MOUNTAIN-PLAINS CONSORTIUM

MPC 22-489 | M. Ahmed, S. Gaweesh and A.K. Bakhshi

ASSESSING THE EFFECTIVENESS OF THE WYOMING CONNECTED VEHICLE PILOT PROGRAM: NEW TRAFFIC SAFETY RESEARCH PERSPECTIVES

MOUNTAIN-PLAINS CONSORTIUM REGION 8

A University Transportation Center sponsored by the U.S. Department of Transportation serving the Mountain-Plains Region. Consortium members:

Colorado State University
North Dakota State University
South Dakota State University
University of Colorado Denver
University of Denver
University of Utah
Utah State University
University of Wyoming

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
MPC-598		

4. Title and Subtitle	5. Report Date
Assessing the Effectiveness of the Wyoming Connected Vehicle Pilot Program: New Traffic Safety Research Perspectives	September 2022
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Mohamed M. Ahmed, Ph.D., PE Sherif Gaweesh, Ph.D., PE Arash Khoda Bakhshi, Ph.D.	MPC 22-489

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Department of Civil and Architectural Engineering University of Wyoming 1000 E. University Ave Laramie, WY	11. Contract or Grant No.

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Mountain-Plains Consortium North Dakota State University PO Box 6050, Fargo, ND 58108	Final Report
	14. Sponsoring Agency Code

15. Supplementary Notes

Supported by a grant from the US DOT, University Transportation Centers Program

16. Abstract

Considering traffic safety concerns and challenging driving conditions on Interstate 80 in Wyoming, the United States Department of Transportation (USDOT) and Federal Highway Administration (FHWA) selected an I-80 400-mile freeway corridor as one of the three sites in the United States to develop, test, and deploy a suite of connected vehicle (CV) applications. The Wyoming Connected Vehicle Pilot Deployment Program (WYDOT CV Pilot) utilizes real-time communication technologies to provide warnings and advisories regarding various road conditions to heavy trucks and light vehicle drivers. One of the ultimate goals of this pilot is to alleviate the traffic safety concerns on the I-80 corridor in Wyoming. Hence, the safety performance assessment of the pilot is pivotal for the WYDOT and FHWA strategic goals.

This report provides a new traffic safety perspective for the safety performance evaluation of the WYDOT CV Pilot through advanced statistical modeling, machine learning, data mining applications, safety data visualizations, high-fidelity driving simulator experiments, and traffic microsimulation modeling. To this aim, the procedure and the analytical inference for developing a baseline and analysis, modeling, and simulation (AMS) framework are presented based on using two distinct but complementary approaches: conduct a before/after analysis to explore crash/crash severity causations during CV pre-deployment as a baseline, and the AMS framework in with/without the CV technology to quantify drivers' behavioral alteration under the effect of various CV applications.

Results unveiled statistically significant real-time traffic-related factors contributing to crash and critical crashes during CV pre-deployment. In the with/without analysis and based on the calibrated and validated AMS framework, the impact of several CV applications was analyzed. These applications included spot weather impact warning (SWIW), distress notification (DN), situational awareness (SA), CV variable speed limit (CV-VSL), work zone warning (WZW), forward collision warning (FCW) and rerouting applications. According to the quantification of drivers' behavioral alterations under various CV notifications utilizing trajectory-level analyses, results affirmed promising safety effects of CV applications. The surrogate measures of safety analysis in microsimulation modeling indicated an enhanced traffic safety performance under various CV market penetration rates (MPR).

17. Key Word	18. Distribution Statement
connected vehicles, dedicated short range communications, evaluation and assessment, microsimulation, performance measurement, pilot studies, traffic safety, trucking safety, truck traffic, weather conditions	Public distribution

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	85	n/a

Assessing the Effectiveness of the Wyoming Connected Vehicle Pilot Program: New Traffic Safety Research Perspectives

Mohamed M. Ahmed, Ph.D., PE
Sherif Gaweesh, Ph.D., PE
Arash Khoda Bakhshi, Ph.D.

Department of Civil and Architectural Engineering
University of Wyoming
1000 E. University Ave, Laramie, WY

September 2022

Acknowledgments

The funding for this study was provided by the WYDOT to the Mountain-Plains Consortium (MPC). All statements and opinions presented in this report are the sole responsibility of the authors and may not necessarily reflect those of WYDOT.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

ABSTRACT

Considering traffic safety concerns and challenging driving conditions on Interstate 80 in Wyoming, the United States Department of Transportation (USDOT) and Federal Highway Administration (FHWA) selected an I-80 400-mile freeway corridor as one of the three sites in the United States to develop, test, and deploy a suite of connected vehicle (CV) applications. The Wyoming Connected Vehicle Pilot Deployment Program (WYDOT CV Pilot) utilizes real-time communication technologies to provide warnings and advisories regarding various road conditions to heavy trucks and light vehicle drivers. One of the ultimate goals of this pilot is to alleviate the traffic safety concerns on the I-80 corridor in Wyoming. Hence, the safety performance assessment of the pilot is pivotal for the WYDOT and FHWA strategic goals.

This report provides a new traffic safety perspective for the safety performance evaluation of the WYDOT CV Pilot through advanced statistical modeling, machine learning, data mining applications, safety data visualizations, high-fidelity driving simulator experiments, and traffic microsimulation modeling. To this aim, the procedure and the analytical inference for developing a baseline and analysis, modeling, and simulation (AMS) framework are presented based on using two distinct but complementary approaches: conduct a before/after analysis to explore crash/crash severity causations during CV pre-deployment as a baseline, and the AMS framework in with/without the CV technology to quantify drivers' behavioral alteration under the effect of various CV applications.

Results unveiled statistically significant real-time traffic-related factors contributing to crash and critical crashes during CV pre-deployment. In the with/without analysis and based on the calibrated and validated AMS framework, the impact of several CV applications was analyzed. These applications included spot weather impact warning (SWIW), distress notification (DN), situational awareness (SA), CV variable speed limit (CV-VSL), work zone warning (WZW), forward collision warning (FCW) and rerouting applications. According to the quantification of drivers' behavioral alterations under various CV notifications utilizing trajectory-level analyses, results affirmed promising safety effects of CV applications. The surrogate measures of safety analysis in microsimulation modeling indicated an enhanced traffic safety performance under various CV market penetration rates (MPR).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

The United States Department of Transportation (USDOT) and the Federal Highway Administration (FHWA) selected three sites in the United States for the first wave of connected vehicle (CV) pilots to showcase the value and spur the adoption of CV technology: New York City, Tampa, Florida, and Interstate 80 (I-80) in Wyoming (Figure 1.2). CV technology goes around level-0 and level-1 of driving automation [1], where drivers are in the safety loop, contributing to more than 90% of motor vehicle crashes [2]. Hence, the safety benefits of CV technology are mostly gained from altering driver behaviors based on the assistance from the real-time CV warnings provided. Under the Vehicular Ad-hoc NETwork (VANET) and dedicated short-range communication (DSRC) or satellites, CV technology improves drivers' awareness regarding unforeseeable roadway hazards through vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), and infrastructure-to-vehicle (I2V) real-time notifications. This ability enables drivers to react proactively and avoid potential traffic crashes in urban or rural areas such as I-80.

Figure 1.1 Connected Vehicle Pilot Deployment Programs in the United States (Source: WYDOT)

I-80 in Wyoming is a segment of an east/west rural corridor in the U.S. northwest, supporting the movement of over 32 million tons of freight per year (at 16 tons per truck). As a mountainous rural freeway, I-80 mostly operates at the level of service (LOS) A or B. Truck volume ranges from 30% to 55% of the total traffic stream on an annual basis, with seasonal variations that can make up as much as 70% of the traffic volume. All elevation is above 6,000 feet, with the highest point reaching 8,640 feet (2,633 m) above sea level at Sherman Summit. The corridor is also characterized by severe weather conditions, i.e., strong winds, heavy snow and fog, severe blowing snow, and low visibility. Due to Wyoming's adverse winter weather conditions, such as snowstorms, strong crosswinds, icy road surfaces, and low visibility from blizzards, and the presence of work zones, there have been remarkable traffic crash records along I-80 in Wyoming, resulting in fatalities, road closures, and tremendous economic loss [3].

To mitigate these safety concerns, the USDOT FHWA selected I-80 in Wyoming (WYDOT CV Pilot) as the only rural site among the mentioned three CV pilots in the country. The Wyoming CV systems and applications developed are expected to enable CV drivers to have improved awareness of potential hazards when driving on I-80, help fleet managers better manage their freight operations, and support WYDOT Traffic Management Center staff to implement more effective traffic control strategies. In this regard, the CVs' performance evaluation is vital to USDOT's strategic goals. Accordingly, the WYDOT CV Pilot team developed a performance measurement and evaluation support plan [4], including 21 performance measures. The major performance categories represent the primary activities and outcomes of the Wyoming CV pilot system. These categories focus on improvements to efficiency, safety, and mobility. Quantitative and qualitative measures were proposed to evaluate the Wyoming CV project, focusing on understanding the extent and impact of the benefits described above.

In this regard, the traditional safety evaluation methodologies presented in the Highway Safety Manual (HSM) [5] might not be appropriate approaches due to the following limitations. First, statistical analysis performed utilizing historical data is the core of the approaches presented in the HSM. Being in an early development phase of the WYDOT CV Pilot, assessing the obtained CVs' safety benefits is challenging and still unclear due to the unavailability of enough data during post-deployment, hindering the use of traditional HSM methodologies. Secondly, at the very beginning stage of the WYDOT CV Pilot, market penetration rate (MPR) of CVs is very low, which is among the key factors affecting the safety benefits obtained from CVs. Thus, to adopt the HSM methodologies in the evaluation process, a significant proportion of vehicles on the roadways should be equipped with CV technology. With these concerns, new innovative approaches are required to investigate the safety effectiveness of this newly introduced technology.

This document presents alternative methodologies, including real-time risk assessment and an analysis, modeling, and simulation (AMS) framework, to assess newly emerging CV technology by integrating real-time information obtained from the WYDOT CV Pilot into high-fidelity driving simulator experiments and microsimulation analysis levels. To this aim, a variety of advanced statistical modeling, machine learning, data mining applications, and safety data visualizations are utilized under two distinct but complementary approaches as follows. The details are presented in the next section:

- Before/after analyses to explore crash/crash severity causations as a baseline during CV pre-deployment under the concept of real-time risk assessment (RTRA) that will be used as a benchmark against the traffic safety performance of I-80 during post-deployment under any CV-MPR.
- With/without analyses under a calibrated and validated AMS framework to quantify drivers' behavioral alterations due to different CV applications and measure CVs' effect on the entire traffic stream.

2. GENERAL METHODOLOGY

This study presents a framework for evaluating the safety performance of the WYDOT CV Pilot. Figure 2.1 illustrates the analytical approach, structured based on a conflation of before/after and with/without analyses.

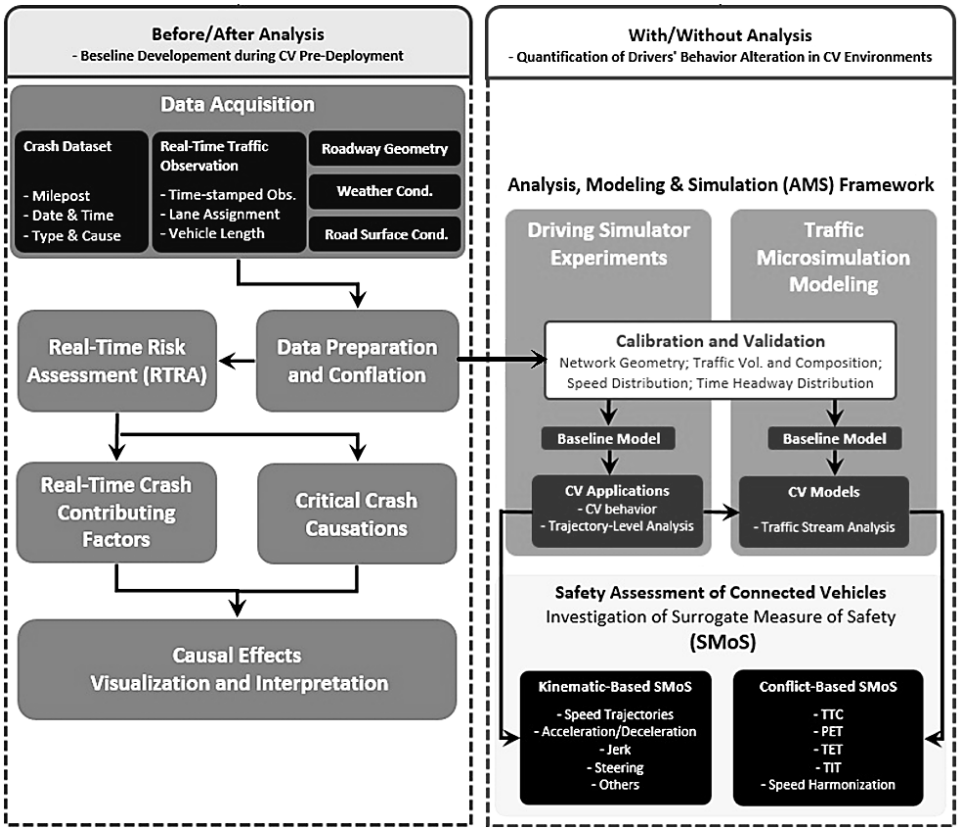


Figure 2.1 The General Analytic Approach

2.1 Before/After Analysis and CV Pre-Deployment Baseline Development

In the before/after analysis, crash contributing factors and causations of critical crashes involving fatal or incapacitating injuries, throughout the 402-miles of I-80 in Wyoming during CV pre-deployment, were explored under RTRA. RTRA investigates crash/crash severity causations during CV pre-deployment that will be used as a benchmark against defined crash causalities during CV post-deployment, unveiling the safety performance of CVs under any MPR. According to the literature, CV-MPR is one of the important macro-level factors affecting the safety and operational performance of the entire mixed traffic stream. Hence, under the concept of before/after analysis, it is essential to conduct a reliable baseline showing the traffic safety condition on I-80 before deploying CV technology. The mentioned requirement was satisfied based on the result of RTRA, focusing on real-time traffic-related crash contributing factors that might be affected by CVs during post-deployment. Under this approach, precursors of individual crashes were analyzed using advanced statistical modeling and machine learning techniques on a conflation of aggregated environmental data with disaggregated high-resolution real-time traffic

observations for the sake of exploring the crash/critical crash causations. In addition, the preprocessed dataset in the RTRA part was utilized to calibrate and validate the AMS framework in the with/without analysis.

2.1.1 Pre-Deployment Dataset

Four databases were conflated to be used in the RTRA during CV pre-deployment and for calibrating and validating the AMS framework. The primary databases comprised crash reports, aggregated roadway geometry characteristics, weather conditions, and real-time traffic observation through the 402 miles of I-80 in Wyoming provided by the WYDOT. These databases were used individually or were conflated with others depending on the purpose of the analysis.

The crash dataset, from January 2017 to July 2017, was used to explore real-time traffic-related crash contributing factors by following matched-case control design (MCCD) to predict the crash likelihood based on analyzing crash precursors and comparing them with normal traffic patterns before non-crash instances. It was preferred to analyze the crash precursors for those crashes that were under the impact of traffic flow oscillations. Therefore, those crashes that were due to drugs or alcohol usage and unknown reasons were eliminated from the dataset. To analyze the crash precursors, the high-resolution real-time traffic observations provided by the WYDOT using Wavetronix speed sensors installed on the corridor were conflated with the mentioned crash dataset. In this respect, 10,735,339 real-time traffic observations were reduced to 70,930 observations corresponding to 203 crashes and 284 non-crash cases utilizing observations from 51 speed sensors on the 402-miles of I-80 in Wyoming.

