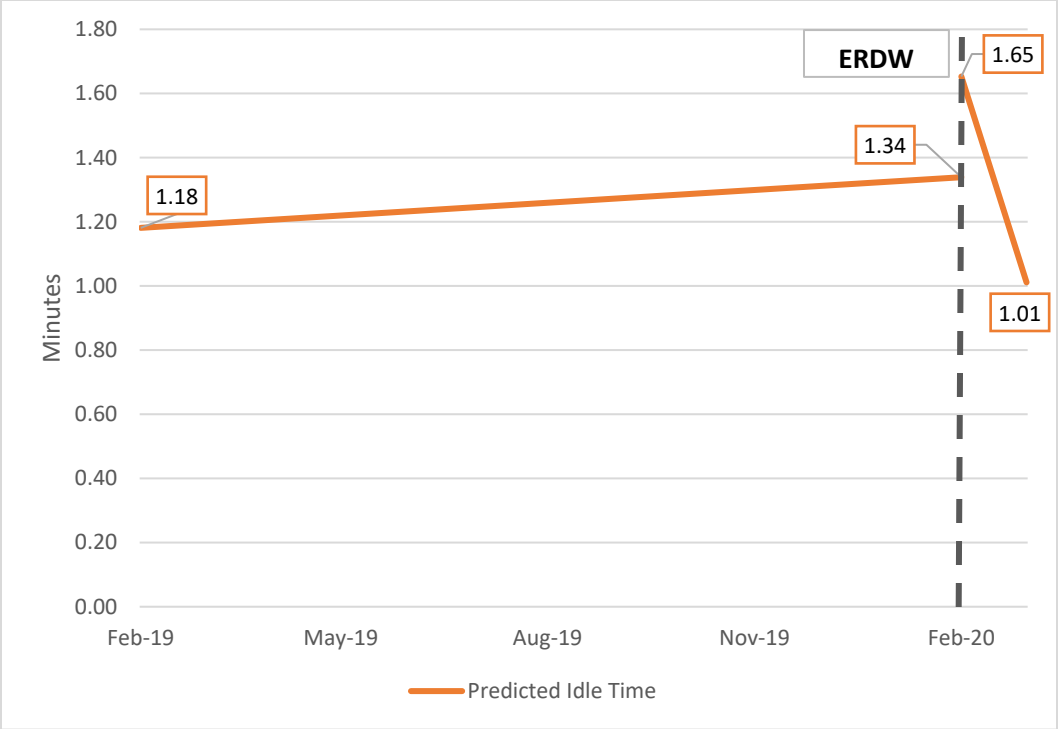
Variable	Regression Models		
	Model 1	Model 2	Model 3
Time Trend (*T*)	0.000710***	0.000628***	0.000624***
	(3.92)	(3.45)	(3.43)
Treatment (*X*)	0.373***	0.379***	0.334***
	(5.47)	(5.61)	(4.78)
Post-intervention (*XT*)	-0.0249***	-0.0247***	-0.0201***
	(-8.87)	(-8.84)	(-5.76)
Tuesday†	0.0607	0.0562	0.0552
	(1.45)	(1.34)	(1.32)
Wednesday	0.0181	0.0224	0.0213
	(0.45)	(0.55)	(0.53)
Thursday	-0.0186	-0.0211	-0.0220
	(-0.46)	(-0.52)	(-0.54)
Friday	-0.468***	-0.471***	-0.471***
	(-13.41)	(-13.46)	(-13.46)
Rain		0.366**	0.384**
		(2.28)	(2.38)
Covid-19 Effect			-0.182**
			(-2.08)
Constant Term	1.112***	0.814***	0.800***
	(30.33)	(5.84)	(5.71)
Sample Size	1573	1573	1573
Adjusted R-square	0.182	0.185	0.185

*Standard errors in parenthesis: * p<0.05, **p<0.01, ***p<0.001*

†*Weekday baseline is Monday*

Figure 7-9 reports the adjusted predictions of the effect of the ERDW deployment on travel times at specific values of the sample. The marginal effects are reported assuming no adverse weather conditions (i.e., no rain), limiting the impact before the onset of the pandemic, and Tuesday as the representative weekday.

At the beginning of the analysis period in February 2019, the estimated average idle time is about 1.2 minutes. After the deployment of the ERDW application in March 2020, the estimated idle time is about one minute.



Source: CUTR, 2020

Figure 7-9. Adjusted Predictions of ERDW Impact on Idle Time

7.1.2.2 Impact on Queue Length

One of the goals of deploying CV applications during UC1 was to address weekday morning peak hour delays resulting from queueing problems at the end of the REL ramp. In this context, the deployment of the ERDW was conceived as a CV-based solution by inducing speed harmonization on travelers to improve flow and throughput.

The original ERDW design called for the joint deployment of the I-SIG and MMITSS applications to estimate queue length and to produce additional performance metrics for evaluation (delay at intersection, percent of arrival on green, and throughput). As detailed in the section describing the ERDW application, the use of MMITSS presented queue estimation constraints that led to a redesign and improvement of the ERDW, with queue length estimated by using participant vehicles as probe vehicles. In the absence of MMITSS-produced output performance measures, the analysis in this report develops a queue length measurement method suitable for performance evaluation. The approach follows and expands the steps described in the ERDW technical documentation to reproduce the BSM-based methodology for queue length measurement.

This section discusses the development of queue length measurement and the result of the impact assessment of the ERDW application in reducing queueing formation.

7.1.2.2.1 Queue Length Estimation Methodology

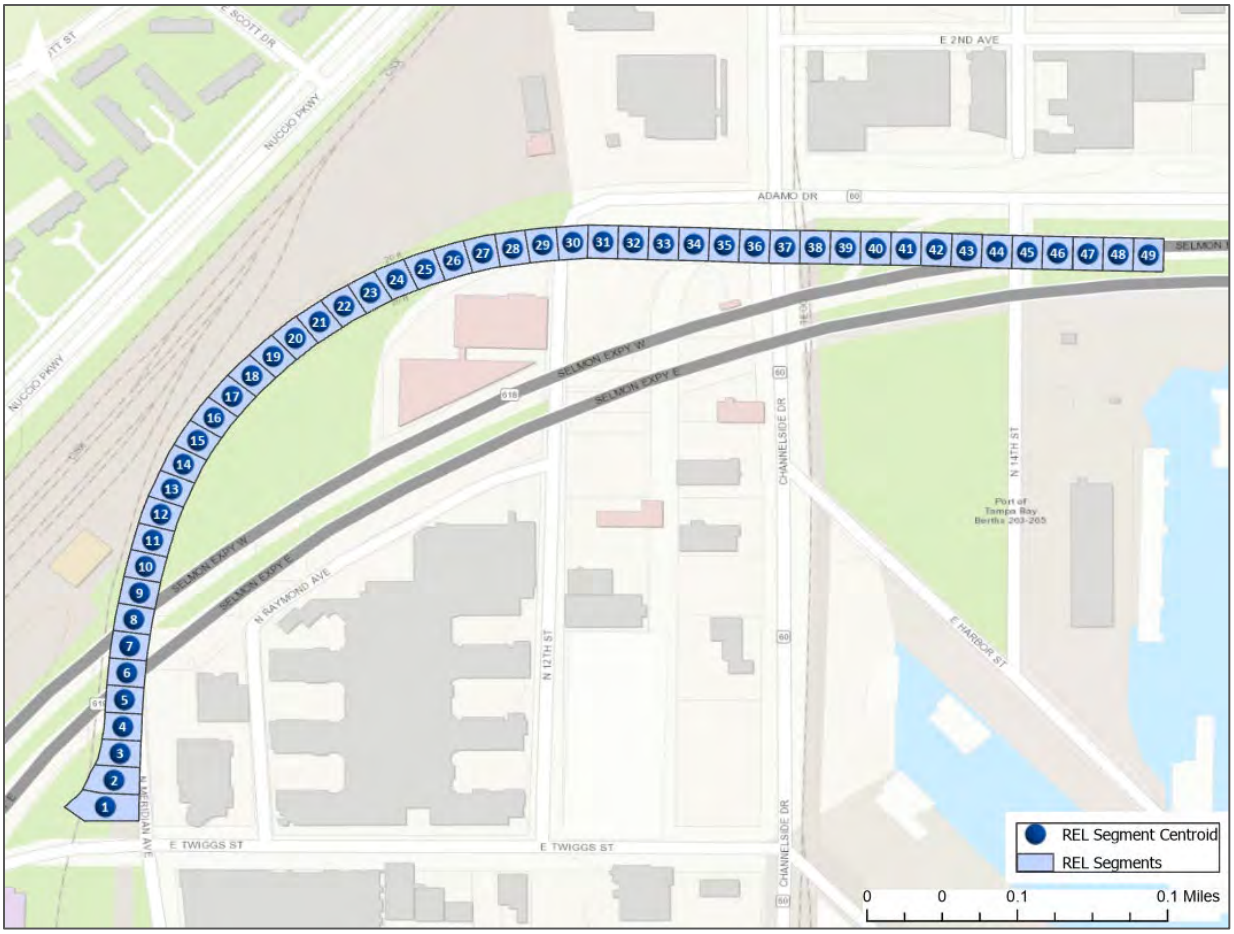
The adopted resolution to ERDW operation relies on queue length measurements obtained using participants' BSMs collected by roadside units as the vehicles travel to the REL. To measure queue length, the REL is divided into a sequence of equally spaced points, each having their geocoded coordinates indexed to form a

sequence of nodes extending from the Twiggs Street intersection to the REL ramp. Then a dynamic process is applied to measure changes in vehicle speeds to detect queuing and to use this information to feed the appropriate speed advisories to the vehicles via the ERDW application.

Since the newly adopted ERDW method entered operation in February 2020, the research team replicated the process to establish a historical baseline of queue length measurements and to backward estimate queue length to cover the entire analysis period (February 4, 2019, to March 20, 2020). As with the ERDW queue estimation approach, the method uses participants' RSU BSMs. Since the goal is to assess the impact of the ERDW deployment on queue length during peak hour weekday travel conditions, the queue length measurement is done at a more aggregate level by estimating queue length over the morning peak hours (7–9 a.m.) at 10-minute intervals. The process consists of the following steps:

1. Divide the section of the REL comprising UC1 into 49 polygons having their centroids equally spaced at 16 meters.[6] Order the sequence to identify Segment 1 as the closest to the Twiggs Street/Meridian Avenue intersection and Segment 49 as the closest to the REL posted 40-mph sign (Figure 7-10).
2. For each polygon or REL segment, estimate mean speed using polynomial local smoothing regression.
3. If the estimated speed is less than or equal to 7 mph, define the polygon as congested.
4. Over the 10-minute interval, determine the highest ordered segment classified as congested.
5. Estimate queue length (meters) as the product of the segment number times 16 meters.

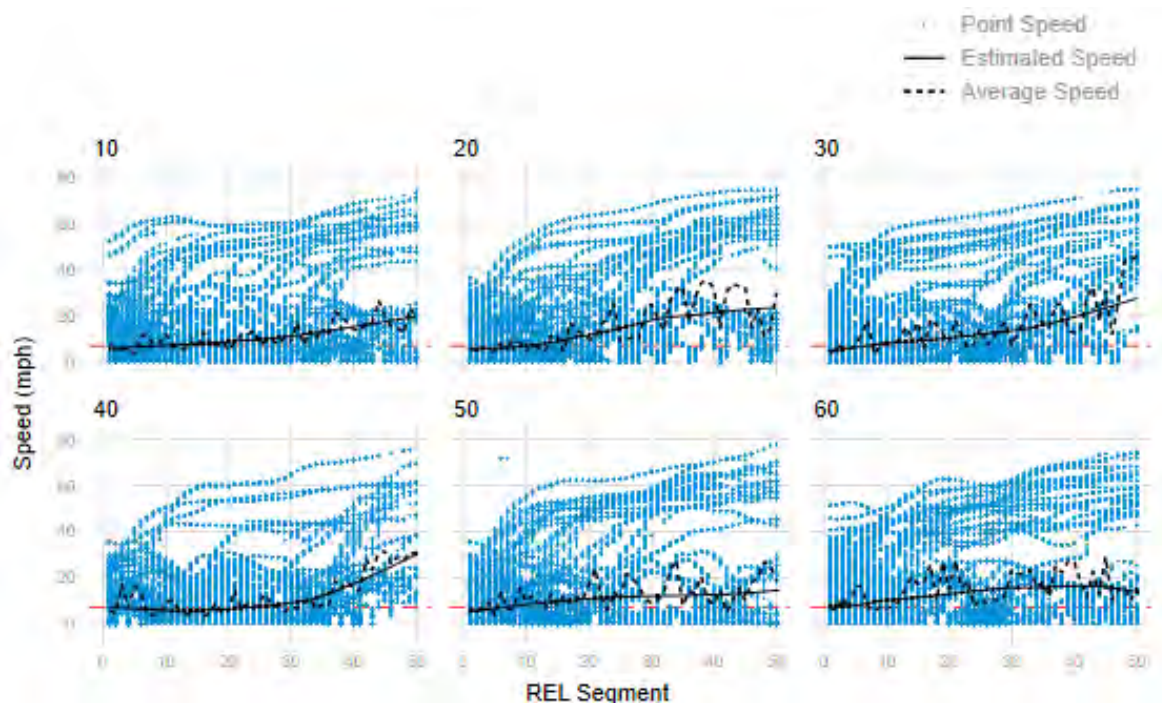
[6] The adopted ERDW resolution uses a total of 43 nodes equally distanced at 8–11 meters. The research team extended the number to 49 nodes solely to measure the observed queue length up to the posted 40-mph sign (node 49).



Source: CUTR, 2020

Figure 7-10. UC1 REL Segment Locations

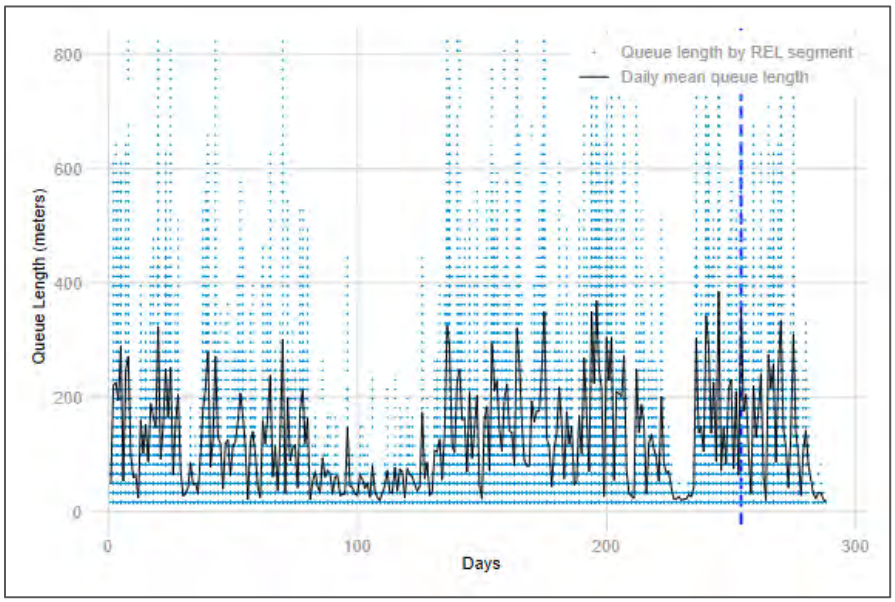
It is important to note the use of nonparametric locally weighted regression to produce the 10-minute speed estimates. By leveraging on the smoothing, the kernel-smoothed speed estimate provides a better measure in the presence of BSM point speed outliers, or whenever there is not enough information available within a given segment and time bin. Figure 7-11 plots the kernel estimated speed (solid line) and compares it with the sample mean speed (dotted line).



Source: CUTR, 2020

Figure 7-11. Ten-Minute Speed Estimates at the REL Segment Level

Figure 7-12 reports the queue length estimation over the analysis period. Each point on the graph represents a queue length value estimated in step 5 of the process. The black line represents the average queue length observed daily. The vertical dash line indicates the start of the ERDW deployment (February 3, 2020).



Source: CUTR, 2020

Figure 7-12. Queue Length Before-After ERDW Deployment

7.1.2.2.2 Interrupted Times Series Analysis Results

The final dataset consists of 1,519 daily measurements on queue length data culled at 10-minute intervals during the period February 3, 2019, through March 20, 2020. The 10-minute interval binning allows merging the travel time dataset with weather data (only available at 10-minute intervals) to check the impact of adverse weather on travel that is not captured by the time trends. Table 7-5 reports the descriptive statistics of the estimation dataset.

Table 7-5. Queue Length Estimation (Meters) – Descriptive Statistics

ERDW Deployment	Sample Size	Average	Min	Max	St. Dev
Before	1,347	189.8	16.5	823	184.8
After	172	223.5	16.5	823	207.9
Overall	**1,519**	**193.6**	**16.5**	**823**	**187.8**

Table 7-6 reports the results of the regression analysis, with three model specifications consistent with those used to regress travel time and travel time reliability. Similarly, the impact of weather is accounted for by using 10-minute interval measurements on the presence and intensity of rain. The impact of the onset of COVID-19 is measured by a dummy indicator set equal to one beginning on March 13, 2020, or one week earlier than the operational changes enacted in response to the pandemic.

Referring to Model 3 as the preferred model, the results show that at the beginning of the analysis period, the estimated maximum queue length (10-minute frequency) is about 174 meters measured from the Twiggs Street/Meridian Avenue intersection. As indicated by the statistically significant parameter associated with the time trend variable (*T*), the average maximum length exhibits a slight upward trend each day.

The relevant parameters associated with ERDW deployment (treatment and post-intervention) are all statistically significant. At the beginning of the ERDW application's deployment on February 3, 2020 (treatment variable), the estimated maximum queue length is lower than its baseline and demonstrates a downward trend as indicated by the parameter associated with post-intervention (-4.87). Except for weather, the controls for weekday travel (baseline Monday) and the COVID-19 pandemic are all statistically significant. The onset of the pandemic is associated with a marked decrease in maximum queue length as indicated by the negative sign and magnitude of the estimated parameter. This is substantiated by the observed reduction in the number of participant vehicles and overall travel times discussed in previous sections of this report.

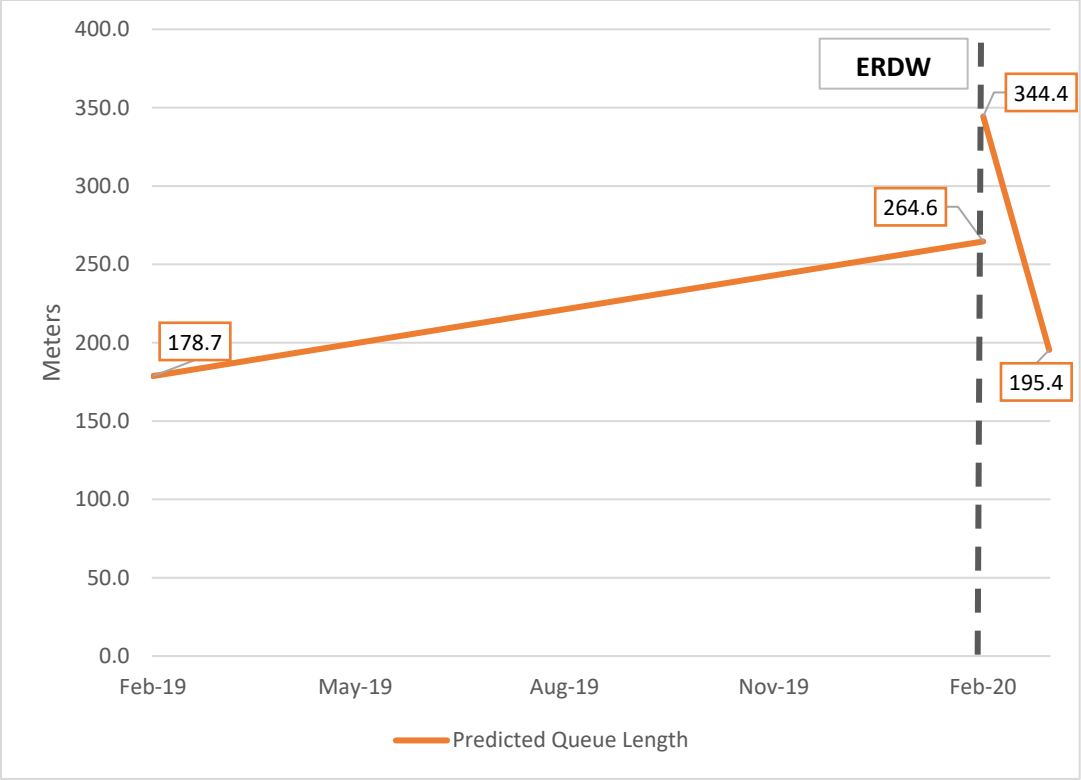
Table 7-6. Regression Model Results

Variable	Regression Models		
	Model 1	**Model 2**	**Model 3**
Time Trend (*T*)	0.245***	0.342***	0.340***
	(3.94)	(5.02)	(4.99)
Treatment (*X*)	116.7***	106.8***	84.65**
	(4.23)	(3.76)	(2.79)
Post-intervention (*XT*)	-7.458***	-7.265***	-4.866**
	(-6.71)	(-6.33)	(-3.09)
Tuesday†	33.30**	25.99	25.32
	(2.20)	(1.61)	(1.57)
Wednesday	-18.08	-15.48	-16.11
	(-1.38)	(-1.08)	(-1.13)
Thursday	-8.071	-13.77	-14.49
	(-0.57)	(-0.94)	(-0.99)
Friday	-137.0***	-134.2***	-133.7***
	(-12.38)	(-11.19)	(-11.13)
Rain		58.24	67.58
		(1.04)	(1.20)
Covid-19 Effect			-108.3***
			(-3.48)
Constant Term	174.1***	103.4**	96.17**
	(14.65)	(2.20)	(2.04)
Sample Size	1775	1416	1416
Adjusted R-square	0.130	0.140	0.143

*Standard errors in parenthesis: * p<0.05, **p<0.01, ***p<0.001*

†*Weekday baseline is Monday*

To relate the impact of the ERDW application, Figure 7-13 reports the adjusted predictions of the effect of the ERDW deployment on queue length at specific values of the sample. The marginal effects are reported assuming no adverse weather conditions (i.e., no rain), limiting the impact before the onset of the pandemic, and Tuesday as the representative weekday.



Source: CUTR, 2020

Figure 7-13. Adjusted Predictions of ERDW Impact on Queue Length

At the beginning of the analysis period in February 2019, the estimated queue length (10-minute frequency) is about 178.7 meters. Over time, the queue length shows an upward increase, which is estimated at about 264.6 meters at the beginning of the ERDW treatment. At the end of the evaluation period and through the ERDW application in March 2020, the predicted queue length decreases to 195.4 meters.

The estimation of Model 3 using a natural log transformation of the dependent variable (not reported here) reveals that the ERDW contributes to a 1.8 percent reduction in queue length with respect to the baseline (pre-intervention period).[7]

7.1.3 Safety Impact

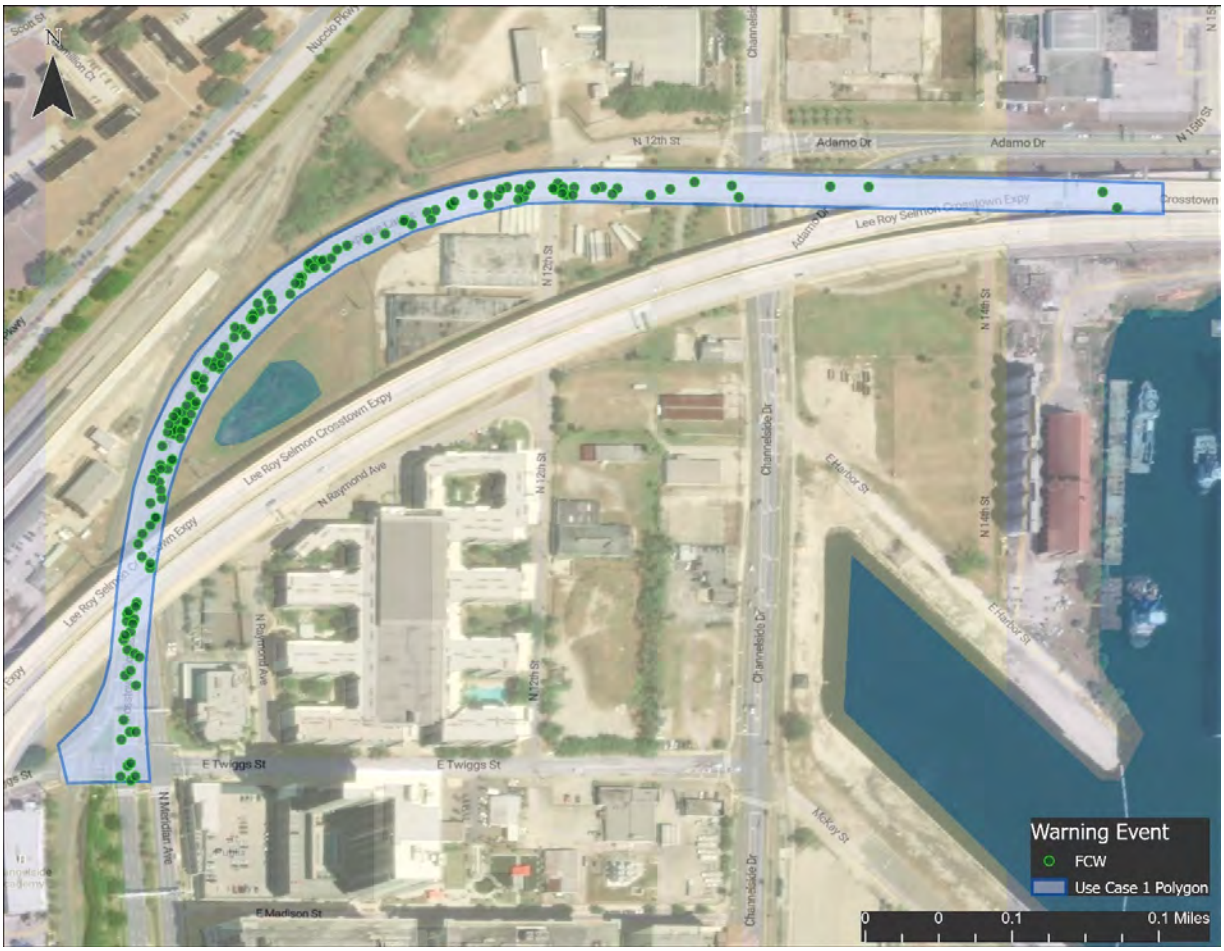
7.1.3.1 Crash Analysis

The crash analysis and reporting system [15] used to collect crash data reports no crashes on the REL for the five years prior to the analysis period (2014–2018).

[7] The natural log transformation of the dependent variable produces an estimate of the post-intervention parameter of -0.0181. Applying the proportional change transformation [(exp(-0.0181) -1] x 100, the estimated proportional change is -1.79 percent.

7.1.3.2 FCW Observed False Positives

During the analysis period, 61 participant vehicles deployed 150 Forward Collision Warnings within UC1. Figure 7-14 maps the FCW events in the UC1 segment. The research team created a BSM event profile for each warning, consisting of 30 seconds before and after the moment of warning. The events can be replayed and analyzed via the THEA CV Pilot Dashboard as a precursory step to data-driven evaluation [16].



Source: CUTR, 2020

Figure 7-14. Map of FCW Events

The first step of the evaluation involves PCE, where each warning event is analyzed and checked for conformity to the default application's operational parameters listed in Chapter 6. In the THEA CV Pilot Dashboard, this step is fully automated and runs daily.

The PCE analysis classified 49 events as false positives, 92 events as potentially true positives, and 9 events as not tested due to missing remote vehicle data. The next step is the visual inspection of the 92 FCW events identified as potentially true positives. Table 7-7 summarizes the results of the false positive analysis.

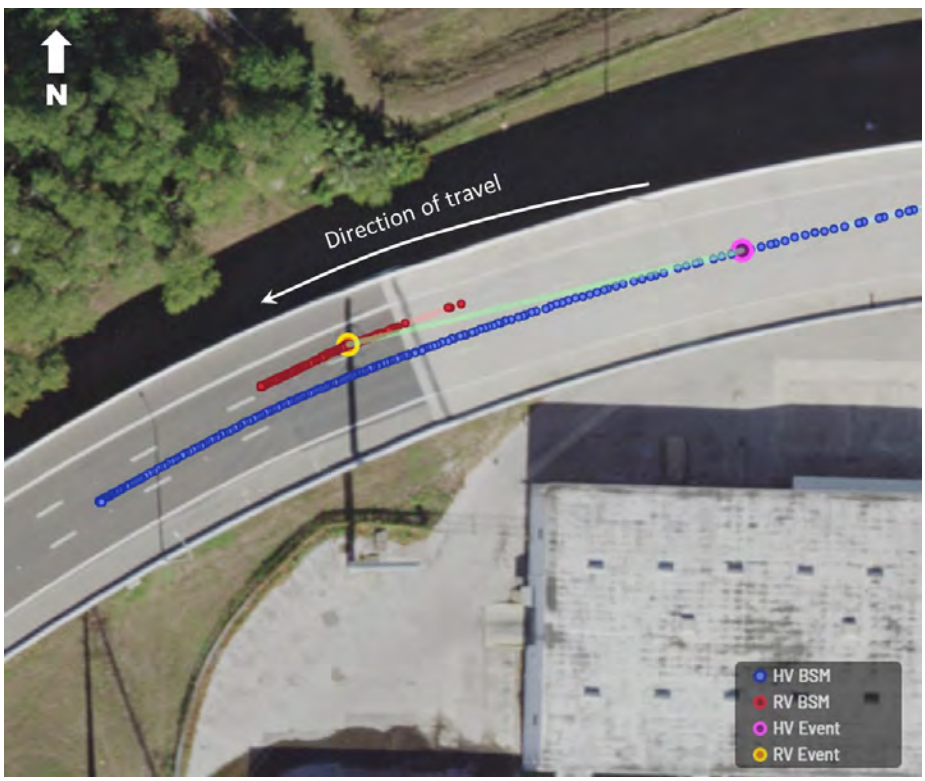
Table 7-7. FCW Analysis – False and True Positives

Classification	Count	Share (%)	Test Performed
False Positive	49	32.7	Automated PCE
False Positive	83	55.3	Visual Inspection
True Positive	9	6.0	Visual Inspection
Not Tested Due to Missing Data	9	6.0	
Total	**150**		

7.1.3.2.1 Factors Associated with FCW False Positives

A visual inspection was conducted to determine if the host and remote vehicles were in the same lane at the moment of warning according to the SAE J2945/1 "ahead in-lane" zone [5]. Because the REL segment used in this analysis is curved, most Forward Collision Warnings were categorized as FP because the OBU could not correctly determine the ahead in-lane zone of the RV (i.e., a warning was triggered but the RV was not in the same lane as the HV). Figure 7-15 illustrates a warning triggered due to the curvature of the road when the RV was in the adjacent lane instead of the same lane as the HV.

The visual inspection revealed that 80 (or 86.9%) of the 92 FCWs that passed the initial parameter check for FP occurred while the host and remote vehicles were traveling in different lanes. These were reclassified as false positives. Lane determination seems to be a major limitation of the application at REL curvature.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-15. FCW in Adjacent Lane

GPS shift due to loss of signal is another underlying cause of six FCWs classified as FP after visual inspection. The loss of signal seems to occur under the Selmon Expressway overpass as vehicles slow down on the REL exit lanes. Figure 7-16 shows an example of the host vehicle seemingly losing signal under the overpass and therefore shifting position closer to the remote vehicle. These instances did not occur often, but if the HV or RV position has a shift, then the application triggers warnings falsely.

Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-16. GPS Shift under the Overpass

7.1.3.3 EEBL Observed False Positives

The approach to evaluate Electronic Emergency Brake Light follows the same steps adopted to evaluate FCW. A BSM event profile was created for each EEBL warning to visually check the conditions under which it was triggered. To deploy the EEBL warning, the host vehicle determines if the remote vehicle sending the hard-braking information is in the lane ahead, the left lane, or the right lane. This allows for a width of three lanes instead of one, which is how the FCW works. During the analysis period and location of this use case, only four EEBL warnings were triggered and recorded. Figure 7-17 maps the EEBL warnings in the UC1 study area.

Source: CUTR, 2020

Figure 7-17. Map of EEBL Events

Three warnings were determined to be FP since they did not meet the application specification parameters. One warning was classified as TP since it met all parameters, and the RV was within the three lanes ahead of the HV. Table 7-8 summarizes the results of the FP analysis.

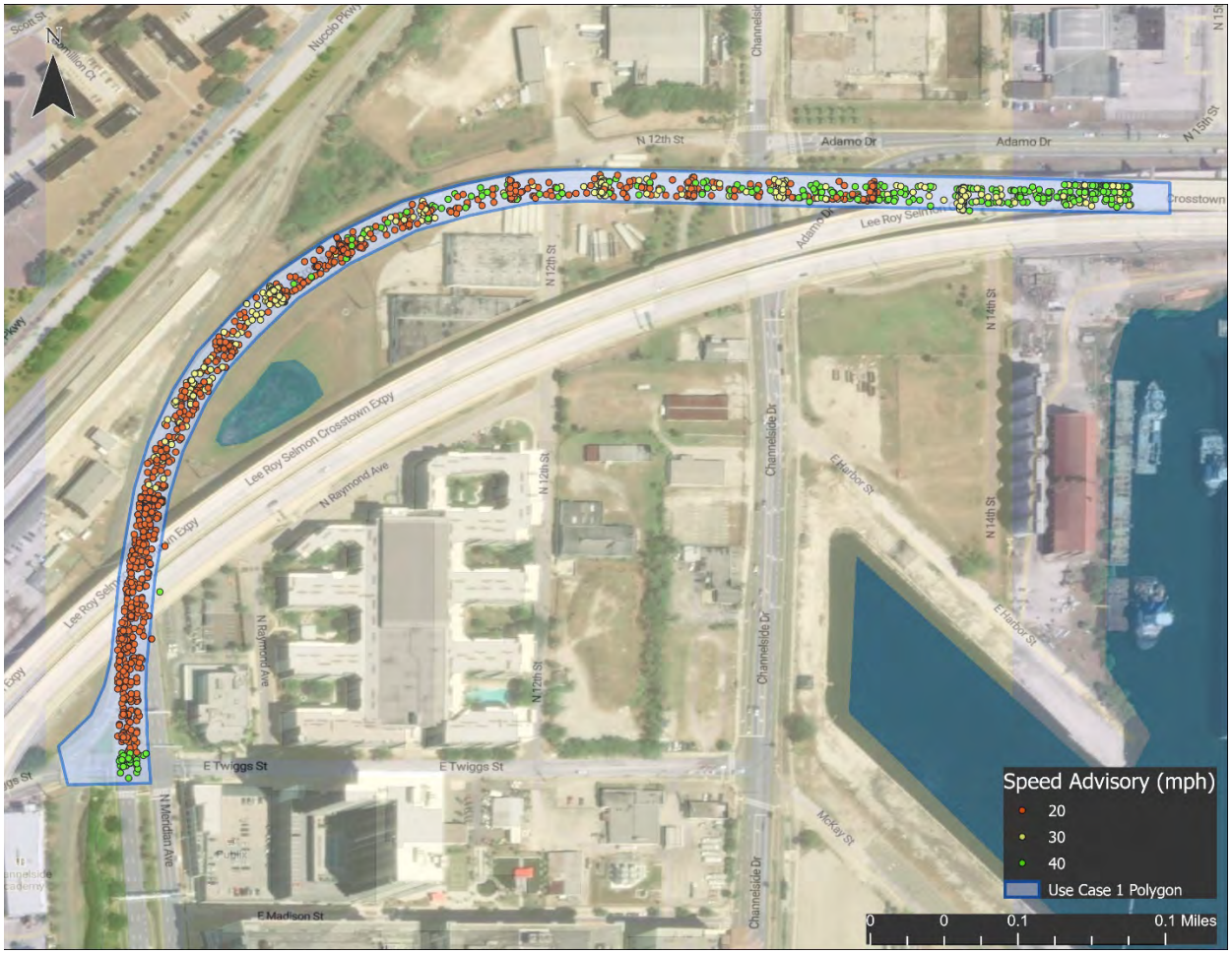
Table 7-8. EEBL Analysis – False and True Positives

Classification	Count	Share (%)	Test Performed
False Positive	3	75.0	Automated PCE
True Positive	1	25.0	Visual Inspection
Total	**4**		

7.1.3.4 ERDW Observed False Positives

The analysis focuses on data generated by the improved ERDW queue length estimation method, beginning February 3, 2020, and ending March 20, 2020, when the REL operational direction was set eastbound in response to the COVID-19 pandemic. February 3, 2020, represents the first day of deployment of the

improved ERDW application. During this period, 129 unique participant vehicles triggered 1,959 ERDWs. Figure 7-18 maps the warnings by speed advisory.[8]



Source: CUTR, 2020

Figure 7-18. Map of ERDW Events by Speed Advisory

To determine the presence and number of false positives, the host vehicle BSM point speeds at the moment of warning were compared against the vehicle's received speed advisory. All warnings triggered while the vehicle was traveling at a speed lower than the posted speed advisory were flagged as FP. Table 7-9 shows that 57.4 percent of ERDWs were categorized as TP and 42.6 percent as FP.

[8] While the analysis of the V2V application uses March 1, 2019, as the starting date of the baseline period, the ERDW assessment uses one additional month to capture any similarities in travel conditions between February 2019 and February 2020.

Table 7-9. ERDW FP and TP Counts

Description	Count	Share (%)
Unique Vehicles	129	--
ERDW	1,959	--
False Positives	835	42.6
True Positives	1,124	57.4

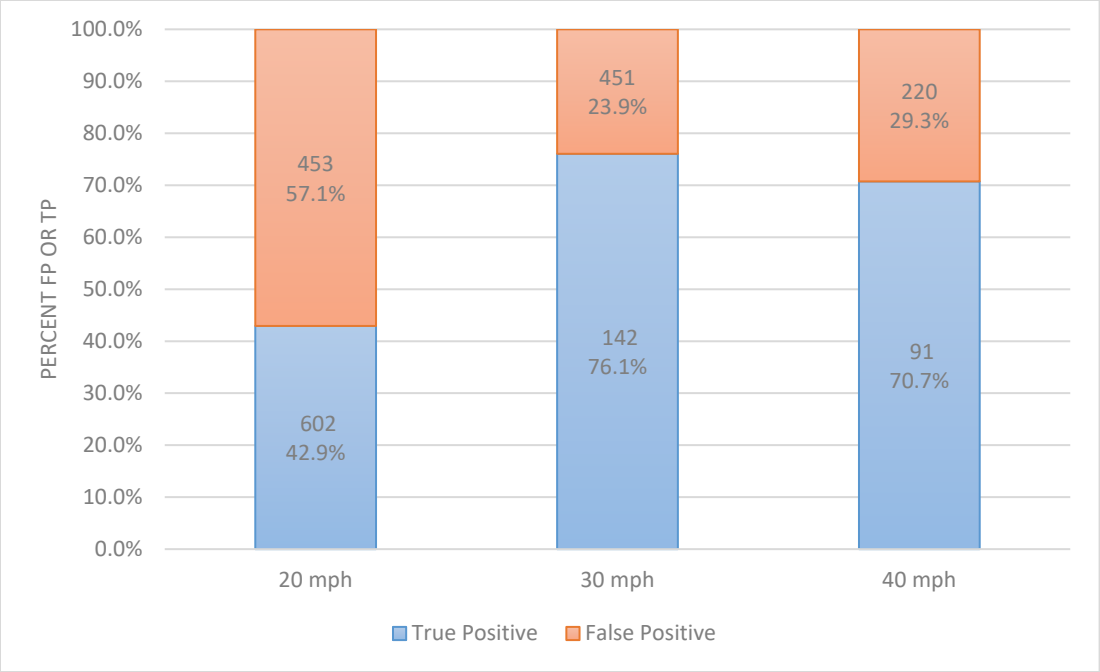
Table 7-10 splits the ERDW counts by FP status and by experimental design group. Of the 129 participant vehicles, 77 belong to the treatment group and 51 to the control group (one vehicle could not be assigned). While the total number of warnings issued by the treatment group is higher than the control group, testing equality of means of average ERDW events per vehicle confirms the two groups issued on average the same number of ERDWs per vehicle. On the other end, the tests confirm the treatment group vehicles issued on average more ERDWs classified as TP per vehicle than the control group (p-value = 0.0029). There is no statistically significant difference between the two groups in terms of average ERDWs per vehicle classified as FP (p-value = 0.4763).