Critical crash causations, including fatal or incapacitating injuries, were analyzed using the crash dataset conflated with real-time traffic observations provided for RTRA. This analysis only focuses on crash cases because it aims to unveil the difference between critical crash precursors versus non-critical ones. Note that the crashes occurred on different mileposts, roadway conditions, and weather conditions. As opposed to MCCD, which could account for confounding factors across the crash and non-crash cases, it was required to consider the effect of environmental factors on crash precursors in analyzing critical crashes. Thus, the mentioned crash dataset and its associated real-time traffic observations were conflated with aggregated roadway geometry characteristics and the weather conditions dataset at the crash scenes.

2.2 With/Without and AMS Framework Development

In addition to the baseline development, the research team conducted a well-calibrated AMS framework based on a three-pronged approach: RTRA, high-fidelity driving simulator experiments, and traffic microsimulation modeling. The conducted AMS is specifically beneficial for the WYDOT CV Pilot to quantify drivers' behavioral alterations due to CV notifications by modeling vehicles' trajectories in comprehensive driving simulator experiments and utilizing kinematic-based surrogate measure of safety (SMoS). More importantly, the mentioned quantification is conflated with microsimulation modeling to reveal the impact of CV application on the entire traffic stream under varying CV-MPRs using conflict-based SMoS. Through this technique, different critical traffic safety events are regarded, and the safety performances of different CV applications are analyzed.

2.2.1 Surrogate Measure of Safety (SMoS)

The SMoS approach, which has gained acceptance as a proactive approach in the literature, was considered for evaluating CVs' safety performance [6]. SMoS can improve the understanding of failure mechanisms and chain of events resulting in traffic critical safety events. The most crucial advantage of SMoS is to perform a reliable safety performance assessment within a shorter time period than traditional performance measurement approaches. SMoS can be divided into two main categories: conflict-based SMoS (C-SMoS) and kinematic-based SMoS (K-SMoS).

C-SMoS, also called temporal and/or spatial proximity-based conflict indicators, analyzes traffic safety performance, assuming that the closer vehicles are to each other, the nearer they are to a collision because conflicts will precede all collisions [6]. The quantitative assessment provided by C-SMoS is objective with an interpretable measure in terms of closeness to a collision. This approach requires traffic observations from real-life data or microsimulation modeling to show traffic conflicts using a variety of proximal temporal measures, such as time-to-collision (TTC), time exposed TTC (TET), time integrated TTC (TIT), modified time-to-collision (MTTC), and post-encroachment time (PET). The research team followed this approach to unveil the CVs' safety effect on the entire mixed-traffic stream under varying CV-MPRs, particularly within traffic microsimulation modeling.

K-SMoS, however, relies on detailed driving information, delving into individual driving behavior [7], [8]. K-SMoS has been rarely used in the safety domain because of the unavailability of vehicle trajectories, which have been become available in CV environments. In CV environments, the large-scaled individual driving data embedded in basic safety messages (BSMs) provide continuous real-time access to vehicle trajectories being affected by driver behaviors. There is a high correlation between aggressive and defensive driving styles with less traffic-safe conditions. Hence, given the vehicle trajectories, it is possible to assess the traffic stream's safety performance by quantifying driver volatility using K-SMoS. Accordingly, well-characterized K-SMoS aided this research in spatiotemporally quantifying individual driving styles and assessing the safety performance of CV advisory/warning messages in a series of with/without analyses.

2.3 Database for Calibration and Validation of the AMS Framework

Conflated high-resolution real-time traffic observations with aggregated roadway geometry and weather information throughout the 402-mile I-80 corridor were used to calibrate and validate the AMS framework. The AMS framework mainly analyzes CV safety performance in mitigating the risk of different types of traffic crashes on the I-80 rural corridor, such as work-zone-related crashes, horizontal curve-related crashes, and secondary crashes. The baselines for each of these investigations (i.e., no CVs in the traffic stream) were calibrated and validated by statistical comparisons between the real-life speed distributions with speed distributions from microsimulation modeling and/or driving simulator studies. Therefore, to analyze CVs' safety performance in work zone areas, it was required to combine real-life speed observations with work zone information on I-80. This conflation was performed by modeling speed distributions from the nearest speed sensors to work zone areas between Laramie and Cheyenne on the I-80 corridor.

3. BASELINE DEVELOPMENT AND IDENTIFICATION OF REAL-TIME CRASH CONTRIBUTING FACTORS

3.1 Introduction

During the WYDOT CV Pilot on I-80, reducing the rate of traffic crashes was an important measure for the CVs' safety performance assessment. The explored crash causations were expected to be affected under various CV applications during CV post-deployment. Hence, this part of the document presents a procedure for the baseline development to indicate crash contributing factors on I-80 during CV pre-deployment, specifically in terms of real-time traffic-related contributing factors.

To this aim, real-time crash prediction models (CPMs) are required to unveil the statistical linkage between traffic flow characteristics and the probability of crashes [9]. In the RTRA arena, it is known that traffic crashes can be predicted by investigating crash precursors during a period preceding crash occurrence [9]–[12]. Under RTRA, most previous analyses investigated RTRA in a particular section of a corridor or limited length of a segment [9]. However, the safety performance assessment of disruptive applications such as CVs on I-80 in Wyoming required looking into a long corridor. Accordingly, this analysis dealt with two main problems in conducting a unique CPM on the 402 miles of I-80. First, it was essential to deal with nonlinear predictors due to a remarkable variation in traffic patterns throughout the 402-mile corridor. Secondly, the study had to account for the small number of real-time traffic observations within a predefined time window of crash precursors on I-80 with comparatively less traffic volume. Under the MCCD concept, these matters were addressed through the use of advanced statistical modeling and proper feature engineering.

Regarding the first problem, the data-driven non-parametric statistical approach allows data to learn from their distributions by minimizing assumptions of pre-specified transformations for nonlinear predictors. In this regard, under logistic regression, three types of statistical modeling, generalized linear model (GLM), generalized non-linear model (GNM), and generalized additive model (GAM), were employed to deal with nonlinear predictors. Relating to the second matter, initial investigations showed many of the crash contributing factors proposed in the literature could not significantly increase the model's discriminative ability to measure traffic fluctuations before crashes. Hence, the predictors were characterized to capture traffic oscillations, both laterally, by considering individual lanes and the whole traffic stream, and longitudinally, by comparing differences of traffic-related variables in space concerning crash locations. The dimensionality reduction of the multidimensional space of variables was performed using three feature selection (FS) techniques: mean decrease accuracy (MDA), mean decrease impurity (MDI), and corrected-impurity importance (CII).

3.2 Variable Description in Matched-Case Control Design (MCCD) Approach

The MCCD was followed to explore real-time traffic-related crash contributing factors by predicting crashes based on analyzing crash precursors and comparing them with normal traffic patterns before non-crash instances within the same timeframe. This technique is a practical one that controls confounding factors such as roadway geometry, driver population, seasonal traffic

variation, and, to some extent, weather conditions. To this aim, real-time traffic observations were considered within a 15-minute time window before crashes and their corresponding non-crash cases from the same speed sensors to hold the same distance between the location of sensors and the locations of the crash and non-crash cases. Real-time traffic observations were conflated with the crash dataset. The non-crash data were reduced through the same day of a week and the same time window of the day that the corresponding crash precursors were investigated. Therefore, two datasets for crash and non-crash conditions were prepared and conflated with real-time traffic observations. The real-time traffic observations dataset comprises the time of the observations, traffic volume, vehicle speed, vehicle length, and lane assignment. If a vehicle's length was less than 30 feet, it was considered a passenger car; otherwise, it was counted as a truck.

Accordingly, the design matrix was structured by characterizing real-time traffic-related variables to capture traffic oscillations laterally and longitudinally. Most of the characterized predictors were processed variables. They were introduced according to the speed variance, mean speed, and regression slope of the speed profile within the 15-minute time windows before the crash and non-crash cases. In total, 29 variables were used to cluster crash and non-crash cases, which are described in Table 3.1.

Table 3.1 Description of Explanatory Variables to Explore Crash Contributing Factors

Continuous Variables [a]	Description	S.D.	Mean	Min.	Max.
T_SpMean	Spatial Difference in Mean Speed of Total Traffic Volume in both Lanes	7.09	-0.10	-34.88	35.52
T_SpVar	Spatial Difference in Speed Variance of Total Traffic Volume in both Lanes	58.05	-11.78	-578.05	458.17
T SpVARoMEAN	Spatial Difference in Speed Variance Divided by Mean Speed for Total Traffic Volume in both Lanes	1.21	-0.16	-17.27	9.79
T_SpSlop	Spatial Difference in the Slop of Speed Regression (from Speed Profile) for Total Traffic Volume in both Lanes	0.02	0.00	-0.03	0.26
T_SpVarDiff	Spatial Difference in Subtraction of Speed Variance in HSL[b] from Speed Variance in LSL[c]	100.79	12.63	-1227.42	533.89
T Volume	Spatial Difference in Total Traffic Volume	28.45	10.90	-88.00	138.00
T_TrP	Spatial Difference in Truck Proportion in Total Traffic Volume in both Lanes	0.17	0.01	-0.57	0.57
T_TrPDiff	Spatial Difference in Subtraction of Truck Proportion in HSL from Truck Proportion in LSL	0.29	-0.04	-1.00	0.84
T_VolumeDiff	Spatial Difference in Subtraction of Traffic Volume in HSL from Traffic Volume in LSL	19.05	6.12	-75.00	72.00
HSL SpMean	Spatial Difference in Mean Speed in HSL	11.55	1.59	-32.85	84.07
HSL_SpVar	Spatial Difference in Speed Variance in HSL	77.20	1.04	-752.67	537.92
HSL_SpVARoMEAN	Spatial Difference in Speed Variance Divided by Speed Mean in HSL	1.47	0.01	-19.68	10.59
HSL SpSlop	Spatial Difference in Slop of Speed Regression (from Speed Profile) for HSL	0.18	-0.01	-3.80	0.27
HSL_Volume	Spatial Difference in Traffic Volume in HSL	14.22	-0.23	-54.00	111.00
HSL_TrP	Spatial Difference in Truck Proportion in HSL	0.30	-0.01	-1.00	1.00
LSL SpMean	Spatial Difference in Mean Speed in LSL	17.7	4.17	-35.20	79.50
LSL_SpVar	Spatial Difference in Speed Variance in LSL	65.67	-11.59	-576.08	475.12
LSL_SpVARoMEAN	Spatial Difference in Speed Variance Divided by Mean Speed in LSL	1.30	-0.18	-17.07	10.52
LSL_SpSlop	Spatial Difference in Slop of Speed Regression (from Speed Profile) for LSL	0.02	0.00	-0.14	0.24
LSL Volume	Spatial Difference in Traffic Volume in LSL	23.86	11.13	-70.00	99.00
LSL_TrP	Spatial Difference in Truck Proportion in LSL	0.27	0.06	-0.45	1.00

Categorical Variables [a]	Description	Number of Positive (Pct.)	Number of Negative (Pct.)
D_ T_SpVARoMEAN	Dummy Variable Representing T_SpVARoMEAN (1: Negative (Reference Level), 0: Positive)	297 (60.9 %)	190 (39.1 %)
D_ T_SpSlop	Dummy Variable Representing T_SpSlop (1: Negative (Reference Level), 0: Positive)	251 (51.5 %)	236 (48.5 %)
D T SpMean	Dummy Variable Representing T SpMean (1: Negative (Reference Level), 0: Positive)	240 (49.2 %)	247 (50.8 %)
D_ T_VolumeDiff	Dummy Variable Representing T_VolumeDiff (1: Negative (Reference Level), 0: Positive)	205 (42.1 %)	282 (57.9 %)
D_HSL_SpSlop	Dummy Variable Representing HSL_SpSlop (1: Negative (Reference Level), 0: Positive)	244 (50.1 %)	243 (49.9 %)
D HSL SpMean	Dummy Variable Representing HSL SpMean (1: Negative (Reference Level), 0: Positive)	212 (43.5 %)	275 (56.5 %)
D_ LSL_SpSlop	Dummy Variable Representing LSL_SpSlop (1: Negative (Reference Level), 0: Positive)	248 (50.9 %)	239 (49.1 %)
D_LSL_SpMean	Dummy Variable Representing LSL_SpMean (1: Negative (Reference Level), 0: Positive)	217 (44.5 %)	270 (55.5 %)

[a] Each of the Continuous variables (C), at Upstream (U) and Downstream (D). Afterward, for all of the observations, the corresponding values of the continuous variables were calculated by subtracting U from D (i.e., C=D - U). If a continuous variable was negative, the corresponding dummy variable took the value of one; otherwise, it took the value of zero.
[b] **HSL**: High-Speed Lane/ (i.e., Left Lane)
[c] **LSL**: Low-Speed Lane/ (i.e., Right Lane)

3.3 Crash Prediction Models (CPMs)

In total, nine CPMs were conducted based on a combination of three types of statistical modeling and three FS techniques. The best model was selected according to the best within-sample and out-of-sample predictive performances across the nine CPMs. On this point, each model was conducted within two main steps. First, the important predictors were selected using MDA, MDI, and CII as FS techniques offered by random forest (RF). Secondly, for each set of the important predictors obtained from the first step, the GLM, GNM, and GAM were conducted.

3.3.1 Random Forest Feature Selection Using MDA, MDI, and CII

FS has been widely used based on the RF technique that aims to remove less essential variables from the model, enhancing its accuracy [9]. An RF model was conducted over all predictors to select important features properly. There are different FS algorithms offered by RF, among which two methods are popular: MDA and MDI [9].

MDA measures the variable importance according to the increase or decrease of error when the associated link between the predictor of interest and response variable is broken by randomly permuting the values of the predictor. Studies have proven that this measure is biased because of overestimating the importance of highly correlated variables [13]–[16]. For a high dimensional space of predictors, MDA is computationally intensive and is not as robust as MDI to highly perturbed space [17]. MDI measures variable importance by adding the weighted impurity for those nodes that use the variable of interest in RF models. The concept of MDI is biased in favor of highly impure features [17]. In other words, adding weighted impurity to define variable importance provides an unfair advantage for those variables with a large number of amounts (such as continuous or high-cardinal categorical variables) [13], [14], [18]. To overcome the mentioned imperfections in MDA and MDI, Nembrini et al. (2018) proposed unbiased impurity-based variable importance for classification and regression problems within the CII concept. Despite MDI and MDA, CII computes unbiased importance for features outperforming MDA and MDI in terms of statistical power [14], [17], [19].

3.3.2 Logistic Regressions Using GLM, GNM, and GAM

Different types of logistic regression could be conducted to deal with binary outcomes (i.e., crash and non-crash cases) using the logit link provided by generalized models, including GLM, GNM, and GAM. Linearity is an essential presumption in GLM. If the linearity for some variables is not met, those should be excluded from the model. Although it is arguable to transform nonlinear predictors to attain linearity, some issues would remain. First, transforming a variable means applying a unique pattern over the entire range of a variable, where the analyst imposes the data to follow a particular predetermined curve regardless of their nature. Secondly, the primary criterion for selecting the best type of transformation could be a daunting task. Lao, Y., et al. proposed utilizing GNM [20]. GNM handles nonlinear variables; however, the situation is similar to GLM, where the analyst defines transformations. The only difference is that, in GNM, the transformations are applied within the model; whereas, in GLM, nonlinear variables should be transformed in advance. Then the transformed variables could be used in the model.