Table 7-10. ERDW Counts by Experimental Design Group

Experimental Group	Vehicles	Total FP	Total TP	Total ERDWs	Average ERDWs per Vehicle	Average ERDWs per Vehicle FP	Average ERDWs per Vehicle TP
Treatment	77	480	762	1,242	26.2	27.9	25.1
Control	51	351	357	708	25.0	27.2	22.8
Combined	**128**	**831**	**1,119**	**1,950**	**25.7**	**27.6**	**24.4**

**One vehicle and nine ERDW events could not be assigned to either group.*

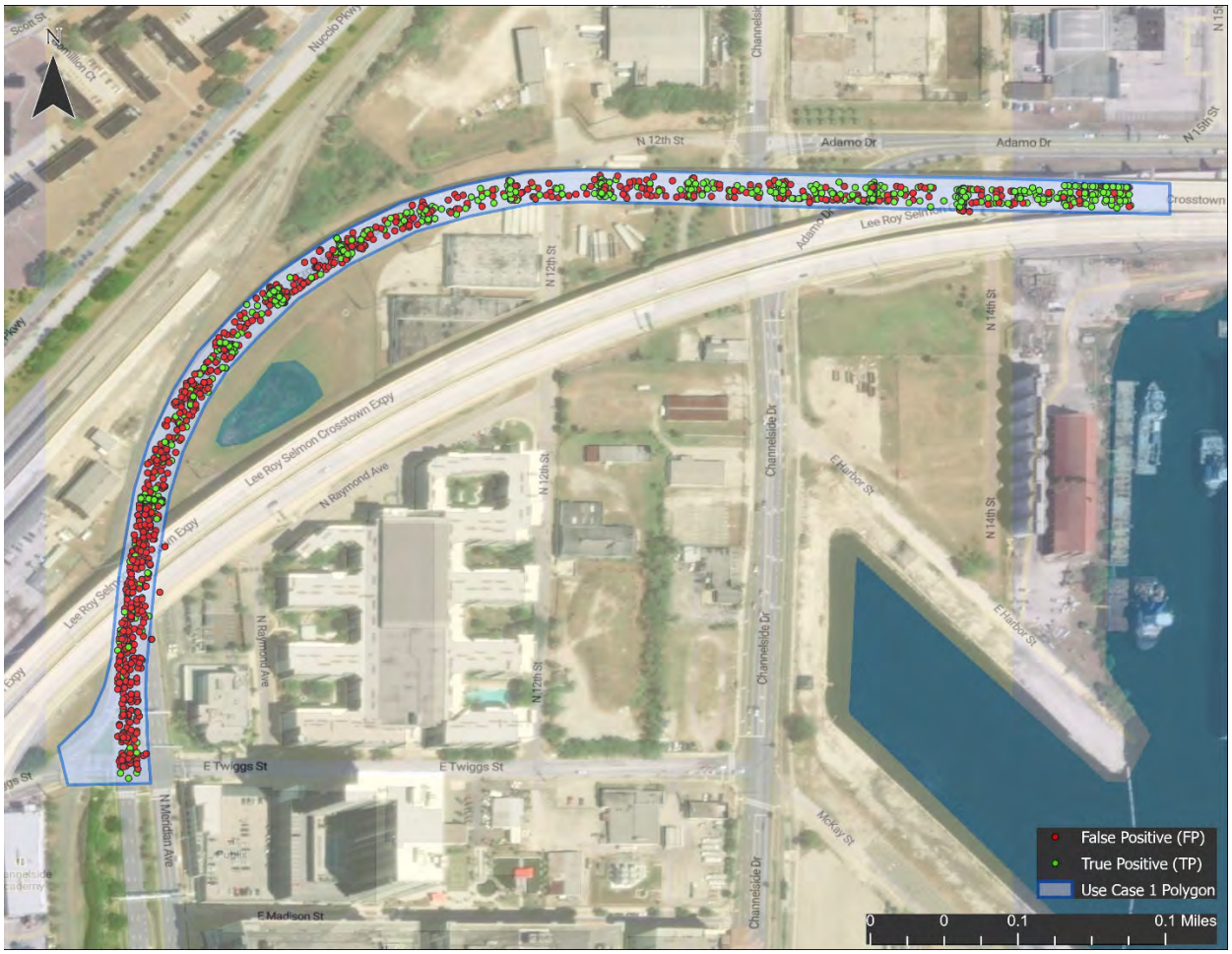
Figure 7-19 shows the breakdown of FP and TP counts by speed advisory to denote any relationship between false positives and speed advisories. The counts reveal a high share of true positives for the 30 to 40 mph speed advisories, 76.1 and 70.7 percent, respectively. Vehicles approaching the more congested segments of the REL are prone to receive a higher share of speed advisories (57.1%) that are not relevant (i.e., FP because the vehicle is already traveling at a lower speed than the posted advisory).



Source: CUTR, 2020

Figure 7-19. ERDW TP and FP Counts and Shares by Speed Advisory

Figure 7-20 maps the observed FPs and TPs to further indicate the clustering of a higher share of FP warnings toward the REL exit at Twiggs Street. A potential explanation of underlying causes can be ascribed to the speed advisories issued via localized broadcasting to all vehicles at the zone level as they approach the end of the queue.



Source: CUTR, 2020

Figure 7-20. False Positive Locations

In summary, the ERDW provided speed advisory information to participant vehicles during the congested hours of travel in the relevant sections of the REL ramp exiting at the Twiggs Street and Meridian Avenue intersection. While the application deployment and data collection time frame were affected by the system engineering changes applied to queue length estimation and by the onset of the COVID-19 pandemic, the abundance of data prior to the application deployment allows a before-after impact assessment of key mobility measures at a more aggregate level, as detailed in section 7.1.4.1.

7.1.3.5 V2V Interactions and Conflict Assessment

Table 7-11 reports the complete count of hours spent in the area, estimated number of interactions and conflicts between connected vehicles, and the FCW event classification. Nine warnings could not be tested due to missing RV information. During the analysis period, a total of 382 vehicles capable of recording data logs spent 485 hours traveling through the area defined for Use Case 1, with an estimated 12,450 interactions that produced 77 conflicts conformable to FCW deployment. Of those 77 conflicts, 9 also recorded a Forward Collision Warning.

Table 7-11. FCW Movement Classifications and Rates

Description	Count	Rate (%)
Unique Vehicles	382	--
Unique Vehicles Deploying Warnings	61	--
Time Spent in Area (Hours)	485	--
FCW (TP + FP + Not Tested)	150	--
V2V Interactions	12,450	--
Conflicts	77	--
True Positives (TP)	9	11.7
False Negatives (FN)	68	88.3
Non-conflicts	12,373	--
True Negatives (TN)	12,241	98.9
False Positives (FP)	132	1.1

Based on the figures of Table 7-11, the overall FP rate is

$$FP\ rate = \frac{132}{(132 + 12{,}241)} \times 100 = 1.1\%$$

The FN rate is

$$FN\ rate = \frac{68}{(68 + 9)} \times 100 = 88.3\%$$

Table 7-12 reports the estimated interactions and conflicts as well as warnings issued following the approach detailed in section 6.4. During the analysis period, a total of 382 vehicles spent 485 hours in the study area of the use case. There were 4,955 interactions that led to 43 conflicts conformable to EEBL deployment. As a result, the overall FP rate of the application is estimated at 0.1 percent, while the FN rate is estimated at 97.7 percent.

Table 7-12. EEBL Movement Classifications and Rates

Description	Count	Rate (%)
Time Spent in Area (Hours)	485	--
EEBL (TP + FP)	4	--
V2V Interactions	4,955	--
Conflicts	43	--
True Positives (TP)	1	2.3
False Negatives (FP)	42	97.7
Non-conflicts	4,912	--
True Negatives (TN)	4,909	99.9
False Positives (FP)	3	0.1

Table 7-13 reports the warnings generated and displayed to participants by the experimental design group. Out of 150 FCWs, 9 were not tested due to missing remote vehicle data, 43 percent were shown to drivers, and 57 percent were not shown. Only one EEBL was displayed to a participant.

Table 7-13. Warning Visibility by Participant Group

Application	Group	HMI Disabled (Warnings Not Displayed)			HMI Enabled (Warnings Displayed)			Grand Total
		TP	FP	Share (%)	TP	FP	Share (%)	
FCW	Control	4	41	31.9			0.0	
	Treatment			0.0	3	42	31.9	
	Treatment (Silent)	1	34	24.8	1	15	11.3	
	FCW Total	**5**	**75**	**56.7**	**4**	**57**	**43.2**	**141**
EEBL	Control		2	50.0			0.0	
	Treatment			0.0		1	25.0	
	Treatment (Silent)	1		25.0			0.0	
	EEBL Total	**1**	**2**	**75.0**	**0**	**1**	**25.0**	**4**

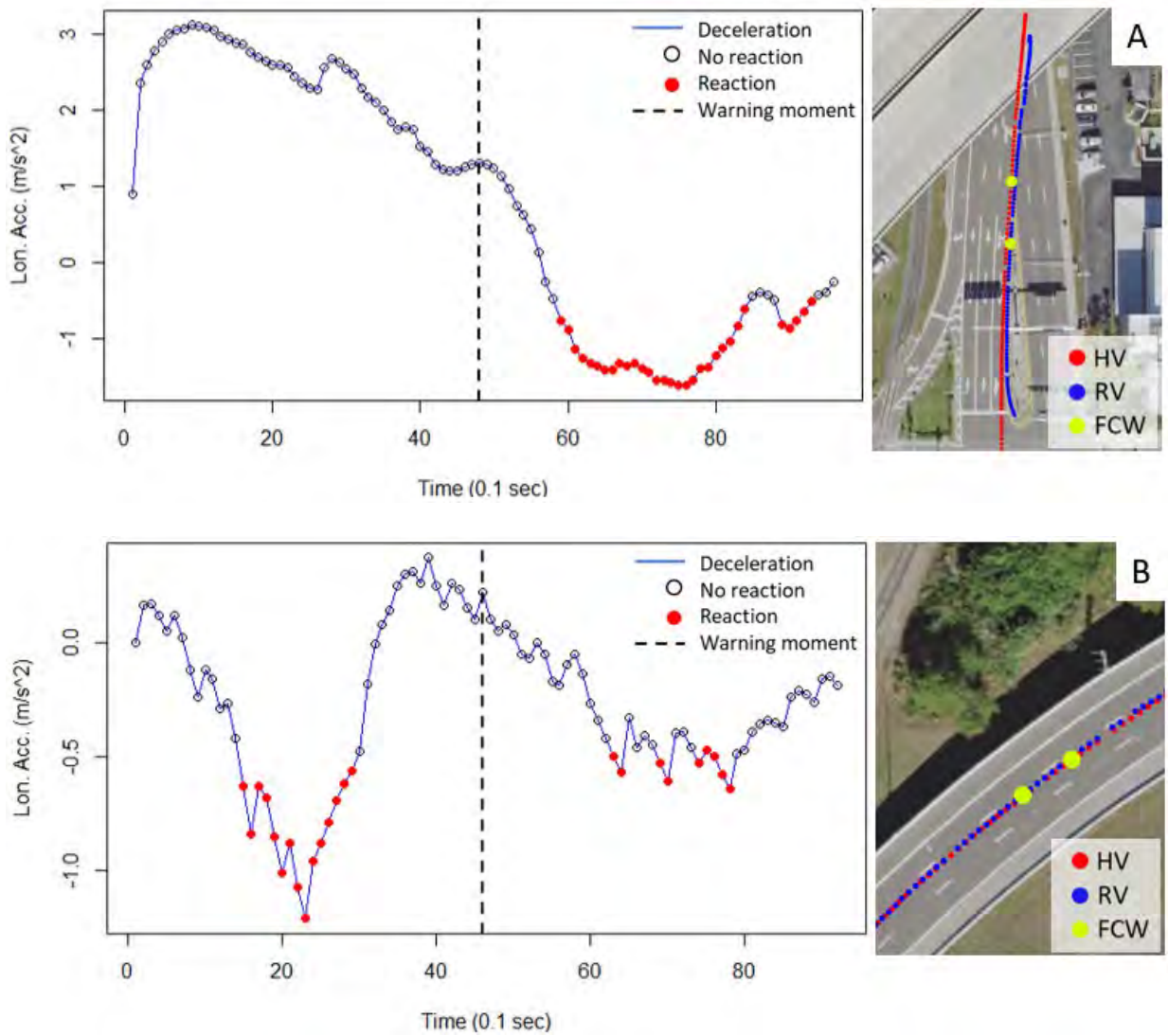
Next, the analysis assessed the behavioral responses by comparing reactions of the treatment and control groups and responses within the treatment group (HMI-disabled versus HMI-enabled). Driver reaction was investigated separately for each V2V application. The conflict identification algorithm identifies a reaction to a conflict if the driver decelerates at a rate below 0.5 mps^2 after the moment of warning. Since drivers facing an FCW or EEBL situation could also change lanes instead of decelerating, this was also visually checked to ensure they did not change lanes with no deceleration. No drivers changed lanes at or after the moment of warning.[9]

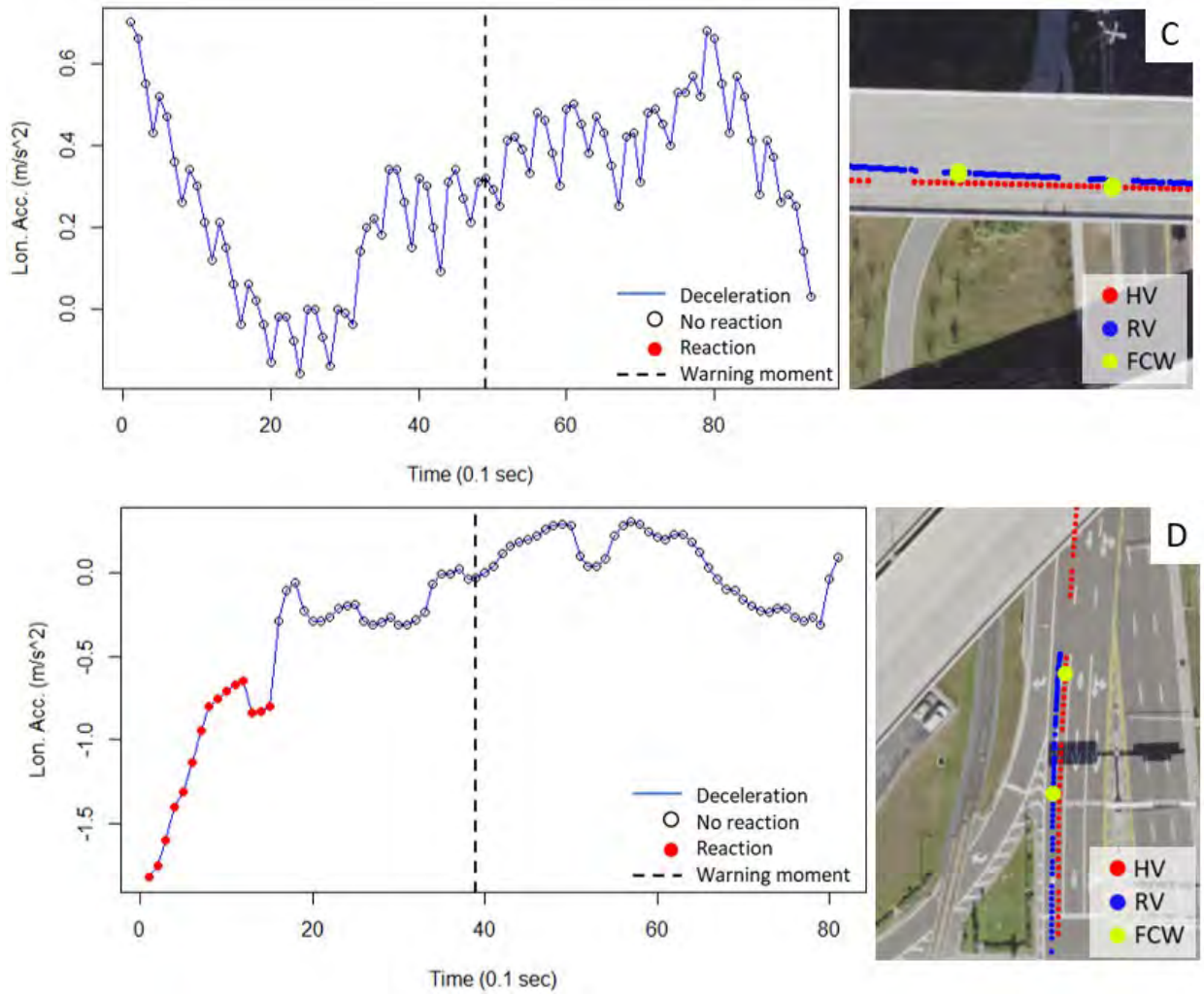
[9] While validating the performance of the algorithm, the visual inspection of the profiles revealed that because of the curvature of the REL, the algorithm identified some lane changes as evasive maneuvers while vehicles were simply following the road geometry. Therefore, the algorithm was only applied to longitudinal acceleration in these locations and the profiles were subject to visual inspections.

7.1.3.5.1 Treatment vs. Control Group Reaction to TP Warnings

The control group generated a total of five warnings classified as true positive events based on the identification steps outlined in Chapter 6. OBUs recorded nine Forward Collision Warnings classified as TP. Of those, four events were triggered by an HMI-enabled unit (treatment) and four by an HMI-disabled unit (control). One true positive FCW was not displayed to a participant in the treatment group during silent mode.

Figure 7-21 (A-D) displays the results of the data mining algorithm presented in Section 6.4.4. The left diagram shows the deceleration profile of each FCW event classified as TP and generated by an HMI-enabled vehicle using the 10-second BSM profile. The map next to the diagram shows the host and remote vehicle trajectories and location of the warning moment. The participant in Figure 7-21 (A) reacted after the warning; the participant in Figure 7-21 (B) reacted before and after the warning; in Figure 7-21 (C), the participant did not react; and in Figure 7-21 (D), the participant reacted before the warning.

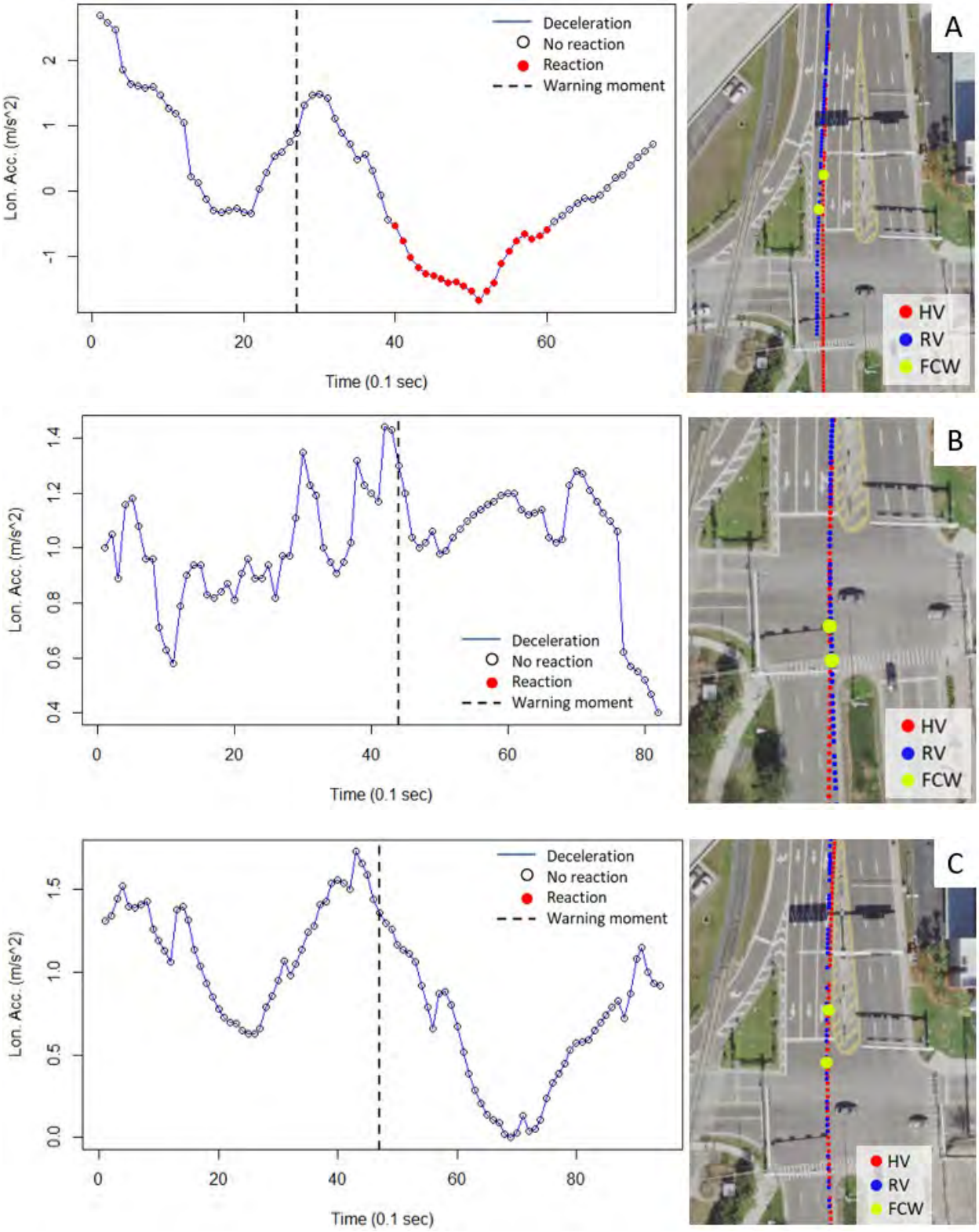


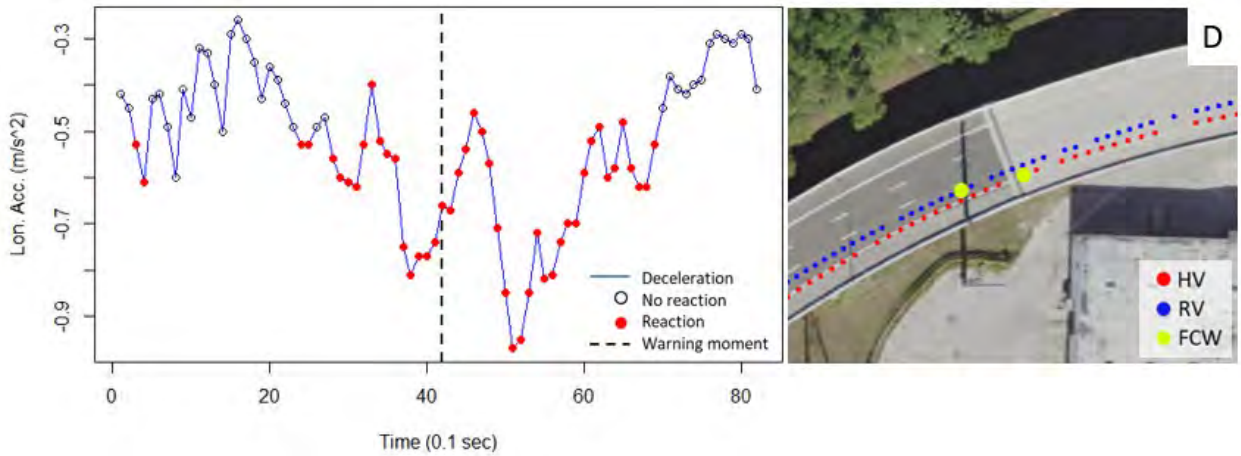


Source: CUTR, 2020

Figure 7-21. Driver Reaction to TP FCW with HMI Enabled

Figure 7-22 (A-D) displays the results of the data mining algorithm applied to the four TP Forward Collision Warnings triggered by participant vehicles assigned to the control group with the HMI disabled. The driver in Figure 7-22 (A) reacted after the warning moment, the drivers in Figure 7-22 (B-C) did not react, and in Figure 7-22 (D), the driver reacted before the warning moment. The mixed reaction results could be due to the specific conditions under which the warning was generated (yet not displayed). The speed might have been at the lowest threshold turning the application into operational mode, but the driver did not perceive it as a conflict situation. Also, the perception-reaction time of drivers differs and therefore the reaction to conflicts might vary from driver to driver.

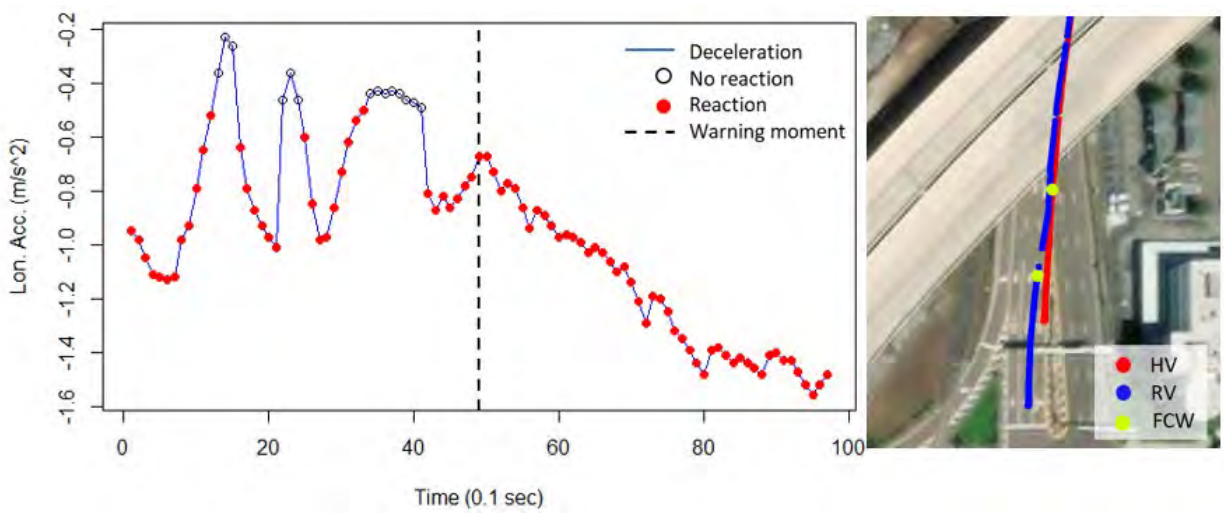

Deceleration
No reaction
Reaction
Warning moment
Lon. Acc. (m/s^2)
Time (0.1 sec)
A
B
C
HV
RV
FCW



Source: CUTR, 2020

Figure 7-22. Driver Reaction to TP FCW with HMI Disabled

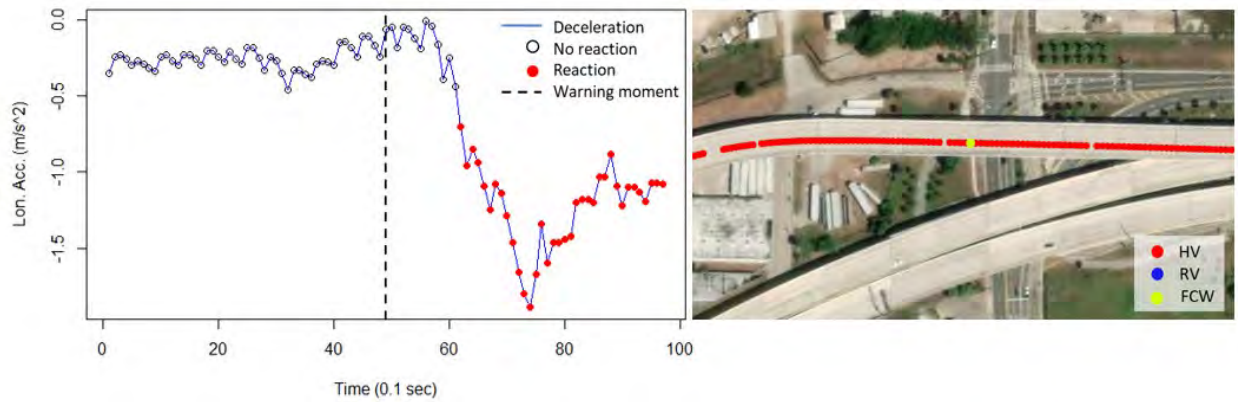
Only one EEBL warning was generated, classified as TP, and attributed to a participant from the treatment group in silent mode (HMI disabled). Figure 7-23 presents the true positive EEBL recorded. Based on the reaction protocol, the driver started reacting before the warning moment and continued reacting after the warning.



Source: CUTR, 2020

Figure 7-23. Driver Reaction to TP EEBL with HMI Disabled

For the ERDW application, the same procedure was used to determine if drivers reacted to the warnings shown to them. An example of a reaction to a visible warning is shown in Figure 7-24, where the driver decelerated after the warning displayed the suggested speed advisory.

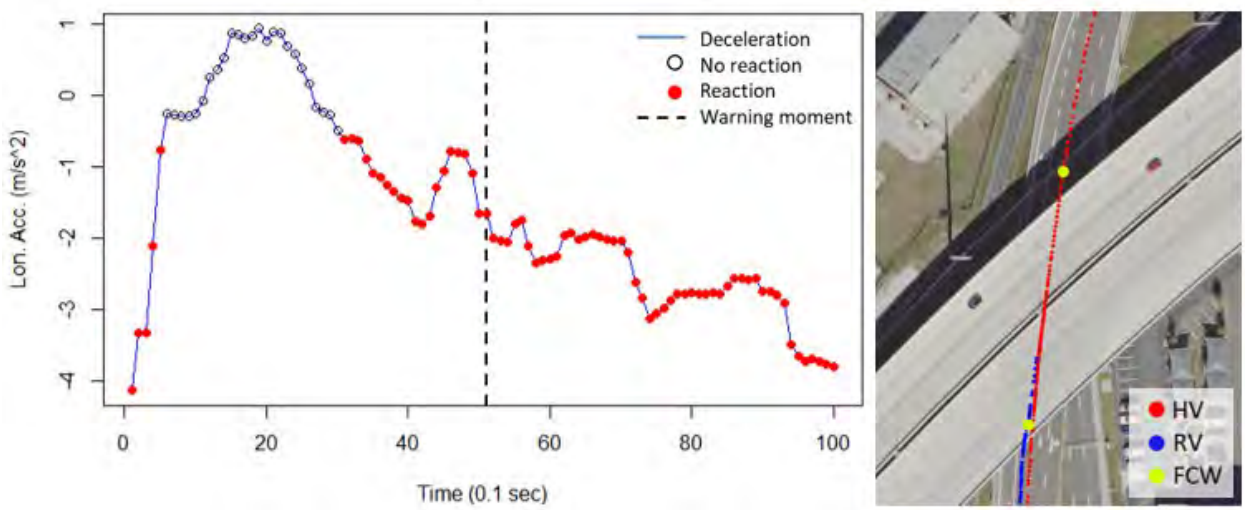


Source: CUTR, 2020

Figure 7-24. Driver Reaction to TP ERDW with HMI Enabled

7.1.3.5.2 Within Group Reaction to TP Warnings

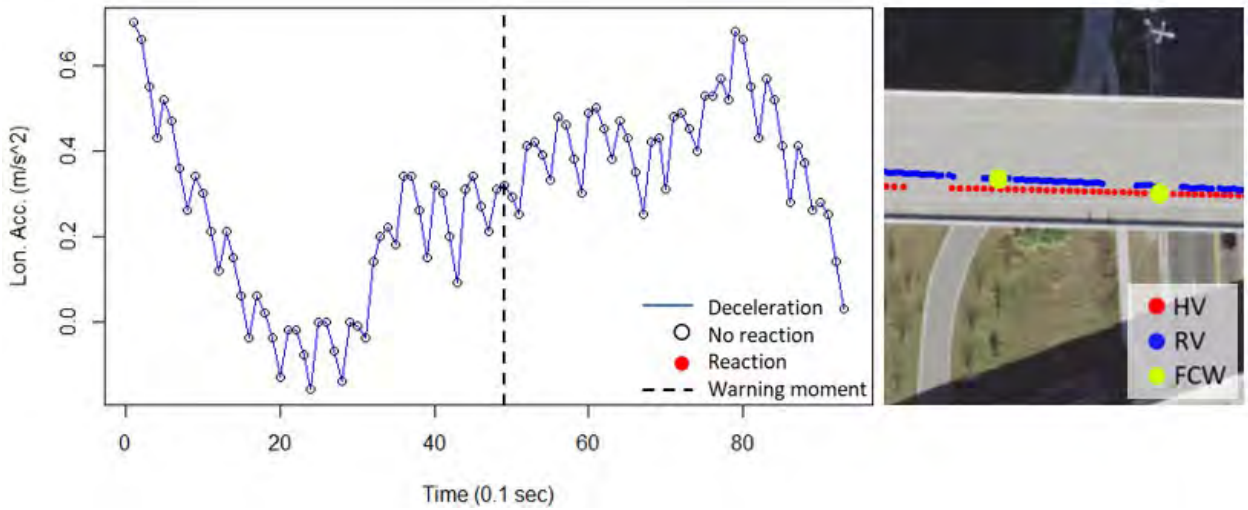
To assess the effect of the warnings on the treatment group, the silent mode was used to determine if drivers reacted differently in TP events with HMI enabled and disabled. Only two true positive FCW events were recorded with this classification. Figure 7-25 shows the FCW with HMI disabled during the silent period. This driver began reacting before the warning moment and hard braking reached values of -4 mps^2 or 0.4 G (g-unit).



Source: CUTR, 2020

Figure 7-25. Driver Reaction to TP FCW with HMI Disabled

The only true positive FCW with HMI enabled after the silent period ended is presented in Figure 7-26. The driver did not exhibit reactions as defined previously.



Source: CUTR, 2020

Figure 7-26. Driver Reaction to TP FCW with HMI Enabled

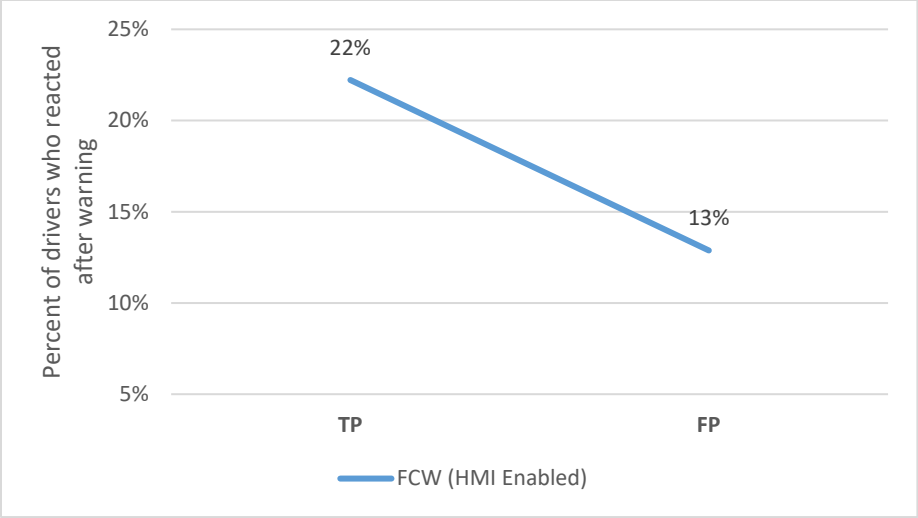
7.1.3.5.3 Driver Reaction to All Visible Warnings (HMI Enabled – All Groups)

The warnings shown to participants were generated under two types of conditions. The first condition involves a TP with another vehicle determined by the parameters and the method explained in the previous sections for each application. The second condition involves a FP where the warning is triggered when there is no conflict with another vehicle. It is possible for the participant to respond under both conditions, albeit differently. Next, the analysis pools all the warnings recorded while the HMI was enabled to evaluate any difference in behavioral response by type of conflict (TP vs. FP).

This analysis is relevant because of the travel conditions underlying the generation of these warnings. During the weekday morning travel, participants driving on the REL can come near one or more participants while at the same time being surrounded by other non-equipped vehicles. This means that a participant seeing a warning (even if classified as FP), might generate an observable reaction since there could be a non-CV ahead causing a conflict situation.

Figure 7-27 illustrates the difference in the share of drivers who reacted after they received a visible warning, grouped by TP and FP classification. Twenty-two percent of participants reacted to FCWs classified as TP compared to only 13 percent who reacted to warnings classified as FP. This indicates that drivers were more likely to react if the conditions around them were classified as TP (a conflict) compared to the conditions determined as FP (no conflict).

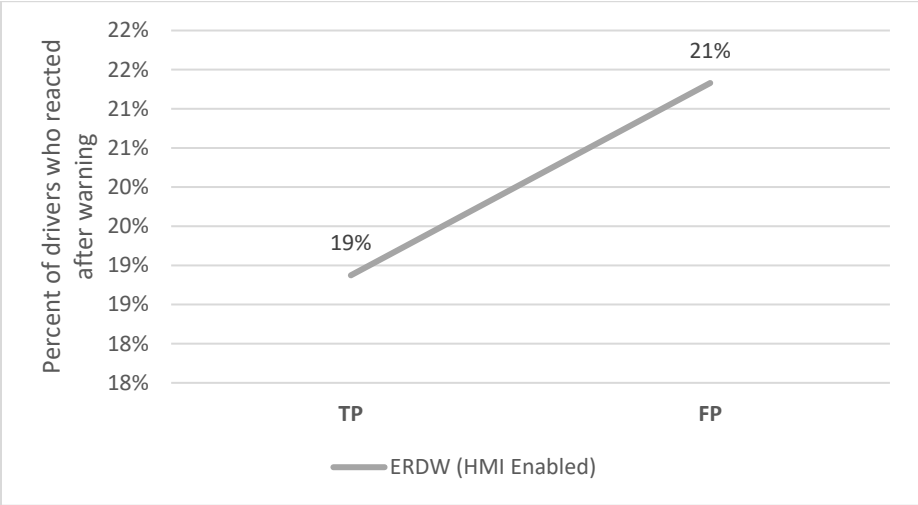
Since the number of EEBL warnings was very low (four), a comparison was not made for the two groups.



Source: CUTR, 2020

Figure 7-27. Proportions of Drivers Reacting to FCW with HMI Enabled

In the case of the ERDW, the reaction is less nuanced; 19 percent of drivers reacted to speed advisories classified as TP compared to 21 percent who reacted to warnings classified as FP (Figure 7-28). This comparison does not demonstrate any statistically significant difference between the two groups, which could be due to speed of travel perception versus the speed advisory suggestion.



Source: CUTR, 2020

Figure 7-28. Proportions of Drivers Reacting to ERDW

7.1.3.5.4 Predictive Value of Warnings

Since crashes and conflicts are rare traffic events, it is useful to examine two additional metrics in the assessment of V2V safety applications: the positive and negative predictive values of warnings. Given the many factors affecting the deployment and efficacy of CV safety applications, these measures evaluate the conditional probability of being in a dangerous situation when a warning is triggered (positive predictive value) or not being in a dangerous situation when a warning is not triggered (negative predicted value).