A simple basis of splitting data into different sub-samples, used by GAMs, can solve the problem of nonlinearity [21]–[23]. Indeed, GAM is a kind of authorized nonparametric GLM where the model is allowed to learn the nonlinearity between response and predictor(s). GAM utilizes a data-driven approach, instead of model-driven, where variables are permitted to determine their trends instead of being imposed by predetermined transformations [24], [25]. The distinction between GNM and GAM is that, in GAM, variables are not supposed to be linear; instead, the applied smoothing functions are to be linear. GAM solves the nonlinearity problem by using smoothing functions like splines as continuous combinations of piecewise polynomials that are fitted locally for a specific sub-section of observations without being affected by other sub-sections. This mechanism assists the variable of interest in finding its pattern more accurately. The terms of

additive in GAM comes from adding smoothing functions that are driven using the data. GAM estimates parameters using a double loop of iteration. The inner one estimates the parameters of the smoothing functions, and the outer one is used for convergence of GAM parameters [25].

For each of the three sets of selected variables from FS techniques, GLM, GNM, and GAM were conducted. In GLMs, according to the statistical assumptions, only linear variables were included. The transformed nonlinear variables and linear ones were used to develop GNMs. Finally, GAMs were conducted based on linear variables and additive smoothing functions of nonlinear variables. Equations 3.1 to 3.3 present the final crash prediction models based on GLMs, GNMs, and GAMs.

Generalized Linear Model (GLM): $$E_Y(y_i|x_i) = \frac{e^{\beta_0+\sum_{j=1}^{J}\beta_j x_{ij}}}{1+e^{\beta_0+\sum_{j=1}^{J}\beta_j x_{ij}}}$$ Equation 3.1

Generalized Nonlinear Model (GNM): $$E_Y(y_i|x_i) = \frac{e^{\beta_0+\sum_{j=1}^{J}\beta_j x_{ij}+\sum_{t=1}^{T}\beta_k x_{it}}}{1+e^{\beta_0+\sum_{j=1}^{J}\beta_j x_{ij}+\sum_{t=1}^{T}\beta_k x_{it}}}$$ Equation 3.2

Generalized Additive Model (GAM): $$E_Y(y_i|x_i) = \frac{e^{\beta_0+\sum_{j=1}^{J}\beta_j x_{ij}+\sum_{k=1}^{K}f_k(x_{ik})}}{1+e^{\beta_0+\sum_{j=1}^{J}\beta_j x_{ij}+\sum_{k=1}^{K}f_k(x_{ik})}}$$ Equation 3.3

Where β_0, β_k, and β_j are coefficients to be estimated, Ey is the expected value of y_i, x_{ij} are the linear predictors, x_{it} are the transformed nonlinear predictors, x_{ik} are the nonlinear predictors, and f_k are splines to be estimated. It is worth noting that, in all models, $\beta_0 + \sum_{j=1}^{J}\beta_j x_{ij}$ is the common term, which indicates all the models have the same structure of GLM, and the only difference is how the models deal with nonlinear predictors.

3.4 Significant Real-Time Traffic-Related Crash Contributing Factors

Table 3.2 compares the performance of nine developed models. According to Table 3.2, it is clear that the combination of CII and GAM outperformed the other models by achieving minimum Akaike information criterion (AIC) and maximum area under the curve (AUC). Hence, the combined GAM and CII was used as the CPM to predict traffic crashes on I-80 during CV pre-deployment and to explore causations of real-time traffic-related crash contributing factors.

Table 3.2 Performance Comparison of Nine Developed Crash Prediction Models

	CPM-1	CPM-2	CPM-3	CPM-4	CPM-5	CPM-6	CPM-7	CPM-8	CPM-9
Feature Selection Technique	MDA			MDI			CII		
Types of Logistic Regressions	GLM	GNM	GAM	GLM	GNM	GAM	GLM	GNM	GAM
AIC	604.04	596.22	589.8	614.94	605.74	594.48	608.6	602.12	579.64
AUC	0.63	0.65	0.66	0.61	0.63	0.66	0.63	0.65	0.72
Number of Sig. Predictors	2	3	5	2	3	6	3	3	7

Note:
CPM= Crash Prediction Model; MDA= Mean Decrease in Accuracy; MDI= Mean Decrease in Impurity;
CII= Corrected Impurity Importance; GLM= Generalized Linear Model; GNM= Generalized Nonlinear Model;
GAM= Generalized Additive Model; AIC= Akaike Information Criterion; AUC= Area Under the Curve

Furthermore, the causal effects of crash contributing factors can be found in Table 3.3, where the result of the selected CPM during CV pre-deployment is presented. As mentioned before, GAM is a non-parametric statistical approach. Accordingly, smoothing functions in GAM led to the estimation of effective degree of freedom (EDF) for nonlinear predictors, as presented in Table 3.3, which is difficult to interpret. This difficulty is the downside of non-parametric statistical methods though they can lead to more accurate predictions compared with parametric models. The research team addressed this deficiency by following the concept of post-hoc interpretation.

Table 3.3 Causal Effect of Real-Time Traffic-Related Factors on the Crash Likelihood on I-80 in Wyoming during CV Pre-Deployment

Combined CII and GAM (AIC: 579.64, AUC: 72%)					
Variables	**Est.**	**Z value**	**χ^2**	**p-value**	**Sig.**
(Intercept)	-0.942	-3.149	-	0.002	*
T_SpMean	-0.108	-3.495	-	<0.000	*
HSL_SpSlop	-4.314	-1.023	-	0.306	
T_SpSlop	9.424	0.713	-	0.476	
D_T_SpMean	0.034	0.108	-	0.914	
T_TrPDiff	-0.887	-1.302	-	0.193	
D_HSL_SpSlop	0.308	1.197	-	0.231	
LSL_SpVARoMEAN	0.872	2.758	-	0.006	*
T_SpVarDiff	0.000	0.015	-	0.988	
T_SpVARoMEAN	-0.861	-2.798	-	0.005	*
LSL_SpMean	-0.020	-2.007	-	0.045	*
HSL_SpMean [a]	EDF= 1.000	-	11.403	0.001	*
HSL_TrP [a]	EDF= 1.000	-	6.235	0.013	*
T_TrP [a]	EDF= 5.553	-	22.114	0.002	*

Note:
Description of variables can be found in Table 2;
CII= Corrected Impurity Importance; GAM= Generalized Additive Model;
AIC= Akaike Information Criterion; AUC= Area Under the Curve;
EDF= Effective Degree of Freedom
[a] Nonlinear Predictor
* Represents statistically significant predictors under 95% Confidence Interval.

3.5 Post-Hoc Interpretation

Depending on when the interpretability is obtained, two general approaches have been proposed in the literature to interpret a prediction model: intrinsic interpretability and post-hoc interpretability [26]–[28]. The former is based on self-explanatory models that incorporate interpretability directly to their structures. The latter, however, needs to develop another independent model, as a shallow model, to interpret the existing model [28], [29]. The post-hoc interpretability can be considered as a distillation process from the highest accurate prediction

model to the highest interpretable model by providing a global vision in a post-hoc manner [29]. The latter technique was applied because it was required to select the highest accurate model for crash prediction on I-80. This selection, however, imposed the choice of GAM as a non-parametric approach, which is not intrinsically interpretable. The following steps were followed:

- Develop a CPM to detect statistically significant real-time traffic-related crash contributing factors.
- Develop a crash interpretation model (CIM) as a shallow model for CPM to interpret and visualize the effect of crash contributing factors on crash risk.

The first step was accomplished based on the combination of CII and GAM. For the second step, the power of visualization tools provided by RF was utilized to visualize the causal effect of the significant real-time traffic-related variables on crash probability.

3.5.1 Random Forest Visualization Tools

After identifying the significant real-time traffic-related predictors by the CPM, these predictors were fed into an RF as the CIM. Four widely used RF visualization tools—partial dependence plot (PDP), individual conditional expectation (ICE), centered-ICE (cICE), and accumulated local effect (ALE) [27], [30]–[34]—were employed to visualize and interpret the causal effect of these factors on crash probability. The details are as follows:

3.5.1.1 Partial Dependence Plot (PDP)

The relationship between the outcome and one or two explanatory variable(s) can be delineated by PDP [30]. PDP estimates the average marginal effect of one or two predictors on the predicted outcomes, which can be a probability in classification problems or a determined value in regression [30]–[32]. To this aim, PDP divides predictors into two sets. The first set involves the feature(s) of interest for which the PDP should be drawn (Xi), and the second set comprises the other features included in the developed model (Xo). Therefore, the combination of Xi and Xo comprises the multi-dimensional space of features based on which the CIM has been developed.

PDP works as a function of one or two variables in Xi and marginalizes the output over the distribution of the features in Xo. In other words, to find the response variable at a specific given value of x_i, PDP generates synthetic observations where the values of the feature of interest are permuted by the x_i across all actual observations. Afterward, the response variable at x_i is calculated by averaging the obtained outcomes over all the synthetic observations. Equation 3.4 parametrizes the overall approaches used in PDP.

$$\hat{f}_{X_i}(x_i) = \frac{1}{n}\sum_{j=1}^{n} \hat{f}\left(x_i, x_o^j\right) \quad \text{Equation 3.4}$$

Where $\hat{f}_{X_i}$ is the partial function for the predictor of interest that estimates the average marginal effect at a given value of x_i, n is the number of observation, x_o^j is the actual value of those predictors that are not under investigation for jth observation, and $\hat{f}\left(x_i, x_o^j\right)$ is the model output for the jth observation at x_i.

3.5.1.2 Individual Conditional Expectation (ICE)

PDP uses the averaged marginal effect of a predictor on response variable and plots one curve through the range of the predictor of interest. PDP, however, is unable to bring sufficient insight into the heterogeneity across observations that come from interactions between predictors. To address this issue, Goldstein et al. (2015) proposed the ICE concept to dig into heterogeneity across observations [33].

As it is apparent from the definition, for a specific predictor of interest, ICE draws one curve per each of the individual synthetic instances while holding the value of the other predictors unchanged. The estimating function ($\hat{f}_{X_i}$) would depend on values of the predictor of interest, and by permuting this value through the range of the predictor, one curve per one instance is obtained. ICE repeats this procedure for all the observations. The outcome is a combination of curves through the domain of the predictor of interest that can provide sufficient insight into the interaction of predictors and heterogeneity across observations. The average of all the drawn curves plotted by ICE is exactly the PDP for the predictor of interest.

3.5.1.3 Centered ICE (cICE)

Since the curves in ICE plots start from different prediction points, it is difficult to assess whether ICE curves differ from one instance to another. ICE curves can be centralized at a certain prediction point that demonstrates only the difference in predicted outcomes to that certain point. The obtained plot is called a centered ICE (cICE) plot [31]. This certain point is usually chosen at the lower bound of the range of the predictor of interest. Equation 3.5 illustrates this procedure.

$$\hat{f}_{cent}^{(i)} = \hat{f}^{(i)} - \mathbf{1}\hat{f}\left(x^a - x^{(i)}\right) \qquad \text{Equation 3.5}$$

Where $\mathbf{1}$ is a vector of 1's and $\hat{f}_{cent}^{(i)}$ is the centralized estimating function for the *ith* observation ($x^{(i)}$) centered at x^a [33]. Note that since cICE are centralized, the outcome probability can take negative values. However, this does not affect visualizing the overall trend [31]. ICE and c-ICE can only present a one-dimensional relationship compared with PDPs, which can show the combined effect of two features on outcomes based on two-dimensional plots.

3.5.1.4 Accumulated Local Effects (ALE)

The procedure used in PDP and ICE assumes that the other predictors have the same marginal distributions for any level of the predictor of interest that can be a false assumption for highly correlated predictors [31], [34]. Apley and Zhu introduced accumulated local effects (ALE) to tackle this issue by averaging differences in conditional distribution predictions instead of marginal distribution predictions [34]. ALE divides the range of the predictor of interest (X_j) into K equal intervals. This division blocks the effect of the correlated variable(s) and the generation of unrealistic synthetic observations. Calculating the average over differences in predictions will lead to the pure main effect of X_j on outcomes [34]. Equation 3.6 shows the uncentered ALE for at $X_j = x$.

$$\hat{\tilde{f}}_j(x) = \sum_{k=1}^{k_j(x)} \frac{1}{n_j(k)} \sum_{i:x_{i,J} \in N_j(k)} [f(z_{k,j}, x_{i,J}) - f(z_{k-1,j}, x_{i,J})] \qquad \text{Equation 3.6}$$

Where j indices the predictor of interest, $k_j(x)$ is the kth interval of Xj within which x lies ($k_j(x) \in \{1, 2, \ldots, K\}$), $n_j(k)$ is the number of observations in $k_j(x)$, $N_j(k)$ is a subset of observations that occur in $k_j(x)$, $z_{k,j}$ and $z_{k-1,j}$, respectively, are the upper and lower bound of X_j in $k_j(x)$, $x_{i,J}$ is the value of other predictors (j ≠ J) for ith observation in $k_j(x)$, and $f(.)$ is the fitted model.

3.5.2 Results and Implications for WYDOT CV Pilot

All the aforementioned visualization tools were used to ensure that the interpretation of causations does not depend on the type of visualization, leading to a universal interpretation. Figure 3.1 narrates the causal effects of significant variables on the crash likelihood throughout the 402 miles of I-80 in Wyoming. Accordingly, a reliable baseline has been conducted to explore the causal effect of real-time traffic-related factors on the crash likelihood before piloting CV technology on the I-80 corridor. The patterns of these causal effects on crash likelihood are expected to be changed and affected due to the impact of CVs' speed adherence and harmonization during post-deployment. This pattern recognition could be attained by following the same approach and conducting real-time CPM during CV post-deployment once CVs reach notable MPRs. The comparison of causation patterns between CV pre- and post-deployment would reveal how CV technology can affect crash likelihood on the I-80 corridor.