In this context, the probability that a warning will be a TP (conflict situation) is the proportion of TP warnings to all warnings (TP + FP). This is a true performance measure of an application since it gauges how likely a driver is to be in a conflict when a warning is received.

For the FCW application, the positive predictive value is calculated as

$$FCW\ positive\ predictive\ value = \frac{TP}{TP + FP} \times 100 = \frac{9}{9 + 132} \times 100 = 6.4\%$$

Given the travel conditions characterizing Use Case 1, this means that if a Forward Collision Warning is triggered, the probability of being in a real dangerous situation is approximately 6 percent.

On the other hand, the probability that a negative (no warning) will be a TN (no conflict) is the proportion of TN to all negatives (TN + FN). The negative predictive value of the FCW application as deployed in UC1 can be calculated as

$$FCW\ negative\ predictive\ value = \frac{TN}{TN + FN} \times 100 = \frac{12241}{12241 + 68} \times 100 = 99.4\%$$

This means that if a Forward Collision Warning is not triggered (no warning), drivers are not in a dangerous situation 99.45 percent of the time, and drivers face a conflict only 0.55 percent of the time (1 in 181). Table 7-14 shows the positive and negative predictive values for the two safety applications as deployed and evaluated specifically for UC1. Being a function of location and travel characteristics, the above values might change if the evaluation is carried out at different locations within the CV Pilot study area.

Table 7-14. Application Predictive Values

Application	Positive Predictive Value (%)	Negative Predictive Value (%)
FCW	6.4	99.4
EEBL	25.0	99.2

7.1.4 Summary of Findings

7.1.4.1 Impact on Mobility

To evaluate the impact of the ERDW on mobility measures, this analysis processed baseline data from 587 unique vehicles generating more than 29 million BSMs and a total of 1,959 ERDW events generated by 129 participants during February–March 2020, comprising the implementation phase of the improvements to the application.

The impact assessment relies on interrupted time-series analysis to compare the effect of the ERDW deployment on the above mean travel times, the 95th percentile travel times, and queue length. The approach

relies on a comparison of time trends in the performance measures before and after deployment of the ERDW application. The goal is to assess if the treatment (i.e., the implementation of the ERDW application) causes a change in patterns upon baseline conditions.

The evaluation of the impact of the speed advisories on these benchmark measures demonstrates that the ERDW positively contributes to mobility improvements. Findings reveal that the ERDW contributes to a 2.1 percent reduction in mean travel times with respect to the baseline (pre-intervention period), a 1.1 percent reduction in idle time, and a 1.8 percent reduction in queue length. At the beginning of the analysis period of February 2019, the estimated TTI is about 2.7. After the deployment of the ERDW application in March 2020, the estimated TTI is about 1.9.

7.1.4.2 Impact on Safety

Assessing the impact of V2V safety applications on improved safety in the context of a deployment presents challenges that are not solely site specific. The complex nature of traffic conflicts and their many contributing factors are significant. Since crashes and conflicts are rare events, developing CV safety applications and fine-tuning their parameters to identify and contribute to preventing rare events presents nontrivial challenges. Even if a near-perfect safety application development and settings are assumed, the potential benefits depend on driver acceptance of the applications, driver reaction to the warnings, and the reaction of other surrounding drivers. For example, a hard-braking reaction to a Forward Collision Warning might prevent a forward collision for the host vehicle, but it can cause several dangerous conflicts or even rear-end crashes behind the vehicle in congested conditions. This is how multi-vehicle rear-end crashes occur. However, if more conflicts are created, then other applications such as the EEBL can potentially trigger more warnings and thus aid in avoiding crashes.

The EEBL application deployed in UC1 triggered four warnings, three of which were false positives. This application is meant to warn drivers of hard-braking vehicles in the lane ahead and in adjacent lanes of travel. The application is only operational if it receives a broadcast of a hard-braking event (lateral acceleration of -4 mps^2) from the remote vehicle. One reason for the low number of warnings may be that drivers who are commuters know the traffic patterns and anticipate congestion, therefore they brake at lower rates than the hard-braking threshold. This leads to fewer hard-braking flags and fewer EEBL warnings. Another reason is that the EEBL does not use time-to-collision as the FCW does, but rather uses the deceleration value for vehicles (as shown in Table 6-1). This might mean that the parameters for EEBL are more tuned for identifying true positives.

To compare the safety impact of the two safety applications, a before-after comparison of conflicts that might have led to warnings given the applications were active was conducted for a period before deployment of the warnings. The before period was set to March 1, 2019, to July 31, 2019 (six months). The after period for analysis was set to August 1, 2019, to March 20, 2020 (seven months, 20 days). As shown in Table 7-15 the two periods experienced similar number of vehicles interacting and in conflict situations. The rate of conflicts per interaction is similar for FCW (0.63% before, 0.62% after) and increased for EEBL (0.49% before, 0.87% after). These rates show that the months before the analysis period exhibited similar characteristics in terms of connected vehicle interactions and conflicts leading to FCW and EEBL deployment.

Table 7-15. UC1 Conflict Rates

	Description	Before Period		After Period	
		Count	Rate (%)	Count	Rate (%)
FCW	Unique Vehicles (in Interactions)	348	--	343	--
	Unique Vehicles (in Conflicts)	56	--	54	--
	V2V Interactions	11,081	--	12,450	--
	Conflicts	70	0.6	77	0.6
	Non-conflicts	11,011	99.4	12,373	99.4
EEBL	Unique Vehicles (in Interactions)	328	--	333	--
	Unique Vehicles (in Conflicts)	35	--	26	--
	V2V Interactions	6,543	--	4,955	--
	Conflicts	32	0.5	43	0.9
	Non-conflicts	6,511	99.5	4,912	99.1

The development of safety applications must balance two perspectives: a high true positive rate (i.e., warnings triggered when there is a conflict) and a low false positive rate (i.e., warnings triggered when there is no conflict). Advanced Driver Assistance Systems deployed on automated vehicles have demonstrated that drivers trust them when they work well. If the false positive rate is high, drivers choose to ignore or even turn off the system due to the nuisance and distraction of alerts [17, 18].

In evaluating Use Case 1, the two safety applications generated 10 warnings classified as true positive, but only 4 were shown to the drivers due to the evaluation's experimental design. Had all participants seen the warnings and reacted as expected, they might have avoided as many as 10 crashes due to the deployment of the warnings. In real-world conditions outside the THEA CV Pilot, this could significantly impact the safety of THEA's REL travelers.

7.1.4.3 Lessons Learned

During the deployment of applications for UC1, several challenges were presented and resolved. In summary:

1. The ERDW application depended on the accurate calculation of queue on the REL in the morning peak hours. This was proven to be inaccurate using MMITSS due to penetration rates and the exit ramp toward Twiggs Street. The team changed the method and used BSMs to calculate the queue length, thereby allowing the application to work properly. Since additional time was required to create the new solution, this meant that the period for performance measurement was reduced to less than two months. The lesson learned is to be prepared to implement changes at a fast pace to deploy solutions in a timely manner.

2. The two safety applications (FCW and EEBL) require fine tuning to properly detect the location of the remote vehicle ahead. In the case of FCW, the RV must be in the lane ahead, and in the case of EEBL, the RV must be in the lane ahead or in adjacent lanes. Due to the curve on the REL exit (UC1 segment), the application could not correctly identify the location of the RV, leading to a high FP percentage.

3. In the case of ERDW, the application must be tuned to avoid FPs, delivering warnings for higher speed advisory when the vehicle is already traveling below that advisory. These FPs contribute to lowering participants' trust in the system or ignoring the warnings altogether.

7.2 Use Case 2: Wrong-Way Entry

7.2.1 Analysis Dataset

The analysis uses data collected from participant vehicles between March 1, 2019, and March 20, 2020. On March 20, 2020, THEA set the REL operational direction to eastbound on a 24-hour basis (leaving downtown Tampa toward Brandon) in response to the COVID-19 pandemic.

To consider the possibility of both legal and illegal movements leading to Wrong-Way Entries with respect to the REL operational directions shown in Figure 6-23, the evaluation focuses on weekday travel and the following time periods: 6 a.m. to 10 a.m. (AM movements) and 3 p.m. to 12 a.m. (PM movements). The weekday time between 10 a.m. and 3 p.m. is excluded because of the REL's split operation during those hours. Finally, the dataset excludes national holidays since the REL operates eastbound on a 24-hour basis during those days.

Table 7-16 reports the total number of WWE warnings generated withing a polygon encompassing Use Case 2 and identifies the number of unique WWE events generated by OBU data logging enabled units during the period of March 1, 2019, through March 20, 2020. Unique events are identified by grouping the WWE warnings issued by the host vehicle within one minute of each other. This sample is used in the ensuing analysis presented in the remaining sections of this report.

Table 7-16. Data Sample for WWE Analysis

REL Operation	Total WWE Warnings	Unique WWE Events
REL Westbound AM (6:00 to 9:59)	906	687
REL Eastbound PM (3:00 to 11:59)	5,070	4,137
Total	**5,976**	**4,824**

7.2.2 Mobility Impact

UC2 did not generate quantifiable mobility measures directly attributable to the WWE application deployment, as detailed in the next section.

7.2.3 Safety Impact

7.2.3.1 Analysis of AM Movements and WWE Events

During the REL westbound morning operations, between 6 a.m. and 10 a.m., the allowed movement is from the REL westbound toward East Twiggs Street and Meridian Avenue. Any OBU-equipped vehicle traveling on the REL and egressing at the Twiggs intersection should not be issuing WWE warnings.

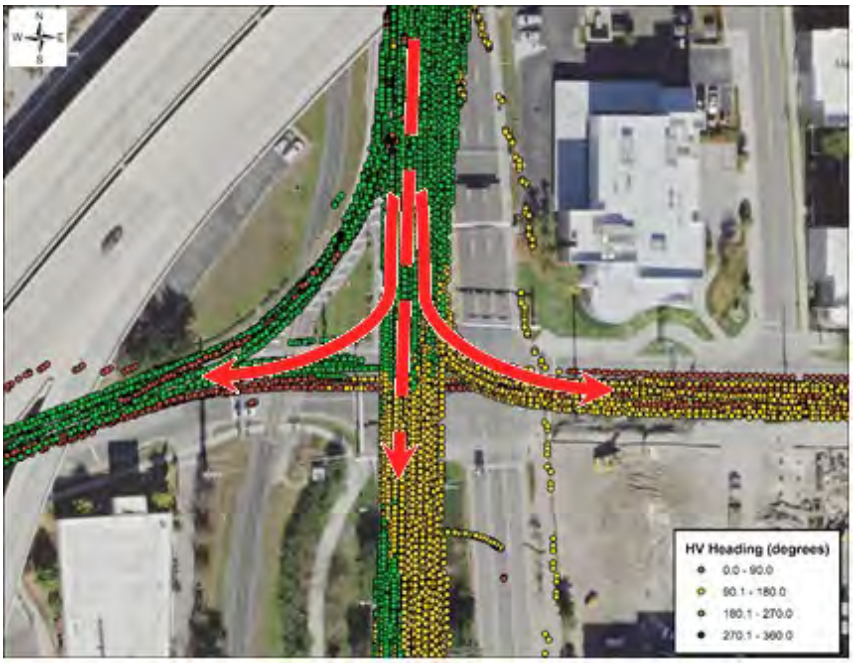
Table 7-17 reports the total number of WWE warnings during the analysis time frame and Figure 7-29 maps their location. A total of 906 warnings were issued to 127 unique vehicles, resulting in 687 unique WWE events (i.e., warnings issued by the HV within one minute of each other during one turning movement). Of these, 665 were unique WWE events issued to vehicles traveling westbound on the REL and exiting at the Twiggs intersection, as shown in Figure 7-30. The remainder (22) were unique WWE events associated with turning movements in other directions, as shown in Figure 7-31.

Table 7-17. WWE Event Sample – REL AM Operation

Description	WWE	Unique WWE Events
REL to Twiggs Westbound	882	665
Other	24	22
Total	**906**	**687**

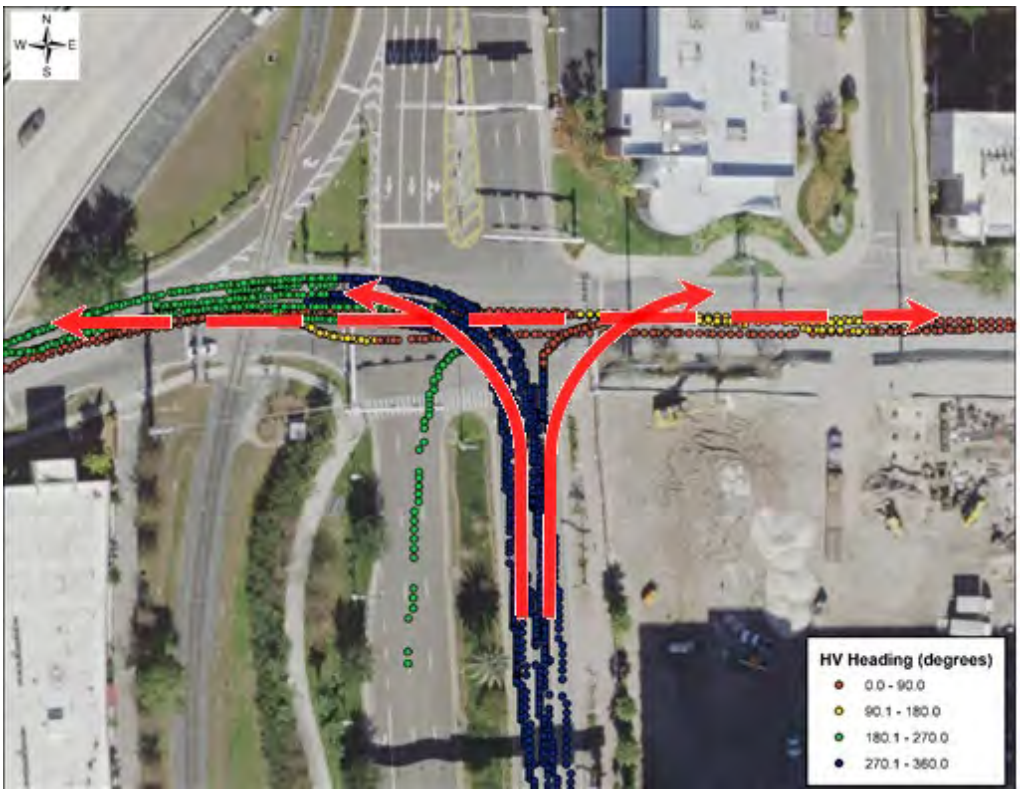
Source: CUTR, 2020

Figure 7-29. WWE Events – REL AM Operation



Source: CUTR, 2020

Figure 7-30. AM WWE Events Turning Movements – REL Westbound to Twiggs St.



Source: CUTR, 2020

Figure 7-31. AM WWE Events Turning Movements – All Other Directions

7.2.3.1.1 Observed False Positives

CUTR analyzed the 665 WWE unique events associated with REL eastbound movement to Twiggs Street and determined they were allowed movements incorrectly identified as violations by the WWE application. The profile analysis of the remaining 22 WWE events revealed that no vehicle attempted to enter the REL the wrong way. Thus, all 687 unique WWE events were classified as false positives. The factors associated with these FPs are outlined in section 7.2.3.1.3.

Table 7-18 reports the complete count of turning movements and the WWE event classification. Note that over the same period, no true positives (wrong-way intrusions) were recorded, resulting in zero false negatives and a false negative rate of zero. The overall false positive rate during the AM REL operational profile is three percent.

Table 7-18. WWE Movement Counts – REL AM Westbound

Description	Turning Movements	WWE Warnings	Unique WWE Events	True Positive Events	False Positive Events	True Negative Events	False Negative Events
REL Westbound to Twiggs St. & Meridian Ave.	22,094	882	665	0	665	20,856	0
Other	2,287	24	22	0	22	2,248	0
Total	**24,380**	**906**	**687**	**0**	**687**	**23,104**	**0**

** Includes: Meridian to E. Twiggs, Meridian to W. Twiggs, E. Twiggs to Meridian, W. Twiggs to Meridian, E. Twiggs to W. Twiggs, W. Twiggs to E. Twiggs and movements that could not be assigned.*

Table 7-19 reports the number of WWE unique events, turning movements, and classification of conflicts (actual WWE) and non-conflicts (no WWE) with the associated rates.

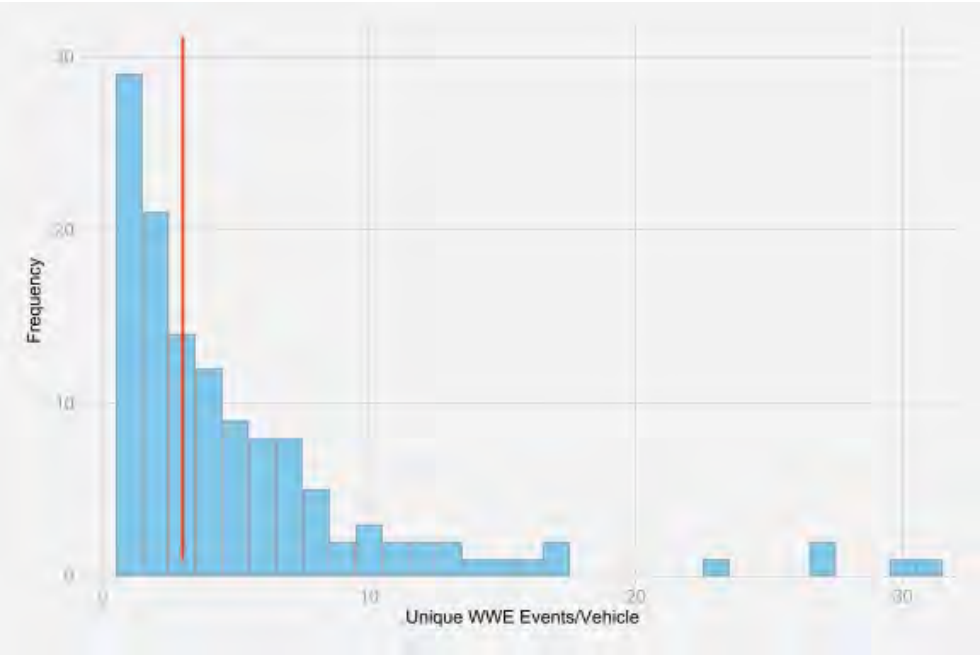
Table 7-19. WWE Movement Classification and Rates – REL AM Westbound

Description	Count	Rate (%)
Unique Vehicles	475	--
Unique Vehicles Deploying Warnings	115	--
WWE Events (PTP + FP)	687	--
Turning Movement	24,380	--
Conflicts	0	--
Potentially True Positives	0	0.0
False Negatives	0	0.0
Non-conflicts	24,380	--
True Negatives	23,693	97.2
False Positives	687	2.8

7.2.3.1.2 False Positives by Vehicle

During the AM REL operation, 665 unique WWE events were issued by 127 vehicles. The higher the number of FP warnings, the higher the potential to decrease a participant's trust in the technology and their acceptance of the application. While the WWE application false positive rate for turning movements during the AM REL

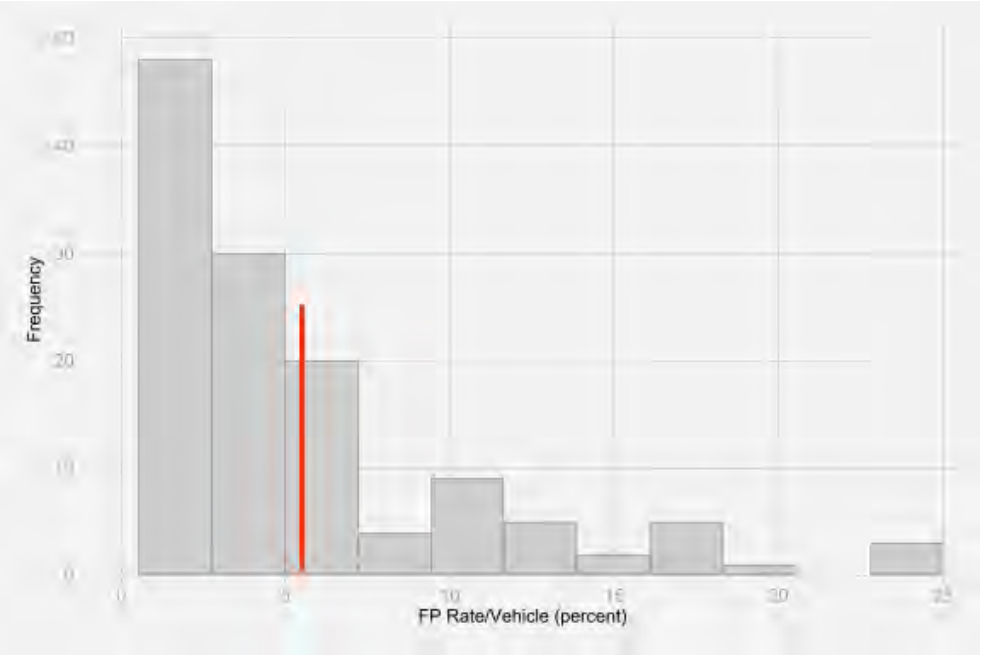
operation is estimated to be three percent, several factors might contribute to a high variability in FP rates across the participants. For example, some participants travel more often than others to the study area, thus increasing the total number of warnings generated during the analysis time frame. Figure 7-32 shows a histogram of unique WWE events per vehicle. While 50 percent of the vehicles generated 3 unique events (red line, or median), some vehicles generated up to 30 unique events.



Source: CUTR, 2020

Figure 7-32. Unique WWE Events per Vehicle during REL AM Operation

Using the approach described above, CUTR estimated the total number of turning movements for each vehicle to compute their false positive rates. Figure 7-33 shows the distribution of FP percentages. A total of four vehicles exhibited a FP share ranging from 19.2 to 25.0 percent. The average FP share is 5.5 percent as marked by the red line.



Source: CUTR, 2020

Figure 7-33. False Positive Share per Vehicle during REL AM Operation

7.2.3.1.3 Factors Associated with False Positives

7.2.3.1.3.1 OBU Last Known Vehicle Heading

During the AM REL operation, vehicles traveling westbound on the REL and exiting at the Twiggs Street intersection should report a vehicle heading of about 180 degrees from north. An investigation of the host vehicle BSM vehicle heading values revealed that when WWE warnings were issued, about 37.5 percent of the warnings (243) had a vehicle heading value of zero degrees.[10] This causes the WWE application to interpret the vehicle's trajectory as traveling northbound and in violation of the allowed movement, thus triggering a false positive. Figure 7-34 maps 882 WWE warnings using a reference color scale for vehicle heading values, with the red dots indicating vehicle headings equal to zero degrees.

[10] Note that for ease of interpretation, vehicle headings are reported in degrees from north. In the raw dataset, the BSM vehicle headings are recorded in conformity to March 2016 SAE J2735 standards and unit of measure (p.192).

Source: CUTR, 2020

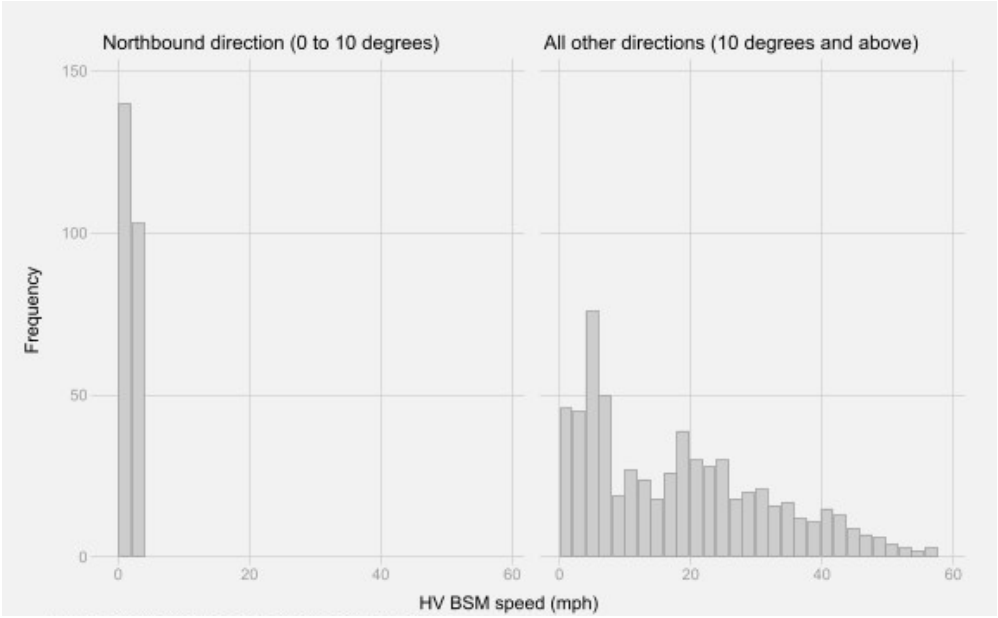
Figure 7-34. HV Heading Values of Southbound Vehicles – REL AM Operation

One of the potential causes leading the OBU to report zero heading values can be the traffic conditions characterizing the morning commute at the end of the REL ramp. During the weekday hours of 7 to 9 a.m., the end of the ramp tends to become heavily congested, with vehicles forming a queue that might extend several feet up the ramp. Figure 7-35 reports BSM point speed during this time, showing the majority of BSMs with a speed varying between zero and 10 miles per hour. This situation is described in greater detail in the performance evaluation assessment of Use Case 1 (Morning Backups).

Source: THEA CV Pilot Dashboard, 2020

Figure 7-35. REL Eastbound Travel Speed during AM Weekday Peak (7 to 9 a.m.)

As vehicles enter the queue and their speed slows to zero, it is possible that the OBU records the last known vehicle heading value as zero (i.e., northbound), thus triggering the WWE application. The analysis of host vehicle headings confirms this hypothesis. Figure 7-36 shows frequency histograms of HV heading versus HV speed, grouped by northbound (0 to 10 degrees from north) and all other headings (>10 degrees from north) for the 882 WWE warnings generated by vehicles traveling the REL westbound to Twiggs Street.



Source: CUTR, 2020

Figure 7-36. Vehicle Heading vs. Speed of Southbound Vehicles – REL AM Operation

7.2.3.1.3.2 OBU with Incorrect MAP Information

A second factor associated with false positives is related to the identification and communication of the allowed lanes via Map Data messages. Figure 7-37 overlays the WWE warnings in relation to the allowed lanes depicted in Figure 6-21, which provides MAP instructions for application deployment. The figure reveals that the OBU is incorrectly classifying an allowed lane as a violating one.

Source: CUTR, 2020

Figure 7-37. MAP Allowed Lanes and WWE Events – REL AM Operation

7.2.3.2 Analysis of PM Movements and WWE Events

During PM weekday travel, the REL operational direction serves commuter traffic flowing from downtown Tampa to east Brandon. Vehicles can access the REL via east and west Twiggs Street and via North Meridian Avenue. The WWE application is intended to warn drivers of illegal maneuvers against the REL operational directions as detailed in section 6.2.4. Table 7-20 reports the WWE counts by turning movement and Figure 7-38 maps their location.

Table 7-20. WWE Event Sample – REL PM Operation

Movement	WWE Warnings	Unique WWE Events
W. Twiggs St. to REL Eastbound	2,033	1,310
E. Twiggs to REL Eastbound	145	139
N. Meridian Ave to REL Eastbound	2,434	2,279
Other*	458	409
Total	**5,070**	**4,137**

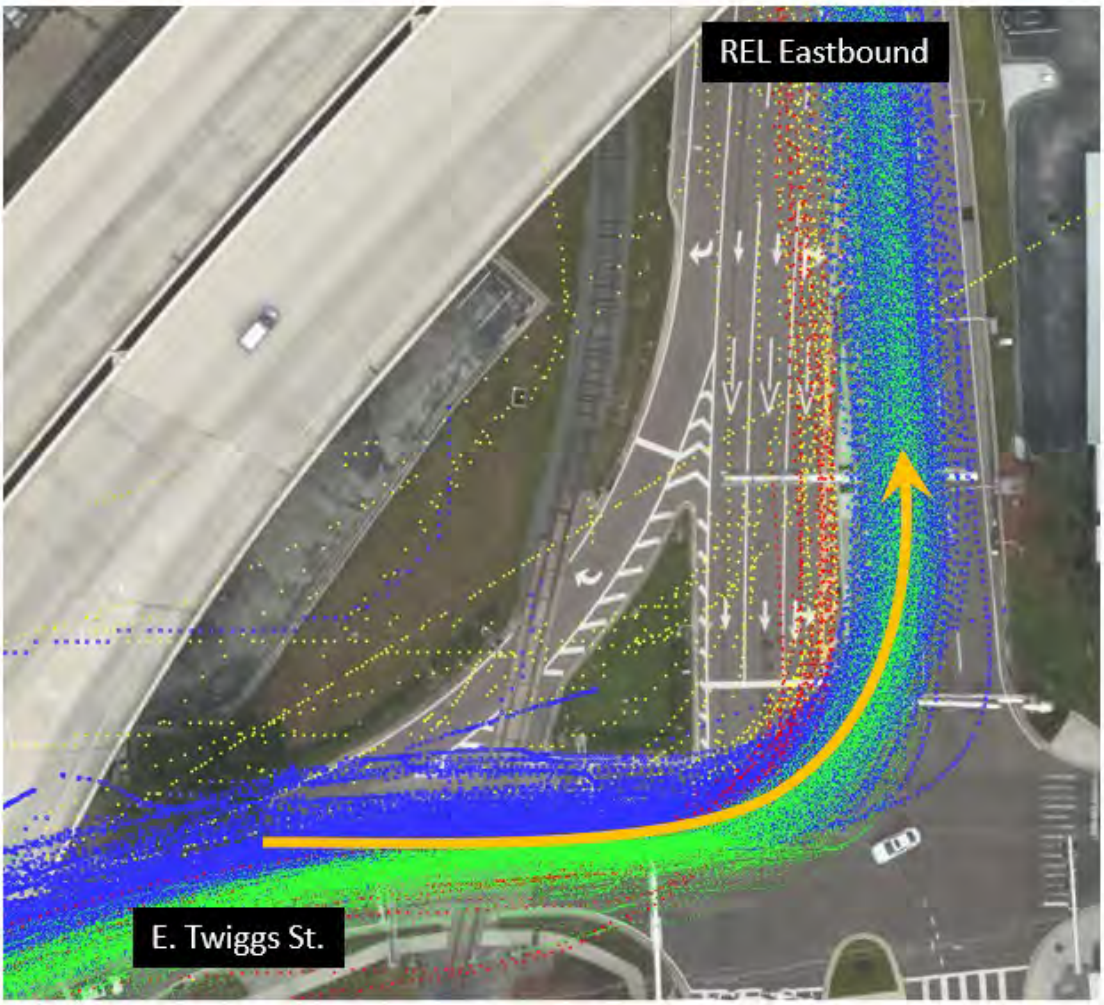
** Includes: Meridian to E. Twiggs, Meridian to W. Twiggs, E. Twiggs to Meridian, W. Twiggs to Meridian, E. Twiggs to W. Twiggs, W. Twiggs to E. Twiggs movements could not be assigned.*

Source: CUTR, 2020

Figure 7-38. WWE Warnings – REL PM Operation

7.2.3.2.1 Observed False Positives

Due to the complexity of turning movements and the WWE application design to capture vehicle trajectories in violation of the REL operational direction, CUTR established a procedure for the WWE event profile assessment to identify false and true positives. The procedure controls for additional confounding factors introduced by existing roadway infrastructure that can induce GPS signal drifting and affect vehicle positional accuracy. The infrastructure present is a bridge underpass that can cause GPS signal drifting to vehicles traveling from west Twiggs Street and making a left turn onto the REL, as shown in Figure 7-39.



Source: CUTR, 2020

Figure 7-39. Turning Movements from W. Twiggs St. to REL Eastbound

7.2.3.2.1.1 Profile Assessment of Movements from Twiggs Street to the REL Eastbound

To analyze the vehicle movements, a series of polygons were used to determine and interpret the vehicle trajectories associated with WWE warnings and relate them to the allowed turning movements. The shape of the polygons is empirically data driven by the RSU BSM data analysis and intended to identify turning movements in the presence of signal drifting.

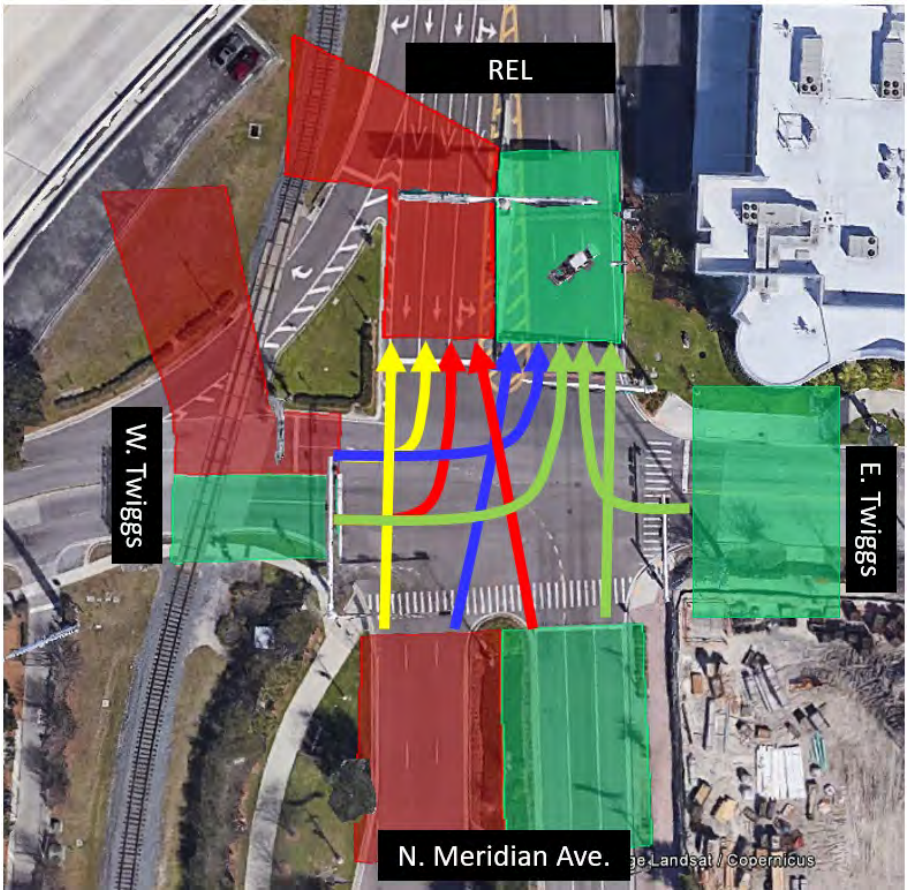
Figure 7-40 maps the polygons at the Twiggs Street intersection and uses four color-sequenced arrows to identify the following instances:

A. The green arrow identifies movements with no GPS signal drift on both the origin and end points of the turning movement, with the green color also indicating an allowed turning movement.

B. The blue arrow identifies movements that start with signal drift only at the origin of the turning movement.

C. The yellow arrow identifies movements where the vehicle is on the wrong side of the road at the beginning and end of the movement. These movements are potentially characterized by GPS signal drift both at the origin and end points of the turning movements (see Figure 7-39).

D. The red arrow identifies movements with no GPS signal drift at both the origin and end points of the turning movement, with the red color also indicating a turning violation.



Source: CUTR, 2020

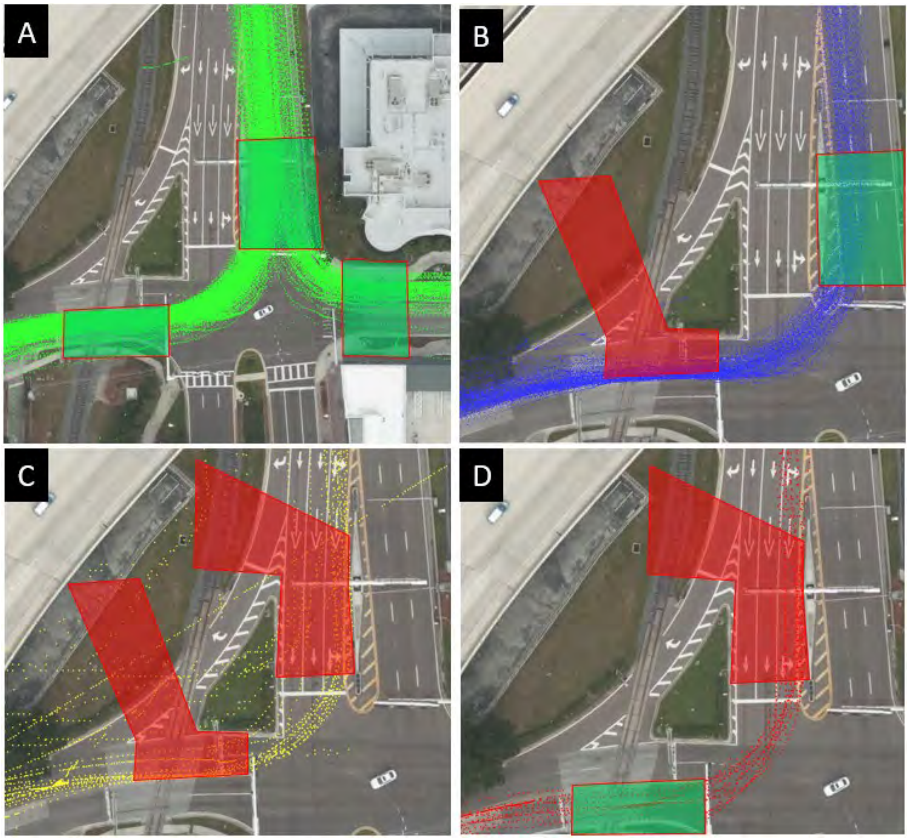
Figure 7-40. Polygons Identifying Movements from Twiggs St. and N. Meridian to REL Eastbound

Figure 7-41 shows all turning movements identified from East Twiggs Street to the REL using color codes consistent with the arrows of Figure 7-40 to identify the instances A through D, specifically as follows:

- **Figure 7-41.A –** These movements can be considered accurate because both their origin and end points do not exhibit any GPS signal drift. In addition, these are allowed movements during the PM REL operations.
- **Figure 7-41.B –** These are allowed turning movements with accurate GPS trajectories. While the movements exhibit some drift at the beginning of the turn, the movement end points are correctly positioned on the allowed lanes (i.e., entrance to the REL).
- **Figure 7-41.C –** These movements exhibit significant GPS signal drift at both the origin and end points. The full 60-second profile shows drifting. The WWE application interprets these turning movements as violations and issues a warning. For some of these movements, WWE warnings might

represent a true event. Some movements could have partially accurate trajectories but there is no mechanism to verify their true path (e.g., a video camera).