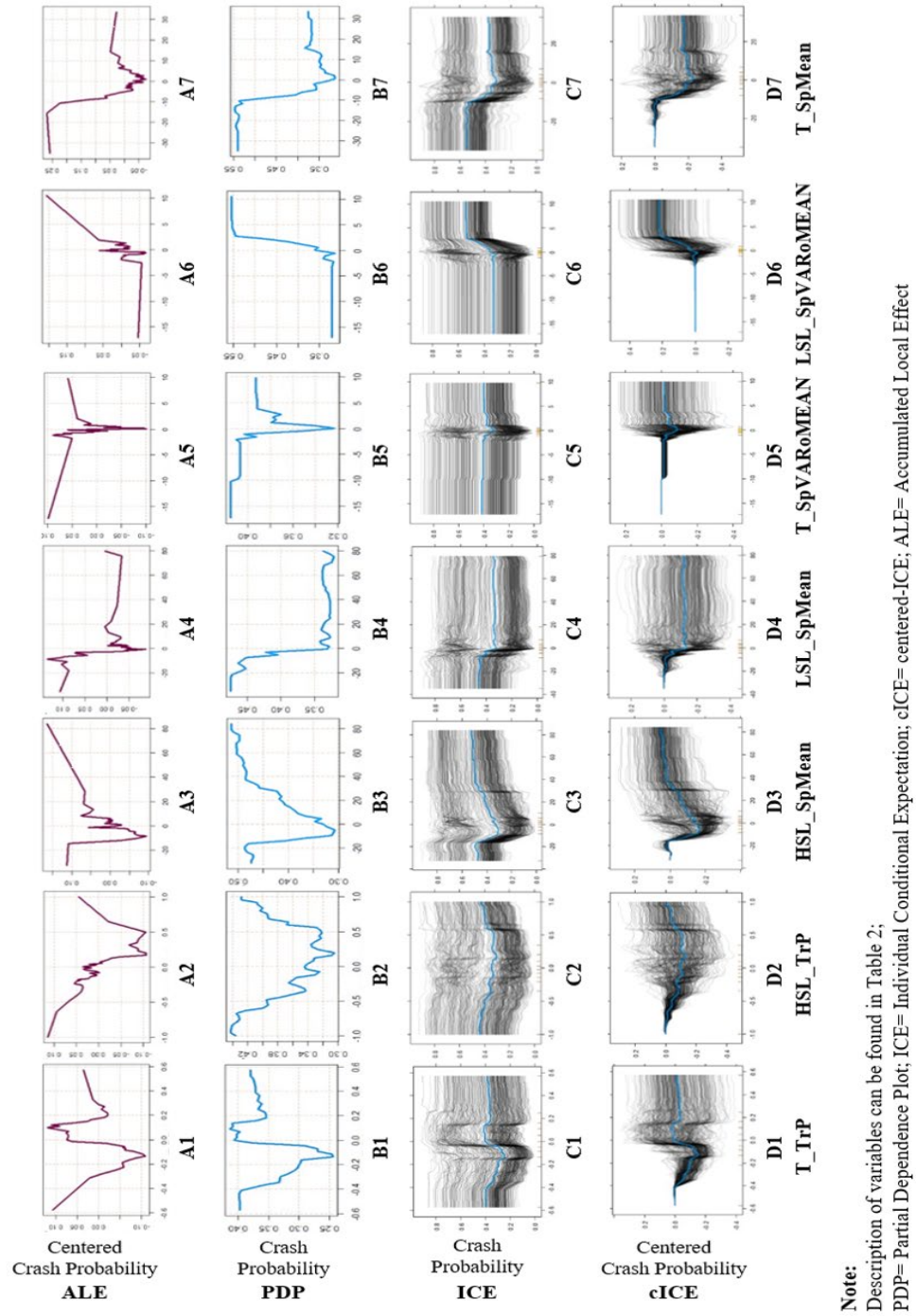


Figure 3.1 Causality Effect Visualization of Crash Contributing Factors on the Crash Probability during CV Pre-Deployment

4. ACCOUNTING FOR UNOBSERVABLE HETEROGENEITY IN IDENTIFICATION OF CRITICAL CRASH CAUSATIONS DURING CV PRE-DEPLOYMENT

4.1 Introduction

Reducing the rate of critical crashes is one of the ultimate goals of the WYDOT CV Pilot on the I-80 corridor. This goal necessitates the investigation of critical crash precursors to explore and identify influential factors affecting their risk compared with non-critical crashes throughout the lengthy 402 miles of I-80. This consideration imposes unobservable factors into crash severity prediction models [35], referred to as unobservable heterogeneity.

Unobservable heterogeneity includes those crash contributing factors that are not observable. If these unobservable factors correlate with the dependent variable and observable predictors in the models, the results would be biased [36]–[39]. Unobserved heterogeneity is a result of nesting in the design matrix and causes dependency across observations [39], [40], while independence is the fundamental assumption of statistical modeling. Violating this assumption would result in unreliable parameter estimate. The random parameter approach, more specifically random intercept, under the concept of hierarchical modeling could account for this issue [41].

A majority of previous studies used a single random parameter, including random intercept or random slop, within two-level hierarchical modeling to account for unobserved heterogeneity [9]. On this point, note that if individual crash observations in the design matrix are nested within more than one random factor, those additional random factors must be accounted for by hierarchical modeling [37]. Accordingly, this part of the documents shows that critical crash occurrence on I-80 in Wyoming is simultaneously nested in different longitudinal grade categories and road surface conditions. Hence, it is essential to consider these two random factors at the same time in the crash severity prediction model. To this aim, the research team took the practical advantage of crossed classified random effect modeling (CCREM) [39] to develop a baseline in identifying critical crash causations during the CV pre-deployment.

4.2 Nested Random Factors versus Crossed Random Factors

In a data hierarchy, individual observations at level-1 might be clustered between different groups of factor(s) at higher levels [42]. The terminology of crossed-random intercept in CCREM is presented for the case of additional random factors that are crossing each other as opposed to being nested within each other [43]. According to Figure 4.1, V and U are two factors within which the individuals are clustered. From Figure 4.1A, nested factors can be encoded when any group of the first factor (V) at level 2 appears only in one particular group of the second factor (U) at level 3. Thus, V itself is nested within U. In this situation, the number of levels in hierarchical modeling should be increased (i.e., from two-level modeling to three-level modeling in this example) [43]–[45].

On the other hand, Figure 4.1C presents crossed factors where any given group of V can arise in more than one group of U [43]. Equivalently, individual observations are not only clustered between different groups of U and V but also are grouped between different combinations of U and V. Using crossed factors at level 2 in two-level modeling under the concept of CCREM can

account for this type of grouping [43]–[49]. Figures 4.1B and 4.1D indicate the cross-classification of individual observations for nested factors and crossed factors, respectively. The white cells show there is no observation, and the number of hypothetical observations in black cells is more than the gray cells. Hence, in the case of nested factors, the cross-classification of observations would be almost diagonal [43], [49]. It was found that the individual crashes on the 402 miles of I-80 are nested within both longitudinal grade categories and road surface categories that are crossing each other at level 2 of data hierarchy, necessitating the application of CCREM.

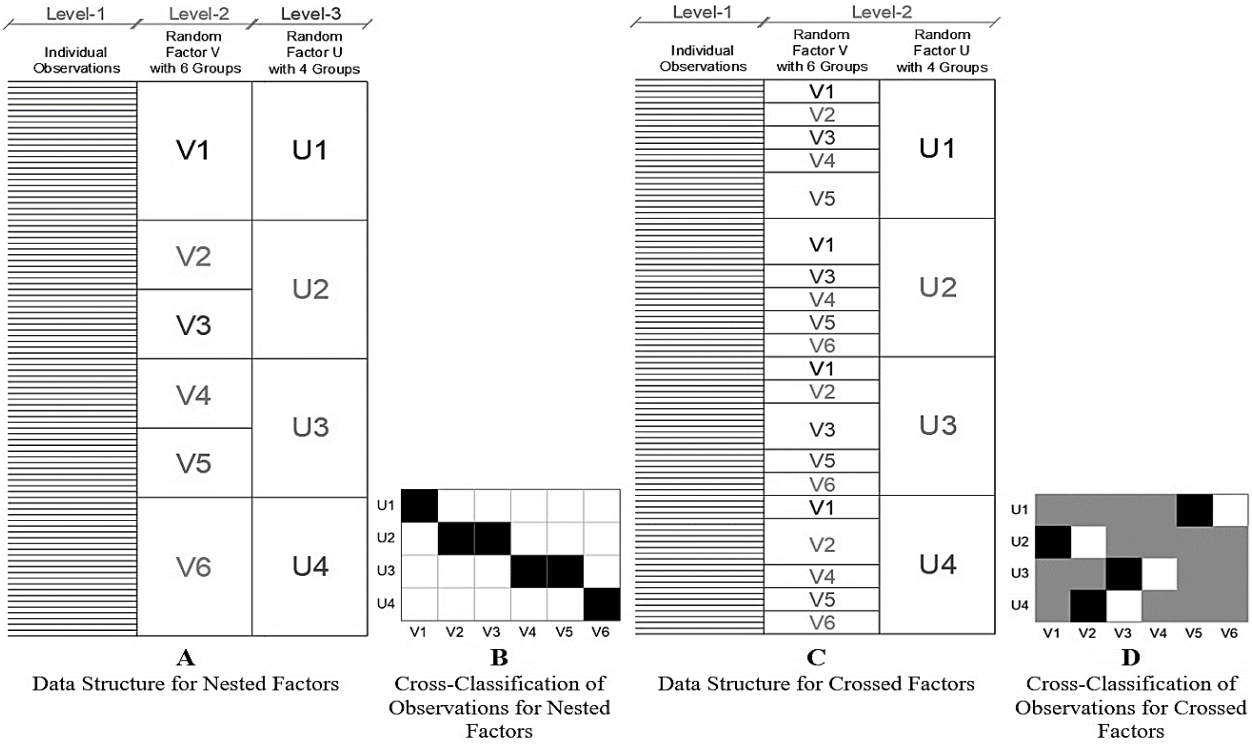


Figure 4.1 Nested Factors versus Crossed Factors

4.3 Variable Description

During the first five months of 2017 (CV pre-deployment), 207 crashes were included for the investigation, among which 28 cases were critical crashes. Four databases were conflated to structure the final dataset used in this study. The primary databases comprised crash reports, aggregated roadway geometry characteristics, weather conditions, and real-time traffic observation through the 402 miles of I-80. For each crash, the final dataset included a variety of predictors corresponding to aggregate non-traffic variables and disaggregate real-time traffic-related variables.

In total, 24 real-time traffic-related variables were characterized following two steps. First, six classes of real-time traffic-related predictors were considered to capture traffic characteristics during 15-minute time windows before crashes. These were 1) mean speed, 2) operating speed (i.e., 85th percentile of speed distribution), 3) standard deviations of speed, 4) traffic volume, 5) truck proportion, and 6) the coefficient of variation in speed. Second, each of these predictor classes was measured based on four types of measurements concerning crash locations. Type 1 and Type 2, respectively, evaluate traffic-related predictors upstream and downstream of crash

locations. Type 3 measured the predictors at the crash locations by averaging its corresponding value between upstream and downstream. Type 4 was allocated to detect the spatial variation of the traffic pattern from upstream to downstream by subtracting Type 1 variables from Type 2. Before initiating the analysis, the important variables were defined using CII as the FS technique, leading to selecting 14 variables described in Table 4.1.

Table 4.1 Description of Explanatory Variables to Explore Critical Crash Causation during CV Pre-Deployment

Continuous Variable	Description	Unit.	Min.	Mean	Max.	SD.
airTemp	Air Temperature	F°	-0.58	28.58	68.72	12.44
roadTemp	Road Temperature	F°	5.00	34.84	97.88	16.09
LshWidth	Width of Left Shoulder	ft	2.00	3.81	4.00	0.59
RshWidth	Width of Right Shoulder	ft	2.00	9.24	10.00	1.43
aveSDSp	Average of Standard Deviation of Speed at Upstream and Downstream	mph	3.40	7.98	20.76	1.95
Sp85U	Operational Speed (The 85th Percentile Speed Distribution) at Upstream	mph	30.48	70.73	85.13	10.22
SpCVARU	Coefficient of Variation in Speed at Upstream	NA	0.06	0.13	0.72	0.06
diffVol	Spatial Difference in Traffic Volume from Upstream to Downstream	Veh/15min	-122	2.48	138	34.10
aveVol	Average Traffic Volume at Upstream and Downstream	Veh/15min	6.50	55.21	172.50	34.76
TrpctD	Truck Proportion at Downstream	NA	0.00	0.58	1.00	0.19

Categorical Variable	Description	# of Level	Description of the levels	# of Occurrence
lanetype	Types of the Road Pavement	(Reference) Category 1	Asphalt	100
		Category 2	Concrete	107
curveCat	Presence of Horizontal Curve	(Reference) Category 1	Tangent Segment	144
		Category 2	Horizontal Curve	63
gradeCat	Longitudinal Grade (G)	(Reference) Category 1	$G \leq -2\%$	45
		Category 2	$-2\% \leq G \leq 0\%$	71
		Category 3	$0\% \leq G \leq +2\%$	55
		Category 4	$+2\% \leq G$	36
RoadCond	Road Surface Conditions	(Reference) Category 1	Dry	58
		Category 2	Wet	35
		Category 3	Ice/Frost	76
		Category 4	Snow	38

4.4 Hierarchical Modeling

Hierarchical modeling can account for the hierarchical structure of the dataset as a result of unobserved heterogeneity when individual observations are nested in some random factors at a higher level of data structure [37]. Specifically relating to this analysis, observations at level 1 (e.g., critical and non-critical crash occurrences) might be nested in one or more factors at level 2 (e.g., road surface conditions and longitudinal grades), making correlation across observations within clusters of the random factors. This hypothesis was assessed by investigating the within-cluster correlation of observations using the intraclass correlation coefficient (ICC). ICC decomposes the variation of the outcomes into within-cluster and between-cluster variations [50] by taking any positive values from 0 to 1 that quantifies the proportion of between-cluster variation to the total variation of outcomes [50]. Equation 4.1 presents ICC for logistic regression, where σ_0^2 is the between-cluster variance (i.e., the random intercept variance) and the term $(\pi^2/3)$ refers to the level 1 variance component in standard logistic distribution [50].

$$ICC = \frac{\sigma_0^2}{\sigma_0^2 + \pi^2/3} \qquad \text{Equation 4.1}$$

4.4.1 Hierarchical Logistic Regression

Four models were developed, and their performances were compared using three performance measures: the Watanabe-Akaike information criterion (WAIC), area under the curve (AUC), and Akaike weights based on WAIC. ICCs were investigated to examine whether or not the crash observations are nested within longitudinal grade categories and road surface conditions.

Model 1. The first model serves as a baseline where the assumption of having statistically independent observations was assumed to be met in the crash dataset. All the independent variables were treated as crash contributing factors (i.e., fixed parameters), and a simple Bayesian logistic regression was conducted. Equation 4.2 parametrizes Model-1, where p_i is the probability of critical crash occurrence for i^{th} observation among n instances, β_0 is the intercept, $\boldsymbol{\beta}$ is the vector of coefficients corresponding to predictors, and $\boldsymbol{X}$ is the vector of predictors.

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \boldsymbol{\beta X} \qquad \text{Equation 4.2}$$

Model 2. This model hypothesizes that the crash occurrences are nested within road surface conditions and treats this factor as a random intercept with four clusters at level-2 of the data structure. Model-2 was developed based on Equation 4.3. In Equation 4.3, β_{00} is the fixed intercept, R_{0j} is the deviation of the fixed intercept due to different road surface conditions where $R_{0j} \sim Normal(0, \sigma_r^2)$, p_{ij} is the probability of critical crash occurrence for i^{th} observation in j^{th} road surface condition, $\boldsymbol{X}$ is the vector of the rest of the predictors, and $\boldsymbol{\beta}$ is the vector of coefficients corresponding to these predictors.

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = (\beta_{00} + R_{0j}) + \boldsymbol{\beta X} \qquad \text{Equation 4.3}$$

Model 3. This model follows Equation 4.4, which has a procedure similar to Model 2; however, it hypothesizes that the critical crash occurrence is nested within four longitudinal grade clusters at level 2 of the data structure. In Equation 4.4, β_{00} is the fixed intercept, G_{0j} is the deviation of the fixed intercept due to different longitudinal grade categories where $G_{0j} \sim Normal(0, \sigma_g^2)$, p_{ij} is the probability of critical crash occurrence for i^{th} observation in j^{th} longitudinal grade category, $\boldsymbol{X}$ is the vector of the rest of the predictors, and $\boldsymbol{\beta}$ is the vector of coefficients corresponding to these predictors.