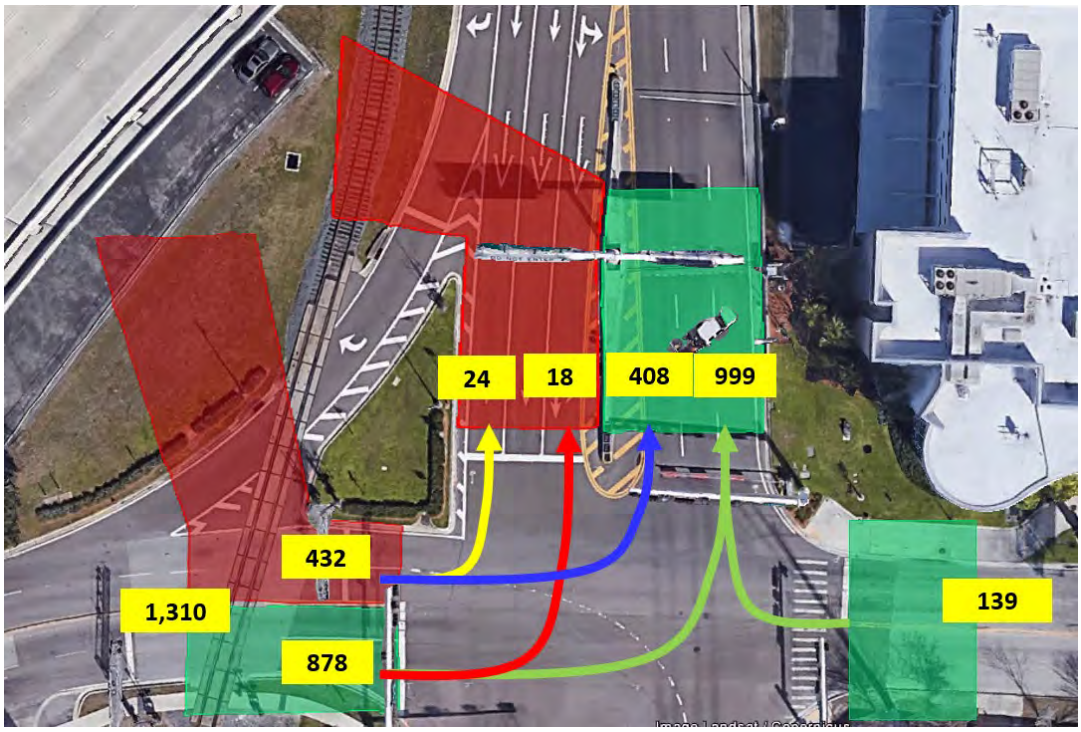
- **Figure 7-41.D –** These movements are considered accurate since they originate with no signal drift and end on the REL while causing a movement violation. Any WWE warnings triggered by these movements can be considered potentially true positives (i.e., vehicles entering the REL the wrong way).



Source: CUTR, 2020

Figure 7-41. Turning Movement Instances from Twiggs St. to REL Eastbound

The turning movement classification leads to identifying the unique WWE event associated with instances A through D, as reported in Figure 7-42. Based on the polygon-based approach, the analysis finds that out of a total of 1,310 unique WWE events, 18 can be considered potentially true positives and the remainder are false positives.



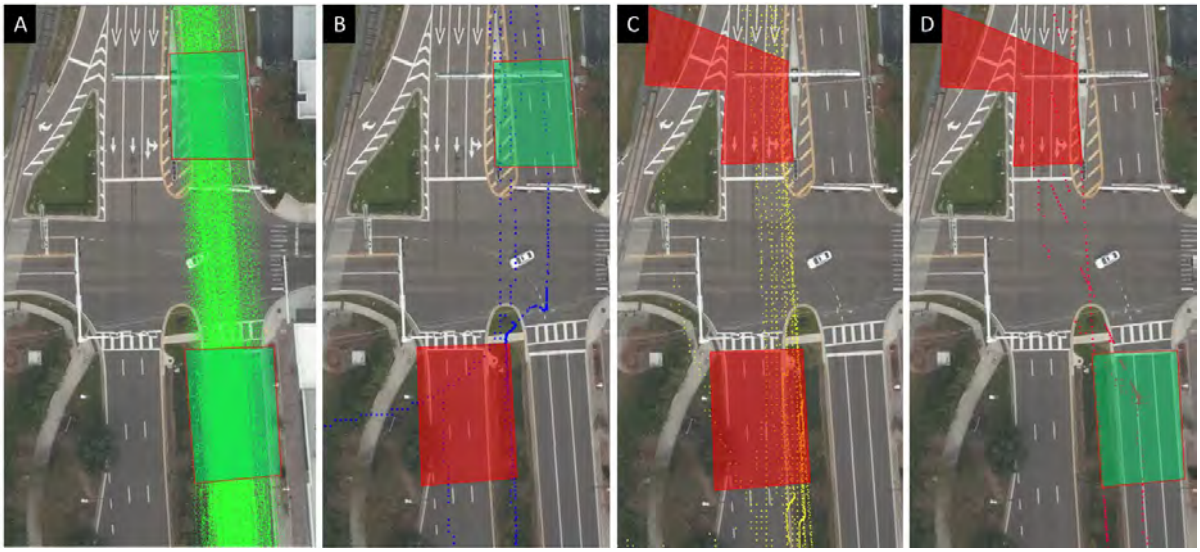
Source: CUTR, 2020

Figure 7-42. WWE Unique Events – Twiggs St. to REL Eastbound

7.2.3.2.1.2 Profile Assessment of Movements from North Meridian Avenue to REL Eastbound

The event profile assessment method of the previous section was applied to all warnings issued to vehicles traveling from North Meridian to the REL eastbound. Figure 7-43 maps and color codes the BSMs to identify turning movement instances A through D, specifically as follows:

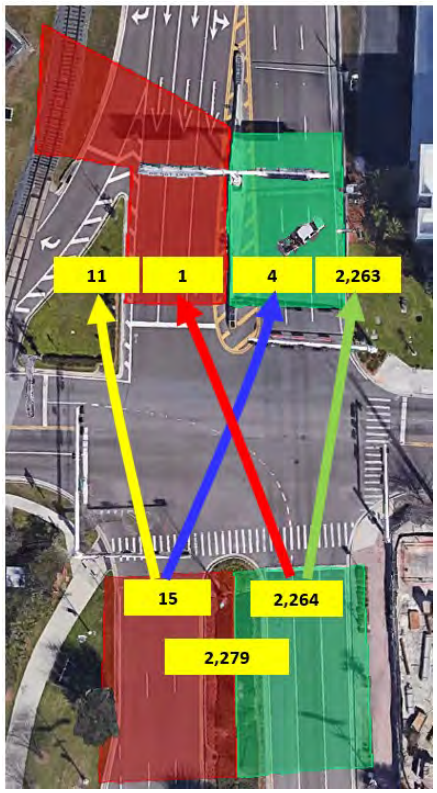
- **Figure 7-43.A –** These movements can be considered accurate because both their origin and end points are allowed movements and do not exhibit any GPS signal issue.
- **Figure 7-43.B –** These movements can be considered accurate because they start with some drift at the origin of the turning movement, but the end point is on the correct side of the road (i.e., entrance to the REL).
- **Figure 7-43.C –** These movements can be considered inaccurate since both the origin and end points exhibit GPS signal drift. Some could have partially accurate trajectories but there is no mechanism to verify their true path (e.g., a video camera). For some of these movements, WWE warnings might represent a true positive.
- **Figure 7-43.D –** These movements are considered accurate since they originate with no signal drift and end on the REL while causing a movement violation. Any WWE warnings triggered by these movements can be considered potentially true positives (i.e., vehicles entering the REL the wrong way).



Source: CUTR, 2020

Figure 7-43. Turning Movement Instances from Meridian Ave. to REL Eastbound

Figure 7-44 shows the analysis results of all the WWE warnings for this movement and the unique WWE events associated with instances A through D.



Source: CUTR, 2020

Figure 7-44. WWE Unique Events – Meridian Ave. to REL Eastbound

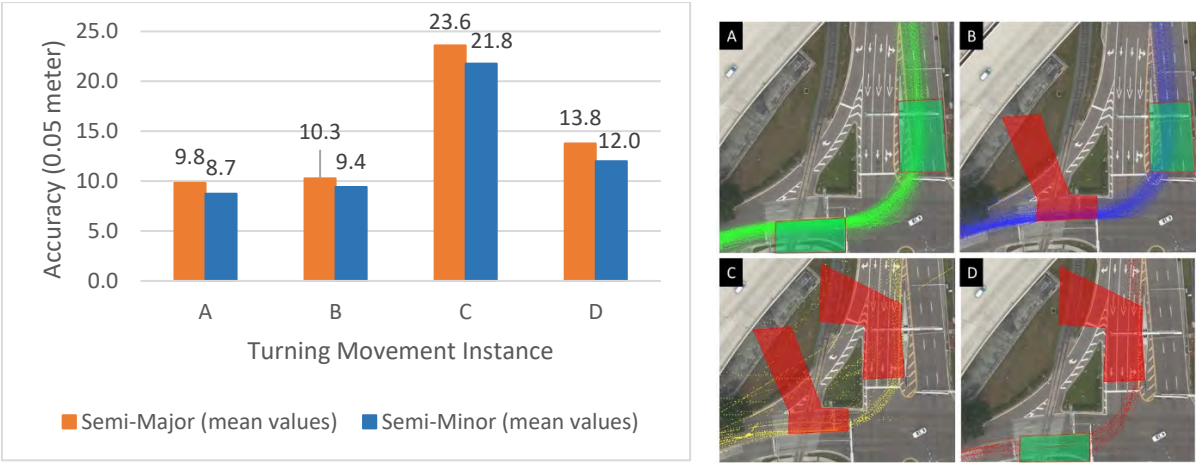
7.2.3.2.2 Factors Associated with False Positives

7.2.3.2.2.1 GPS Signal Inaccuracy

Vehicle positional accuracy is critical when implementing connected vehicle technology and using GPS for safety applications. Data elements within the BSM sent through the event profile provide information to assess vehicle positional accuracy.

Figure 7-45 shows the mean values of semi-major[11] and semi-minor[12] axis accuracy variables from the before-after one-minute profiles of the WWE warnings issued at the East Twiggs Street turn onto the REL. Accuracy is measured in units of 0.05 meters, with a value of 254 indicating an axis length of 12.70 meters or greater and a value of 255 indicating the axis value is unavailable.

The figure shows that the turning movement instances C and D, which lead to WWE warnings due to the application interpreting vehicle trajectories as REL movement violations, are characterized by higher semi-major and semi-minor values. Instance C shows the highest GPS inaccuracy, which is confirmed by BSM drifting at the beginning and end points of the turning movements. Therefore, without a validation of the vehicle trajectory (i.e. onboard video), the trajectories of these instances cannot be verified and labeled as TPs.



Source: CUTR, 2020

Figure 7-45. Vehicle Positional Accuracy – REL PM Operation W. Twiggs St. Turning Movements

7.2.3.2.2.2 WWE Warning Sequence

As described previously, the application issues multiple warning levels. The data do not distinguish between the different levels and use one flag for all warning types. However, analyzing the sequence of warnings

[11] Data element used to express the radius (length) of the semi-major axis of an ellipsoid representing the accuracy that can be expected from a Global Navigation Satellite System (GNSS) in 5-cm steps, typically at a one sigma level of confidence (SAE J2735).

[12] Data element used to express the radius of the semi-minor axis of an ellipsoid representing the accuracy that can be expected from a GNSS in 5-cm steps, typically at a one sigma level of confidence (SAE J2735).

makes it possible to infer if the warnings are correctly issued in terms of per-spec sequence. In particular, the first “Do Not Enter” warning is based on vehicle path estimation. It is possible that if the OBUs are unable to correctly identify the vehicle path, they issue “Do Not Enter” warnings at a high rate. To test this hypothesis, an analysis of the warning sequence was conducted to identify movements where only one warning (most likely “Do Not Enter”) was issued to the vehicle.

Table 7-21 shows the number of warnings issued during a given WWE event characterized by warnings issued by the same vehicle within one minute of each other. The results show that there was only one potentially true positive event with one warning, but there were 3,950 false positives (81.9% of total events) with only one warning. This illustrates a potential failure of the “pre-warning” due to the wrong identification of a vehicle’s true path.

Table 7-21. WWE Sequence Analysis

Warnings Issued per WWE Event	Total Number of WWE Events	Potentially True Events		False Positives	
		Count	Share (%)	Count	Share (%)
1	3,951	1	0.0	3,950	81.9
2	685	5	0.1	680	14.1
3	126	6	0.1	120	2.5
4	42	4	0.1	38	0.8
5	14	3	0.1	11	0.2
6	3	0	0.0	3	0.1
7	3	0	0.0	3	0.1
Total	**4,824**	**19**	**0.4**	**4,805**	**99.6**

To further verify this hypothesis, the warnings were mapped and color coded by sequence. Figure 7-46 shows the warnings at the intersection. During the warning sequence, the first warning (presumably “Do Not Enter”) should be issued before the vehicle enters the REL to travel eastbound. As shown in Figure 7-46, the green color warnings are not all located south of the REL. Approximately 90 percent of single warnings were issued before the vehicle entered the REL, which is a good indication that these were in fact “Do Not Enter” warnings.

Source: CUTR, 2020

Figure 7-46. Map of WWE Warnings by Sequence

7.2.3.2.3 False Positive Rate

Using the approach detailed in Chapter 6, the analysis considered all the turning movements made by participants during the analysis time frame traveling through the WWE application's deployment area and having OBUs with data logging capability enabled. Table 7-22 reports the complete count of turning movements and the appropriate WWE event classification. The overall false positive rate during PM REL operation is 35.7 percent.

Table 7-22. WWE Classification – REL PM Operation

Description	Turning Movements	WWE Warnings	Unique WWE Events	True Positive Events**	False Positive Events	True Negative Events	False Negative Events**
W. Twiggs St. to REL Eastbound	4,169	2,033	1,310	18	1,292	2,303	6
E. Twiggs to REL Eastbound	704	145	139	0	139	359	0
Meridian Ave to REL Eastbound	7,499	2,434	2,279	1	2,278	3,645	1
Other*	2,265	458	409	0	409	1,114	0
Total	**14,637**	**5,070**	**4,137**	**19**	**4,118**	**7,421**	**7**

** Includes: Meridian to E. Twiggs, Meridian to W. Twiggs, E. Twiggs to Meridian, W. Twiggs to Meridian, E. Twiggs to W. Twiggs, W. Twiggs to E. Twiggs and movements that could not be assigned.*
*** A complete determination of TP cannot be made. FN determination is affected as a result.*

The analysis also identified 19 events classified as *potentially* true positives (PTPs), or events that present the characteristics of a true positive but for which a final TP determination cannot be made. These are discussed in detail in the following section. The proportion of PTP events with respect to all unique WWE events is 0.3 percent.

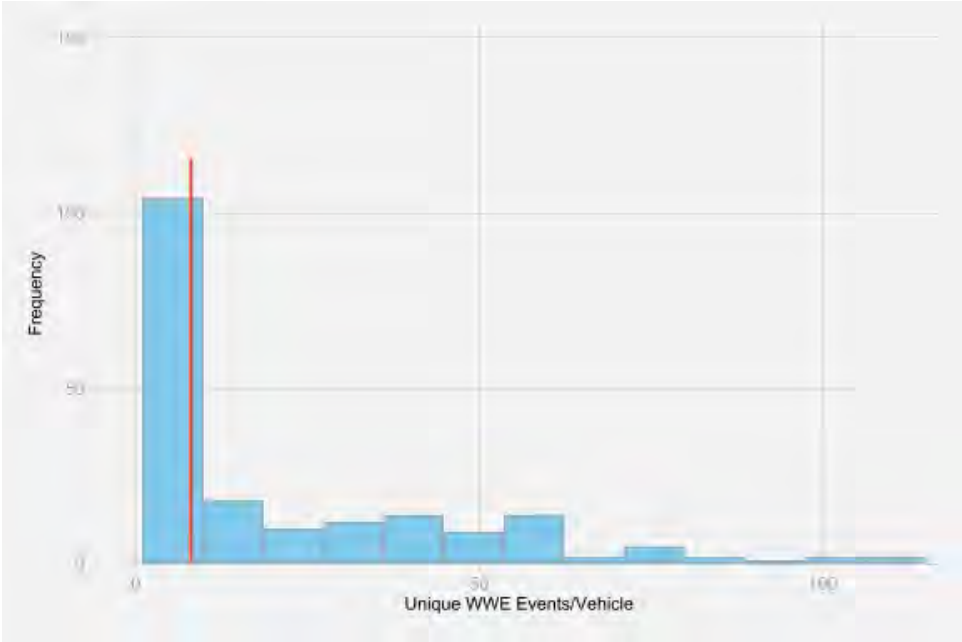
Table 7-23 reports the number of WWE unique events, turning movements, their classifications as FP and PTP, as well as their rates. For all the movements comprising the PM REL profile, the FP rate is calculated at 28.1 percent.

Table 7-23. WWE Movement Classification and Rates – REL PM Westbound

Description	Count	Rate (%)
Unique Vehicles	459	--
Unique Vehicles Deploying Warnings	201	--
WWE Events (PTP + FP)	4,137	--
Turning Movement	14,637	--
Conflicts	30	--
Potentially True Positives	19	63.3
False Negatives	11	36.7
Non-conflicts	14,637	--
True Negatives	10,519	71.9
False Positives	4,118	28.1

7.2.3.2.4 False Positive Rate by Vehicle

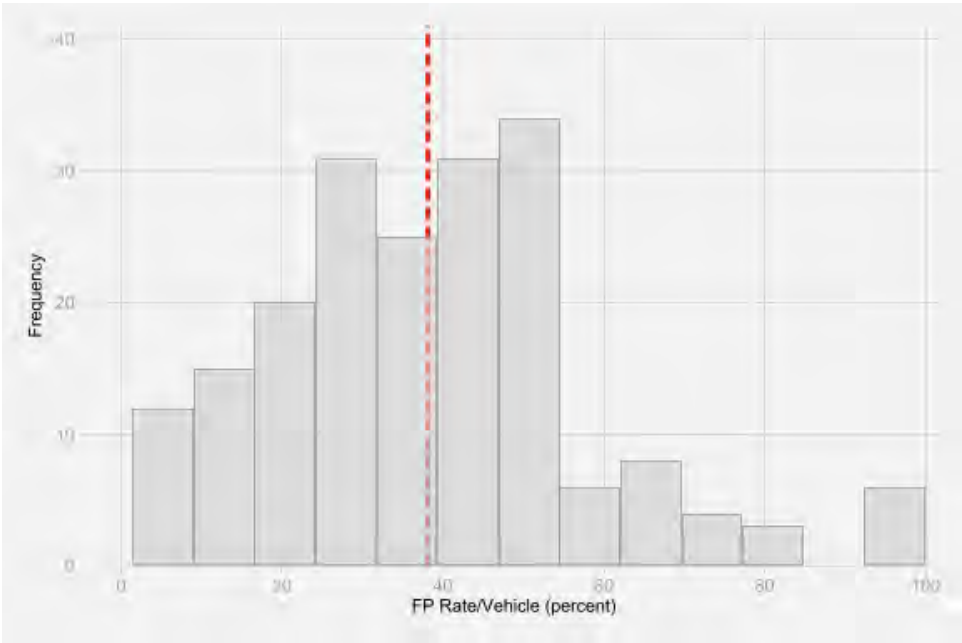
During the REL PM operation, 201 vehicles generated 4,137 unique WWE events. Figure 7-47 shows the distribution of unique WWE events per vehicle. During the analysis time frame, some participants traveled more often than others in and out of the study area, thus increasing the total number of warnings generated. While 50 percent of the vehicles issued 8 unique WWE events or less, four vehicles generated more than 100 unique WWE events.



Source: CUTR, 2020

Figure 7-47. Distribution of WWE Unique Events per Vehicle

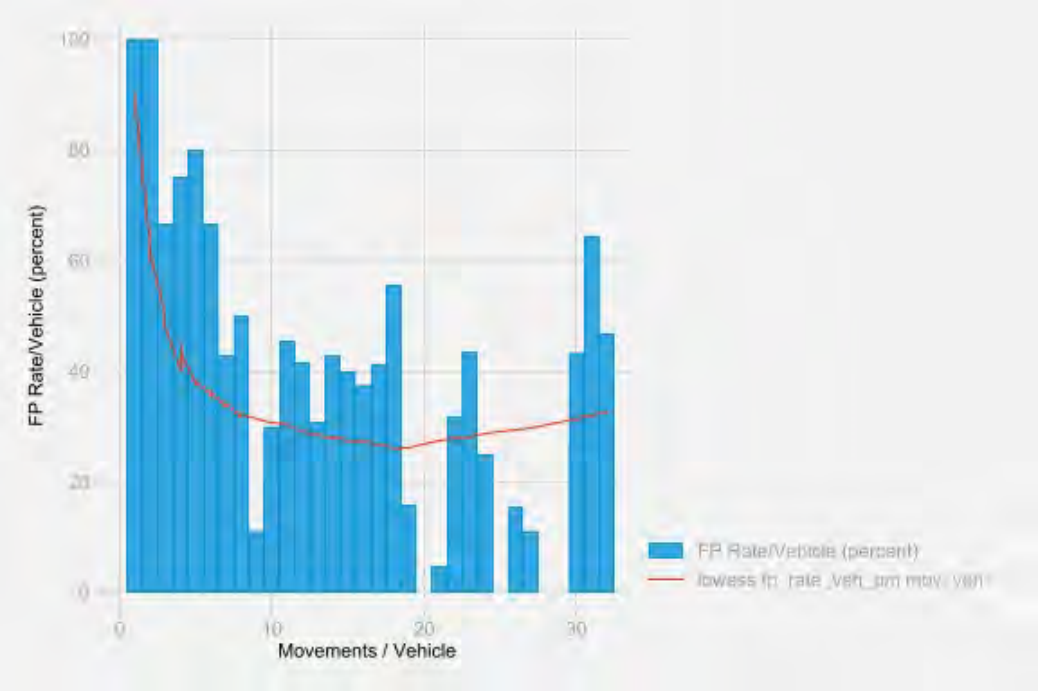
The heterogeneity in the amount of travel done over the study time frame can contribute to a high variability in FP rates across the participants. Figure 7-48 shows the distribution of false positive rates. The average FP share is 38.1 percent, as marked by the red line. Four vehicles exhibited a FP share of 100 percent and these vehicles made only one movement each during the study time frame.



Source: CUTR, 2020

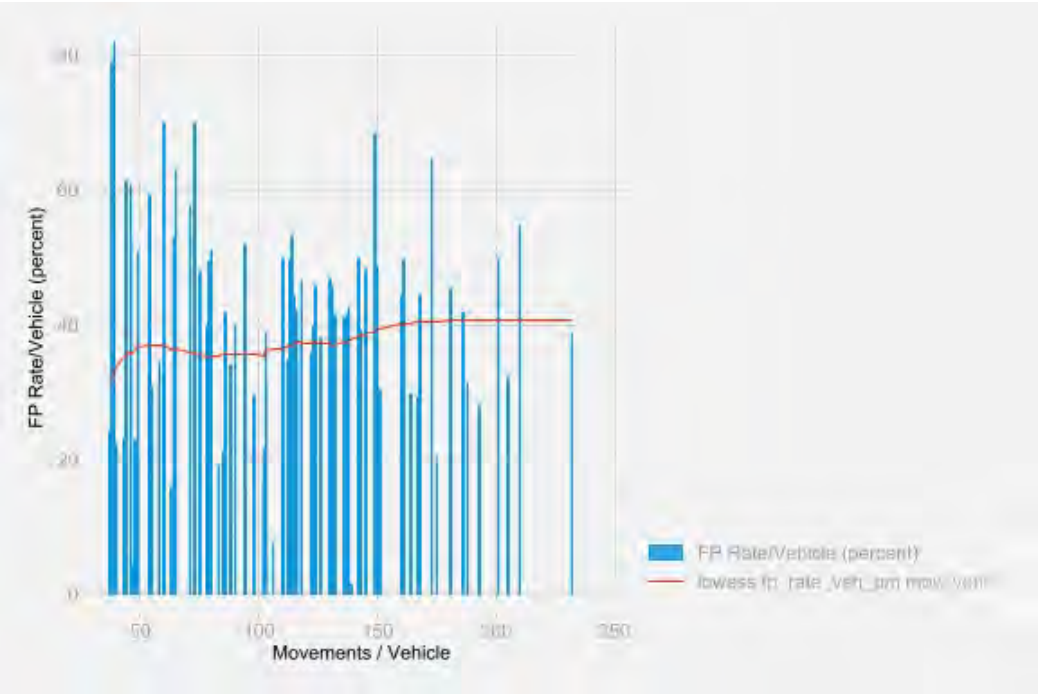
Figure 7-48. FP Share Per Vehicle during REL PM Operation

Figure 7-49 and Figure 7-50 overlay the false positive rate distribution over the number of turning movements performed during the study time frame, distinguishing between participants who made less than the average number of movements and those who made more than the average number of movements. The red line represents a nonlinear estimate of the FP rate with respect to turning movements performed. The figures show that participants making less than 10 turning movements over the study time frame have a higher false positive rate than those making more movements than the average (i.e., traveling more often through the study area).



Source: CUTR, 2020

Figure 7-49. FP Rate per Vehicle – Participants with 36 or Less Movements



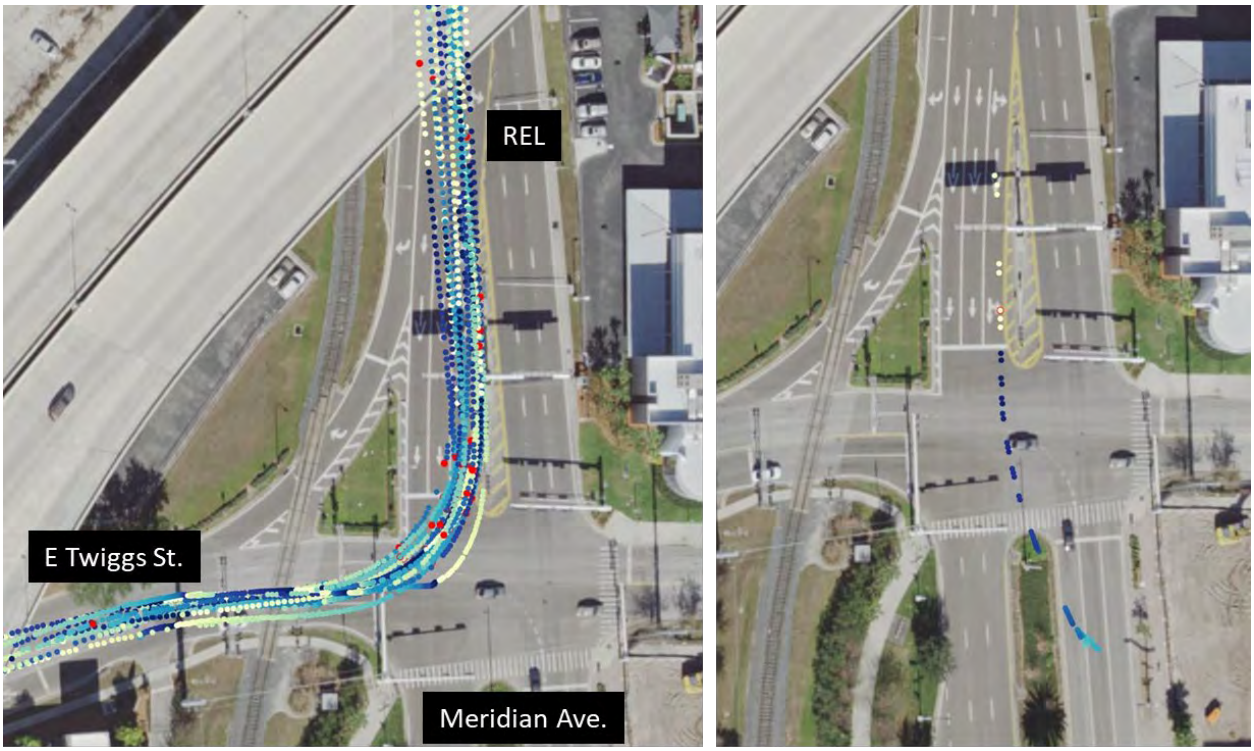
Source: CUTR, 2020

Figure 7-50. FP Rate per Vehicle – Participants with More than 36 Movements

7.2.3.2.5 Potentially True Positives

During the PM REL operation, 56 warnings were flagged as potentially true positives (PTPs). These warnings passed a visual check where the vehicle trajectory was confirmed through the wrong-way lanes into the REL. These warnings were issued to 14 individual vehicles and represent 19 unique events where a vehicle received one or more sequential WWE warnings within one minute of each other.

Figure 7-51 shows the trajectories of vehicles passing through the left side of the entrance to the REL. Fifty-four warnings (96%) were issued by vehicles making a left turn from East Twiggs Street (Figure 7-51, left) and two were issued to a vehicle traveling on North Meridian Avenue (Figure 7-51, right).



Source: CUTR, 2020

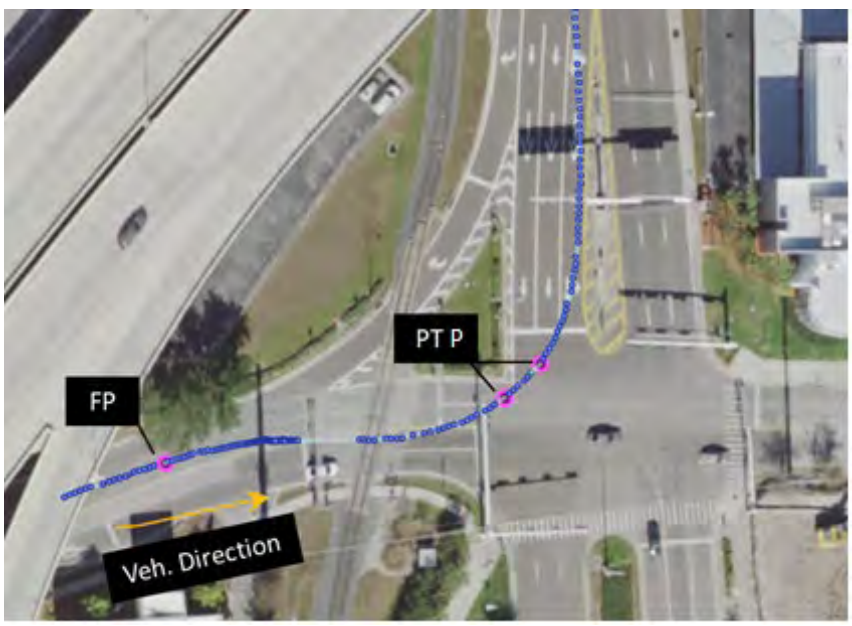
Figure 7-51. Turning Movements of Potentially True Positives

Table 7-24 lists the 19 unique events and warning sequence, sorted by descending number of PTPs. As shown in the table, the first warning for event numbers 8, 12, 13, and 19 was a false positive, but the sequential warnings were potentially true positives. This is because warnings were evaluated separately and not as part of a sequence. For these events, the vehicle was traveling eastbound on East Twiggs Street and far from the REL entrance. The four FPs were most likely “Do Not Enter” warnings as described in the previous section.

Figure 7-52 shows one instance where a vehicle received the first warning on East Twiggs Street and it was determined to be FP. Then the vehicle turned left on the REL in seemingly the wrong way, thus it received two WWE warnings that were determined to be PTP based on the methodology adopted in this study. The label of FP was only applied because in examining the warning at the moment issued, the driver was not showing an intention of entering the REL the wrong way.

Table 7-24. Potentially True Positive (PTP) Warning Sequence

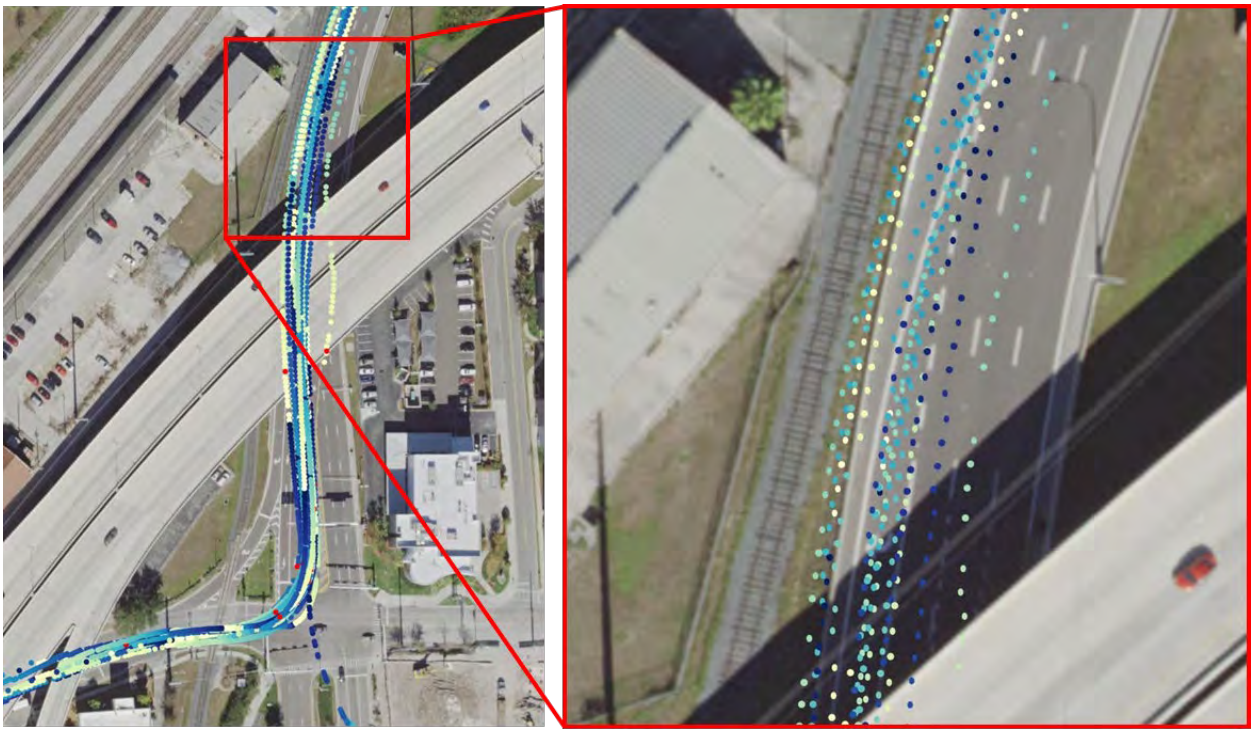
Unique WWE Event #	Warning Sequence					No. PTP Warnings
	1	2	3	4	5	
1	PTP	PTP	PTP	PTP	PTP	5
2	PTP	PTP	PTP	PTP	PTP	5
3	PTP	PTP	PTP	PTP	PTP	5
4	PTP	PTP	PTP	PTP		4
5	PTP	PTP	PTP	PTP		4
6	PTP	PTP	PTP	PTP		4
7	PTP	PTP	PTP			3
8	FP	PTP	PTP	PTP		3
9	PTP	PTP	PTP			3
10	PTP	PTP	PTP			3
11	PTP	PTP	PTP			3
12	FP	PTP	PTP			2
13	FP	PTP	PTP			2
14	PTP	PTP				2
15	PTP	PTP				2
16	PTP	PTP				2
17	PTP	PTP				2
18	PTP					1
19	FP	PTP				1



Source: CUTR, 2020

Figure 7-52. WWE Event Comprising of FP and PTP Warnings

Further analysis of the PTP turning movements of vehicles traveling north on the REL and beyond the underpass shows evidence of GPS drift as previously described. A closer examination of the trajectories of these events reveals that even though the vehicles' position seems accurate at the intersection, the BSMs exhibit GPS drifting when they travel beyond the underpass. Figure 7-53 shows the PTP trajectories and the inset map shows a close-up on the REL where the trajectories seem to be shifted to the left side of the road, and some are even outside its boundaries.



Source: CUTR, 2020

Figure 7-53. GPS Shift of PTP Trajectories – REL PM Operations

In addition, the two PTPs associated with the vehicle traveling north on Meridian Avenue show an irregular pattern that suggests GPS shift (perhaps due to loss of signal) while the vehicle is stopped at the intersection, potentially at a red signal.

Out of the 14 vehicles that experienced a PTP event, only 1 was assigned to the treatment group. This vehicle was issued five sequential WWE warnings classified as one PTP event. The onboard unit's HMI was disabled during this event, so no data exist to assess the driver's reaction. For all vehicles that potentially traveled the wrong way, no vehicle was observed to make a U-turn to change direction since the REL direction was eastbound in response to a dangerous situation.

In summary, this analysis categorizes these WWE unique events as *potentially* true positives because there is no alternative method to validate that they are true positives. Table 7-25 summarizes the PTPs along with unique events and vehicles. These findings and the evidence of vehicle trajectories further suggest that some of the unique WWE events flagged as PTP based on trajectory assessment may not have been flagged correctly. In other words, some of these warnings might be position tracking errors and some might be false positives.

Table 7-25. PTP Warning Breakdown

Description	Turning Movements	Sequential WWE (PTP)	Unique WWE Events	Unique Vehicles
Twiggs to REL Eastbound	24	54	18	13
Meridian to REL Eastbound	2	2	1	1
Total	**26**	**56**	**19**	**14**

7.2.4 Summary of Findings

The results of the Wrong-Way Entry application deployment were compared with a before period (i.e., before the WWE application was implemented) using the Safety Analysis Methodology discussed in Chapter 6, RSU BSMs collected between May 1, 2018, and February 28, 2019, and the AM and PM breakdown consistent with REL operations.

7.2.4.1 AM REL Operations

During the morning REL operations, between 6 and 10 a.m., the allowed movement is westbound toward East Twiggs Street and Meridian Avenue. Any OBU-equipped vehicle traveling on the REL and egressing at the Twiggs intersection should not be issuing WWE warnings. A total of 906 warnings were issued by 127 vehicles, resulting in 687 unique WWE events. All these events were characterized as false positives, or instances where there was no conflict. Over the same period, no true positives (wrong-way intrusions) were recorded, resulting in zero false negatives and a false negative rate of zero.

The overall false positive rate during the AM REL operational profile is 2.8 percent. The analysis identifies incorrect MAP information and the aftermarket OBU's ability to predict vehicle trajectories in the presence of traffic congestion as the two main causes of false positives. When considering all the movements performed by participants who received the warnings, the analysis finds substantial heterogeneity in false positive rates.

During the before period, no conflicts (WWE movements) were recorded using RSU BSMs collected from 482 vehicles performing 21,957 movements, as shown in Table 7-26.

Table 7-26. WWE Before-After Movement and Conflict Comparison – AM

Description	Before Period		After Period	
	Count	Share (%)	Count	Share (%)
Unique Vehicles	482		475	
Turning Movements	21,957		24,380	
Conflicts	0	0	0	0
Non-conflicts	21,957	100	24,380	100

7.2.4.2 PM REL Operations

During afternoon/evening weekday travel, between 3 p.m. and 6 a.m., the REL direction is eastbound to serve commuter traffic flowing from downtown Tampa east to Brandon. Vehicles can access the REL via East Twiggs Street (eastbound and westbound) and via North Meridian Avenue. The WWE application is intended to warn drivers of illegal maneuvers against the REL operational direction. A total of 5,070 warnings were issued by

201 vehicles, resulting in 4,137 unique WWE events, the majority issued to vehicles traveling on North Meridian Avenue toward the REL (55.1%).

A detailed examination of all turning movements identified 19 potentially true positives, or instances of a warning issued for a potential conflict or violation of the allowed entry movement in the study area. All these instances occurred during the PM REL operation. The analysis shows inconclusive evidence in identifying false negatives, or instances where the application failed to deploy when a conflict was present.

The overall false positive rate is 28.1 percent, with some turning movements leading to higher false positive rates. While GPS signal accuracy is present, it is not the main contributor to these rates. Only about 1.8 percent of the turning movements from East Twiggs Street to the eastbound REL and 0.5 percent of turning movements from North Meridian Avenue to the eastbound REL exhibited signal drift.

The main contributor to the false positives is the triggering of the first warning in the WWE warning sequence, or the "pre-warning." Although there is no method to truly establish if vehicles receiving only one WWE warning received the "Do Not Enter" pre-warning, the analysis suggests that for approximately 90 percent of cases the first warning was issued before the vehicle entered the MAP-defined "no entry zone," thereby pointing to a challenge in determining the correct vehicle trajectory.

Comparing with the before period as defined above, more conflicts were identified in the after period (most due to the increased number of participants and greater number of movements). The overall rate of conflicts per turning movement is 0.2 percent, compared to the before period with 0.1 percent. Table 7-27 shows that compared to the before period, the after period recorded more vehicles, turning movements, and conflicts (potentially true positive WWEs) as defined in the detailed analysis.