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = (\beta_{00} + G_{0j}) + \boldsymbol{\beta X} \qquad \text{Equation 4.4}$$

Model 4. Investigation of the crossed effect of longitudinal grade categories and road surface conditions within the concept of CCREM at level 2 of the data structure is performed by Model 4 following Equation 4.5. In Equation 4.5, $\beta_{0(00)}$ is the fixed intercept, $G_{0(0k)}$ and $R_{0(j0)}$ are, respectively, the deviation of the fixed intercept due to different longitudinal grades and varying road surface conditions, where $G_{0(0k)} \sim Normal(0, \sigma_g^2)$ and $R_{0(j0)} \sim Normal(0, \sigma_r^2)$. $p_{i(jk)}$ is the probability of critical crash occurrence for i^{th} observation in j^{th} road surface condition occurred at k^{th} longitudinal grade categories, $\boldsymbol{X}$ is the vector of the rest of the predictors, and $\boldsymbol{\beta}$ is the vector of coefficients corresponding to these predictors. Note that there are different alternative notations in the literature for CCREM.

$$\log\left(\frac{p_{i(jk)}}{1-p_{i(jk)}}\right) = (\beta_{0(00)} + R_{0(j0)} + G_{0(0k)}) + \boldsymbol{\beta X} \quad \text{Equation 4.5}$$

4.5 Bayesian Inference

All models were conducted under Bayesian inference utilizing no-U-turn sampler (NUTS) for sampling [51]. Using three Markov Chain Monte Carlo (MCMC) simulations for the four models, 4,000 samples were generated within which the first 1,000 samples were considered as the warmup samples and eliminated from the posterior distribution. The continuous variables were scaled and centered, which can improve the efficiency of MCMC [51]. Among the 3,000 samples after warmup, there was no divergence in sampling and no warning regarding the maximum tree depth that shows the samples had an acceptable efficiency [51]. Furthermore, the same priors using non-informative distributions were considered under the logit scale to perform a fair comparison among the four models. Equations 4.6 and 4.7 present the chosen priors for the fixed intercepts (i.e., β_{00} and $\beta_{0(00)}$) and coefficients ($\boldsymbol{\beta}$) in all developed models, respectively.

$$\beta_{00}\ and\ \beta_{000} \sim Normal(-1.85,\ 2) \quad \text{Equation 4.6}$$

$$\beta \sim Normal(0,\ 2) \quad \text{Equation 4.7}$$

For the hierarchical models (i.e., Model 2, 3, and 4), the extent of the deviations of cluster-specific intercepts from the fixed intercepts (R_{0j}, $R_{0(j0)}$, G_{0j}, and $G_{0(0k)}$) are supposed to obtain posterior distributions. Hence, σ_r and σ_g require prior distributions because $R_{0j} \sim Normal(0, \sigma_r^2)$, $R_{0(j0)} \sim Normal(0, \sigma_r^2)$, $G_{0j} \sim Normal(0, \sigma_g^2)$, and $G_{0(0k)} \sim Normal(0, \sigma_g^2)$. Both of σ_r and σ_g play the same role as σ_0 in Equation 4.1 and demonstrate how much of the variability of the outcomes is due to its systematic variation at level 2 of the data structure. Since σ_r and σ_g are standard deviations, they can only take positive values and follow half-normal distributions. Hence their non-informative prior distributions are, respectively, presented by Equations 4.8 and 4.9. Similarly, the standard deviations were set to be 2, to have a large value under the logit scale.

$$\sigma_r \sim HalfNormal(0,\ 2) \quad \text{Equation 4.8}$$

$$\sigma_g \sim HalfNormal(0,\ 2) \quad \text{Equation 4.9}$$

4.6 Results Interpretation and Implication

Table 4.2 presents the results and compares the performances of the four models using WAIC, AUC, and Akaike weight. In Model 2, the standard deviation of the random road surface intercept was estimated to be 0.58, leading to the corresponding ICC=0.09 that indicates almost 9% of the variation of the outcomes lies between different clusters of this factor. Similarly, Model 3 estimated the standard deviation of random longitudinal grade intercept to be 0.58 which resulted in ICC=0.09, implying that 9% of the variation of the outcomes is due to the different groups of road surface conditions. These ICCs confirmed that the critical crash occurrences are nested within these two factors at level 2 of the data structure. Note that in Model 4, the corresponding ICC cannot be reported since it is not specified for models with crossed random effects [43].

Table 4.2 Models' Estimations for Identification of Critical Crash Causations during CV Pre-Deployment

		Model-1				Model-2				Model-3				Model-4			
Variables		Est.	2.5% CI	97.5% CI	Sig.	Est.	2.5% CI	97.5% CI	Sig.	Est.	2.5% CI	97.5% CI	Sig.	Est.	2.5% CI	97.5% CI	Sig.
Intercept		-1.10	-2.37	0.08		-1.38	-2.72	-0.17	*	-1.51	-2.83	-0.26	*	-1.83	-2.89	-0.78	*
scaleaveVol		0.40	-0.11	0.92		0.39	-0.14	0.91		0.37	-0.15	0.89		0.35	-0.16	0.87	
scaleLshWidth		-0.02	-0.54	0.55		-0.01	-0.51	0.55		-0.05	-0.55	0.51		-0.04	-0.53	0.51	
scalediffVol		-0.46	-0.96	0.03		-0.44	-0.93	0.02		-0.45	-0.93	0.03		-0.46	-0.90	-0.02	*
scaleroadTemp		-0.36	-1.43	0.63		-0.36	-1.40	0.60		-0.34	-1.36	0.62		-0.32	-1.38	0.67	
scaleSp85U		0.33	-0.40	1.12		0.34	-0.38	1.08		0.31	-0.40	1.09		0.34	-0.35	1.05	
scaleaveSDSp		-0.83	-1.66	-0.08	*[b]	-0.82	-1.65	-0.04	*	-0.80	-1.60	-0.06	*	-0.80	-1.58	-0.06	*
scaleSpCVARU		0.78	-0.13	1.65		0.75	-0.19	1.63		0.74	-0.17	1.61		0.71	-0.14	1.54	
scaleTrpctD		-0.01	-0.51	0.51		-0.01	-0.51	0.49		-0.02	-0.52	0.50		-0.03	-0.51	0.46	
scaleairTemp		0.60	-0.40	1.63		0.60	-0.36	1.60		0.57	-0.40	1.59		0.55	-0.44	1.58	
scaleRshWidth		-0.46	-0.88	-0.04	*	-0.48	-0.90	-0.05	*	-0.46	-0.87	-0.04	*	-0.48	-0.89	-0.07	*
lanetype	Cat. 2	-0.83	-1.88	0.21		-0.85	-1.87	0.12		-0.75	-1.77	0.23		-0.77	-1.79	0.21	
curveCat	Cat. 2	-0.61	-1.71	0.42		-0.60	-1.68	0.44		-0.60	-1.68	0.40		-0.57	-1.65	0.40	
Road Surfac Cond.	Cat. 2	-0.93	-2.68	0.60						-0.97	-2.74	0.57					
	Cat. 3	-0.12	-1.48	1.25						-0.21	-1.56	1.14					
	Cat. 4	-0.75	-2.60	0.96						-0.79	-2.62	0.87					
Long. Grade	Cat. 2	-0.94	-2.12	0.24		-0.92	-2.10	0.24									
	Cat. 3	-0.23	-1.42	0.94		-0.18	-1.36	0.99									
	Cat. 4	-0.68	-2.30	0.88		-0.66	-2.30	0.88									
σ_r						**0.58**								**0.46**			
ICC for R_{0j}						**0.09**								**NA[a]**			
σ_g										**0.58**				**0.46**			
ICC for G_{0j}										**0.09**				**NA**			
WAIC		**183.4 (23.5)[b]**				**179.9 (22.8)**				**180.4 (23.0)**				**176.1 (22.2)**			
AUC		**69.5**				**73.8**				**71.1**				**76.6**			
Akaike Weight		**0.02**				**0.12**				**0.09**				**0.77**			

Note:

Description of variables can be found in Table 4; CI= Credible Interval; ICC= Intraclass Correlation Coefficient; WAIC= Watanabe-Akaike Information Criterion; AUC= Area Under the Curve; σ_r= The standard deviation of the random road surface intercept; σ_g= The standard deviation of random longitudinal grade intercept; R_{0j}= The deviation of the fixed intercept due to different road surface conditions; G_{0j}= The deviation of the fixed intercept due to different longitudinal grade categories.

* Representative of statistically significant predictor under 95% CI.

[a] The corresponding ICC cannot be reported since it is not specified for models with crossed random effects.

[b] The number in parentheses are the associated standard error.

Model 4, however, outperformed all others and achieved the minimum WAIC, maximum AUC, and maximum Akaike weight by 176.1, 76.6%, and 0.77, in the same order. These results reveal the potential of considering the crossed effects of two random intercepts at level 2 of the data hierarchy by capturing more variation of the outcome. Model 4 obtained minimum standard error of WAIC, which means Model 4 not only has superior predictive performance, but it also could minimize the uncertainty in WAIC estimates.

4.6.1 Analysis of Crossed-Random Intercepts

Based on Equation 4.5, since both the road surface and the longitudinal grade have four clusters, 16 crossed-random intercepts can be generated under CCREM. These crossed-random intercepts explain the average of the logit of critical crash probability under any combined effect of

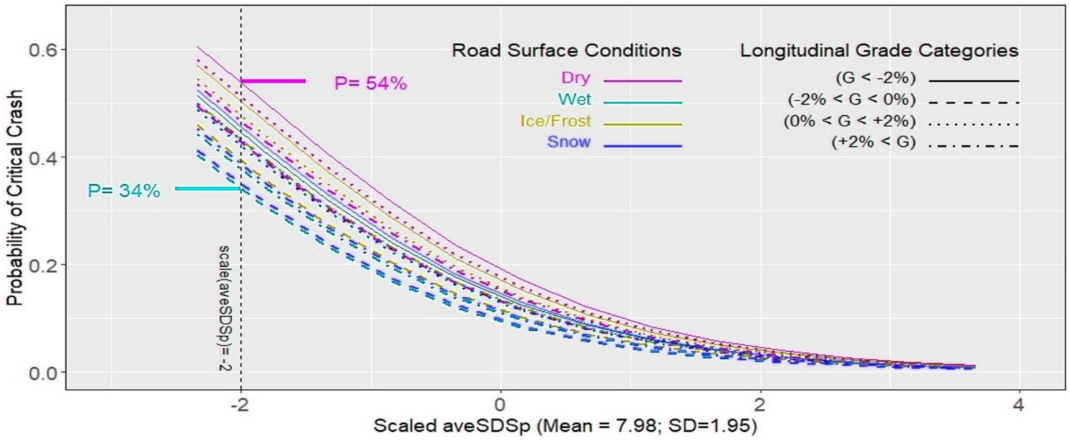
longitudinal grades and road surface conditions. Relevant to Equation 4.5, Figure 4.2 gives 16 rounded crossed-random intercepts under crossing longitudinal grade categories ($G_{0(0k)}$) and road surface conditions ($R_{0(j0)}$) [i.e., $(\beta_{0(00)} + R_{0(j0)} + G_{0(0k)})$].

Figure 4.2 Crossed-Random Intercepts for Different Combinations of Road Surface and Longitudinal Grade Categories

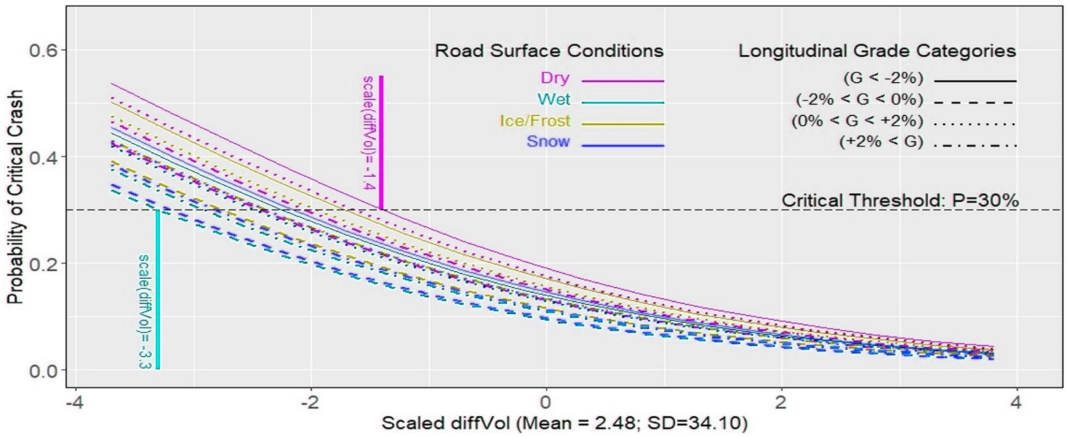
Figure 4.2 depicts that on steepest downgrades (G < -2%) when the road surface is dry, the crossed-random intercept can obtain the maximum possible value by -1.42, leading to the maximum probability of critical crashes compared with other situations. On the other hand, the combined effect of the wet road surface and less steep downgrades (i.e., -2% < G < 0%) has resulted in the minimum crossed-random intercept by -2.23, corresponding to the minimum critical crash risk in comparison with other possible conditions.

4.7 Recommendations for WYDOT

Figure 4.3 illustrates the effect of the 16 crossed-random intercepts on the probability of the critical crashes through accounting for two traffic-related predictors (i.e., *diffVol* and *aveSDSp*). The variation of critical crash probabilities according to the different combined effects of the road surface and categorized longitudinal grade at any given value of these real-time predictors can be seen. For instance, from Figure 4.3A, when the scaled *aveSDSp* is -2, the critical crash probability is 34% on the wet road and less steep downgrade. This probability, however, at the same value of *aveSDSp*, can increase up to 54% on the dry and steepest downgrade. Figure 4.3B demonstrates a practical application of the crossed-random intercept in CCREM for active traffic management (ATM) on the I-80 corridor. Suppose 30% in critical crash probability is set to be the hypothetical threshold. This threshold, however, can be met at different values of real-time traffic-related predictors according to various combinations of environmental conditions and roadway geometry characteristics. For instance, Figure 4.3B elucidates that through the range of scaled *diffVol*, the mentioned threshold can be met when this variable is -1.4 on a dry and steepest downgrade, or even at -3.3 on a wet road surface and less steep downgrade. This conclusion draws a valuable insight for ATM in timely application interventions not only based on real-time traffic-related predictors, but also considering the environmental factors.



A- Marginal Effect of Average of Standard Deviation of Speed at Upstream and Downstream of the Crash Location (aveSDSp) under Varying Crossed Random Intercepts



B- Marginal Effect of Spatial Difference in Traffic Volume from Upstream to Downstream of the Crash Location (diffVol) under Varying Crossed Random Intercepts

Figure 4.3 Crossed Random Effect of Longitudinal Grade and Road Surface Intercepts on Critical Crash Risk through the Range of Significant Traffic-Related Variables

5. ANALYSIS, MODELING, AND SIMULATION (AMS) FRAMEWORK

5.1 Introduction

The safety benefits of the WYDOT CV Pilot would be gained mostly from the change in driving behavior due to the effect of real-time CV warnings provided under VANET. Therefore, there is an essential need to quantify and measure CV drivers' behavioral alterations under the effect of CV notifications. Accordingly, comprehensive high-fidelity driving simulator experiments were conducted at the University of Wyoming Driving Simulator Lab (WyoSafeSim), which consists of truck and passenger car driving simulators. In the developed experiments and under the concept of with/without analysis, the effect of several CV applications on truck drivers' behavior was tested. The main goal of this procedure was drawn to develop a well-calibrated and validated analysis, modeling, and simulation (AMS) framework to be used for the safety performance assessment of the WYDOT CV Pilot based on the quantification of the drivers' behavioral alteration. Under the AMS framework, the research team assessed CV safety performance based on two techniques: first, the analysis of vehicles' trajectory-level observations from the driving simulator to quantify drivers' behavioral alteration due to CV notifications utilizing K-SMoS; second, the conflation of driving simulator experiments with microsimulation modeling to unveil the CVs' safety effect on the entire traffic stream under varying MPRs using C-SMoS.