Table 7-27. WWE Before-After Movement and Conflict Comparison – PM

Description	Before Period		After Period	
	Count	Share (%)	Count	Share (%)
Unique Vehicles	411	--	459	--
Turning Movements	9,420	--	14,637	--
Conflicts	10	0.1	30	0.2
Non-conflicts	9,410	99.9	14,607	99.8

7.2.4.3 Lessons Learned

Findings indicate that GPS signal inaccuracy, while being a contributor to potentially false positive warnings (i.e., WWE triggered when drivers traveled on allowed side of the road), is not the main cause of the high false positive rates generated by the WWE application. Overall, it can be concluded that:

1. Even though the pre-warning that triggered the "Do Not Enter" warning was designed to warn drivers not to enter the wrong way before they did, it was executed in a manner that needs more fine tuning. This warning depends on estimating the vehicle's trajectory well in advance of the intersection, leading to many FP events since the vehicles would not actually enter the wrong way.

2. The loss of heading on the OBU seems to be one of the culprits behind the FPs in the AM period. The characteristics of the urban environment and overpasses when vehicles traveled under them created GPS signal loss, thus rendering the vehicle's heading to report zero degrees (which means

northbound direction thus triggering the WWE app). The recommendation is to use the last known heading or improve the algorithm and use other data to ensure that the OBU maintains its location and heading even if there is a loss of GPS signal.

7.3 Use Case 3: Pedestrian Conflicts

7.3.1 Analysis Dataset

The analysis uses data collected from OBU data logs for two periods: (1) Participant vehicles traveling between March 1, 2019, and October 31, 2019, and (2) THEA CV Pilot test vehicles conducting tests between June 1, 2020, and August 31, 2020.

Due to system operational issues described in Chapter 6, the system underwent a change of pedestrian detection technology after October 2019. The new thermal sensor-based system became operational on August 5, 2020. At the time of this report and due to the COVID-19 pandemic, no Pedestrian Collision Warning data have been collected from participant vehicles. The pandemic negatively impacted driving behavior and pedestrian traffic in the study area. In summary, the analysis uses the following data:

1. Warning event data collected and stored in OBU data logs.
2. BSM data collected and stored in OBU data logs.
3. Pedestrian Safety Message (PSM) data generated by RSU No. 3 using the pedestrian identification algorithms from:
    a. LiDAR sensors for the period prior to August 5, 2020
    b. Thermal camera sensors for the period after August 5, 2020.

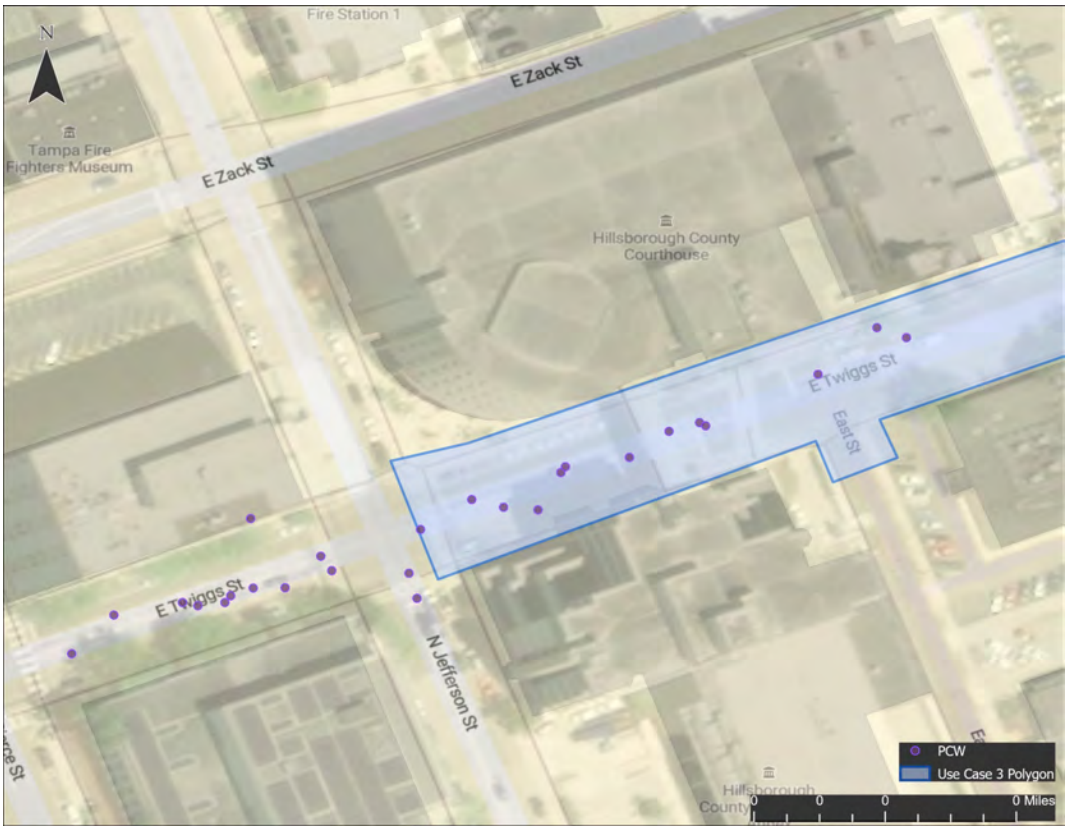
7.3.2 Mobility Impact

The PMESP did not consider analyzing the impact on mobility.

7.3.3 Safety Impact

7.3.3.1 Observed False Positives under the LiDAR System (Old System)

Although the LiDAR system failed to accurately identify pedestrians, it intermittently operated during Phase 3 and issued 27 PCWs to participant vehicles between March 1 and October 31, 2019. Figure 7-54 maps the warning events. These Pedestrian Collision Warnings were manually visually inspected (instead of the automated process used for other applications) because during this time, the PCW operational and configuration parameter were set to test mode.



Source: CUTR, 2020

Figure 7-54. Map of PCW Events with LiDAR

Of the 27 PCWs, 4 were classified as true positive. Table 7-28 summarizes the results of the visual inspection. In summary, 85 percent of the triggered PCWs with the LiDAR system were false positive and 15 percent were true positive.

Table 7-28. PCW Analysis of LiDAR System – False and True Positives

Description	Count	Share (%)	Test Performed
PCW (TP + FP)	27	--	--
False Positives	23	85.2	Visual Inspection
True Positives	4	14.8	Visual Inspection

7.3.3.1.1 Factors Associated with False Positives with LiDAR

7.3.3.1.1.1 Large Distance between HV and Pedestrian

One cause of FPs was that the warning was issued when the distance between the HV and the pedestrian was large (greater than 70 meters), and therefore no conflict was present. Figure 7-55 shows an example of a false positive where the HV is on a side street and has yet to turn onto Twiggs Street. In this case, there is no conflict between the vehicle and the pedestrian crossing. This, however, was due to a test parameter that used 300 meters as the reference distance instead of the 70 meters used later under the thermal cameras system.

Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-55. FP PCW Due to Large Distance and Different Road

Figure 7-56 shows another example of a large distance between the HV and the pedestrian (greater than 70 meters). No conflict exists between the vehicle and the pedestrian.

Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-56. FP PCW Due to Large Distance

7.3.3.1.1.2 Pedestrian Standing on Sidewalk

Another cause of false positives was the system's inability to correctly identify pedestrians who had the intention of crossing at the crosswalk. Figure 7-57 shows an example of a FP due to the pedestrian being on the sidewalk but not crossing at the crosswalk. Warnings triggered during this situation were labeled as FP events since there was no conflict between the HVs and the pedestrians.

Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-57. FP PCW Due to Pedestrian on the Sidewalk

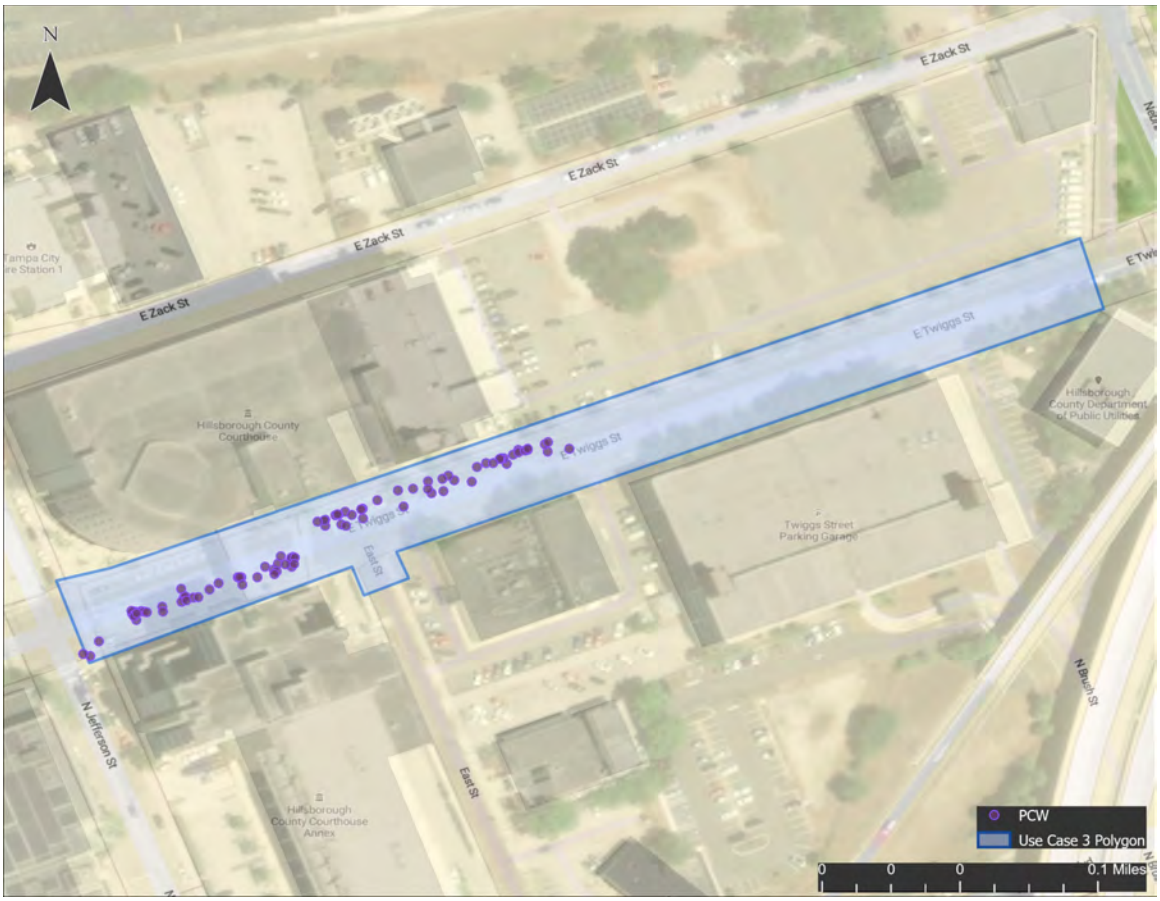
7.3.3.2 Observed False Positives with Thermal Camera (Current System)

Prior to entering full operation on August 5, 2020, the thermal camera system was subject to several tests to ensure functionality and data generation meeting the Phase 2 system design and system architecture specification. The data generated were subject to analysis for final approval to deploy. The following results are based on the analysis of the test datasets.

As shown in Figure 7-58, test vehicles successfully triggered 87 PCWs and generated the relevant warning moment events and Pedestrian Safety Messages conducive to assessment. Using the PCE approach adopted to evaluate other CV applications, the research team was able to automate the event profile generation associated with the PCW applications. As participants generate warnings, these events can be replayed and analyzed via the THEA CV Pilot Dashboard as a precursory step to data-driven evaluation.

Parameter Conformity Evaluation is the first step of the evaluation, where each warning event is analyzed and checked for conformity with respect to the default application's operational parameters shown in Table 6-5.

The PCE analysis of PCW events flagged 62 warnings as FP and 17 as potentially true positives (PTP), while PCE could not be performed on 8 events due to missing PSM data. The next step was a visual inspection of the PCWs identified as potentially true positives. Out of 17 PTPs that were visually inspected, 1 warning failed and was thus flagged as FP. Table 7-29 summarizes the results of the warning analysis via PCE and then via visual inspection on PTPs. In summary, 72.4 percent of triggered PCWs were FP and 18.4 percent were TP.



Source: CUTR, 2020

Figure 7-58. Map of PCW Events with Thermal Camera

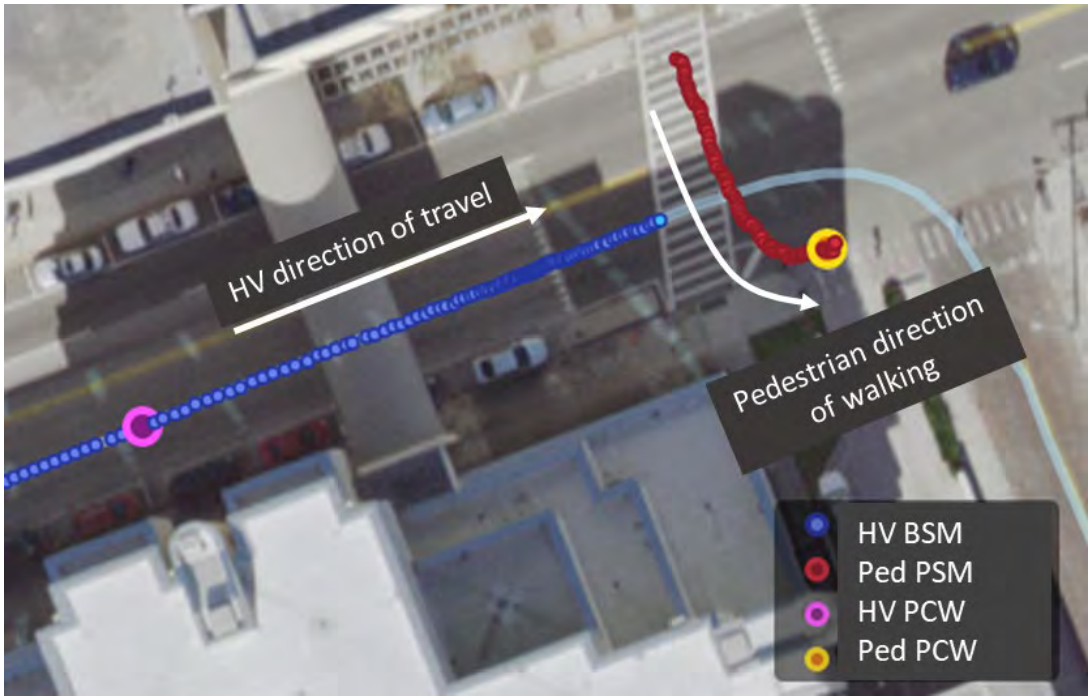
Table 7-29. PCW Analysis of Thermal Camera – False and True Positives

Description	Count	Share (%)	Test Performed
PCW (TP + FP + Not Tested)	87	--	--
Not Tested	8	--	--
False Positives	63	72.4	PCE Flagged 62 PCWs as FP. Visual Inspection of PTPs Changed 1 PTP to FP.
True Positives	16	18.4	PCE & Visual Inspection

7.3.3.2.1 Factors Associated with False Positives with Thermal Camera

The PCE flagged non-dangerous encounters of HVs and pedestrians as FP. However, a warning passing PCE is not guaranteed to be a TP and requires visual inspection. Vehicle trajectory inspection was conducted to determine if the HV and pedestrian were in the intended paths according to the specs defined in the System Design Document for the PCW to trigger warnings [8]. Even though the application relies on operational parameters to determine if the HV and pedestrian are in a collision trajectory, the visual inspection shows that

out of 17 warnings flagged as conflicts, only 1 was in fact a FP. Figure 7-59 shows the HV path of trave toward the crosswalk while the pedestrian (i.e., a researcher conducting the test) is standing at the corner of the curb after having crossed the road.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-59. PTP PCW Changed to FP (Non-conflict between HV and Pedestrian)

7.3.4 Summary of Findings

The initially deployed LiDAR-based Pedestrian Collision Warning system faced several deployment challenges resulting in reliability issues and failure to meet the required deployment specifications. The system integrator replaced the system with a thermal sensor to accurately detect and track pedestrians. After testing, the new system became operational on August 5, 2020. During the operational time of the LiDAR sensors, the PCW application triggered 27 warnings that consisted of 85 percent FPs due to the sensors' inability to correctly identify pedestrians and triggering warnings at large distances between the HV and pedestrian. The large distance was due to the loosened operational parameters of the system at the time.

The change of sensors from LiDAR to thermal camera shows an overall improvement in pedestrian identification and tracking. During testing scenarios, the system was able to correctly identify pedestrians on the sidewalk and not trigger warnings, and it was able to correctly identify pedestrians on the crosswalk and trigger warnings as intended. The test data cannot be used as an overall reliability of the system as several scenarios were purposefully testing the operation of the new sensors.

Due to the COVID-19 pandemic that began in March 2020 and its impact on the participants' travel in the area, no PCW warning data have been recorded from participant vehicles at the time of this report. The new system became officially operational on August 5, 2020. Further data collection in subsequent months can provide information as to the effectiveness of the PCW application.

Use Case 3 was a very anticipated application of advanced systems that utilized LiDAR sensors to detect and track pedestrians at a crosswalk. The use of LiDAR has proven successful in other contexts but seems to have failed to provide the necessary consistency in data and tracking of pedestrians. The CV Pilot THEA team adopted changes to the system with new thermal and visual camera sensors to replace the LiDARs.

7.4 Use Case 4: Transit Signal Priority

As described in section 6.2.3, the Transit Signal Priority application underwent a change in operations and therefore has not produced data for performance evaluation as of the date of this report. The TSP is currently undergoing testing. The results of these tests will be documented by the system engineers with USDOT. Use Case 5: Streetcar Conflicts

7.4.1 Analysis Dataset

The analysis uses data collected from participant and streetcar vehicles between March 1, 2019, and August 24, 2020. Use Case 5 considers all daily travel occurring when the TECO streetcar is in service. This is because the Vehicle Turning Right in Front of Transit Vehicle application is designed to warn drivers of participant vehicles and streetcar operators of imminent collisions when the vehicle makes a right turn in front of the streetcar.

7.4.2 Mobility Impact

The PMESP did not consider mobility assessment for UC5.

7.4.3 Safety Impact

7.4.3.1 Observed False Positives

During the analysis period, 13 unique participant vehicles and seven unique streetcars deployed 61 VTRFTV warnings. Out of 61 warnings, 69 percent were shown to drivers (i.e., treatment group). Per application specifications, it is expected that VTRFTV warnings are triggered by both streetcar and participant vehicles involved in a VTRFTV conflict. The warning event analysis revealed a mismatch between the number of unique streetcars and participant vehicles that triggered warnings. During visual inspection, it became clear that not all warnings were present for both streetcars and vehicles for the same event. Figure 7-60 maps the warning events. The research team created a BSM event profile for each warning, consisting of 30 seconds before and after the moment of warning. The events can be replayed and analyzed via the THEA CV Pilot Dashboard as a precursory step to data-driven evaluation.

Source: CUTR, 2020

Figure 7-60. Map of VTRFTV Events

The PCE classified 54 warnings (89%) as false positives and 7 as potentially true positives. The visual inspection of these events revealed that the PCE failed to correctly classify some events as PTP, therefore a visual inspection was carried out for all warning events regardless of PCE classification. Table 7-30 summarizes the results of the false positive analysis, first via PCE and then via visual inspection.

Table 7-30. VTRFTV Analysis – False and True Positives

Classification	Count	Share (%)	Test Performed
False Positive	54	88.5	Automated PCE
True Positive	7	11.5	Automated PCE
Total	**61**		
False Positive	52	85.2	Visual Inspection
True Positive	9	14.8	Visual Inspection
Total	**61**		

The VTRFTV can trigger multiple warnings per conflict instance as long as the parameters are met. The visual inspection classified nine warnings as true positives. Those TP warnings occurred in five unique conflict events. Table 7-31 presents the sequence of warnings in the events where TP warnings were recorded. The FP classification was conducted independently, therefore some warnings are FP and some are TP in one conflict between a vehicle and a streetcar. Each event was classified as TP regardless of the FP warnings in the sequence.

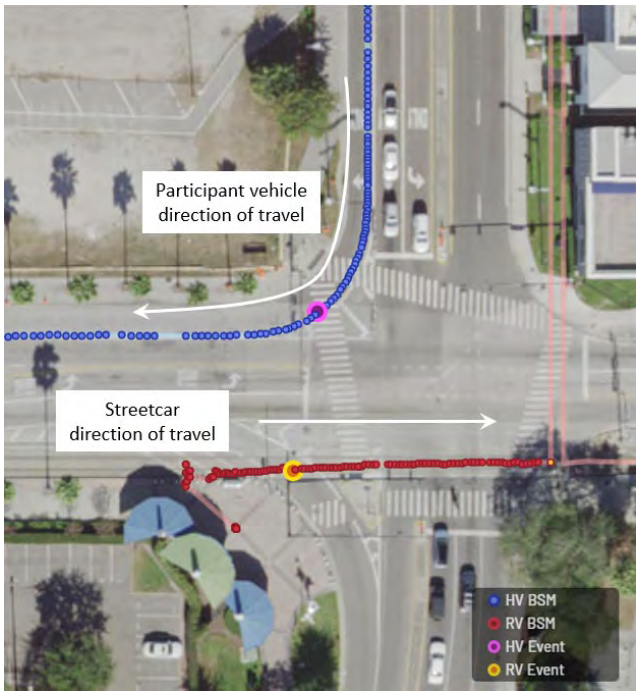
Table 7-31. VTRFTV TP Event Warning Sequence

Unique VTRFTV Event #	Warning Sequence					
	1st	2nd	3rd	4th	5th	6th
1	TP					
2	FP	FP	TP	TP	TP	
3	TP	TP				
4	FP	FP	TP	TP	FP	FP

7.4.3.2 Factors Associated with False Positives

7.4.3.2.1 Vehicle Trajectories

A vehicle trajectory inspection was conducted to determine if the HV and RV were in the intended paths according to the specifications defined in the System Design Document for the VTRFTV to trigger warnings [10]. Even though the application relies on operational parameters to determine if the two vehicles are in a collision trajectory, the visual inspection revealed that the streetcar and participant vehicle did not meet the required paths to trigger the warnings in 45 events classified as false positive (87%). Figure 7-61 illustrates one of these instances where the two vehicles are traveling in opposite directions and thus not on a path leading to a conflict and warning generation. Figure 7-62 shows another instance where the HV and RV are traveling in the same direction but on opposite sides of the road. It seems that at times, the VTRFTV application cannot discern correctly the relative HV-RV locations and their intended paths.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

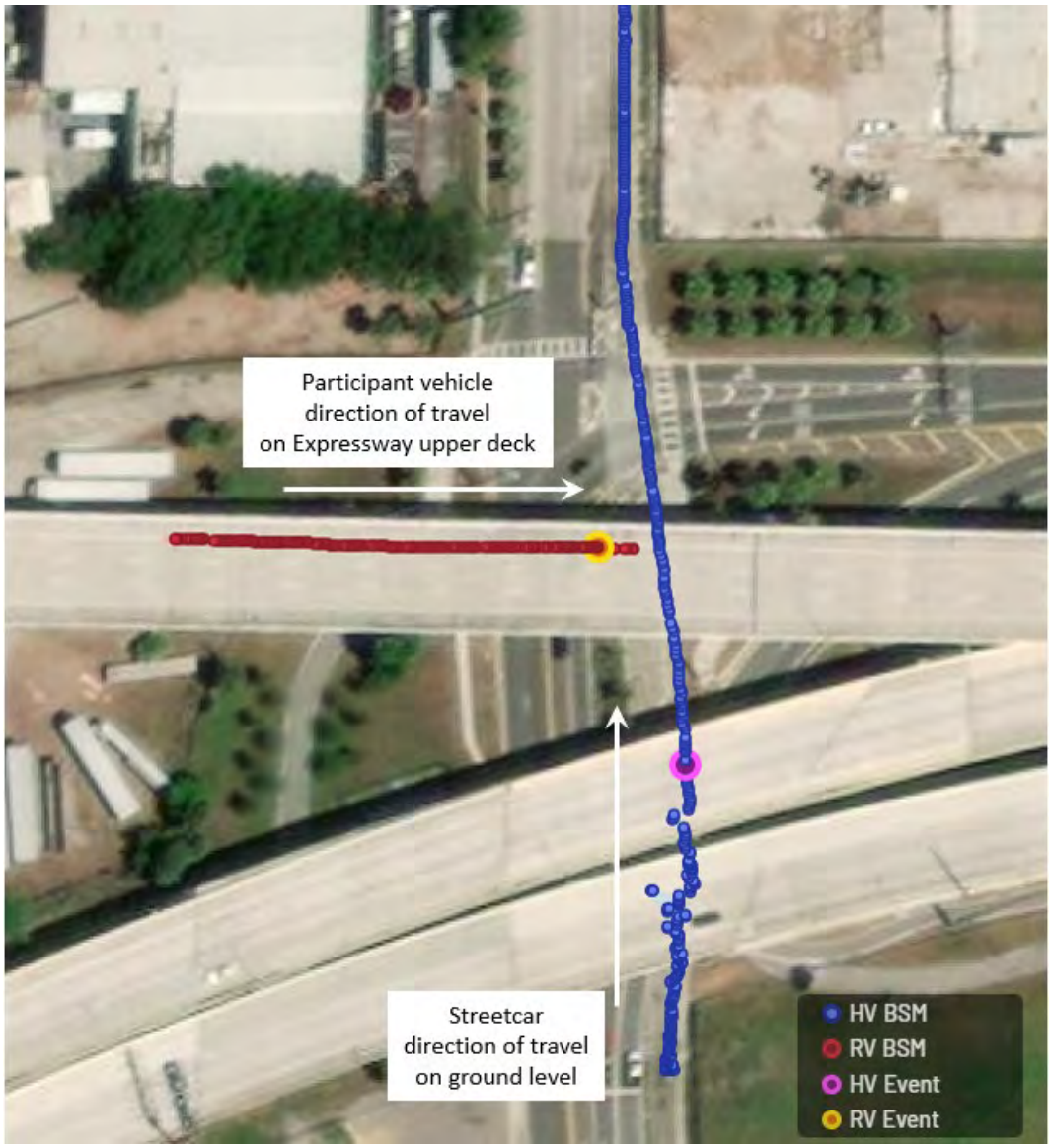
Figure 7-61. FP VTRFTV (Opposite Direction)

Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-62. FP VTRFTV (Same Direction)

7.4.3.2.2 Roadway Elevations

Another cause of false positive events is related to determining the correct elevation difference between the host and remote vehicles. This is relevant at the Channelside Drive and Adamo Drive intersection due to the upper decks of the Selmon Expressway's Reversible Express Lanes (REL), which can produce GPS interference. Figure 7-63 demonstrates a false positive event where the RV is traveling on the REL and the HV (streetcar) is traveling north on Channelside Drive. This problem was later corrected by adding a delta elevation parameter to the VTRFTV application to check the elevation differential between vehicles before issuing a warning.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-63. FP VTRFTV (Different Elevation)

7.4.3.3 V2V Interactions and Conflict Assessment

Due to the roadside unit's location in the study area, there is a segment of the streetcar route along Channelside Drive that is not covered and therefore no BSMs were recorded. Since the methodology described above uses RSU BSMs, this presents gaps. Therefore, the interactions and conflicts estimated between vehicles and streetcars only applies to segments with available data. Figure 7-64 displays the streetcar route and the RSU coverage gap.

Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-64. Study Area RSU Coverage

Table 7-32 reports the estimated number of interactions and conflicts and the VTRFTV event classification. During the analysis period, an estimated 7,167 interactions that produced 64 conflicts conformable to VTRFTV deployment were identified. Of those 64 conflicts, 3 also had a VTRFTV generated and recorded in the OBU data logs.

Table 7-32. VTRFTV Movement Classifications and Rates

Description	Count	Rate (%)
VTRFTV Unique Events (TP + FP)	34	--
V2V Interactions	7,167	--
Conflicts	64	--
True Positives (TP)	4	6.2
False Negatives (FN)	60	93.8
Non-conflicts	7,103	--
True Negatives (TN)	7,073	99.6
False Positives (FP)	30	0.4

Based on the figures of Table 7-32, the overall FP rate is

$$FP\ rate = \frac{30}{30 + 7073} \times 100 = 0.4\%$$

The FN rate is

$$FN\ rate = \frac{60}{60 + 4} \times 100 = 93.8\%$$

7.4.3.4 Multiple Level Assessment

The focus of Use Case 5 and the deployment of the VTRFTV application was to increase safety [4]. This section details the safety impact evaluation of the VTRFTV application and the warnings deployed by participants and streetcars while driving. The assessment is carried out on multiple levels. The two periods for the before-after study were selected as follows: (a) the before period was set to March 2014 to February 2019 (five years before the application deployment period), and (b) the after period was set to March 2019 to August 24, 2020.

7.4.3.4.1 Crash Analysis

Only specific types of crashes are relevant to this analysis in that the crashes need to be avoidable using the VTRFTV application. Right turn crashes for vehicles and crashes caused by right-turning vehicles on the streetcar are deemed avoidable using the deployed VTRFTV application. Police-reported traffic crashes were accessed via the Signal Four Analytics portal, which reports these crashes for the state of Florida [15].

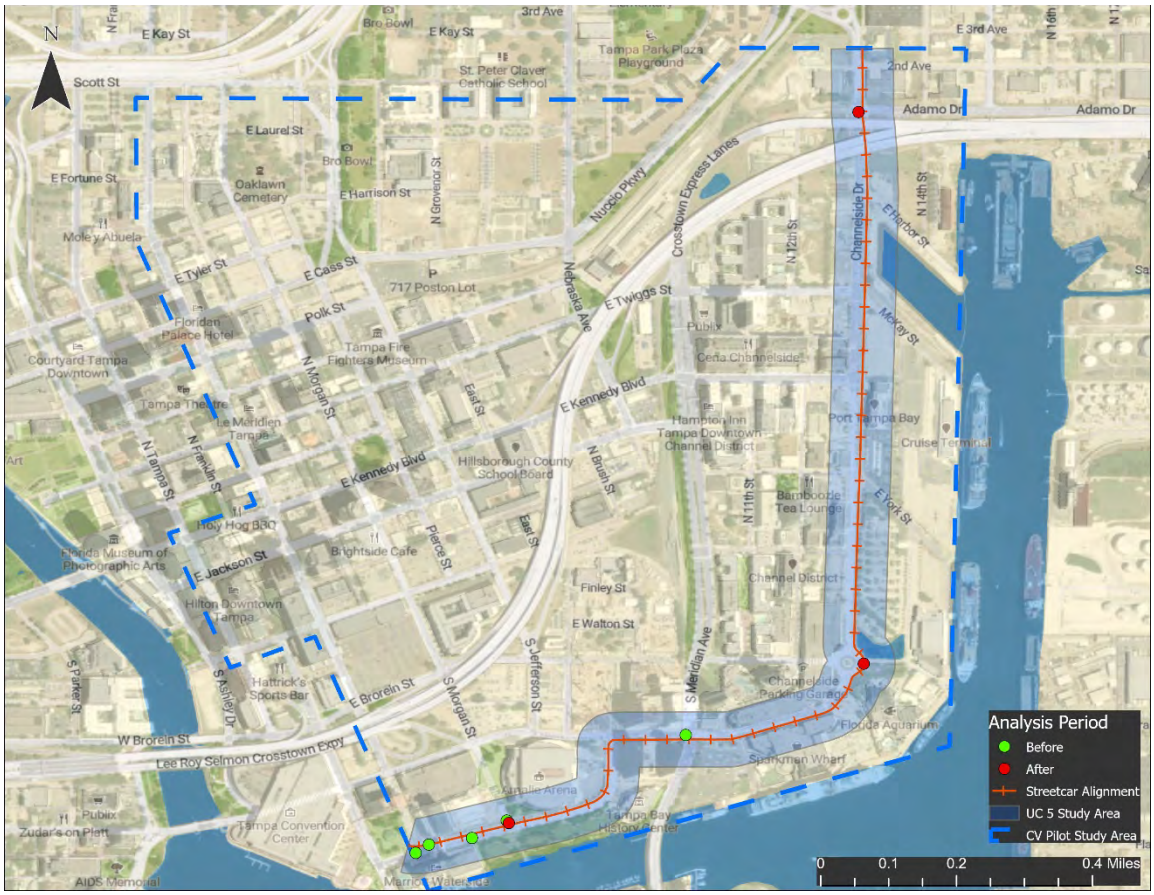
Table 7-33 reports the crashes and calculated crash rates. Figure 7-65 maps their locations along the streetcar route. The before period reported six crashes and the after period reported three crashes. However, upon reviewing the crash reports, researchers determined that only two crashes in the before and one in the after period were avoidable using the VTRFTV (i.e., vehicle turning right in front of a transit vehicle). The other

crashes included left turn in front of a transit vehicle, stopped vehicle on the rail tracks, and one pedestrian-streetcar crash. Only the pedestrian-involved crash reported injury, whereas all the other crashes reported property damage only.

Table 7-33. Crash Rates for UC5

Period	Dates	No. of Months	No. of Avoidable Crashes	Crash Rate (Crashes/Month)
Before	3/2014–2/2019	60	2	0.033
After	3/2019–8/2020	18	1	0.056

Based on the limited number of crashes, the crash rate per month shows an increase in the after period. Crash analysis requires longer periods of time since crashes are rare events; usually a few years are needed to build a crash history profile for specific locations or routes. The result of the crash analysis for UC5 is inconclusive due to the short time frame in the after period and the low number of crashes in the before period.



Data Source: Signal 4 Analytics, 2020

Figure 7-65. Streetcar Related Traffic Crashes

7.4.3.4.2 V2V Interactions and Conflict Assessment

The Performance Measurement and Evaluation Support Plan's recommended approach to safety evaluation was the before-after period or an interrupted time series approach since streetcars cannot be randomly assigned to the study. The participant recruitment conceived for the Pilot grouped participants into treatment and control groups. In this context, the treatment group includes those participants having the HMI enabled (warnings displayed on the mirror), while the control group includes those participants with fully functioning OBUs and all CV applications installed, but with the HMI disabled (warnings not displayed on the mirror).

Ultimately, the goal of the conflict assessment to measure safety performance was to compare the behavioral responses of the treatment and control groups under similar conditions. In addition, the deployment considered three months of data collection during which the treatment group was placed in silent mode. During this time, all warning events generated by the treatment group were recorded in OBUs but not displayed via HMI. At the end of silent mode, the HMI was turned on (using OTA protocols) for the entire duration of Phase 3 of the CV Pilot. This staged approach allowed additional insight into the before-after behavioral responses of the same subjects.

Table 7-34 reports the warnings generated and displayed to participants by experimental design group. Out of 61 VTRFTV warnings, 68.9 percent were shown to drivers (treatment group) and 31.1 percent were not shown (control and silent mode).

Table 7-34. VTRFTV Warning Visibility by Participant Group for UC5

Group	HMI Disabled (Warnings Not Displayed)			HMI Enabled (Warnings Displayed)			Grand Total
	TP	FP	Share (%)	TP	FP	Share (%)	
Control	3	4	11.5			0.0	7
Treatment			0.0	2	9	18.0	11
Treatment (Silent)	2	3	8.2		4	6.6	9
Streetcar		7	11.5	1	26	44.3	34
Total	**5**	**14**	**31.1**	**3**	**39**	**68.9**	**61**

Next, the analysis assessed the behavioral responses by comparing the reactions of the treatment and control groups and the response within the treatment group (HMI disabled versus HMI enabled). The driver reaction was investigated separately for each V2V application. The conflict identification algorithm identifies a reaction to a conflict if the driver decelerates at a rate below 0.5 mps^2 after the moment of warning. Since drivers facing a VTRFTV situation could also change lanes instead of decelerating, this was also visually checked to ensure they did not change lanes with no deceleration. No drivers changed lanes at or after the moment of warning.[13]

7.4.3.4.2.1 Treatment vs. Control Group Reaction to TP Warnings

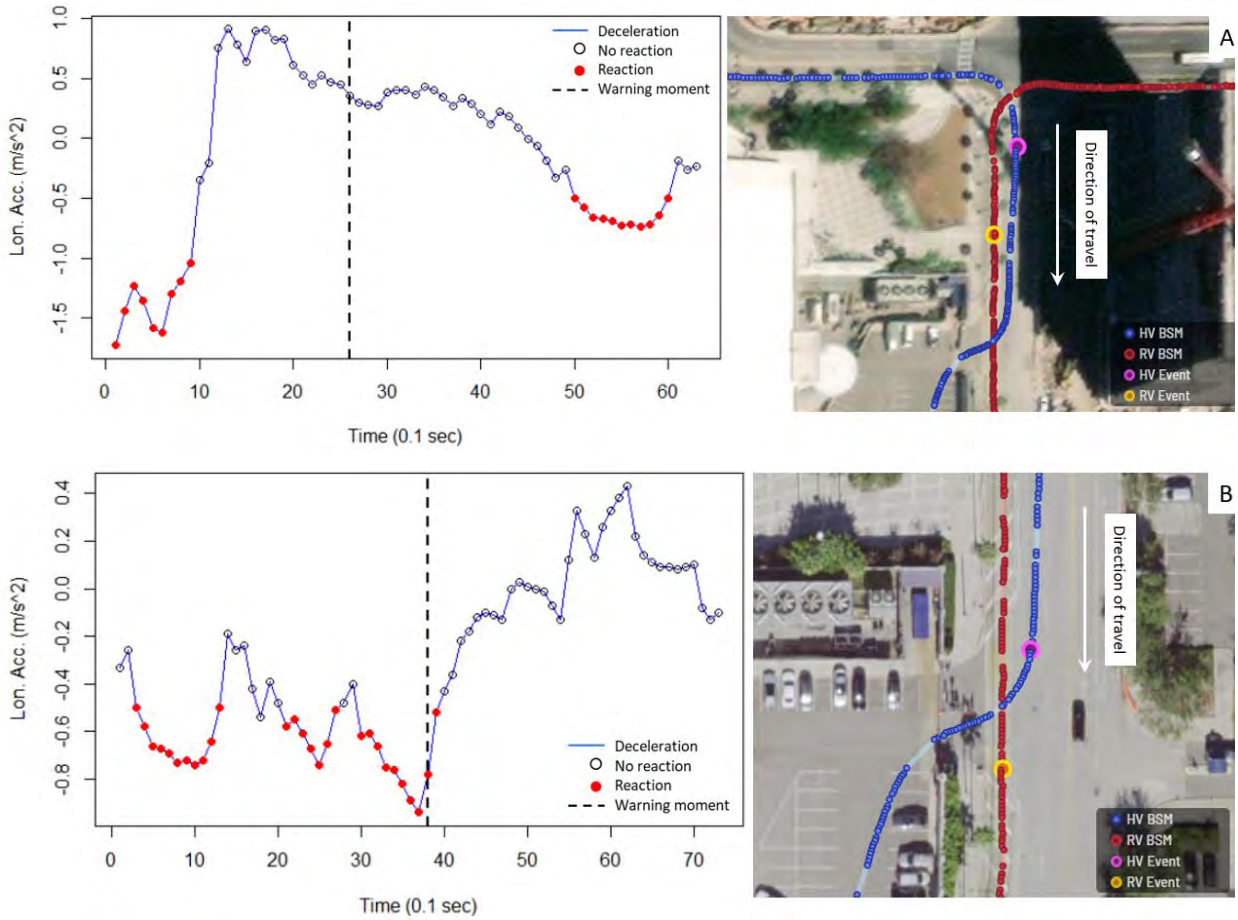
The onboard units recorded eight VTRFTV warnings classified as true positive events. Of those, two events were triggered by an HMI-enabled unit (treatment) and three by an HMI-disabled unit (control). Two true

[13] To simplify the results of the reaction algorithm, the lateral reaction was removed since all warning event trajectories were visually inspected to confirm their trajectory.

positive VTRFTV warnings were not displayed to participants in the treatment group during silent mode. In addition, one TP was recorded by streetcar with HMI enabled. The streetcars were not designed to have the HMI disabled, but errors in the unit setup configuration disabled the HMI warnings on some streetcar OBUs. Figure 7-66 (A-B) displays the results of the conflict identification algorithm.