5.2 Experimental Approach

At the WyoSafeSim lab, driving simulator analyses with a focus on truck drivers were performed through two main parts: the CV training phase and comprehensive truck driving simulator experiments for with/without analysis. A summary of the procedure follows.

5.2.1 CV Training Phase

The proposed CV training program contains two major components: an e-learning module and a hands-on driving simulator training module. The research team determined the contents of the e-learning module and simulation scenarios of the driving simulator training module based on the requirements of USDOT's training tasks. Professional snowplow truck drivers from WYDOT were invited to participate and evaluate the updated training materials and procedure. The framework of the CV training program is presented in Figure 5.1.

The training started with a program background introduction and completion of the consent form. In the next step, a pre-training questionnaire was conducted to collect the demographic profile, driving experience, crash history, and experience of each participant. In the e-learning module, participants were introduced to the basic components of the Wyoming CV system (e.g., human machine interface [HMI]), the concept and function of each CV warnings, and the proper response they should implement.

5.2.1.1 Human Machine Interface (HMI) Design

One of the key components of the CV system is the onboard HMI that delivers the received CV warnings to drivers, as illustrated in Figure 5.2. The participants learned from the e-learning module that the main function of the CV system is only to provide timely warnings to inform the

driver with real-time information about upcoming hazardous roadway and weather conditions, and that the CV system will not control his/her vehicle or affect his/her driving.

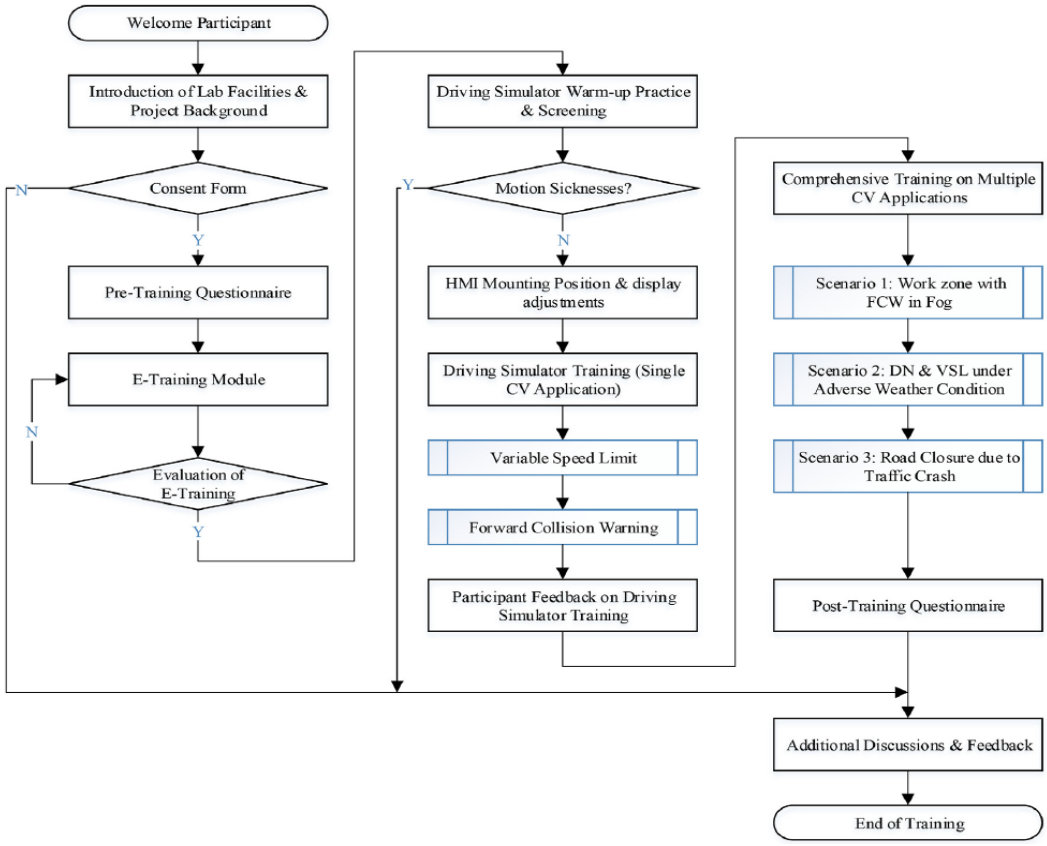


Figure 5.1 Flow Chart of the Developed Wyoming CV Training Framework

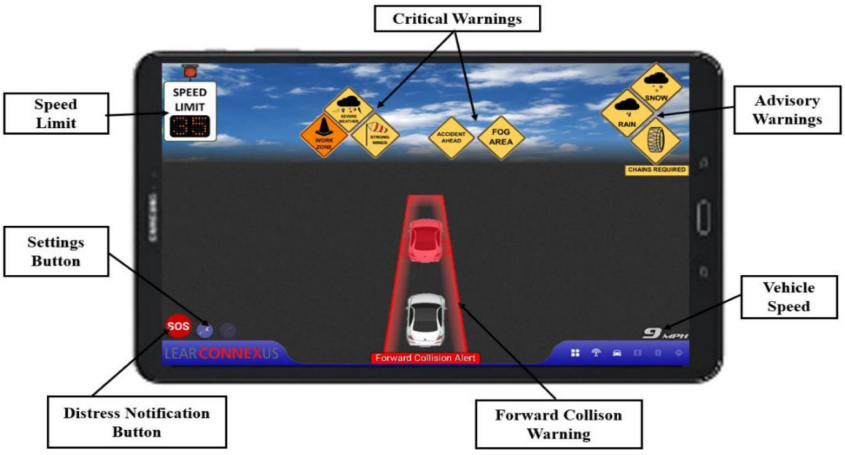


Figure 5.2 The Layout of Human–Machine Interface Display

The HMI was designed in a 10.1-inch android tablet and integrated into the truck driving simulator. The displayed TIMs' location mimics the designed HMI provided in the WYDOT CV Pilot [52], which was based on the importance of the communicated message. Participants were also educated about the priority level of CV warnings and how to recognize the most urgent alert in case of multiple alerts appearing simultaneously on the HMI.

5.2.2 Hands-on Driving Simulator Training Module

5.2.2.1 Participants

A total of 24 professional truck drivers with a commercial driver's license (CDL) were recruited for the experiment. The recruited drivers were familiar with traffic operation performance on I-80 in Wyoming under varying weather conditions and roadway characteristics employed by WYDOT. The minimum required sample size was calculated based on Equation 5.1 [53], where Z is the z-score for the defined confidence level, p is the population proportion, e is the margin of error, and N is the population size.

$$Sample\ Size = \frac{\frac{Z^2 \times p \times (1-p)}{e^2}}{1+\left(\frac{Z^2 \times p \times (1-p)}{e^2 \times N}\right)} \quad \text{Equation 5.1}$$

An initial confidence interval of 90% with the corresponding z-score of 1.65 (Z) and a 10% margin of error (e) were determined to calculate the sample size. Prior knowledge of the population proportion (p) was used to calculate the optimum sample size. However, it is challenging to have reliable prior information about the population's acceptance of the CV in advance. Due to the known benefits of the CV applications in the literature, it was assumed that 90% of participants would affirm the CV's benefits. This assumption was confirmed after conducting the DS experiment, where 95% of participants agreed on the significant safety benefits of the CV [54]. Accordingly, the minimum sample size was calculated to be 22 participants for a population of 200 truck drivers (N) that will have CV technology equipped in their trucks during the WYDOT CV Pilot post-deployment on I-80.

5.2.2.2 Equipment

The CV hands-on training module was performed using a high-fidelity driving simulator lab located at the University of Wyoming (WyoSafeSim Lab), which is a freight truck (2000 Sterling AT9500 18-wheeler semi-trailer) cockpit cab. The simulator has a 140-degree field of view (FOV) housed in a closed room, where temperature, sound, and light are fully controlled. The simulators have an open architecture software with complete source code of simulation creator tool, which offers the flexibility of designed roadways and developing driving scenarios. The simulators allow three degrees of freedom (roll, pitch, and heave) using four linear actuators as the D-Box motion platform. The simulators provide motion cues to immerse the driver into a real driving experience with changes in kinematics, such as velocity, acceleration, and deceleration, which vary according to the passenger car or truck's dynamic performance. Vehicle kinematics from the driving simulator was recorded on a frequency of 60 Hz coupled with the video recordings for the center screen and the HMI using the SimObserver software.

5.2.2.3 Warm-Up Driving Practice

The warm-up driving practice was conducted to let the participants get familiar with the functions of the driving simulator environment. The driving simulator training platform was designed as a two-way four-lane freeway segment with a 75-mph speed limit to represent the basic operating conditions of freeway I-80 in Wyoming. During the warm-up driving practice, each participant drove for five to 10 minutes; additional practice time was provided to participants who needed more time.

5.2.3 Comprehensive Truck Driving Simulator Experiments for With/Without Analysis

Three comprehensive simulation scenarios were developed to simulate different real-world traffic and weather conditions on an I-80-like freeway. Each truck driver drove each of the scenarios two times in CV and non-CV environments (i.e., with/without analysis presented in Figure 5.3), resulting in six main scenarios per driver, narrated in Figure 5.4.

A B

Figure 5.3 With/Without Analysis in the Truck Driving Simulator Experiments

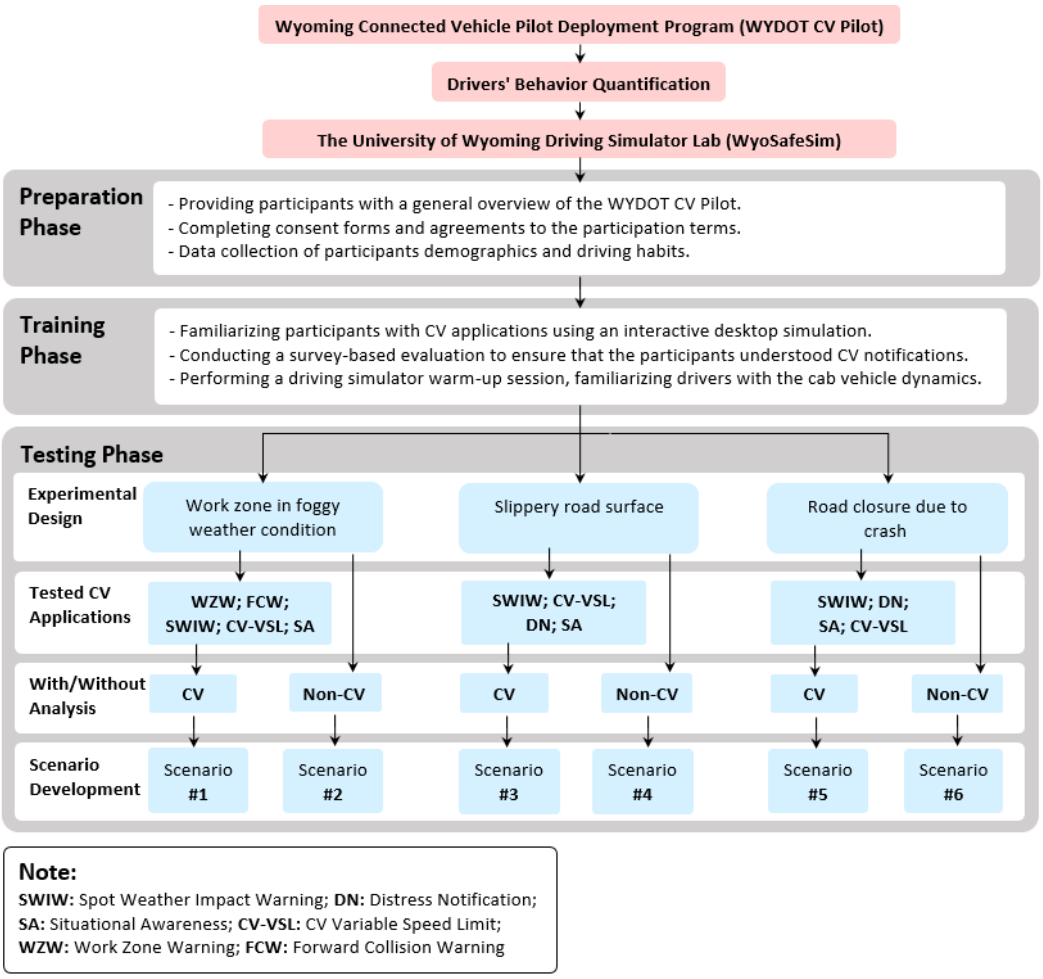


Figure 5.4 Truck Driving Simulator Experiments at the WyoSafeSim

The driving simulator scenario consists of a work zone in fog, a slippery road surface under adverse weather conditions, and road closure because of a crash in severe weather. Under the mentioned driving simulator scenarios, the effect of several CV applications on driver behavior under various traffic critical safety events were tested, including SWIW, DN, SA, CV variable speed limit (CV-VSL), FCW, and WZW.

5.2.4 Evaluation of Driving Simulator Experiments

After experiencing the Wyoming CV application in the driving simulator study, a comprehensive post-drive questionnaire survey was employed to collect participants' qualitative opinions regarding their preferences on different CV warning modalities and the effectiveness of CV technology under various real-world driving conditions. Results show that the majority of participants (96.2%) preferred to have the CV warnings displayed at the combination of visual and auditory modalities. For the auditory-warning modality, it was found that using a simple "beep" sound for advisory warnings and a series of louder "beep" sounds for critical warnings would best draw a driver's attention. For the visual warning modality, results show that, by grouping CV warnings to different priority levels and presenting warnings that have a higher priority closer to

the driver (i.e., left side on the CV HMI), drivers tended to more easily perceive the imminent safety hazard when multiple warnings were displayed on the CV HMI [52], [54]–[56].

Furthermore, the research team investigated driver distraction levels due to the HMI based on vehicles' kinematics analysis, questionnaire surveys, and eye-tracking analysis. The result affirmed the appropriateness of the designed HMI where the quantified drivers' additional cognitive workload due to the HMI was found to be within the acceptable ranges proposed by the existing guidelines [57], [58].

5.2.5 Microsimulation Modeling in the AMS framework

To unveil the effect of the WYDOT CV Pilot on the entire traffic stream, the microsimulation modeling was integrated into the developed AMS. The primary reason for this consideration is being limited by the low number of CVs in a real-world setting. Specifically, under the WYDOT CV Pilot, there are only 400 heavy trucks equipped with the CV technology, which is a negligible proportion of the traffic volume on I-80 with average annual daily traffic (AADT) of 10,000 to 20,000. Accordingly, driving simulator experiments were integrated into VISSIM® traffic microsimulation modeling to evaluate macro- and micro-level factors contributing to CVs' safety effect on I-80's entire traffic stream. In this regard, the important point is the calibration of CV longitudinal and lateral driving behavior (i.e., car-following parameters) in the microsimulation modeling. Most previous studies performed this calibration based on assumptions regarding driver behavior alterations in the CV environment. This research, however, followed a more realistic approach and presented innovative traffic safety research perspectives.

Driver behavior under CV and non-CV environments were quantified by analyzing the results of driving simulator experiments. This quantification was applied in the microsimulation modeling for developing a well-calibrated model. The proposed AMS framework employs the VISSIM simulation with the Surrogate Safety Assessment Model (SSAM) for safety performance evaluation. Accordingly, C-SMoS were employed for the safety performance assessment of the WYDOT CV Pilot.