7.4.3.4.2.2 Participant Reaction to Warnings

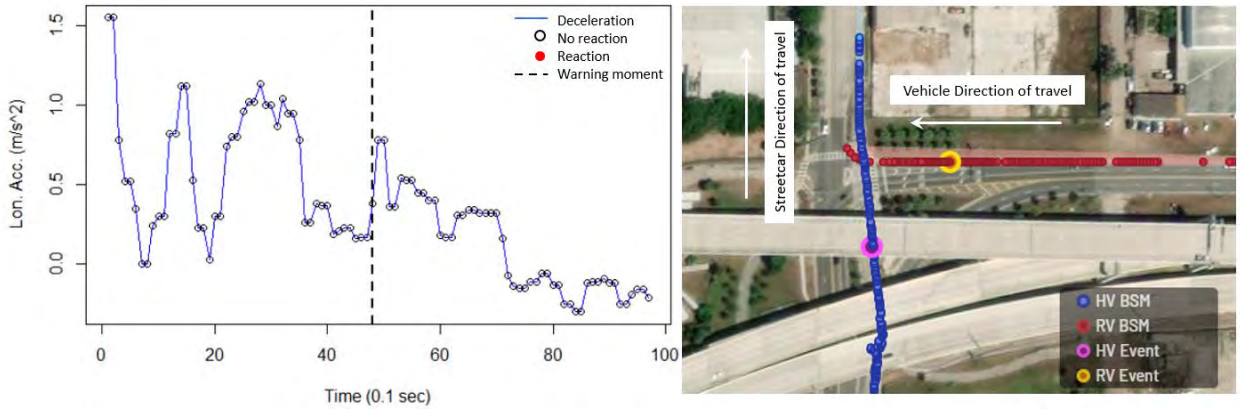
The left diagram shows the deceleration profile of each VTRFTV event classified as true positive and generated by an HMI-enabled vehicle using the 10-second BSM profile. The map next to the diagram shows the host and remote vehicle trajectories and location of the warning moment. The participant in Figure 7-66 (A) decelerated before and after the warning, while the participant in Figure 7-66 (B) reacted before the warning.



Source: CUTR, 2020

Figure 7-66. Driver Reaction to TP VTRFTV with HMI Enabled

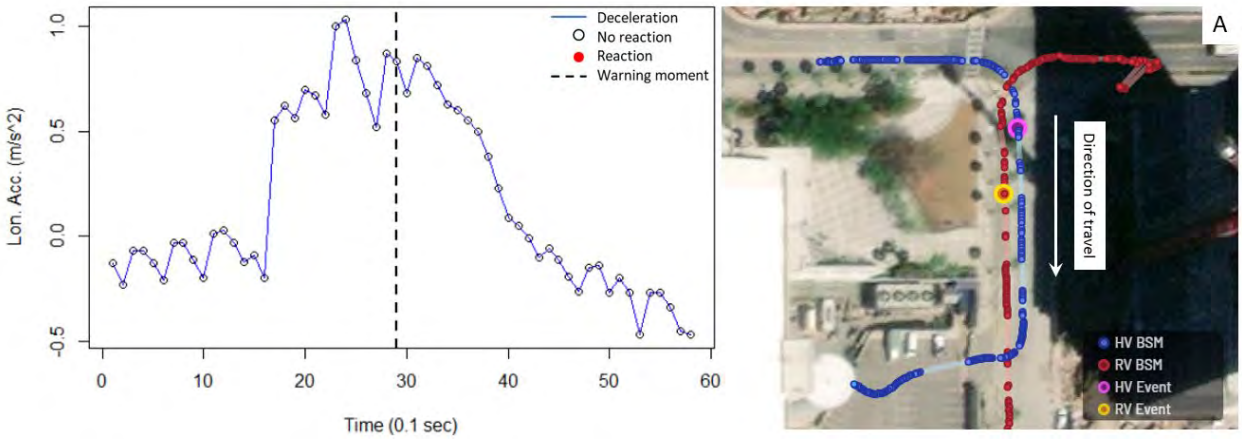
Figure 7-67 shows the profile for the only true positive VTRFTV warning issued to a streetcar. There is no evidence of a response from the streetcar operator.

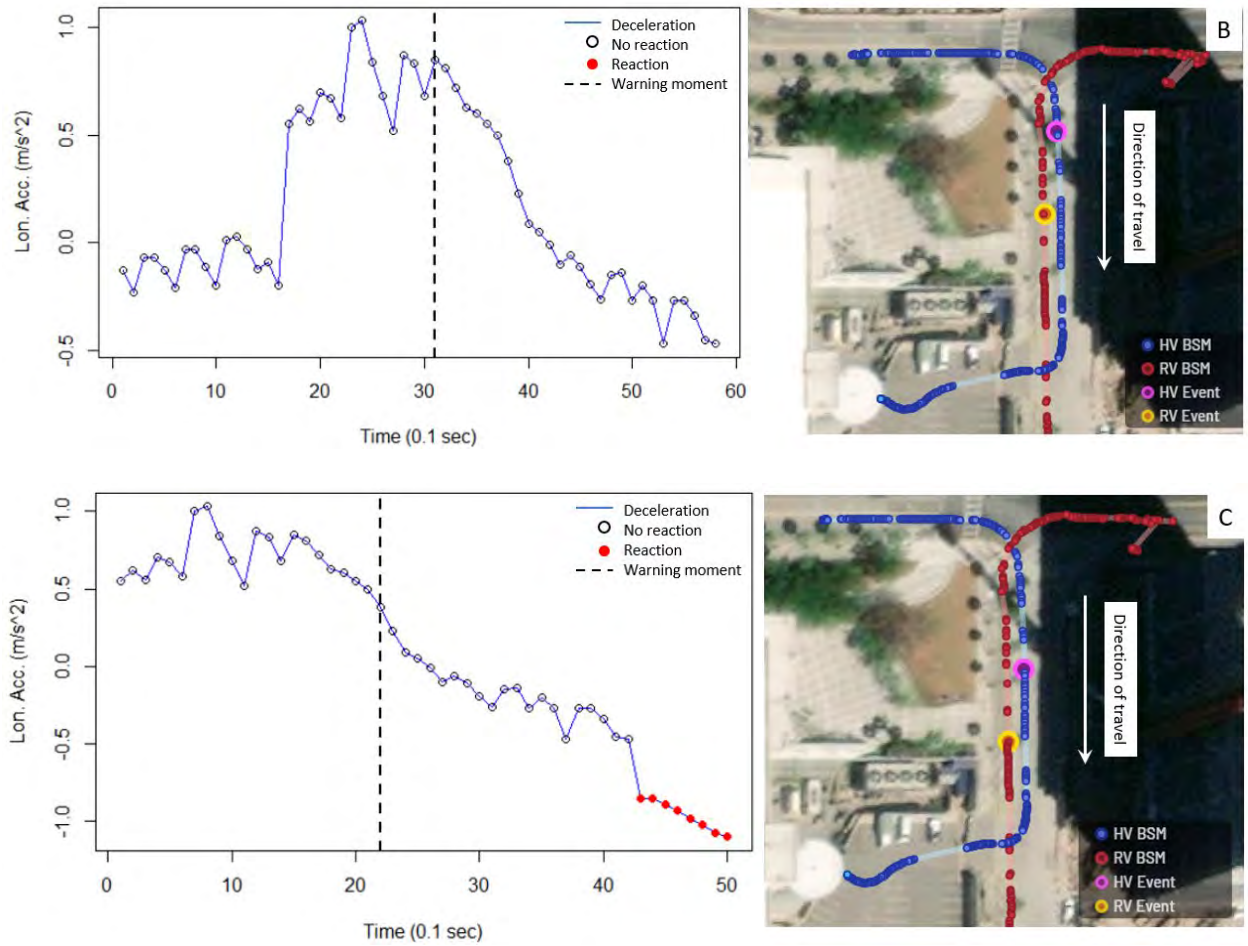


Source: CUTR, 2020

Figure 7-67. Streetcar Operator Reaction to TP VTRFTV with HMI Enabled

Figure 7-68 (A-C) displays the results the algorithm applied to the three TP warnings triggered by a participant vehicle assigned to the control group with the HMI disabled. All three warnings belong to the same vehicle and occur in sequence within seconds of each other. The driver did not react after the first two warnings (A and B) but showed deceleration after the third warning (C). This might have been because the vehicle would make a right turn shortly after this event.





Source: CUTR, 2020

Figure 7-68. Driver Reaction to TP VTRFTV with HMI Disabled

The mixed reaction results could be due to the specific conditions under which the warning was generated (yet not displayed). This application works in slower speeds and on local roads (as opposed to highways), therefore the drivers could be slowing down for a traffic signal turning red, or to make turns, or for other vehicles around them.

7.4.3.4.2.3 Within Group Reaction to TP Warnings

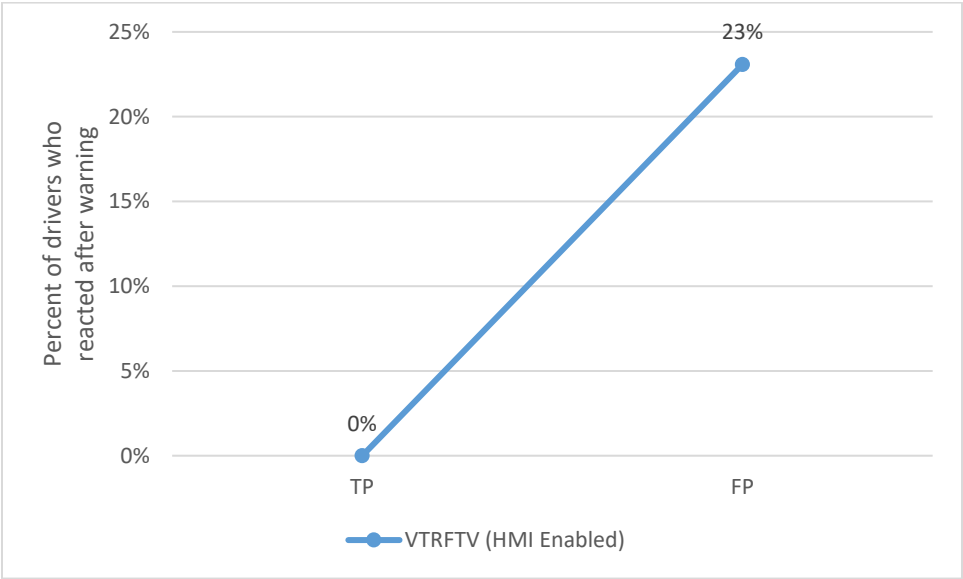
No warnings were issued to participant vehicles in the treatment group with silent mode and HMI enabled. Streetcars, however, had one warning with HMI enabled, although the reactions cannot be compared as the streetcar operator was not the same every time.

7.4.3.4.2.4 Driver Reaction to All Visible Warnings (HMI Enabled – All Groups)

The warnings shown to participants were generated under two types of conditions. The first condition involves a true conflict (TP) with another vehicle determined by the parameters and the method explained in the previous sections. The second condition involves a false positive (FP) where the warning is triggered when

there is no conflict with another vehicle. It is possible for the participant to respond under both conditions, albeit differently.

Next, the analysis pools all the warnings recorded while the HMI was enabled to evaluate any difference in behavioral response by type of conflict (TP vs. FP). Figure 7-69 illustrates the difference in the share of drivers who reacted after they received a visible warning, grouped by classification (TP vs. FP). No participants (or streetcar operators) reacted to VTRFTV warnings classified as TP compared to 23 percent of drivers who reacted to warnings classified as FP. This might be due to the timing of the warning during the sequence of events. In addition, the reaction expected might have been too extreme, where a vehicle decelerating at a normal rate could still avoid the streetcar.



Source: CUTR, 2020

Figure 7-69. Proportions of Drivers Reacting to VTRFTV with HMI Enabled

7.4.3.4.2.5 Predictive Value of Warnings

Since crashes and conflicts are rare traffic events, it is useful to examine two additional metrics in the assessment of V2V safety applications: the positive and negative predictive values of warnings. Given the many factors affecting the deployment and efficacy of CV safety applications, these measures evaluate the conditional probability of being in a dangerous situation when a warning is triggered (positive predictive value) or not being in a dangerous situation when a warning is not triggered (negative predictive value).

In this context, the probability that a warning will be a true positive (conflict situation) is the proportion of TP warnings to all warnings (TP + FP). This is a true performance measure of an application since it gauges how likely a driver is to be in a conflict when a warning is received.

For the VTRFTV application, the positive predictive value is calculated as

$$VTRFTV\ positive\ predictive\ value = \frac{TP}{TP + FP} \times 100 = \frac{4}{4 + 30} \times 100 = 11.8\%$$

Given the travel conditions characterizing UC5, this means that if a Vehicle Turning Right in Front of Transit Vehicle event is triggered, the probability of being in a real dangerous situation is approximately 12 percent.

On the other hand, the probability that a negative (no warning) will be a true negative (no conflict) is the proportion of TN to all negatives (TN + FN). The negative predictive value of the VTRFTV application as deployed in UC5 can be calculated as

$$VTRFTV\ negative\ predictive\ value = \frac{TN}{TN + FN} \times 100 = \frac{7073}{7073 + 60} \times 100 = 99.2\%$$

This means that if a VTRFTV event is not triggered (no warning), drivers are not in a dangerous situation 99.16 percent of the time and if a warning is triggered, the chance of the situation being a conflict is 0.84 percent (1 in 119). Note that these are conditional probabilities given a warning has been triggered or not triggered, which differs from FP and TP rates calculated earlier.

7.4.4 Summary of Findings

The VTRFTV application is intended to warn both drivers and streetcar operators when a vehicle is making a right turn in front of a streetcar. The analysis of the warning events suggests that more work is needed to fine-tune the VTRFTV application. The complexity of intersections and location of tracks created scenarios that were not anticipated during the application design phase. Most of the observed false positives were due to the application's inability to determine a conflicting path with the streetcar. Elevation issues at the Adamo Drive and Channelside Drive intersection also produced false positives, perhaps due to GPS inaccuracy caused by the Selmon Expressway underpass. In addition, a major component of VTRFTV implementation is that it is a V2V application and therefore works anywhere two vehicles interact, regardless of location.

The development of safety applications must balance two perspectives: a high true positive rate (warnings triggered when there is a conflict) and a low false positive rate (warnings triggered when there is no conflict). Advanced Driver Assistance Systems deployed on automated vehicles have demonstrated that drivers trust them when they work well. If the false positive rate is high, drivers choose to ignore or even turn off the system due to the nuisance and distraction of alerts [17-19].

In evaluating Use Case 5, the Vehicle Turning Right in Front of Transit Vehicle application generated 61 warnings that occurred in 34 unique events. The sequential warnings varied from one warning per event to six warnings in one event. Out of these warnings, eight (13%) were classified as true positive during four unique events, but only three warnings (one event) were shown to the driver due to the evaluation's experimental design. The conflict detection algorithm confirmed the participant was not engaged in a conflict with the streetcar.

7.5 Use Case 6: Traffic Progression

7.5.1 Analysis Dataset

7.5.1.1 Mobility Dataset

The analysis of Use Case 6 focuses on Meridian Avenue only. This is because I-SIG was not successfully deployed and because participant vehicles did not generate V2V warnings while traveling on Florida Avenue. The analysis uses data collected from participant vehicles between May 1, 2018, and August 31, 2020. This use case considers travel occurring on weekdays (Monday–Friday) between 6:00 and 9:59 a.m. for morning peak hours and between 3:00 and 6:59 p.m. for afternoon peak hours. During these times, traffic congestion is high, which is the focus of Use Case 6.

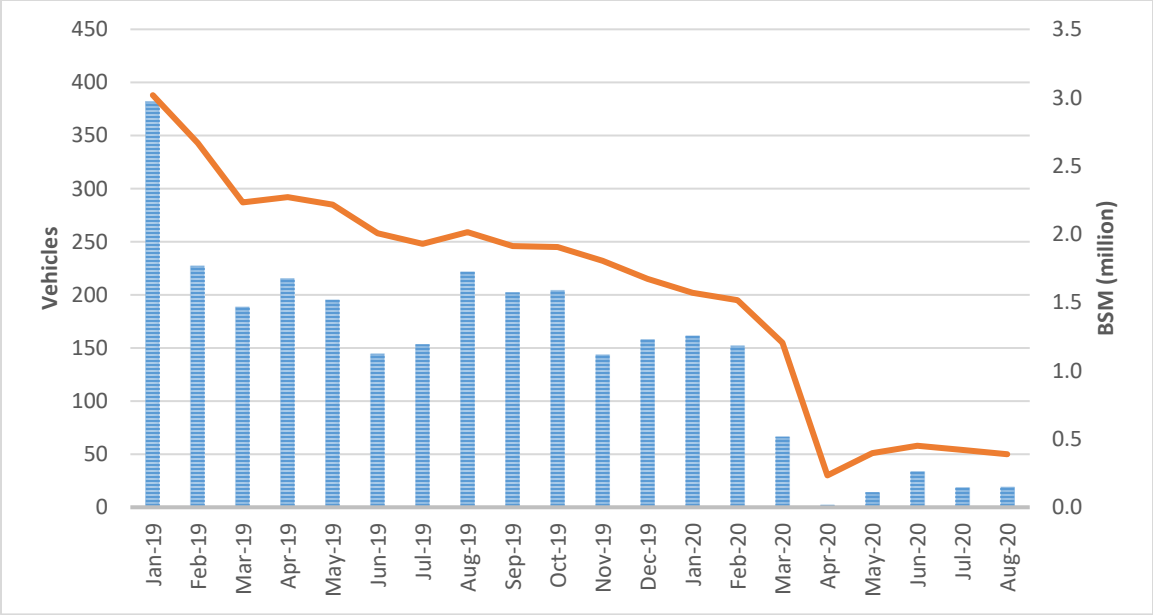
7.5.1.2 Safety Dataset

The dataset used for safety analysis includes data collected between May 1, 2018, and August 31, 2020, and is separated into the morning peak period (6:00–9:59 a.m.) and the afternoon peak period (3:00–6:59 p.m.).

7.5.2 Mobility Impact

7.5.2.1 Impact on Travel Time

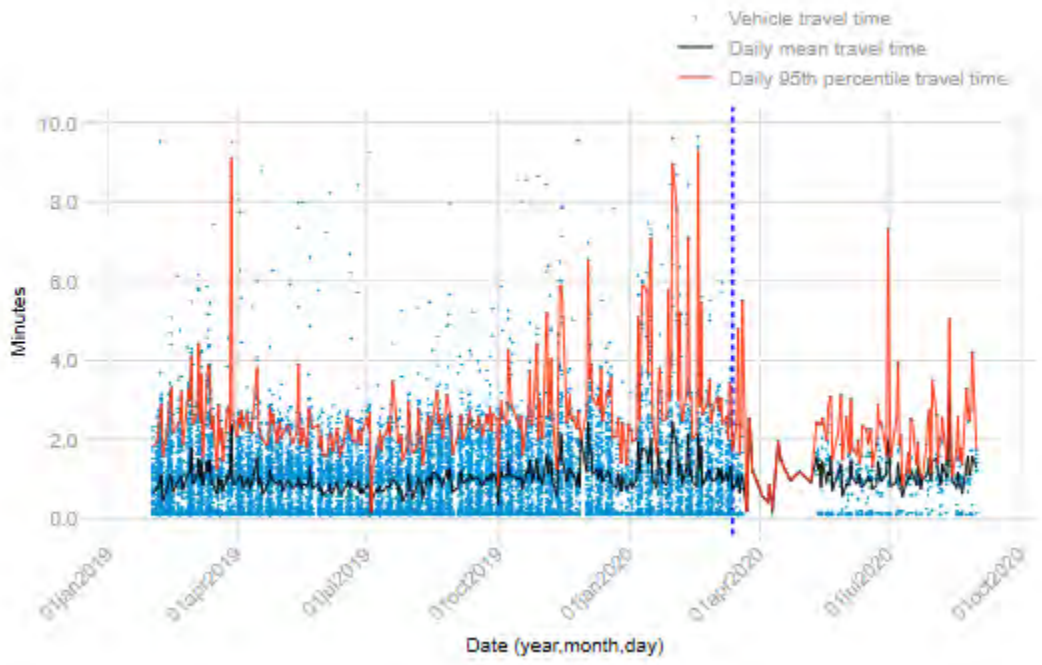
Use Case 6 only generated data conducive to setting up the baseline. As discussed in section 6.2.5, the I-SIG and MMITSS architecture were not deployed and did not generate the required data to conduct a before-after assessment. Nonetheless, the research team collected and analyzed data at the vehicle level using RSU BSMs. The dataset consists of travel time at the vehicle level that was computed using the first and last BSM as each vehicle traveled on Meridian Avenue during the morning and afternoon peak periods. A total of 22.6 million BSMs were collected from 719 unique participant vehicles during the period of February 2019 through August 2020 (Figure 7-70).



Source: CUTR, 2020

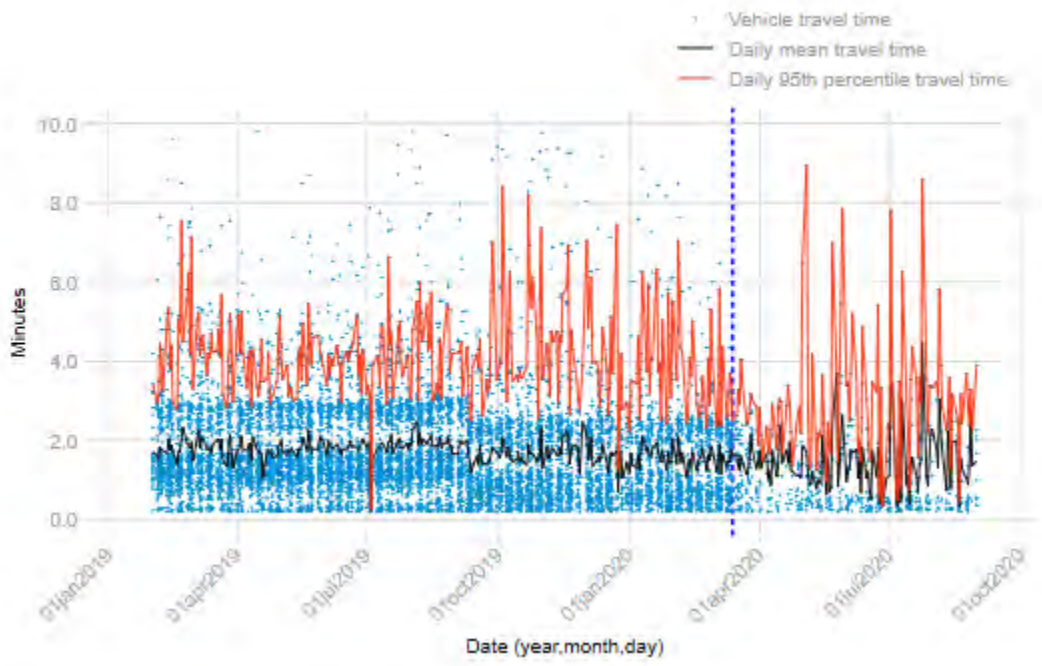
Figure 7-70. Mobility Evaluation Analysis Dataset

Figure 7-71 and Figure 7-72 illustrate scatterplots of participant vehicle travel times over the analysis period, with the black and red lines showing the daily mean and 95th percentile travel times. For reference, the dotted blue line represents the beginning of the pandemic travel restrictions in Tampa (March 20, 2020). The recurring spikes in the graph demonstrate intra-day variability during weekday travel.



Source: CUTR, 2020

Figure 7-71. Travel Times 10-Minute Intervals – Morning Peak (7:00 to 9:59 a.m.)



Source: CUTR, 2020

Figure 7-72. Travel Times 10-Minute Intervals – Afternoon Peak (3:00 to 6:59 p.m.)

7.5.3 Safety Impact

While the focus of Use Case 6 is on the deployment of the I-SIG application to improve traffic progression, there are inherent safety and mobility benefits associated with the deployment of V2V safety applications [4]. This section details the safety impact evaluation of the FCW, EEBL, and IMA applications and their warnings issued to participants while driving on Meridian Avenue. The assessment is carried out on multiple levels.

7.5.3.1 Crash Analysis

Only specific types of crashes are relevant to this analysis in that the crashes need to be avoidable using the safety applications deployed. In UC6, the following crashes may be avoided by deployment of the three applications [5]:

- Forward Collision Warning: rear-end crash.
- Electronic Emergency Brake Light: rear-end crash, sideswipe crash.
- Intersection Movement Assist: angle crash.

To conduct the crash analysis, police-reported traffic crashes were accessed via the Signal Four Analytics portal, which reports both long-form and short-form crashes for the state of Florida [15]. According to the Florida Department of Highway Safety and Motor Vehicles, a long form must be completed and submitted to the department within 10 days after an investigation is completed by the law enforcement officer who, in the regular course of duty, investigates a motor vehicle crash that meets one of the following criteria [20]:

- Resulted in death of, personal injury to, or any indication of complaints of pain or discomfort by any of the parties or passengers involved in the crash.
- Involved a violation of sections 316.061(1) (leaving the scene of crash with an attended vehicle or property) or 316.193 (driving under the influence), Florida Statutes.
- Rendered a vehicle inoperable to a degree that required a wrecker to remove it from the scene of the crash.
- Involved a commercial motor vehicle.

In any crash for which a long form is not required, the law enforcement officer may complete a short-form crash report or provide a driver exchange-of-information form to be completed by all drivers and passengers involved in the crash. A short form usually includes minor crashes only. These crashes are underreported since not all minor crashes involve a law enforcement agency report.

During Phase 1 of the CV Pilot and setup of the use cases, between 2010 and 2013, the UC6 Meridian Avenue segment recorded 20 crashes during the morning (6:00–9:59 a.m.) and afternoon (3:00–6:59 p.m.) periods.

The two periods for the before-after study were selected as follows: (a) the before period was set to February 2014 to February 2019 (five years before the analysis period), weekdays 6:00–9:59 a.m. and 3:00–6:59 p.m., and (b) the after period was set to March 2019 to August 2020, weekdays 6:00–9:59 a.m. and 3:00–6:59 p.m. Table 7-35 reports the crashes and calculated crash rates for Meridian Avenue. The crash rate is not calculated in the traditional crashes per vehicle miles traveled (VMT) because data for 2020 are not finalized. The before period reported 40 crashes, of which 37 were potentially avoidable crashes using V2V applications, and the after period reported 16 crashes, of which 14 were potentially avoidable.

Table 7-35. Crash Rates for UC6

Period	Dates	No. of Months	No. of Avoidable Crashes	Crash Rate (Crashes/Month)
Before	2/2014–2/2019	61	37	0.60
After	3/2019–8/2020	18	14	0.78

Table 7-36 details the crash types for the before and after periods. The highest percentage of crashes in the before period were rear-end crashes (35.1%), whereas rear-end and sideswipe crashes each accounted for 35.7% of total crashes in the after period.

Table 7-36. Crash Types for crashes within UC6

Crash Type	Before	%	After	%
Angle	7	18.9	2	14.3
Left Turn	10	27.0	2	14.3
Rear End	13	35.1	5	35.7
Right Turn	1	2.7	0	0.0
Sideswipe	6	16.2	5	35.7
Total	**37**		**14**	

Table 7-37 conveys the crash severity according to the KABCO injury classification scale. Most of the crashes were property damage only; no incapacitating injury or fatal crashes were reported.

Table 7-37. Crash Severity for crashes within UC6

Injury Severity	Before	%	After	%
Property Damage Only	30	81.1	10	71.4
Possible Injury	6	16.2	3	21.4
Non-incapacitating Injury	1	2.7	1	7.1
Incapacitating Injury	0	0.0	0	0.0
Fatality	0	0.0	0	0.0
Total	**37**		**14**	

Based on the reported number of crashes, the simple crash rate per month shows an increase in the after period. Crash rates per VMT can be calculated once the data are finalized for the after period that includes the year 2020.

Crash analysis requires longer periods of time since crashes are rare events; usually a few years are needed to build a crash history profile for a specific segment. The result of the crash analysis for UC6 indicates that the percentage of rear-end crashes remained similar in the before and after periods, and that sideswipe crashes increased by 20 percent.

7.5.3.2 FCW Observed False Positives

During the analysis period, a total of 55 Forward Collision Warnings were issued for both AM and PM periods. Table 7-38 summarizes the number of FCWs and percentage of false positives.

Table 7-38. FCW Warning Summary

	AM Period	PM Period	Total
No. of Vehicles	26	11	35*
No. of FCWs	38	17	55
No. of FP FCWs	24 (63.2%)	3 (17.6%)	15 (27.3%)
Avg. Warnings per Vehicle	1.46	1.55	1.57

**Some vehicles issued warnings in both periods.*

Figure 7-73 maps the FCW events in the UC6 segment. The research team created a BSM event profile for each warning, consisting of 30 seconds before and after the moment of warning.

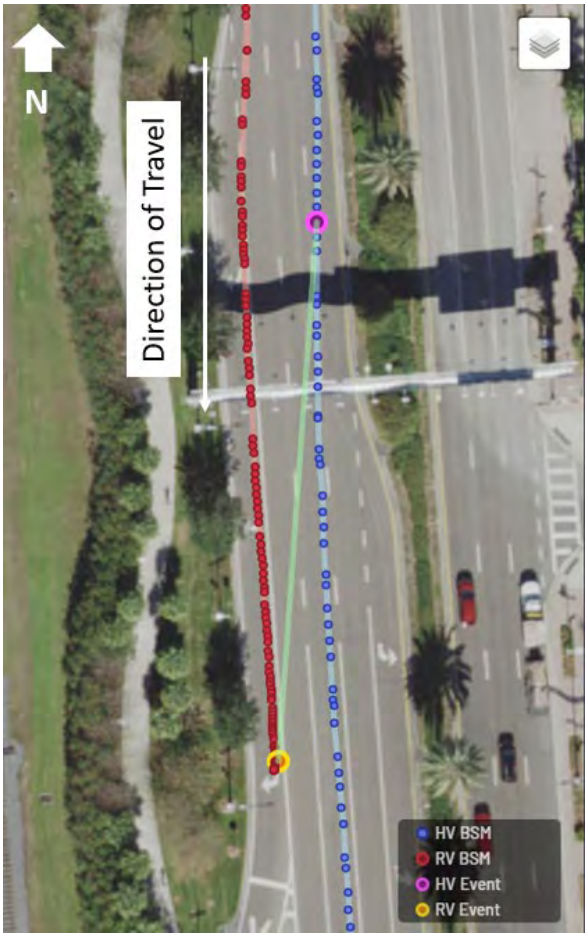
Source: CUTR, 2020

Figure 7-73. Map of FCW Events

Parameter Conformity Evaluation (PCE) is the first step of the evaluation, where each warning event is analyzed and checked for conformity to the default application's operational parameters listed in Chapter 6. The PCE analysis classified 24 events as false positives for both periods and 31 as potentially true positives. The next step is the visual inspection of the 31 FCW events identified as potentially true positives (PTPs).

7.5.3.2.1 Factors Associated with False Positives

A visual inspection was conducted to determine if the host and remote vehicles were in the same lane at the moment of warning according to the SAE J2945/1 “ahead in-lane” zone [5]. Previous analysis of FCW events in UC1 presented false positives due to the inability of the OBU to correctly determine the ahead in-lane zone of the remote vehicle (i.e., a warning was triggered but the RV was not in the same lane as the HV). Figure 7-74 illustrates a warning triggered when the RV was in the adjacent lane instead of the same lane as the HV.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-74. FCW in Adjacent Lane

The visual inspection revealed that 9 (29%) of the 31 Forward Collision Warnings that passed the initial parameter check for FP occurred while the HV and RV were traveling in different lanes. These were reclassified as FPs. Lane determination seems to be one of the application’s limitations even on a relatively straight road segment.

GPS shift due to loss of signal is usually another underlying cause of some false positive warnings. For this use case, GPS shift occurred even though the road segment does not have overpasses as in other use cases. No FP warnings were issued due to GPS loss. Table 7-39 summarizes the results of the false positive analysis.

Table 7-39. FCW Analysis – False and True Positives

Classification	Count	Percent	Test Performed
False Positive	24	43.6	Automated PCE
False Positive	9	16.4	Visual Inspection
True Positive	22	40.0	Visual Inspection
Total	**55**		

Table 7-40 reports the complete count of hours spent in the area, estimated number of interactions and conflicts between connected vehicles, and the FCW event classification separately for morning and afternoon periods. During the analysis period, vehicles capable of recording data logs spent 282 hours in the morning period and 278 in the afternoon period traveling through the area defined for UC6, with an estimated 4,943 interactions that produced 146 conflicts conformable to FCW deployment.

Table 7-40. FCW Movement Classifications and Rates

Description	AM Period		PM Period	
	Count	Rate (%)	Count	Rate (%)
Time Spent in Area (Hours)	282	--	278	--
FCW (TP + FP)	38	--	17	--
V2V Interactions	3,237	--	1,656	--
Conflicts	85	--	61	--
True Positives	14	16.5	8	13.1
False Negatives	71	83.5	53	86.9
Non-conflicts	3,152	--	1,595	--
True Negatives	3,128	99.2	1,586	99.4
False Positives	24	0.8	9	0.6

Based on the figures of Table 7-40, the overall (AM and PM period combined) false positive rate is

$$FP\ rate = \frac{33}{33 + 4{,}714} \times 100 = 0.7\%$$

The overall false negative rate is

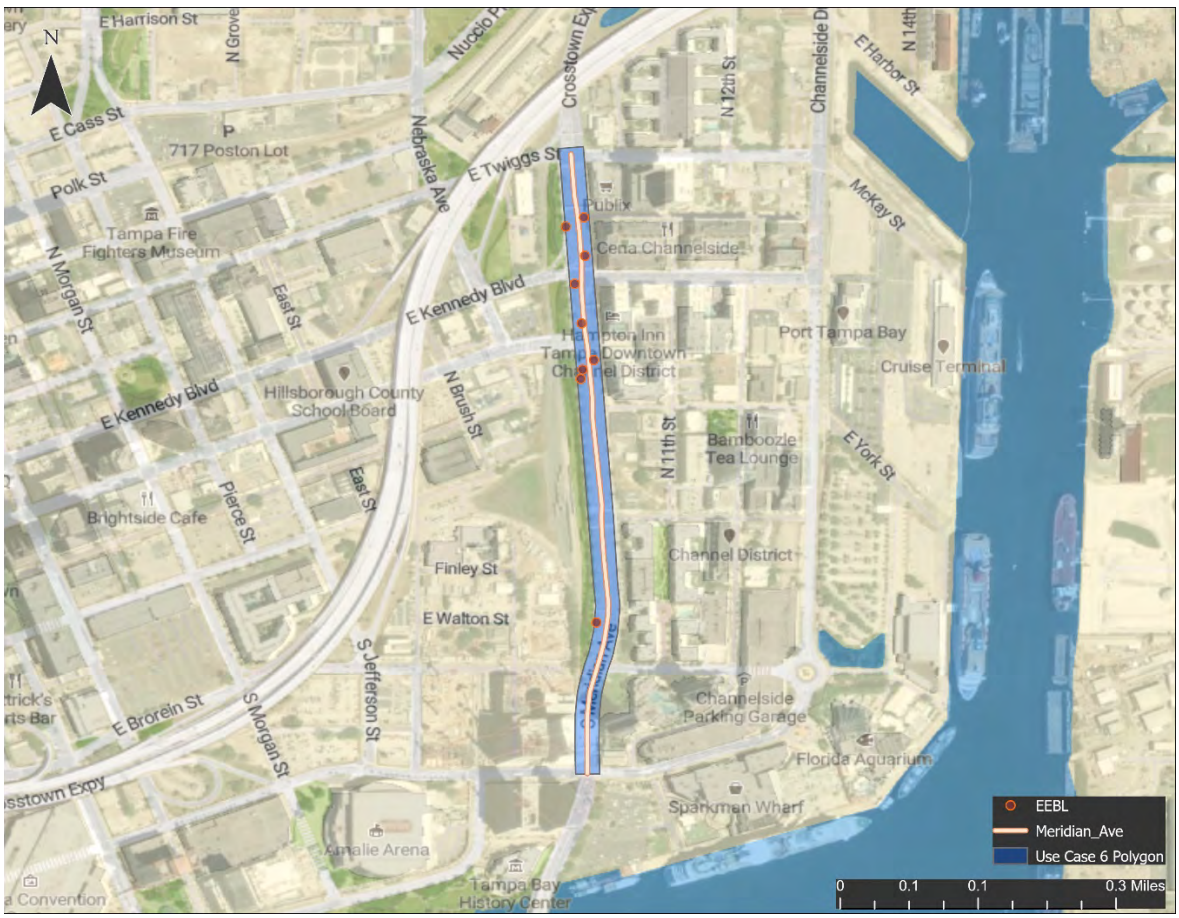
$$FN\ rate = \frac{124}{124 + 22} \times 100 = 84.9\%$$

7.5.3.3 EEBL Observed False Positives

The approach to evaluate Electronic Emergency Brake Light warnings follows the same steps adopted to evaluate FCWs. A Basic Safety Message event profile was created for each EEBL warning to visually check the conditions under which it was triggered. To deploy the EEBL warning, the host vehicle determines if the remote vehicle sending the hard-braking information is in the lane ahead, the left lane, or the right lane. This allows for a width of three lanes instead of one, which is how the FCW works. During the analysis period, nine EEBL warnings were triggered and recorded. Table 7-41 summarizes the number of EEBL warnings, vehicles, and false positives obtained for Use Case 6. Figure 7-75 maps the EEBL warnings in the UC6 study area.

Table 7-41. EEBL Warning Summary

	AM Period	PM Period	Total
No. of Vehicles	6	3	9
No. of EEBLs	6	3	9
No. of FP EEBLs	3 (50%)	2 (67%)	5 (55%)
Avg. Warnings per Vehicle	1	1	1



Source: CUTR, 2020

Figure 7-75. Map of EEBL Events

Three warnings were determined to be false positive since they did not meet the application specification parameters listed in Chapter 6. After visual inspection, four warnings were classified as true positive since they met all the parameters, and the remote vehicle was within one of the three lanes ahead of the host vehicle.

Table 7-42 reports the estimated interactions and conflicts, as well as warnings issued, following the approach detailed in Chapter 6. During the analysis period, an estimated 2,816 interactions led to 22 conflicts conformable to EEBL deployment. As a result, the overall false positive rate of the application is estimated at 0.18 percent, while the false negative rate is estimated at 81.82 percent.

Table 7-42. EEBL Movement Classifications and Rates

Description	AM Period		PM Period	
	Count	Rate (%)	Count	Rate (%)
Time Spent in Area (Hours)	282	--	278	--
EEBL (TP + FP)	6	--	3	--
V2V Interactions	2,517	--	299	--
Conflicts	18	--	4	--
True Positives	3	16.7	1	25.0
False Negatives	15	83.3	3	75.0
Non-conflicts	2,499	--	295	--
True Negatives	2,496	99.9	293	99.3
False Positives	3	0.1	2	0.7

Based on the figures of Table 7-42, the overall (AM and PM period combined) false positive rate is

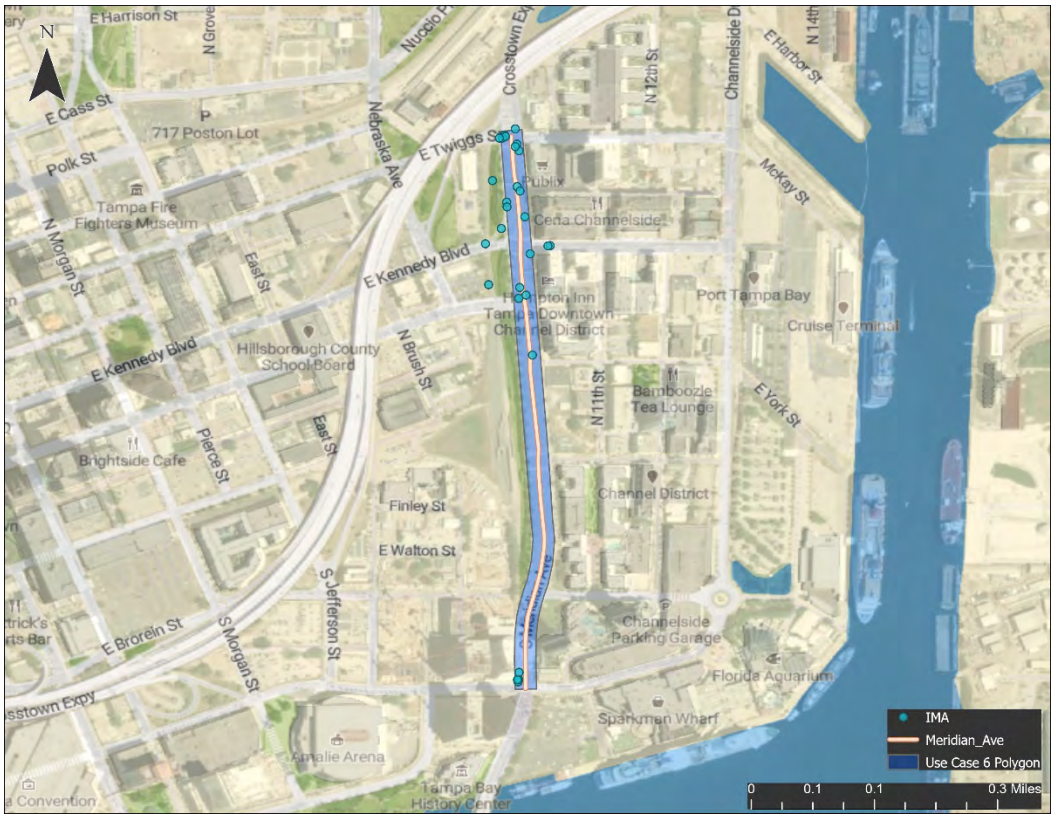
$$FP\ rate = \frac{5}{5 + 2789} \times 100 = 0.2\%$$

The overall false negative rate is

$$FN\ rate = \frac{18}{18 + 4} \times 100 = 81.8\%$$

7.5.3.4 IMA Observed False Positives

Using the same morning and afternoon periods, onboard units recorded 31 Intersection Movement Assist events, as shown in Figure 7-76.



Source: CUTR, 2020

Figure 7-76. Map of IMA Events

Table 7-43 reports the classification by peak period. Three events had missing remote vehicle information and only four were initially determined to be false positives due to the PCE automated process. However, the visual inspection of the remaining 24 warnings showed that all were false positives.

Table 7-43. IMA Warning Summary

	AM Period	PM Period	Total
No. of Vehicles	14	12	24*
No. of IMAs	16	12	28
No. of FP IMAs	16 (100%)	12 (100%)	28 (100%)
Avg. Warnings per Vehicle	1.14	1.0	1.17

**Some vehicles issued warnings in both periods.*

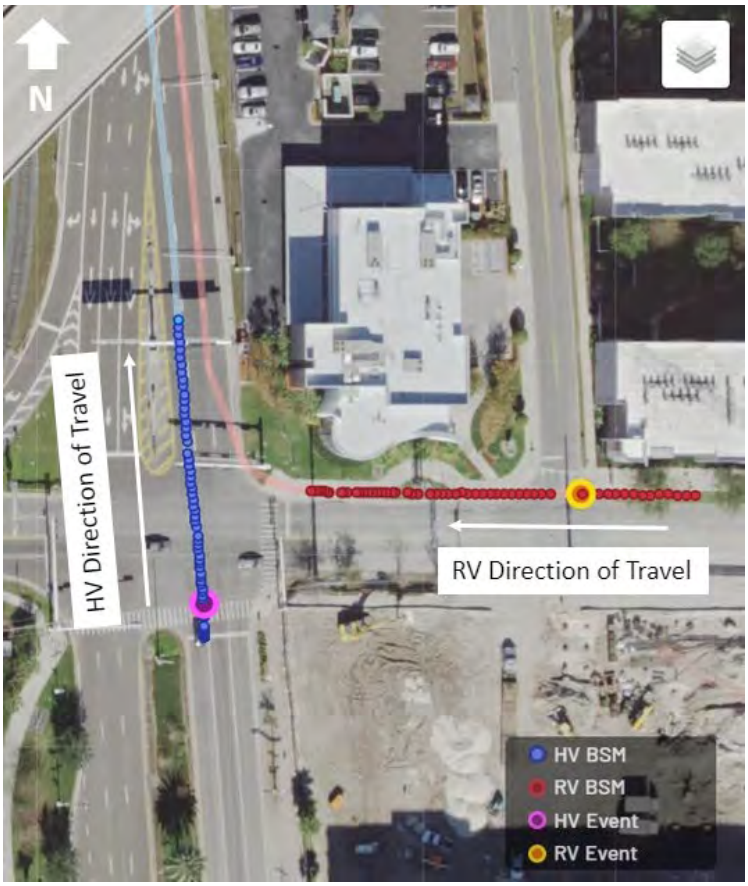
7.5.3.4.1 Factors Associated with False Positives

One reason that IMAs were determined to be false positives via visual inspection was the orientation of the two vehicles. For an IMA warning to be valid, the two vehicles must be approaching each other and not moving away from each other. Figure 7-77 shows an example of an IMA where the two vehicles are traveling initially in the same direction. The remote vehicle makes a left turn and then the host vehicle issues the warning. However, in this case, the remote vehicle is traveling away from the host vehicle without any real conflict. This occurred in 13 (42%) of the IMA warnings issued.

Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-77. FP IMA Due to False Orientation

Another reason for false positive IMAs was the large distance associated with the search zone of the host vehicle. Even though the vehicle paths were appropriate for an IMA and the speed thresholds were met, vehicles approaching intersections in an urban setting usually have to stop due to a traffic signal or stop sign. Figure 7-78 illustrates this type of IMA warning. Even if the parameters are met at the moment of warning, the time to intersection is large and the remote vehicle (or host vehicle) has time to stop for the stop sign or signal, which does not constitute a real danger. This situation occurred in eight (26%) of the IMAs issued to participants of the Pilot.

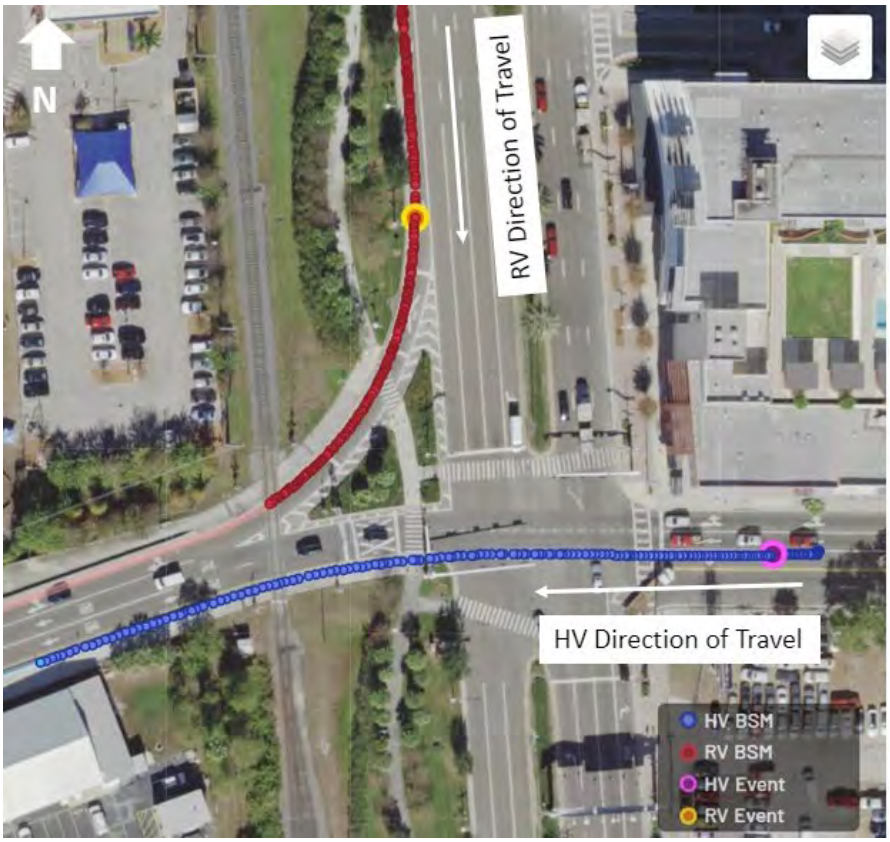


Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-78. FP IMA Due to Large HV-RV Distance

The final reason for false positives observed in the study was the application's inability to capture the nuances of channelization at intersections. In five (16%) of the IMAs, the warning was issued with all parameters met although the vehicles followed a path that did not intersect, such as merging in the same road, or one vehicle turned before the intersection. Figure 7-79 shows an example of such a warning. The two vehicles merge on the same road traveling approximately at the same time, but their paths never intersect, which does not signify a real danger for either vehicle.

The remaining two IMAs where issued to vehicles due to GPS shift (i.e., one of the vehicles was sending the wrong location information). This is evident by examining the vehicle trajectories during a visual inspection.



Source: THEA CV Pilot Performance Evaluation Dashboard, 2020

Figure 7-79. FP IMA Due to No Intersection between HV-RV Path

Researchers used the same process to evaluate IMA as the one adopted for the FCW application. A larger polygon (shown in Figure 7-80) was applied to capture the IMA interactions between vehicles traveling on Meridian Avenue and vehicles traveling on side streets. To avoid the inclusion of interactions where host and remote vehicles traveled in the same direction, the HV-RV heading difference values listed in Chapter 6 were applied. The heading difference range values (90 ± 25 degrees) were chosen by considering Meridian Avenue's orientation (North-South) and the side street angles as they approach Meridian. Having obtained the count of IMA interactions, HV and RV minimum speed values were applied to determine potential IMA conflict counts.

Source: CUTR, 2020

Figure 7-80. Polygon Used for IMA Assessment

The next step in the conflict assessment was to create BSM profiles anchored around potential IMA conflicts and triggered warnings. The profiles were visually checked to examine any HV-RV situations posing an imminent threat of the two vehicles intersecting.

Table 7-44 shows the estimated interactions and conflicts, as well as warnings issued, following the approach detailed above and in Chapter 6 of this report for morning and afternoon hours. Vehicles traveling in the UC6 study area had an estimated 12,513 interactions with six conflicts conformable to IMA deployment with no true positives. As a result, the application's overall false positive rate is estimated at 0.22 percent, while the false negative rate is estimated at 100 percent.

Table 7-44. IMA Movement Classifications and Rates

Description	AM Period		PM Period	
	Count	Rates (%)	Count	Rates (%)
Time Spent in Area (Hours)	282	--	278	--
Number of Vehicles	450	--	452	--
IMA (TP + FP + Not Tested)	16	--	15	--
V2V Interactions	8,490	--	4,023	--
Conflicts	1	--	5	--
True Positives	0	0.0	0	0.0
False Negatives	1	100.0	5	100.0
Non-conflicts	8,489	--	4,018	--
True Negatives	8,473	99.8	4,006	99.7
False Positives	16	0.2	12	0.3

7.5.3.5 V2V Interactions and Conflict Assessment

The Performance Measurement and Evaluation Support Plan's recommended method of safety evaluation considered the use of either a full random or a quasi-experimental design approach. This is because the participant recruitment conceived for the Pilot grouped participants into treatment and control groups. In this context, the treatment group includes participants who have the Human Machine Interface enabled (warnings displayed on the mirror), while the control group includes participants with fully functional OBUs and all CV applications installed, but with the HMI disabled (warnings not displayed on the mirror).

Ultimately, the goal of the conflict assessment to measure safety performance was to compare the behavioral responses of the treatment and control groups under similar conditions. In addition, the deployment considered three months of data collection during which the treatment group was placed in silent mode. During this time, all warning events generated by the treatment group were recorded in OBUs but not displayed via HMI. At the end of silent mode, the HMI was turned on (using OTA protocols) for the entire duration of Phase 3 of the CV Pilot. This staged approach allowed additional insight into the before-after behavioral responses of the same subjects.

Table 7-45 reports the warnings generated and displayed to participants by the experimental design group. Out of 55 FCW events, 33 percent were shown to drivers and 67 percent were not shown. Out of nine EEBL events, 33 percent were shown to drivers and 67 percent were not shown. Out of 31 IMA events, 39 percent were shown to drivers and 61 percent were not shown.

Table 7-45. Warning Visibility by Participant Group

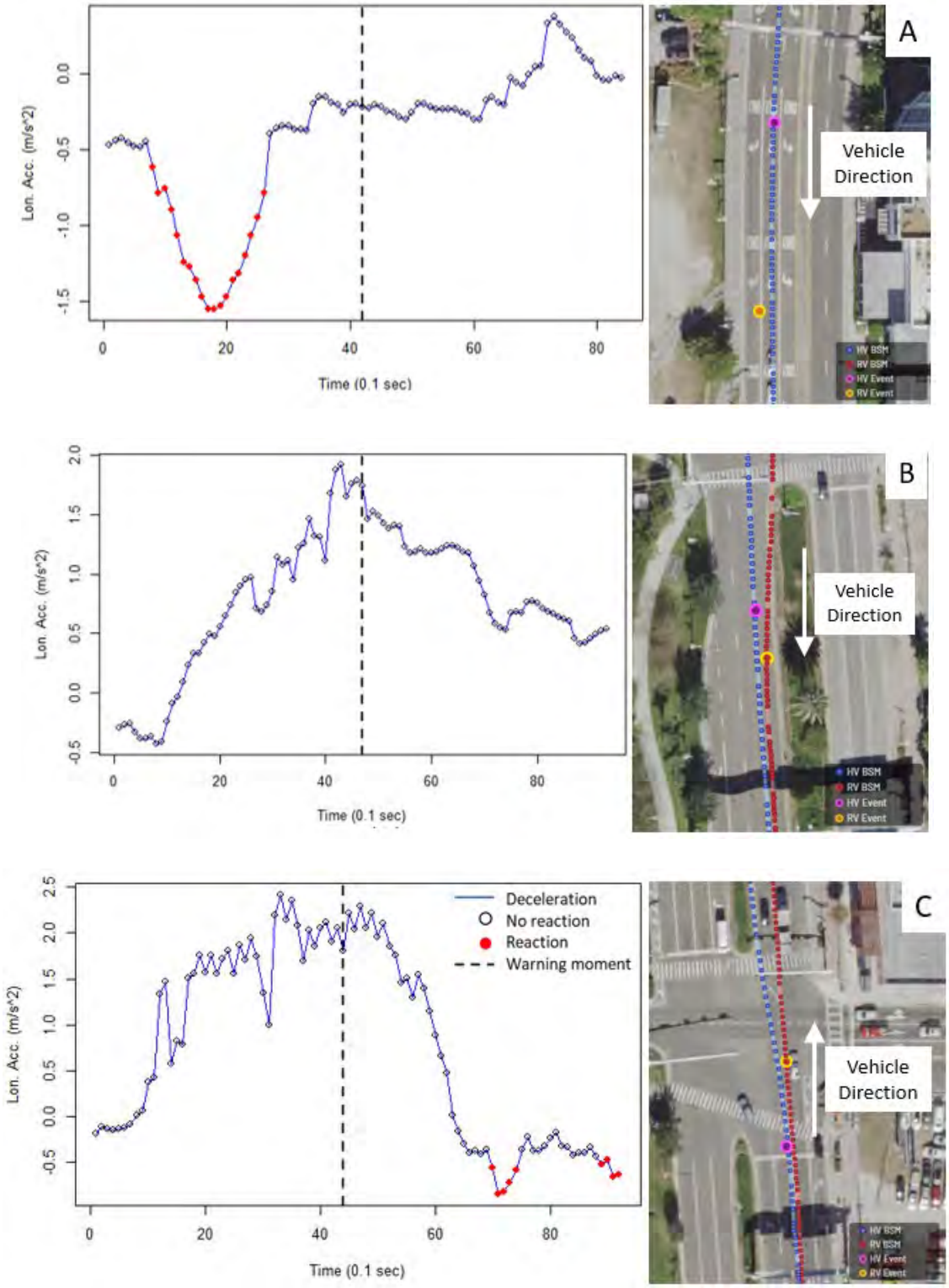
Application	Group	HMI Disabled (Warnings Not Displayed)			HMI Enabled (Warnings Displayed)			Grand Total
		TP	FP	Share %	TP	FP	Share %	
FCW	Control	7	16	41.8	0	0	0.0	
	Treatment	0	0	0.0	6	10	29.1	
	Treatment (Silent)	8	6	25.5	1	1	4	
	FCW Total	**15**	**22**	**67.3**	**7**	**11**	**32.7**	**55**
EEBL	Control	3	1	44.4	0	0	0.0	
	Treatment	0	0	0.0	0	2	22.2	
	Treatment (Silent)	0	2	22.2	1	0	11.1	
	EEBL Total	**3**	**3**	**66.7**	**1**	**2**	**33.3**	**9**
IMA	Control	0	14	45.2	0	0	0.0	
	Treatment	0	0	0.0	0	9	29.0	
	Treatment (Silent)	0	5	16.1	0	3	9.7	
	IMA Total	**0**	**19**	**61.3**	**0**	**12**	**38.7**	**31**

Next, the analysis assessed the behavioral responses by comparing the reactions of the treatment and control groups and the responses within the treatment group (HMI disabled versus HMI enabled). The driver reaction was investigated separately for each V2V application. The conflict identification algorithm identifies a reaction to a conflict if the driver decelerates at a rate below 0.5 mps^2 after the moment of warning. Since drivers facing an FCW or EEBL situation could also change lanes instead of decelerating, this was also visually checked to ensure they did not simply change lanes with no deceleration.

7.5.3.5.1 Treatment vs. Control Group Reaction to TP Warnings

The control group generated a total of 10 warnings classified as true positive events based on the identification steps outlined in Chapter 6. OBUs recorded 22 Forward Collision Warnings classified as TP. Of those, seven events were triggered by an HMI-enabled unit (treatment) and seven by an HMI-disabled unit (control). Eight true positive FCWs were not displayed to participants in the treatment group during silent mode.

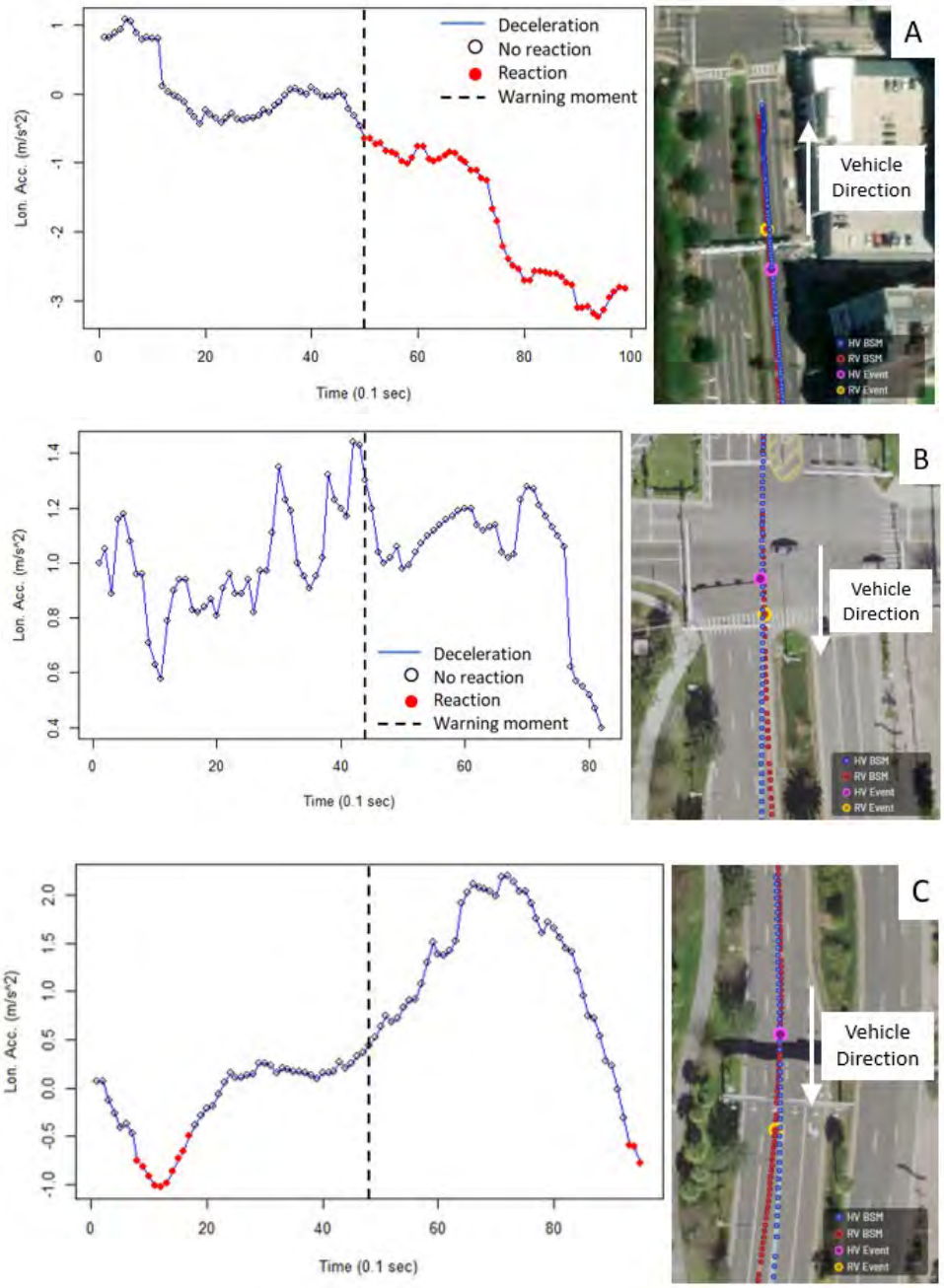
Figure 7-81 (A-C) displays the results of the data mining algorithm presented in Chapter 6. The left diagram shows the deceleration profile of each FCW event classified as true positive and generated by an HMI-enabled vehicle using the 10-second BSM profile. The map next to the diagram shows the host and remote vehicle trajectories and location of the warning moment. The participant in Figure 7-81 (A) reacted before the warning, the participant in Figure 7-81 (B) did not react, and in Figure 7-81 (C), the participant reacted after the warning. Overall, for true positive FCW events, participants reacted in two warnings, two others did not react, and two reacted before the warning.



Source: CUTR, 2020

Figure 7-81. Driver Reaction to TP FCW with HMI Enabled

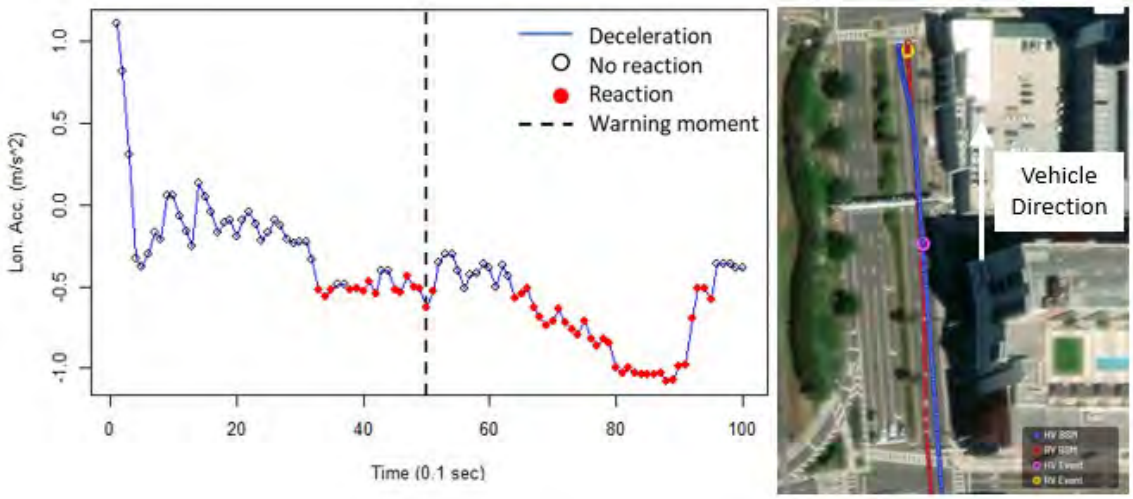
Figure 7-82 (A-C) displays the results of the data mining algorithm applied to the seven true positive warnings triggered by participant vehicles assigned to the control group with HMI disabled. The driver in Figure 7-82 (A) reacted after the warning moment, the driver in Figure 7-82 (B) did not react, and the driver in Figure 7-82 (C) reacted before the warning moment. Out of seven true positive FCWs with HMI disabled, two drivers reacted after the warning, three drivers reacted before, and two did not react.



Source: CUTR, 2020

Figure 7-82. Driver Reaction to TP FCW with HMI Disabled

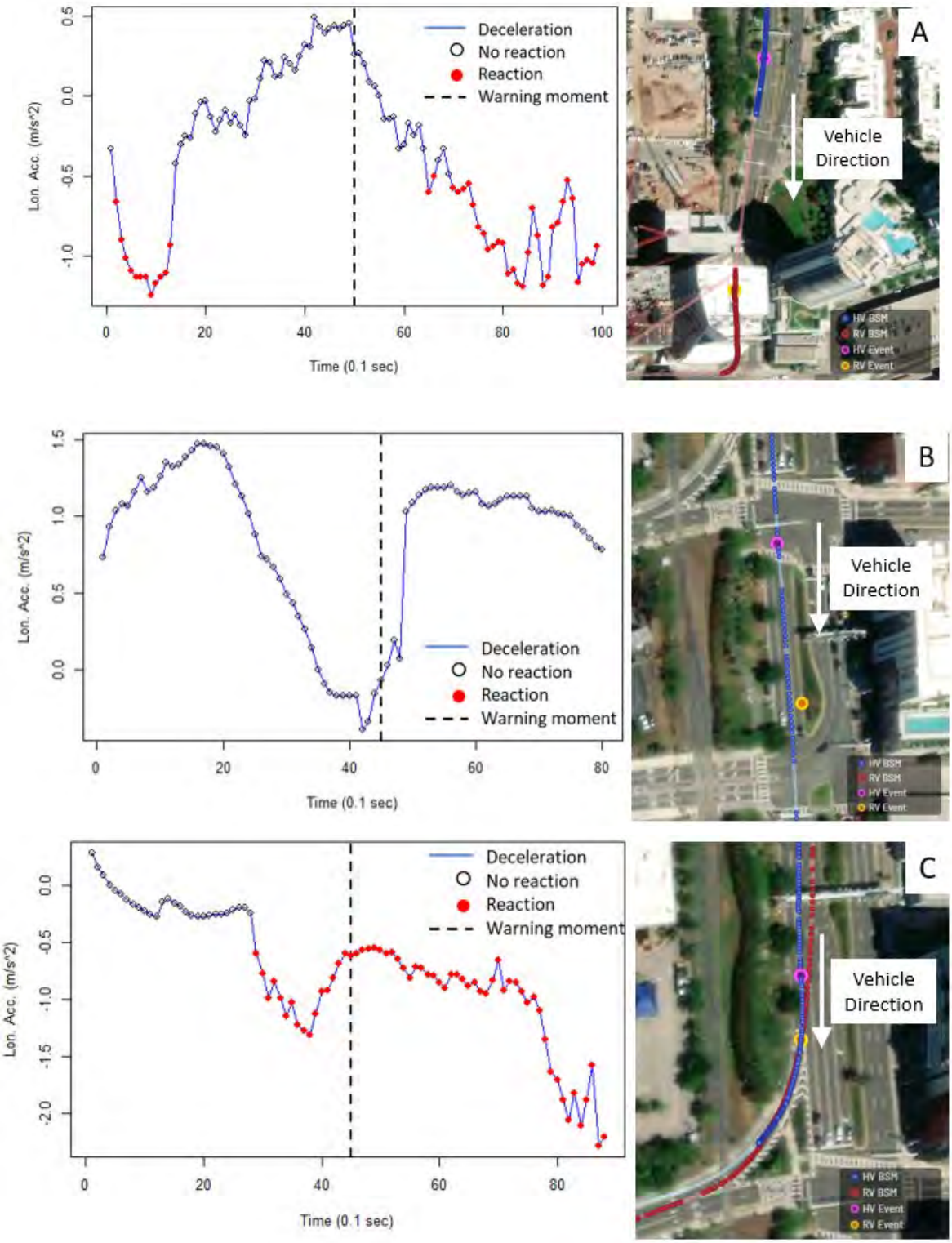
Of the four true positive EEBLs, three were issued to the control group with HMI disabled and one was issued to a participant in the treatment group. Figure 7-83 shows the reaction of the participant in treatment group.



Source: CUTR, 2020

Figure 7-83. Driver Reaction to TP EEBL with HMI Enabled

Figure 7-84 (A-C) shows the driver reactions to the three true positive EEBLs with HMI disabled. In Figure 7-84 (A), the driver reacted before the warning moment, in Figure 7-84 (B), the driver did not react, and in Figure 7-84 (C), the driver reacted before the warning.

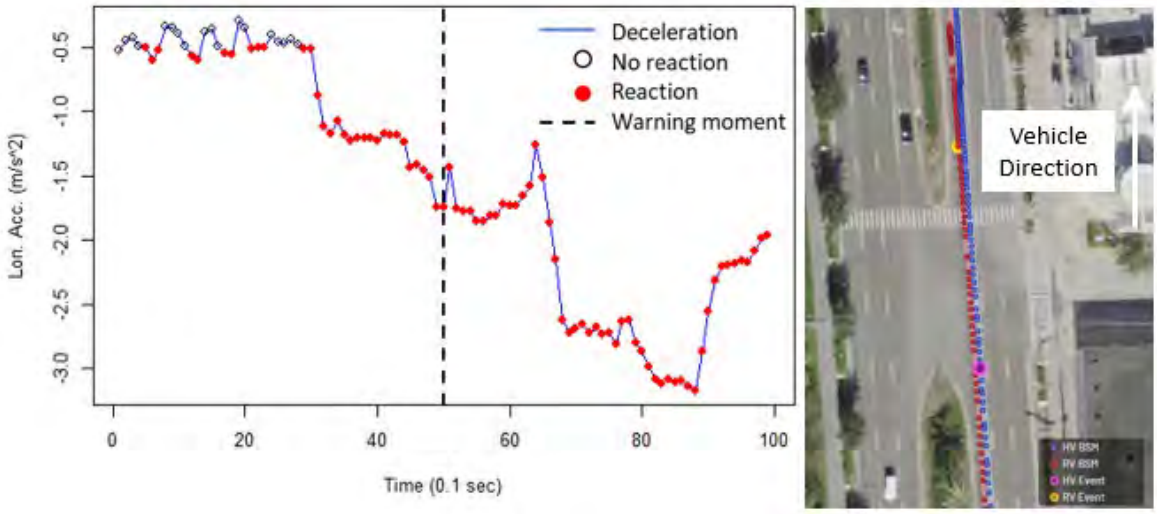


Source: CUTR, 2020

Figure 7-84. Driver Reaction to TP EEBL with HMI Disabled

7.5.3.5.2 Within Group Reaction to TP Warnings

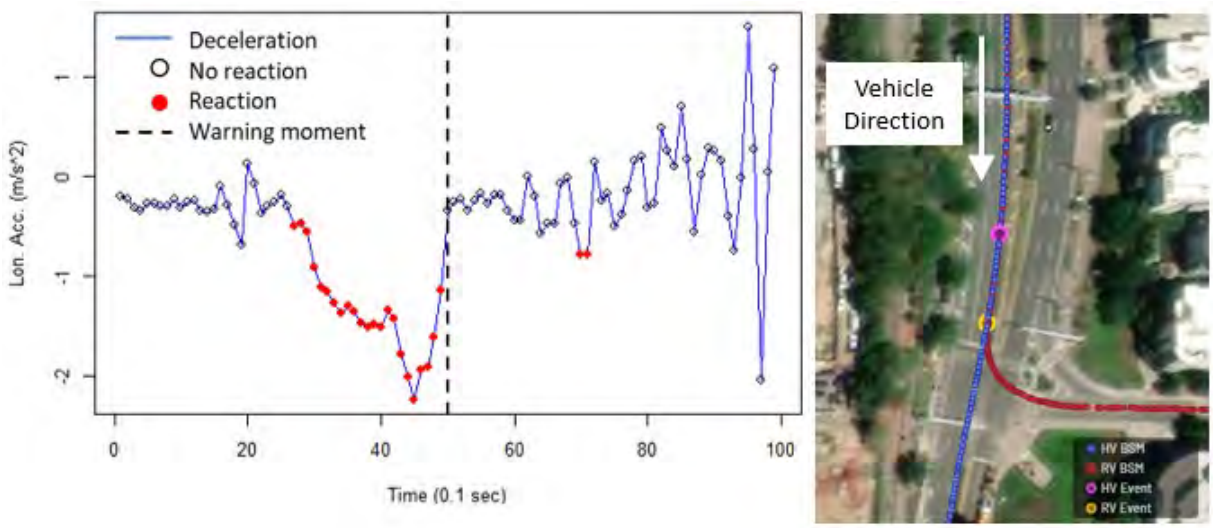
To assess the effect of the warnings on the treatment group, the silent mode was used to determine if drivers reacted differently in true positive events with HMI enabled and HMI disabled. Nine true positive FCW events were recorded with this classification, eight with HMI disabled and one with HMI enabled. Figure 7-85 shows the FCW with HMI disabled during the silent period. This driver began reacting before the warning moment and hard braking reached values of -3 mps^2 or 0.3 G (g-unit). For the eight true positives in silent mode, seven showed reaction before the warning moment and one had missing data for reaction assessment.



Source: CUTR, 2020

Figure 7-85. Driver Reaction to TP FCW with HMI Disabled

The only true positive FCW with HMI enabled after the silent period ended is presented in Figure 7-86. The driver reacted before the warning moment.



Source: CUTR, 2020

Figure 7-86. Driver Reaction to TP FCW with HMI Enabled

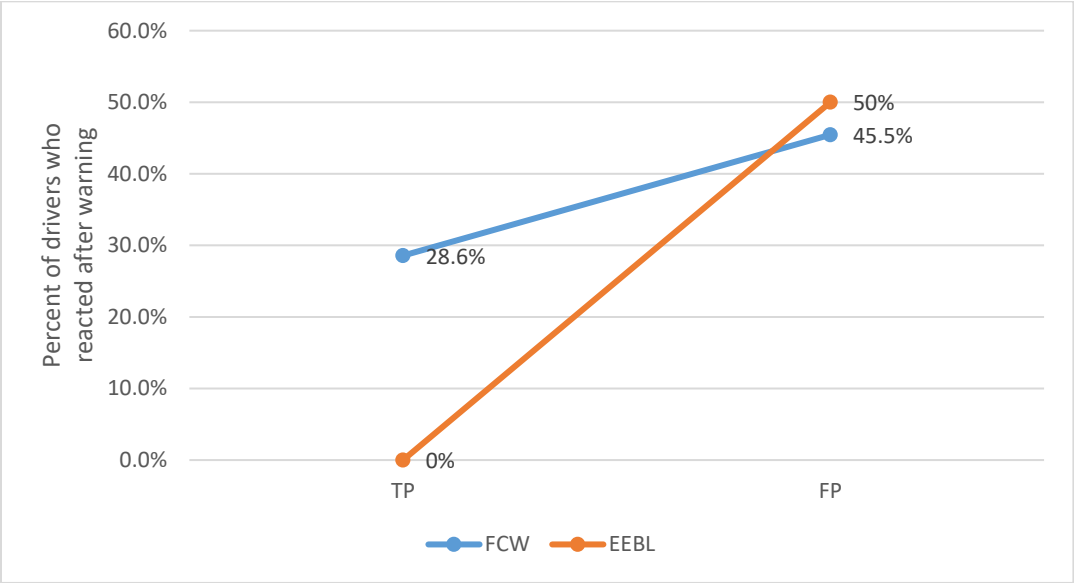
These results illustrate the nuances of having an application determine the conditions correctly and trigger warnings in a timely manner so that drivers can react to them. In the case of true positive FCWs for the treatment group, all drivers began reacting before the warning, most likely to the real conflict situation they faced. In this case, it seems the warnings were delivered late since all drivers began reacting before the warning moment.

7.5.3.5.3 Driver Reaction to All Visible Warnings (HMI Enabled – All Groups)

The warnings shown to participants were generated under two types of conditions. The first condition involves a true conflict (TP) with another vehicle determined by the parameters and the method explained in the previous sections for each application. The second condition involves a false positive (FP) where the warning is triggered when there is no conflict with another vehicle. It is possible for the participant to respond under both conditions, albeit differently. Next, the analysis pools all the warnings recorded while the HMI was enabled to evaluate any difference in behavioral response by type of conflict (TP vs. FP).

This analysis is relevant because of the travel conditions underlying the generation of these warnings. During the weekday morning travel, participants driving on Meridian Avenue can come near one or more participants while at the same time being surrounded by other non-equipped vehicles. This means that a participant seeing a warning (even if classified as FP), might generate an observable reaction since there could be a non-CV ahead causing a conflict situation.

Figure 7-87 shows the difference in the share of drivers who reacted after they received a visible warning, grouped by TP and FP classification. Twenty-nine percent of participants reacted to FCW warnings classified as TP compared to 46 percent who reacted to warnings classified as FP. Fifty percent of participants reacted to false positive EEBLs and zero percent to TPs. No true positive IMAs were recorded.



Source: CUTR, 2020

Figure 7-87. Proportions of Drivers Reacting to FCW with HMI Enabled

7.5.3.5.4 Predictive Value of Warnings

Since crashes and conflicts are rare traffic events, it is useful to examine two additional metrics in the assessment of V2V safety applications: the positive and negative predictive values of warnings. Given the many factors affecting the deployment and efficacy of CV safety applications, these measures evaluate the conditional probability of being in a dangerous situation when a warning is triggered (positive predictive value) or not being in a dangerous situation when a warning is not triggered (negative predictive value).

In this context, the probability that a warning will be a true positive (conflict situation) is the proportion of TP warnings to all warnings (TP + FP). This is a true performance measure of an application since it gauges how likely a driver is to be in a conflict when a warning is received.

For the FCW application, the positive predictive value is calculated as

$$FCW\ positive\ predictive\ value = \frac{TP}{TP + FP} \times 100 = \frac{22}{22 + 33} \times 100 = 40.0\%$$

Given the travel conditions characterizing Use Case 6, this means that if a Forward Collision Warning is triggered, the probability of being in a real dangerous situation is approximately 40 percent.