6. MITIGATING WORK ZONE-RELATED CRASHES IN CV ENVIRONMENTS

6.1 Introduction

With the rapidly aging roadway infrastructure, essential needs for their maintenance, and roadway geometry modification requirements [59], work zones (WZs) have become prevalent in the U.S. transportation network. Statistics show that in 2015 almost 96,626 WZ-related crashes occurred, indicating one WZ crash every 5.4 minutes [60]. Moreover, reduced visibility caused by foggy conditions on I-80 can intensify the risk of crashes in WZs as it imposes an unusual pattern in driving behaviors, where drivers must compensate for the reduced visibility and its associated risks [61]. On average, 21% of annual crashes and 16% of all fatal crashes are weather-related, where crashes in foggy conditions are prone to be more severe and can involve multiple vehicles compared with clear conditions. In WZ areas, drivers' mental workload increases, impacting their performance and increasing their driving behavior errors [62]. Timely and dynamic increases in drivers' situational awareness can mitigate this lack of performance [63].

On this point, CV technology has been proven to effectively enhance drivers' situational awareness [63]. Various CV advisory/warnings notifications, such as the work zone warning (WZW) and spot weather impact warning (SWIW) application, aid CV drivers to perform better driving behavior relating to the chosen headway, late and early merging, speed selection, and lane changing maneuvers in WZs [63]. This part of the document presents the investigation of the safety performance of varying CV-MPRs on I-80 WZs under fog based on a conflation of driving simulator and microsimulation modeling using various C-SMoS.

6.2 Driving Simulator Experimental Design

According to Figure 6.1, the driving simulator experiment involved testing two distinct pilot applications: the SWIW and the WZW. The SWIW application consisted of two notifications communicated before the WZ advance warning area. The first weather notification informed drivers of an upcoming fog. The subsequent weather notification suggested a 55-mph advisory speed for the upcoming foggy condition. Following the notifications, the fog was gradually introduced, and the visibility was reduced to 250 feet. Farther downstream, participants received four distinct WZ advance warning area notifications sequentially under the foggy weather conditions.

Under with/without analysis, drivers had to rely on the roadside static and dynamic signs for guidance in the baseline scenario. In the CV scenario, the first WZ notification informed drivers of an upcoming work zone in one mile. The second notification updated the distance placard on the WZ warning sign to 0.5 miles and alerted drivers of an upcoming reduction in the speed limit to 45 mph. The third notification updated the distance on the WZ warning sign placard to 1,000 feet. A right lane closure (i.e., lane merge) warning and a 45 mph speed limit sign were displayed in the last notification. The 45 mph speed limit warning in the CV scenario was communicated shortly in advance of the corresponding static road speed limit sign to allow drivers to discern both the static sign and the HMI warning in a non-conflicting manner. Distinctively, the lane merge warning in the CV scenario was displayed well in advance of the static road sign to give drivers an increased timespan to merge toward the open lane. The early lane merge warning, a prominent

feature of WYDOT Pilot's WZW application, was devised with the hope of reducing late and forced merging behaviors at work zones, particularly for large trucks.

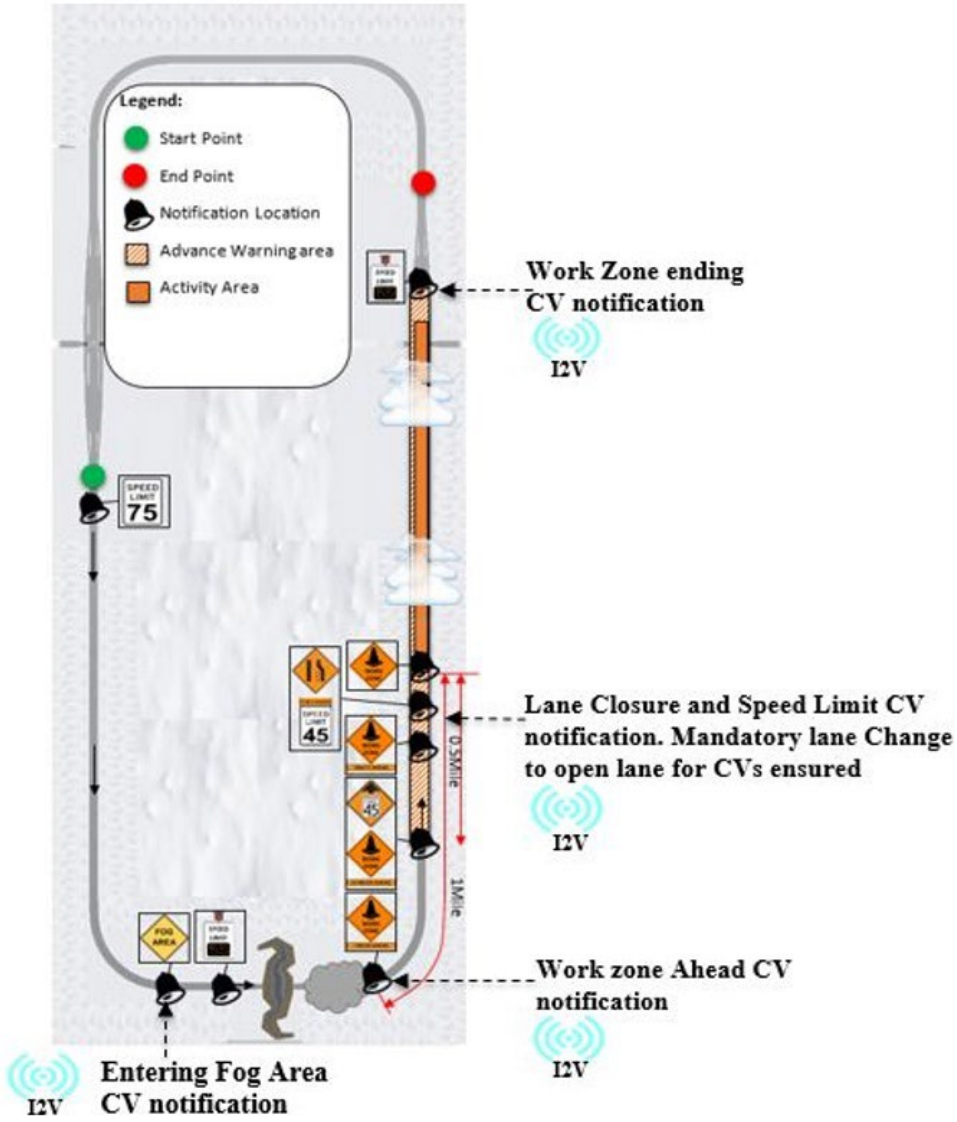


Figure 6.1 The Driving Simulator Work Zone Area Scenario under Foggy Weather Conditions

6.3 Microsimulation Modeling

The research team integrated the result of the driving simulator with microsimulation modeling under VISSIM®. In this regard, various C-SMoS and speed-related variables were investigated to show the safety performance of CV technology in I-80 WZ areas. The traffic information, including traffic volume, traffic composition, and speed distribution, were obtained from WYDOT's speed sensors installed on the I-80 corridor to calibrate and validate the microsimulation modeling.

6.3.1 Baseline Calibration and Validation

For the non-CV scenario (i.e., the baseline), the Wiedemann-99 car-following and lane-changing parameters were adapted according to the literature for a WZ area. In WZs, three parameters are substantial to be adjusted: the standstill distance (CC0), desired time headway (CC1), and the maximum additional following distance beyond the calculated safety distance (CC2) [64]. The most suitable values for these parameters were obtained using the genetic algorithm as an optimization method in MATLAB [65]. Therefore, the minimum of absolute relative error (ARE) was employed as the objective function, which can be achieved when the difference between real-life traffic volume and the traffic volume obtained from the generated parameters becomes less than 1% [65]. In this regard, Equation 6.1 presents the details, where Vol_{obs} is the real-life traffic volume observation and $Vol_{\ Sim}$ is the traffic volume resulting from the microsimulation modeling.

$$\text{ARE} = \left| \frac{(Vol_{Sim} - Vol_{Obs})}{Vol_{Obs}} \right| \qquad \text{Equation 6.1}$$

The model was validated using the Geoffrey E. Havers (GEH) statistic test [66], which shows the errors between simulated and real-time traffic volume based on Equation 6.2. According to the GEH test result, the test statistics were within an acceptable range (i.e., GEH<5 [66]).

$$GEH = \sqrt{\frac{2(Vol_{obs}(n) - Vol_{sim}(n))^2}{Vol_{obs}(n) + Vol_{sim}(n)}} \qquad \text{Equation 6.2}$$

According to Figure 6.2, the model was validated by comparing speed profiles for non-connected vehicles from driving simulator data with simulated data. Five sets of different random sampled trajectories of 23 vehicles in the simulation were selected to minimize bias. In each case, the speeds were compared to the driving simulator data using two-sample t-tests at every 0.1 miles. All comparisons were insignificant (p-value > 0.05) for all five sets, indicating no significant difference between the mean of the speeds of the driving simulator data and microsimulation data.

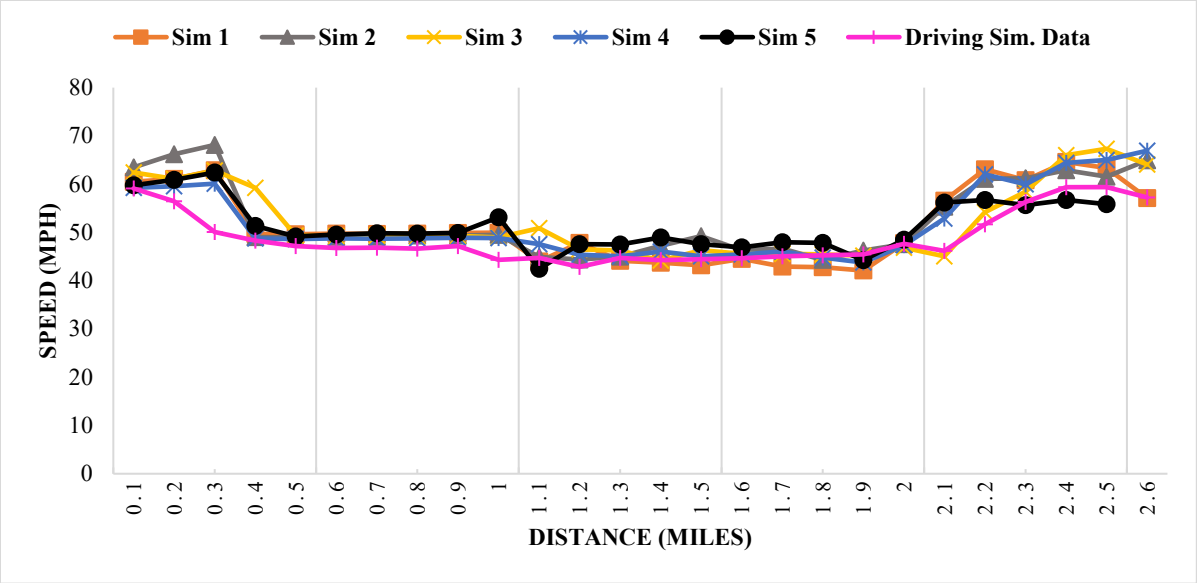


Figure 6.2 Comparison of Simulated Speed Profiles and Driving Simulator Data

6.3.2 CV Scenario Development

Six different scenarios corresponding to 0% (non-CVs), 10%, 20%, 30%, 50%, and 60% MPRs of CVs were considered to assess the safety performance of CVs. The research team utilized the results, of the driving simulator experiment to internally calibrate CV driver behaviors under the impact of the received WZW and SWIW notifications in VISSIM® by changing Wiedemann-99 parameters. The observed changes of CV driver behaviors in the driving simulator were regarded from different points of view and were applied on Wiedemann-99 parameters as follows:

- The cooperative lane change parameter associated with CVs should be "Yes" as CVs perform more cooperative lane-changing maneuvers.
- CC0, CC1, and CC2 parameters should be modified for CVs because they keep smaller distances at non-zero speed. Besides, CVs' CC8 and CC9 parameters should be altered due to much smoother CV acceleration.
- The speed distribution for a CV is more harmonized than non-CVs, with a reduction of 10% and 20% in the mean and standard deviation of speed.
- In comparison with non-CVs, CVs perform early merging in the WZ transition area.

6.4 Conflict-Based Surrogate Measure of Safety (C-SMoS)

The driver speed behavior in various areas of the WZ was evaluated using the trajectory data from the microsimulation modeling utilizing the SSAM. Conflict assessment for WZ safety was conducted using several C-SMoS, which have been extensively used in the literature as an emissary crash in microsimulation models [6]. SSAM provides a range of criteria for defining a collision. In this analysis, different types of C-SMoS were used to measure WZ crash risk as follows:

6.4.1 Time-To-Collision (TTC)

Equation 6.3 formulates time-to-collision (TTC) that is defined as "the time required for two vehicles to collide if they continue at their present speeds on the same path" [6].

$$TTC = \begin{cases} \frac{D_{1-2}}{V_2 - V_1}, & if\ V_2 > V_1 \\ \infty, & Otherwise \end{cases} \quad \text{Equation 6.3}$$

Where D_{1-2} represents the gap distance between the leading and the following vehicle, V_1 and V_2 are the speeds of the leading and following vehicles, respectively. The TTC threshold value was set to be 1.5 seconds in this study [6].

6.4.2 Modified Deceleration Rate to Avoid a Crash (MDRAC)

Modified deceleration rate to avoid a crash (MDRAC) takes the drivers' perception reaction time (PRT) into account, which is essential in the reduced visibility zone. MDRAC measures the effect of CV warnings in the WZ, where the following vehicle must adapt its speed to that of the leading vehicle. Equation 6.4 expresses MDRAC [6].

$$MDRAC = \begin{cases} \frac{V_2 - V_1}{2(TTC - R)}, & if\ TTC > R \\ \infty, & Otherwise \end{cases} \quad \text{Equation 6.4}$$

Where R is the PRT, which is defaulted as 2.5 seconds, and the rest of the variables are the same as Equation 6.3. MDRAC values greater than the threshold of 3.4m/s2 [67] are considered as a potentially risky driving situation and reflect the severity of conflicts based on TTC. MDRAC values greater than 3.4m/s2 were considered for the safety performance assessment of CVs.

6.4.3 Time Exposed Time-to-collision (TET) and Time-Integrated Time-to-collision (TIT)

To better illustrate the risk, the time exposed time-to-collision (TET) and time integrated time-to-collision (TIT) derived from TTC were utilized. TIT and TET measure the severity and exposure to risky situations, respectively [6], which is achieved by setting up various data collection points at various segments of the WZ in the simulation network. The indexes TET and TIT are defined mathematically as:

$$TET(t) = \sum_{n=1}^{N} \delta_t \times \Delta t, \quad \delta_t = \begin{cases} 1, & 0 < TTC\ (t) \le TTC^* \\ 0, & otherwise \end{cases} \quad \text{Equation 6.5}$$

$$TET = \sum_{t=1}^{T} TET(t) \quad \text{Equation 6.6}$$

$$TIT(t) = \sum_{n=1}^{N} \left[\frac{1}{TTC\ (t)} - \frac{1}{TTC^*}\right].\Delta t, \quad 0 < TTC\ (t) \le TTC^* \quad \text{Equation 6.7}$$

$$TIT = \sum_{t=1}^{T} TIT(t) \quad \text{Equation 6.8}$$

Where t is the time ID, n is the vehicle ID, N is the total number of vehicles, δ is the switching variable, Δt is the time step, which was 0.1 seconds in simulation, T is the simulation period, and TTC* is the threshold of TTC, which was set to be 1.5 seconds.