On the other hand, the probability that a negative (no warning) will be a true negative (no conflict) is the proportion of TN to all negatives (TN + FN). The negative predictive value of the FCW application as deployed in Use Case 6 can be calculated as

$$FCW\ negative\ predictive\ value = \frac{TN}{TN + FN} \times 100 = \frac{4714}{4714 + 124} \times 100 = 97.4\%$$

This means that if a Forward Collision Warning is not triggered (no warning), drivers are not in a dangerous situation 97.44 percent of the time, and drivers face a conflict only in 2.56 percent of the time. Table 7-46 shows the positive and negative predictive values for the three safety applications as deployed and evaluated specifically for Use Case 6. Being a function of location and travel characteristics, the above values might change if the evaluation is carried out at different locations within the CV Pilot study area.

Table 7-46. Application Predictive Values

Application	Positive Predictive Value (%)	Negative Predictive Value (%)
FCW	40.0	97.4
EEBL	44.4	99.4
IMA	0.0	99.9

7.5.4 Summary of Findings

7.5.4.1 Impact on Mobility

The analysis finds no evidence of an impact on mobility due to a lack of data caused by deployment constraints related to the I-SIG application's implementation.

7.5.4.2 Impact on Safety

Using the procedure described in section 6.4, the interactions and conflicts for all connected vehicles passing through the UC6 area were estimated over a baseline, or "before" period. The two periods for this study are therefore set as follows:

- Before Period: May 2018 – February 2019
- After Period: March 2019 – August 2020.

The before period begins when roadside unit BSMs became available as participant recruitment goals were achieved. The after period is the same as the warning and conflict assessment presented previously and begins at the inception of Phase 3 of the Pilot. Table 7-47 reports a summary of measures for each application.

Table 7-47. Conflict Rates for UC6

	Before Period			After Period		
General Measure	**AM**	**PM**	**Total**	**AM**	**PM**	**Total**
Time Spent in Area (Hours)	255	164	419	282	278	560
Number of Vehicles	453	377	533*	450	452	548*
FCW Measure	**AM**	**PM**	**Total**	**AM**	**PM**	**Total**
V2V Interactions	5369	1429	6798	3237	1656	4893
Conflicts	84	32	116	85	61	146
Conflict Rate (Conflicts/Vehicle/Year)	6.37	4.53	4.55	5.87	4.25	4.17
Conflict Rate (Conflicts/Vehicle)	0.19	0.08	0.22	0.19	0.13	0.27
EEBL Measure	**AM**	**PM**	**Total**	**AM**	**PM**	**Total**
V2V Interactions	4261	750	5011	2517	299	2816
Conflicts	27	2	29	18	4	22
Conflict Rate (Conflicts/Vehicle/Year)	2.05	0.15	2.20	1.37	0.30	1.67
Conflict Rate (Conflicts/Vehicle)	0.06	0.01	0.05	0.04	0.01	0.04
IMA Measure	**AM**	**PM**	**Total**	**AM**	**PM**	**Total**
V2V Interactions	12132	2765	14897	8490	4023	12513
Conflicts	0	1	1	1	5	6
Conflict Rate (Conflicts/Vehicle/Year)	0.00	0.08	0.08	0.08	0.38	0.46
Conflict Rate (Conflicts/Vehicle)	0.00	0.00	0.00	0.00	0.01	0.01

** Vehicles reported interactions and conflicts in both AM and PM periods.*

In evaluating Use Case 6, the three safety applications generated 26 warnings classified as true positive, but only 8 were shown to drivers due to the evaluation's experimental design. CUTR researchers developed and

implemented a data-driven method to evaluate driver reactions to these warnings. The results indicate that out of the eight true positives with HMI enabled, the drivers reacted after the warning was issued in only two instances. For the remainder, the participants show no reaction or reacted before the warning. This illustrates the challenge of fine-tuning the application parameters to first correctly determine a conflict and then issue the warning in a timely manner for the driver to react. The variations in driver characteristics (perception-reaction time) and the vehicle's ability to stop (brake performance) create challenges in establishing global parameters that work for all drivers and a broad set of conditions.

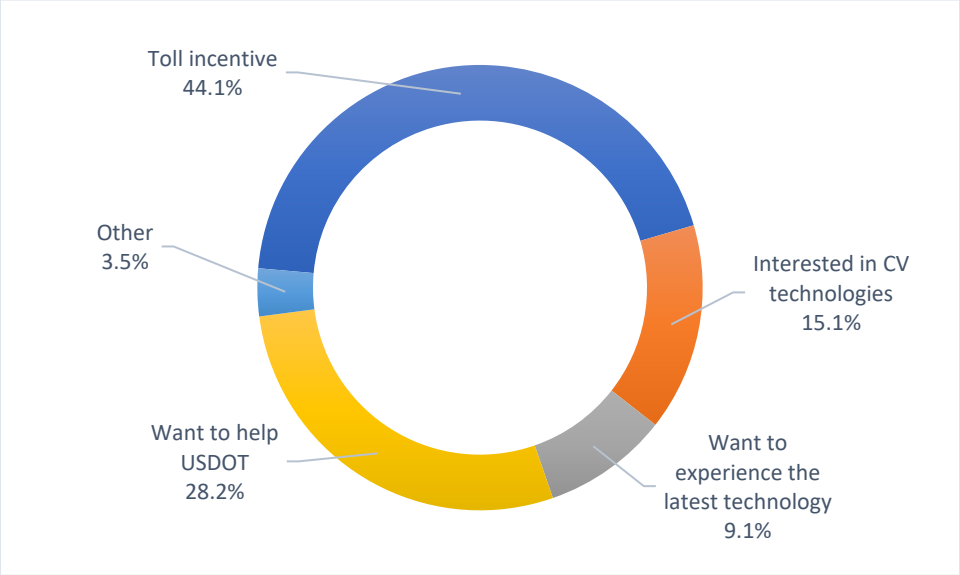
In addition, the IMA application seems to have not worked for this use case as the warnings issued were either too far or for non-conflicting vehicles. The current settings of the IMA application might produce better results in non-urban environments, suggesting the need for further fine-tuning more suitable to the congested roads of a central business district.

7.6 Participant Surveys

7.6.1 Initial Survey

Each participant scheduled for installation completed a survey administered at the installation site. A total of 1,058 responses were collected, including participants who dropped because either they expressed no intension of having equipment installed in their vehicles or were excluded from participation due to the vehicles not meeting the minimum study requirements. The survey collected information on the participants travel profile, their perceived knowledge of CV technologies, and their reasons to participate in the study. Figure 7-88 shows that the toll rebate incentive was the main reason to join (44.1%), followed by the intent to help USDOT advance CV research (28.2%) and an inherent interest in CV technology (15.1%).

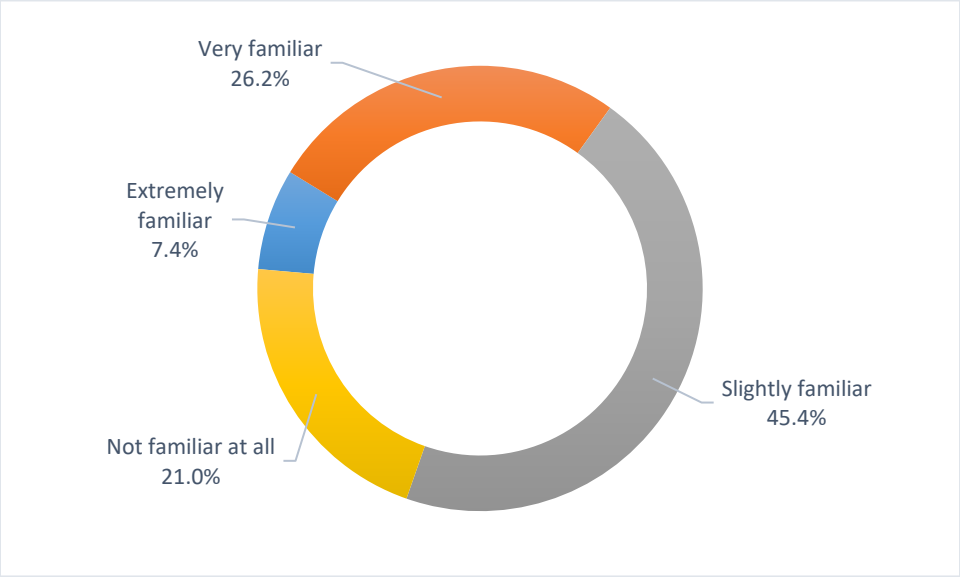
Figure 7-88. Participants' Reasons to Join the CV Pilot



Source: CUTR, 2020

About 21 percent of the participants were not familiar at all about CV technologies, as shown in Figure 7-89. Cross-tabulation of these answers with those of Figure 7-88 reveals that 61 percent of those not familiar with

CV technology joined because of the toll rebate and 72 percent of those familiar with CV technology did so because of their interest in the technology and willingness to experience the latest technology.



Source: CUTR, 2020

Figure 7-89. Knowledge of CV Technologies

In terms of commute patterns, 75 percent of the participants travel at least five days a week and 89 percent commute to downtown during the morning peak hour (6 to 9 a.m.). About 67 percent of the participants use the REL three to five times a week.

7.6.2 Exit Survey

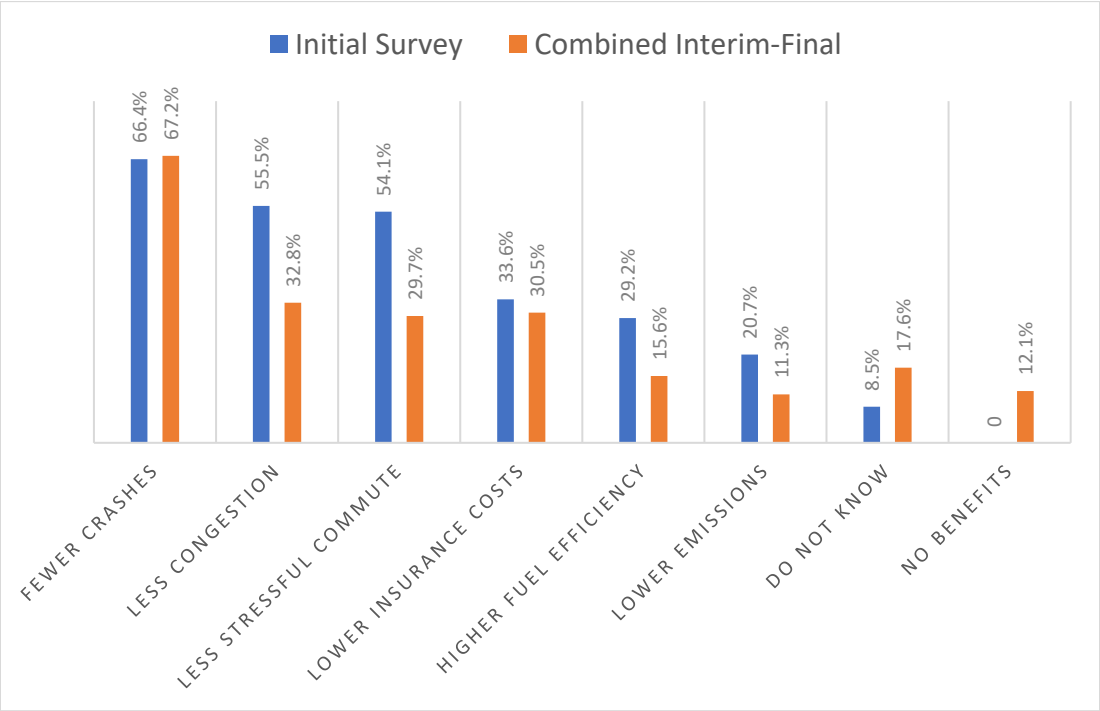
Participants leaving the study at any point during Phase 3 were asked to complete an exit survey. A total of 369 participants dropped since inception and through September 30, 2020. Out of these, 132 participants (35.8%) completed the exit survey. The main reason for leaving the study was personal (i.e., relocating, trading the vehicle, change of job and commute route) as indicated by 42.4 percent of the respondents. About 17.8 percent pointed to CV-related issues as a reason to drop off the study. This included difficulties in using the rearview mirror due to either the application not working (i.e., showing a green screen or bars) and difficulties in seeing at night due to the LCD display.

7.6.3 Interim and Final Surveys

The interim and final surveys were pooled together to analyze the responses to the same questions in both surveys that were geared at understanding the participants' perceived benefits and concerns about CV technologies, as well as to understand how they perceived the efficacy of the V2V and V2I applications deployed. The interim collected responses from 389 participant and the final survey obtained 256 responses for a sample size of 645 participants out of 1,012 initial installations (63.7%).

7.6.4 Participants Perceptions about CV Technologies

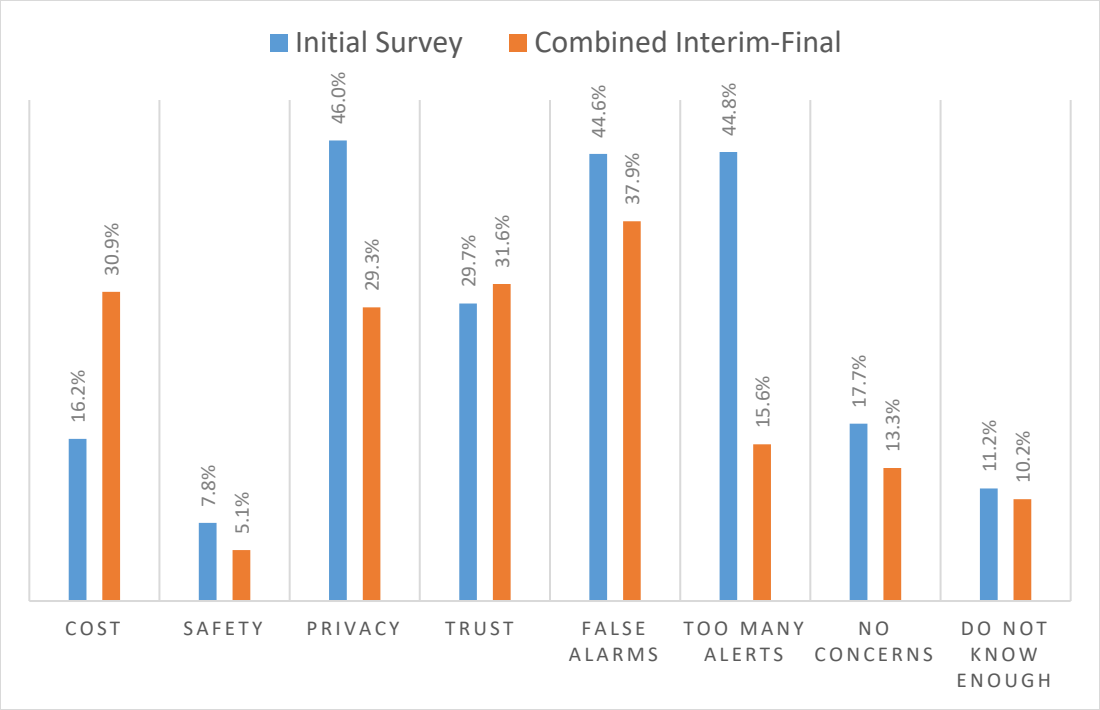
The initial survey at installation asked two questions to ascertain the participant's perceived benefits and concerns associated with CV technology. The goals were to subsequently compare the responses as participants enrolled into the study and were exposed to the technologies. Figure 7-90 compares the responses about perceived benefits between the initial and the interim/final surveys. Participants perceived safety as the greatest benefit of CV technologies and this perception slightly changed as they participated in the study. On the other hand, respondents' perceptions about the impact of CV technologies in reducing congestion and improving the commute experience decreased. This could be due to other factors playing a role in determining these perceptions, such as the currently low CV penetration rates and the localized impact of the speed harmonization application (ERDW). This is confirmed by the ensuing analysis of the perceived effectiveness of each application.



Source: CUTR, 2020

Figure 7-90. Perceived Benefits Before-After Participation

Participants were asked to express their opinion about a series of concerns associated with CV technologies. Before actively participating in the study, about 46 percent of the participants expressed concerns about the impact on their privacy. As they participated in the study, these concerns reduced, as about 29 percent of participants expressed some concern. Other relevant changes between the two periods were related to the reduced concerns about the applications generating false alarms and too many alerts. On the other hand, the concern about the costs of CV technologies increased with about 31 percent of respondents being concerned in the interim and final phase of the project compared to 16.2 percent at the beginning of the study (see Figure 7-91). These responses might have been affected somewhat by reliability issues associated with the aftermarket units requiring participants to be called back for repairs and OBU swaps.



Source: CUTR, 2020

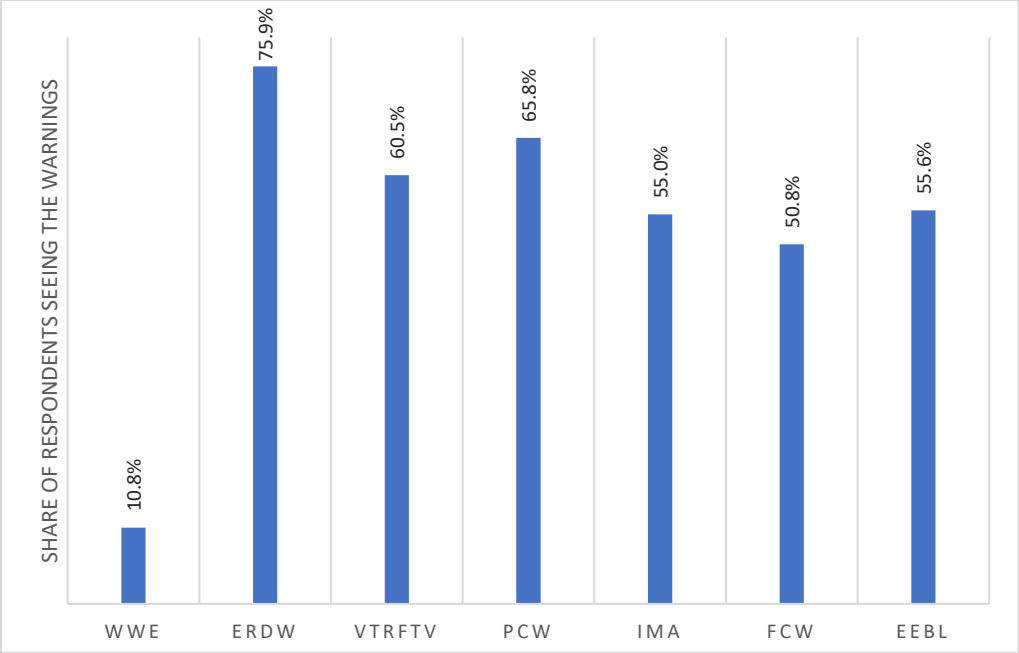
Figure 7-91. Perceived Concerns Before-After Participation

7.6.5 Participants Experience with V2V and V2I Applications

In the interim and final surveys, participants were asked questions related to the deployment of the applications with the intent to gauge how they benefitted from each app and how each app performed. The following figures are based on the answers from participants assigned to the treatment group who had the HMI turned on and were able to see and hear the application warnings.

Figure 7-92 reports the percent of respondents who answered yes to the question about each application helping in avoiding a conflict situation. For example, a total of 130 respondents in the combined interim/final survey stated they experienced the WWE application and about 10.8 percent (14 participants) determined the WWE to be effective in avoiding a conflict. As discussed in Chapter 5, this share can be explained by the WWE application performance and by the low probability of being exposed to a conflict of this type. The ERDW speed advisory is perceived to be effective by 75.9 percent of those receiving it.

Figure 7-93 shows the share of participants receiving the warnings who answered that it was clear to them they received a warning. On average, about 70 percent confirmed the clear administration of the warning, but with variation among the apps. The WWE and IMA app warnings were relatively less clear. Note that the WWE entry shows multiple icons on the mirrors as the application is deployed, perhaps negatively affecting its delivery.



Source: CUTR, 2020

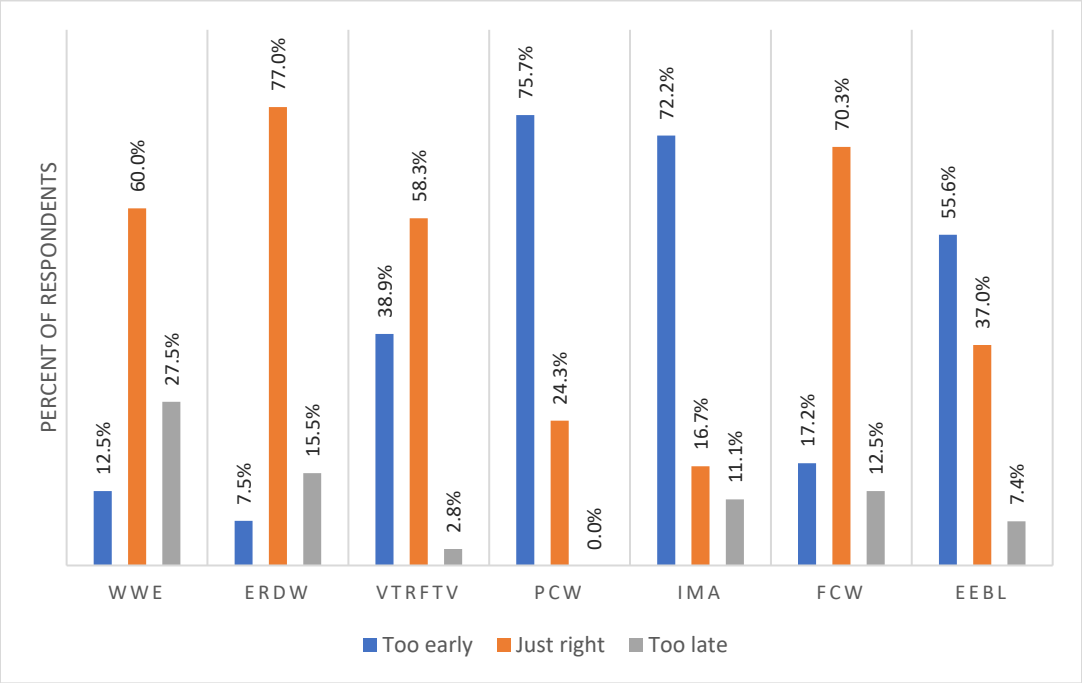
Figure 7-92. V2V and V2I Participant Perceived Application Effectiveness

Source: CUTR, 2020

Figure 7-93. Application Warning Was Clear

In terms of application effectiveness, the participants receiving the warnings were asked about the timely delivery of the warning (Figure 7-94). The responses show variability among the applications. While the WWE, ERDW, and FCW were received at the about the right time, the PCW, IMA, and EEBL were perceived as

deploying too early. This information is relevant when put in context of the warning profile analysis of Chapter 5 and response in the presence of a conflict situation.

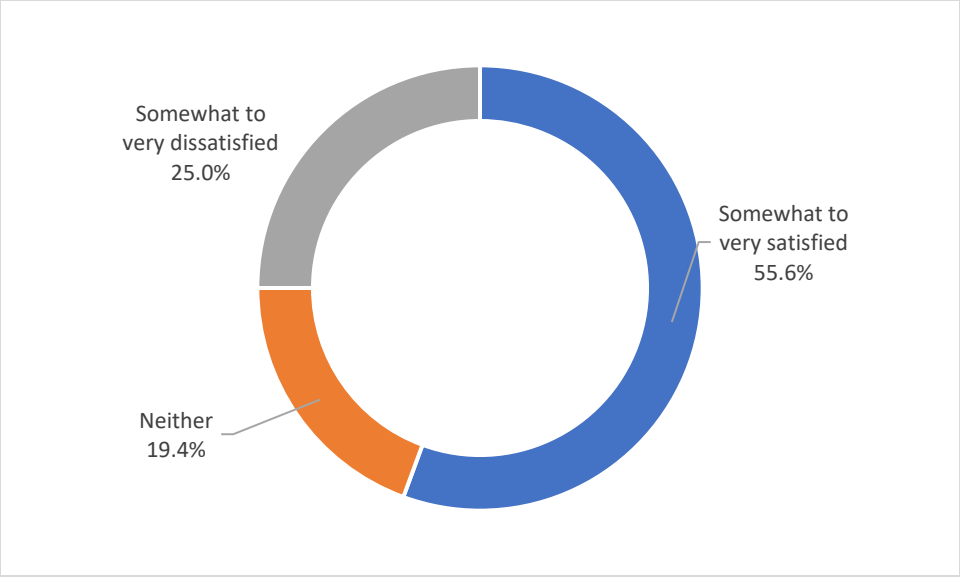


Source: CUTR, 2020

Figure 7-94. Application Timing

7.6.6 Overall Satisfaction with the Study

The exit, interim, and final surveys included a question to gauge participant satisfaction about the entire study. Figure 7-95 shows that about 55.6 percent of the participants were somewhat to very satisfied about the study, with 25.0 percent expressing some form of dissatisfaction. There is an inherent bias that must be accounted for in analyzing the results from this question that is specific to the adopted experimental design that split the participant pool into treatment (HMI on) and control (HMI off) groups. Those participating in the study and part of the control group might not have perceived any incremental benefits because the HMI was turned off, thus affecting their participation to answering these questions. Table 7-48 splits the response rates of Figure 7-95 between the two groups, showing that participants with the HMI on were overall more satisfied about the study than participants who had the HMI off.



Source: CUTR, 2020

Figure 7-95. Overall Participant Satisfaction with the Study

Table 7-48. Participant Satisfaction by Experimental Group

Satisfaction	HMI Off (%)	HMI On (%)	Total (%)
Somewhat to Very Satisfied	46.5	65.8	55.6
Neither	21.8	16.7	19.4
Somewhat to Very Dissatisfied	31.7	17.5	25.0
Total	**100.0**	**100.0**	**100.0**

Chapter 8. Conclusions

At the time of this report, the Tampa CV Pilot is the only site to have deployed CV technologies simultaneously on private vehicles, buses, streetcars, and city fleets. The study implemented a robust experimental design to enroll more than 1,000 private citizens to experience the functionality and impact of several V2V and V2I applications. Participants were split into treatment (HMI enabled) and control (HMI disabled) groups using a randomized two-to-one matching (two treatment to one control) stratified by gender, age, income, and education. The study continuously generated data for over 18 months and currently continues to do so. The extensive data collection and sharing efforts will serve to inform USDOT policy making and the research community on the benefits of these technologies.

The performance evaluation of this study focused on measuring, reporting, and assessing the impact and contribution of V2V and V2I applications toward enhanced mobility and safety by adopting a use case approach. The study evaluated six use cases, each geared at solving mobility and safety issues in specific areas of downtown Tampa and a key section of its roadway network managed by the CV Pilot deployer.

Overall, findings provide evidence that the deployment contributed to the enhanced mobility and safety of travelers and pedestrians. The evaluation also revealed some application-specific issues that can be resolved by improving the currently deployed OBU firmware and further research and development.

This chapter presents the key findings from the use case evaluation, outlines key limitations, and summarizes lessons learned.

8.1 Findings from Use Case Evaluation

8.1.1 Application Effectiveness

To assess the impact on safety, CUTR evaluated each application effectiveness via a false positive assessment. In addition, contingent upon data availability, the team estimated four performance rates: false positive, false negative, true positive, and true negative rates. The applications were evaluated with respect to planned performance measures adopted for each use case.

Table 8-1 lists the applications ultimately deployed in Phase 3 of the Pilot. As noted, Use Case 4 (Transit Signal Priority) did not produce data for analysis as the TSP application was not successfully deployed at the time this report was produced. Also, Use Case 6 (Traffic Progression) did not produce data to compare the mobility benefits on a before-after basis because the I-SIG had issues in working with the traffic controllers and was not deployed at the intersections. Therefore, this study did not evaluate these two applications against the planned performance measures.

Table 8-1. Application Deployment Summary

	UC1	UC2	UC3	UC4	UC5	UC6
EEBL	Deployed					Deployed
ERDW	Deployed					
FCW	Deployed					Deployed
IMA						Deployed
I-SIG	Not Operational					Not Operational
PCW			Deployed			
TSP				Not Operational		
VTRFTV					Deployed	
WWE		Deployed				

8.1.1.1 Impact on Mobility

The system impact evaluation showed marginal improvements on mobility for Use Case 1. The broadcasting of speed advisories via the ERDW application contributed to improvements compared with the baseline (before ERDW deployment) conditions. The ERDW contributed to:

- 2.1 percent reduction in mean travel times with respect to the baseline (pre-intervention period)
- 1.8 percent reduction in idle time or time spent traveling at less than one mile per hour.
- 1.8 percent reduction in queue length
- A travel time index reduction from 2.7 to 1.9.

The above findings can be considered short-term impacts since the evaluation period was only 35 weekdays (Feb 3, 2020 to March 20, 2020) that the improved ERDW application was deployed and limitations in the data induced by the changes in travel behavior caused by the COVID-19 pandemic. Another factor to consider is the relative CV penetration rate, which plays a key role in affecting the magnitude of such benefits. Further data collection and analysis could help assess whether the improvements are temporary or longer lasting.

8.1.1.2 Impact on Safety

The analysis shows variability in how the V2V and V2I applications contributed toward improved safety:

- Use Case 1 showed that the FCW rate of conflicts per vehicle interaction did not change between before and after periods, with a rate of conflicts of 0.6 percent before and after FCW deployment. In the case of EEBL, results showed an increase in the rate of conflicts from 0.5 percent (before) to 0.9 percent (after).
- Use Case 2 demonstrated that the WWE performance varied according to the commuter peak travel flow:

- In the PM peak period, the application correctly warned drivers of entering the wrong way and identified 14 participants of 19 potentially true conflicts. The analysis found that the PM peak period is characterized by complex turning movement to access the REL while driving eastbound. GPS signal inaccuracy, potential participant moral hazard in undertaking turns into the REL, and the complexity of the warning delivery system contributed to a false positive rate of about 28 percent.

- On the other hand, the AM period did not experience a single wrong-way occurrence during the entire deployment, but the application generated a high number of false positives. Most of these false warnings could be reduced by adjusting the OBU firmware that incorrectly flags the MAP message identifying the revoked lanes and the applications' failure to report correct vehicle heading while caught up in the morning queue. The false positive rate for the AM period was 2.8 percent.

- Use Case 3 focused on improving pedestrian safety by deploying the PCW application. The analysis is inconclusive because the application deployment required changing pedestrian detection technology. The adoption of the new thermal sensor cameras came late into Phase 3, preventing the safety evaluation assessment due to a lack of participant data.

- Use Case 5 centered on the safety improvements on public transportation by deploying the VTRFTV application on local streetcars. The analysis found that the VTRFTV deployed 61 warnings, of which 8 (13.1%) were classified as TP during four unique events, but only 3 warnings (one event) were shown to the participant due to the evaluation's experimental design. The conflict detection algorithm confirmed the participant was not engaged in a conflict with the streetcar.

- Use Case 6 deployed FCW, EEBL, and IMA to improve the safety of commuters traveling through a busy arterial characterized by several signalized intersections:

    - The crash analysis showed that the percentage of rear-end crashes remained similar in the before and after periods, and that sideswipe crashes increased by 20 percent.

    - The rates of conflicts per vehicle normalized over time, showed a decrease for FCW (from 4.6 to 4.2), EEBL (from 2.2 to 1.7), and an increase in IMA (from 0.1 to 0.5).

    - The IMA application seems to have not worked for this use case as planned. This is because the warnings issued were either too far or in non-conflicting situations. The current settings of the IMA application might produce better results in non-urban environments, suggesting the need for further fine-tuning more suitable to congested roads of central business districts.

8.1.1.3 Participants Perceptions

In coordination with USDOT's Independent Evaluators, CUTR implemented a series of surveys to collect socio-demographic, travel behavior, and specific feedback from the use of the equipment and exposure to the applications. The survey analysis allowed gauging the overall satisfaction about the system and the contribution of the applications in terms of safety and mobility.

- Overall, participants were somewhat or very satisfied with participation in the study (56%). Nineteen percent were indifferent and 25 percent somewhat dissatisfied. Further stratification revealed that lack

of satisfaction was due to the experimental design, where those in the control group (HMI disabled) expressed more dissatisfaction due to the lack of interaction with the applications.

- At the beginning of the study, the majority of participants (66%) perceived safety as the greatest benefit of CV technologies, followed by expectations about reduced congestion (56%) and a less stressful commute (54%). The perception about safety did not change as they participated in the study, with expectations about reduced congestion and a less stressful commute reduced to 33 and 30 percent, respectively. This could be due to other factors playing roles in determining these perceptions, such as the localized impact of the speed harmonization application (ERDW) and the currently low CV penetration rates.

- With respect to how the applications performed, the survey responses show variability among the applications. While the WWE, ERDW, and FCW were perceived to be received at about the right time, the PCW, IMA, and EEBL were perceived as being issued too early.

- Before actively participating in the study, about 46 percent of participants expressed concerns about the impact on their privacy. As they took part in the study, these concerns lessened as about 29 percent of participants expressed some concerns.

- On the other hand, the concern about the costs of CV technologies increased with about 31 percent of respondents being concerned in the interim and final phase of the project compared to 16.2 percent at the beginning of the study. These responses might have been affected somewhat by reliability issues associated with the aftermarket units requiring participants to be called back for repairs and OBU swaps.

8.2 Study Limitations

As with any large deployment, the Tampa CV Pilot faced several challenges in deploying a technology that is relatively new and supplied by multiple vendors characterized by a high degree of variability in terms of research and development capabilities. The analysis and evaluation therefore have certain limitations as outlined below.

8.2.1 Performance Evaluation Goals

Not all the planned performance evaluation measures (see Chapter 2, Table 2-2) were actually implemented. As discussed in each use case's lessons learned, factors such as the lack or tardiness in implementing an application affected the data collection, thus preventing the measurement of a given performance metric. For example, the evaluation could not extend to consider the impact of CV technologies on the planned environmental metrics.

8.2.2 Study Area and Use Case Specific Evaluation

As outlined in the PMESP [4], the study was designed to evaluate the performance of the CV technologies and application in the context of six use cases. This approach, while instrumental to the Pilot assessment, limited the analysis to consider only vehicles that travel in a limited area concerning the deployment of specific V2I and V2V applications. This method therefore provides a focus for analysis but limits the use of data by narrowing the pool of participants who travel through the use case area. An alternative would be to evaluate

the impact of the application for the entire deployment site. This is one of the goals of the USDOT Independent Evaluators measuring the impact of this and other CV Pilot deployments and beyond the scope of this study.

8.2.3 Technology Readiness

8.2.3.1 Aftermarket OBU Vendors

Over the course of the Pilot deployment, three vendors supplied OBUs and one vendor provided the RSUs. One of the OBU vendors exited the study during the first months of deployment and the units deployed had to be replaced by one of the other two remaining vendors.

In deploying the OBUs, the vendors created a method to update the firmware and configuration parameters over the air. This method was developed so that the units could be updated remotely without having to bring participants to the study's installation facility multiple times. The OBUs could receive the updated firmware and configuration parameters when traveling inside the study area and receiving updates via the RSUs. This method was successfully used for several firmware updates and configuration changes. In some instances, however, the updates did not reach all participants, thus creating a challenge in the collection and OTA transmission of the OBU data logs or affecting successful updates to the latest application configuration and operational parameters.

Notwithstanding the continuous feedback loop established between the performance evaluation team and the system deployers, firmware issues still represent a challenge, and the warning applications continue to require further fine-tuning of the application configuration parameters.

8.3 Lessons Learned

The THEA CV Pilot safety applications were not intended to actively intervene on the vehicles braking system. Rather, they were meant to warn drivers with the goal of analyzing their response to visual and audio warnings. The analysis showed that in several cases, the warnings issued could have a positive effect on safety by reducing potential crashes if the drivers saw the warnings and reacted to them. An example of this is the deployment of applications in Use Case 6. The three applications issued 26 true positive warnings which could have a significant effect in reducing crashes for this corridor.

At the same time, the safety evaluation of this study pinpoints to drawbacks in the actual implementation of V2V applications, mostly due to the setup and implementation of the OTA firmware updates and the applications' operational and functional parameters.

The OBU aftermarket suppliers engaged in the deployment exhibited a high degree of variability in terms of research and development capabilities. This heterogeneity impacted the development, refinement, and level of maturity of some of the THEA CV Pilot applications (i.e., I-SIG, TSP, PCW).

Other factors affecting the application performance are more specific to the site deployment: a highly dense urban environment posing challenges to V2I applications that are more dependent on accurate GPS positioning and signal stability. The findings from this study can inform current and future pilot deployments and point to implementable solutions.

References

1. Waggoner, J., et al., *Connected Vehicle Pilot Deployment Program Phase 1, Concept of Operations (ConOps) – Tampa (THEA)*. 2016. p. 118p.
2. Reichardt, C.S., *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Social Service Review, 2002(3): p. 510.
3. Shadish, W.R., T.D. Cook, and D.T. Campbell, *Experimental and quasi-experimental designs for generalized causal inference*. 2002: Boston : Houghton Mifflin, 2002.
4. Concas, S., A. Kourtellis, and S.L. Reich, *Connected Vehicle Pilot Deployment Program Performance Measurement and Evaluation Support Plan, Phase 2 UPDATE – Tampa (THEA)*. 2019, FHWA: https://rosap.ntl.bts.gov/view/dot/31732.
5. Society of Automotive Engineers, *J2945: On-Board System Requirements for V2V Safety Communications*, S. Standards, Editor. 2016, SAE International.
6. Society of Automotive Engineers, *J2735: Dedicated Short Range Communications (DSRC) Message Set Dictionary*. 2016, SAE International.
7. Siemens Mobility Inc., *UC1 Morning Backup Operational Plan: Improved Queue Length Measurement Operation Description.* 2020.
8. Johnson, S., et al., *Connected Vehicle Pilot Deployment Program Phase 2 - System Architecture Document – Tampa (THEA)*. 2017, U.S. Department of Transportation Intelligent Transportation Systems (ITS) Joint Program Office: Washington, DC.
9. Siemens Mobility Inc., *UC3 Pedestrian Safety Operational Plan, Improved Pedestrian Sensor Operation Description*. 2020.
10. Johnson, S., et al., *Connected Vehicle Pilot Deployment Program Phase 2 - System Design Document (SDD) – Tampa (THEA)*. 2018, U.S. Department of Transportation Intelligent Transportation Systems (ITS) Joint Program Office: Washington, DC.
11. McDowall, D., R. McCleary, and B.J. Bartos, *Interrupted Time Series Analysis*, ed. O.U. Press. 2019, New York, NY.
12. Wooldridge, J.M., *Introductory Econometrics*. 2019, Boston, MA: Cengage Learning.
13. Kamrani, M., S. Concas, and A. Kourtellis, *Systems and methods for detecting the location of debris and unexpected objects on roads*. 2020.
14. Maasoumi, E. and J. Racine, *Entropy and predictability of stock market returns*. Journal of Econometrics, 2002. **107**(1): p. 291-312.
15. *Signal 4 Analytics*. 2020 [cited 2020; Available from: https://s4.geoplan.ufl.edu/.
16. CUTR. *CV Pilot Performance Measurement Dashboard*. 2020; Available from: https://cavdashboard.com.
17. Parasuraman, R., P.A. Hancock, and O. Olofinboba, *Alarm effectiveness in driver-centred collision-warning systems.* Ergonomics, 1997. **40**(3): p. 390-399.
18. Rahman, M.M., et al., *Modelling driver acceptance of driver support systems.* Accident Analysis & Prevention, 2018. **121**: p. 134-147.
19. Donmez, B., et al., *Drivers' attitudes toward imperfect distraction mitigation strategies.* Transportation Research Part F: Traffic Psychology and Behaviour, 2006. **9**(6): p. 387-398.
20. FLDHSMV. *Crash Records and Crash Reporting*. 2020 [cited 2020; Available from: https://www.flhsmv.gov/courts-enforcement/about-crash-records-and-crash-reporting/frequently-asked-questions/.

U.S. Department of Transportation
ITS Joint Program Office – HOIT
1200 New Jersey Avenue, SE
Washington, DC 20590

Toll-Free “Help Line” 866-367-7487

www.its.dot.gov

FHWA-JPO-20-829

U.S. Department of Transportation