6.4.4 Data Analytics

From Table 6.1, the mean MDRAC greater than 3.4 m/s2 and the mean of critical TTCs less than 1.5 seconds indicate a critical situation. The values in the table represent the critical car following situations, where the higher the deviation of SMoS from the thresholds, the greater the crash risk. The number of critical TTCs decreases as MPR increases. Notably, the number of conflicts decreased by 80% from the base scenario, when 30% of vehicles in the traffic volume are connected. This conclusion is important for the WYDOT CV Pilot because trucks comprise 30% of total traffic on I-80 during the summer. Hence, achieving significant safety benefits by 30% MPR on the entire traffic volume shows how CVs on I-80 can improve WZ safety if all trucks are connected.

Furthermore, one of the important aspects of this study is to evaluate how the influence of CVs on driver behavior can affect WZ safety after encountering foggy conditions by considering the various sub-segments of the WZ area. Table 6.1 presents the evaluation of crash risk under varying CV-MPRs according to a summary of TIT and TET for different WZ areas, including the advance warning, transition, activity, and termination areas. The results indicate that 60% of MPR had the lowest crash risk across all areas. Looking at the 30% CV MPR, there is a notable reduction in TET and TIT across all WZ areas compared with the other preceding MPRs (0%, 10%, 20%).

Figure 6.3 shows the relationship between three SMoSs, including the number of critical TTCs, the mean of critical TTC, and the mean of critical MDRAC, and different CV MPRs. It is inferable that the risk of crashes decreases as market penetration increases. According to Figure 6.3C, above 60% MPR, SSAM did not record any conflicts, which stems from dealing with I-80 as a low-traffic volume corridor.

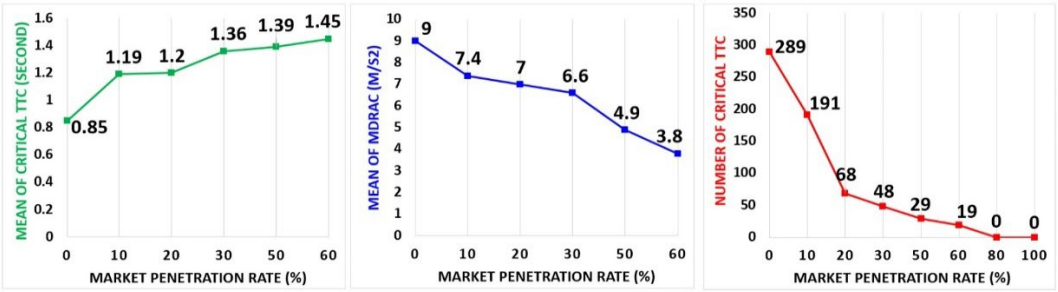
Table 6.1 Safety Assessment of the Entire Work Zone Section

CV-MPR (%)	Number of Critical TTCs	Statistics	Critical TTC (s [a])	Critical MDRAC (m/s²)
0 (baseline)	289	Mean	0.85	9
		Std. Deviation	0.41	1.1
10	191	Mean	1.19	7.4
		Std. Deviation	0.77	0.92
20	68	Mean	1.2	7
		Std. Deviation	0.38	1.2
30	48	Mean	1.36	6.6
		Std. Deviation	0.32	0.55
50	29	Mean	1.39	4.9
		Std. Deviation	0.89	0.99
60	19	Mean	1.45	3.8
		Std. Deviation	0.24	0.28

Safety Assessment of Individual Work Zone Areas

	Advance Warning Area		Transition Area		Activity Area		Termination Area	
CV-MPR (%)	TIT(s)	TET(s)	TIT(s)	TET(s)	TIT(s)	TET(s)	TIT(s)	TET(s)
0 (baseline)	1347	601	210	114	860	318	484	118
10	1051	513	171	62	590	248	377	77
20	755	386	165	58	370	200	278	62
30	311	143	101	41	178	109	150	36
50	254	112	56	29	142	67	45	27
60	162	75	25	12	87	35	22	9

Note:
TTC= Time-To-Collision; MDRAC= Modified Deceleration Rate to Avoid a Crash; TIT= Time-Integrated Time-to-collision; TET= Time Exposed Time-to-collision; MPR= Market Penetration Rate.
[a] Second



A B C

Figure 6.3 Relationship between Conflict-Based Surrogate Measure of Safety (C-SMoS) and Connected Vehicle Market Penetration Rate (CV-MPR)

6.5 Spatiotemporal Analysis of Speed Harmonization

In addition to the considered C-SMoS, speed harmonization in four WZ areas has been investigated under varying CV-MPRs. During the simulation run, the speeds of individual vehicles in the traffic stream were observed using several data collection points located at 0.1-mile intervals in all individual WZ areas. The effect of work zone warning (WZW) under different CV-MPRs on speed harmonization has been evaluated based on the spatiotemporal aggregation of mean speed and standard deviation of speed.

According to Table 6.2, the one-way ANOVA revealed that the mean of mean speed and standard deviation of speeds from different MPR scenarios in all four sub-segments of the WZ during foggy conditions are statistically different from each other under 95% confidence level. Turning to the post hoc analysis, the result of the two-tailed t-test depicted that, for most cases, both standard deviations of speed and mean speed obtained their associated minimum means under 60% MPR of CVs across all sub-segments. However, in some cases, the minimum mean of groups

corresponding to mean speed and standard deviation of speed can be achieved even at less MPR than 60%. In this regard, because almost 85% of all pairwise comparisons (34 out of 40 comparisons) showed that the minimum mean speed and standard deviation of speed could be obtained under 60% MPR, this MPR is suggested as the optimum MPR to reach the maximum safety in a WZ area under foggy conditions on I-80.

Table 6.2 Statistical Analysis of Observed Mean Speed and Standard Deviation of Speed across Varying MPRs for Different Work Zone Areas

		Mean Speed			Standard Deviation of Speed		
		F-Statistic	P-value	Sig.	F-Statistic	P-value	Sig.
Advance Warning Area		94.275	<0.05	*	65.042	<0.05	*
Transition Area		27.948	<0.05	*	159.33	<0.05	*
Active Area		20.365	<0.05	*	114.4	<0.05	*
Termination Area		39.54	<0.05	*	109.32	<0.05	*
Post hoc pairwise comparison between 60 % MPR [a] scenario and other CV-scenarios							
		Mean Speed			**Standard Deviation of Speed**		
		t-Statistic	P-value	Sig.	t-Statistic	P-value	Sig.
Advance Warning Area	0%MPR(Non-CV)	71.377	<0.05	*	22.889	<0.05	*
	10%MPR	17.220	<0.05	*	9.655	<0.05	*
	20%MPR	9.348	<0.05	*	5.857	<0.05	*
	30%MPR	5.945	<0.05	*	0.164	0.869	
	50%MPR	3.181	<0.05	*	-0.228 [a]	0.819	
Transition Area	0%MPR(Non-CV)	76.295	<0.05	*	35.126	<0.05	*
	10%MPR	5.497	<0.05	*	16.877	<0.05	*
	20%MPR	11.014	<0.05	*	9.695	<0.05	*
	30%MPR	7.490	<0.05	*	5.449	<0.05	*
	50%MPR	4.544	<0.05	*	3.526	<0.05	*
Activity Area	0%MPR(Non-CV)	152.216	<0.05	*	33.308	<0.05	*
	10%MPR	6.025	<0.05	*	17.156	<0.05	*
	20%MPR	6.116	<0.05	*	13.987	<0.05	*
	30%MPR	2.086	<0.05	*	8.139	<0.05	*
	50%MPR	2.809	<0.05	*	6.158	<0.05	*
Termination Area	0%MPR(Non-CV)	63.912	<0.05	*	31.792	<0.05	*
	10%MPR	3.279	<0.05	*	8.160	<0.05	*
	20%MPR	1.117	0.265		4.065	<0.05	*
	30%MPR	2.922	<0.05	*	1.353	0.177	
	50%MPR	1.532	0.126		-1.244	0.214	

[*] Representative for statistically significant under 95% Confidence Intervals.
[a] MPR= Market Penetration Rate
[b] Negative t-Statistic indicates the mean of the reference sample (60% MPR) in pairwise comparison is more than the corresponding compared sample.

6.6 Accounting for Human-Related Unobservable Heterogeneity in the Safety Performance of WZW application

From a microscopic perspective, at level-0 and level-1 of driving automation (i.e., no automation) [1], which is the main focus of the WYDOT CV Pilot Program, human drivers are in charge of the execution of steering and acceleration/deceleration as well as monitoring of the driving environment [1]. Hence, human drivers are still in the safety loop, which might add human-related unobservable heterogeneity to the safety performance assessment of CVs [7], [63], [68]–[70]. Drivers adapt their driving behavior according to the received CV notifications that might be affected by CV drivers' characteristics and other vehicles in the traffic stream. Compared with aggressive CV drivers, conservative CV drivers might require more time and distance to execute the advisory messages effectively and safely. Hence, it seems that changing the provided delivery distance of WZW would affect CVs' safety performance. This part of the document explains the

importance of delivery distance on the safety performance assessment of the WZW CV application on I-80 in Wyoming using the same simulation framework presented in the previous sections.

6.6.1 Modification of Microsimulation Modeling

According to Figure 6.4, the geometry of the two-to-one lane dropped WZ layout was modified for different settings of WZW delivery distances. As suggested by the MUTCD, the minimum distance to provide WZW was 1500 feet before the transition area [71]. Three additional delivery distances with 1500-ft increments (i.e., 3000, 4500, and 6000 feet upstream of the transition area) were regarded to analyze the WZW safety performance.

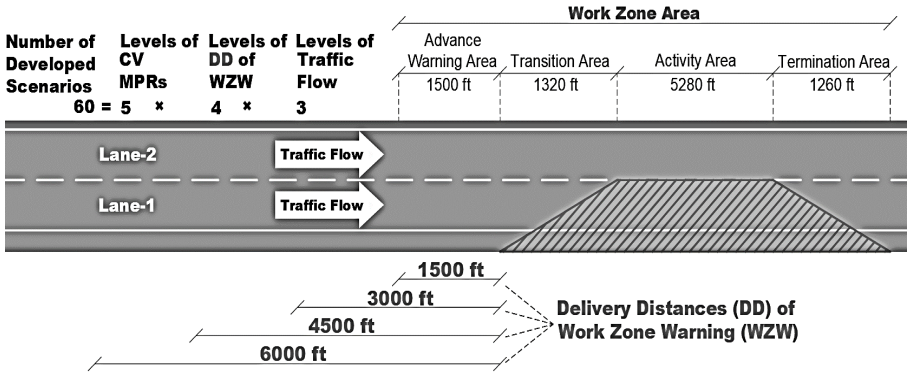


Figure 6.4 Conceptualization of Microsimulation Scenarios

Similar to the previous analysis, CVs were modeled to move into the open lane when they receive a WZW notification, which varies depending on the different delivery distances. Different delivery distances were simulated using various links connected to the transition area via connectors, as presented. Each link has a specific length corresponding to each level of the delivery distance. The CVs' lane-changing behaviors were adjusted by altering the "lane change distance" in the connectors' menu and updating lane-changing parameters in VISSIM.

The combination of four different WZW delivery distances, five levels of CV-MPR, and three types of traffic flow characteristics resulted in 60 microsimulation modeling scenarios. For each scenario, the safety performance of WZW was evaluated in the four WZ areas individually using TTC with a critical value of 2.5 seconds. Accordingly, each scenario resulted in a specific number of critical TTCs. Therefore, because the outcome variable was observed for each of the four WZ areas, the conversion of simulation into statistical observation led to four design matrices, including 60 observations per area. Table 6.3 presents the statistical description of variables.

6.6.2 Visual Inference

Figure 6.5 visualizes the effect of varying WZW delivery distances, CV-MPR, and different traffic volume levels on the number of critical TTCs observed in 60 developed scenarios across all WZ areas in a heat map. It can be concluded that, at any given CV-MPR, the safety performance of CVs in the WZ can be enhanced if the WZW delivery distance increases. These visual investigations have been validated by the result of negative binomial regressions.

Table 6.3 Statistical Description of the Used Variables

	Advance Warning Area				Transition Area				Activity Area				Termination Area			
Variable	Cat.	# of Observation			Cat.	# of Observation			Cat.	# of Observation			Cat.	# of Observation		
Traffic Volume [a]	Low	20	(Ref. Cat.)		Low	20	(Ref. Cat.)		Low	20	(Ref. Cat.)		Low	20	(Ref. Cat.)	
	Mod.	20			Mod.	20			Mod.	20			Mod.	20		
	High	20			High	20			High	20			High	20		
MPR	0	12	(Ref. Cat.)		0	12	(Ref. Cat.)		0	12	(Ref. Cat.)		0	12	(Ref. Cat.)	
	20	12			20	12			20	12			20	12		
	40	12			40	12			40	12			40	12		
	60	12			60	12			60	12			60	12		
	80	12			80	12			80	12			80	12		
Delivery Distance of WZW Application	1500	15	(Ref. Cat.)		1500	15	(Ref. Cat.)		1500	15	(Ref. Cat.)		1500	15	(Ref. Cat.)	
	3000	15			3000	15			3000	15			3000	15		
	4500	15			4500	15			4500	15			4500	15		
	6000	15			6000	15			6000	15			6000	15		
	Min.	Max.	Mean	Var.	Min.	Max.	Mean	Var.	Min.	Max.	Mean	Var.	Min.	Max.	Mean	Var.
Number of Critical TTCs	0	77	23.67	332.29	0	69	17.32	237.37	0	1018	219.9	74582.04	0	37	11.18	80.28

[a] Three categories of real-time traffic volume observations are as follows: Low traffic volume: 140 vph; Moderate traffic volume: 555 vph; High traffic volume: 896 vph.

Note:

MPR= Market Penetration Rate; WZW= Work Zone Warning; Ref. Cat.= Reference Category; Var.= Variance

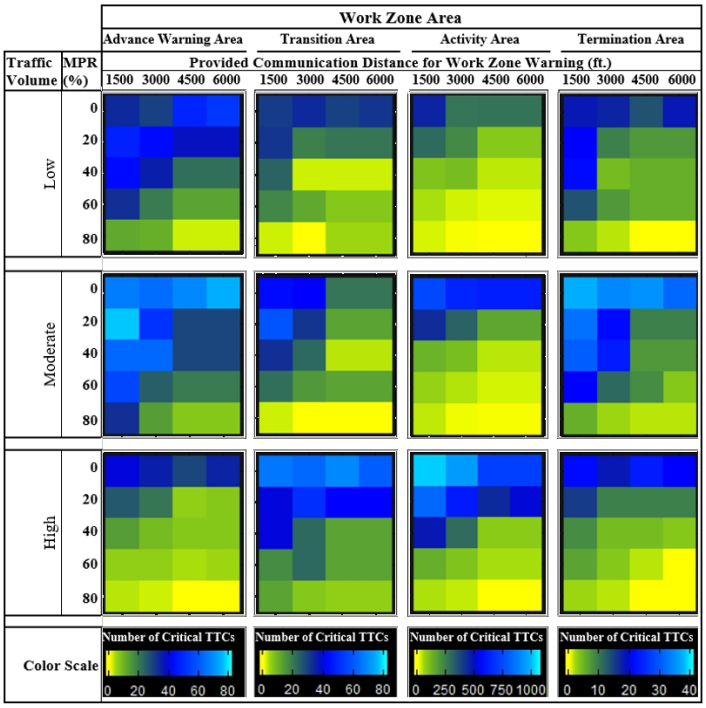


Figure 6.5 Number of Critical TTCs in 60 Scenarios for Varying Traffic Volumes and across Four Work Zone Areas

6.6.3 Hierarchical Negative Binomial Regressions

The effect of WZW delivery distance, CV-MPR, and varying traffic volumes on the number of critical TTCs was assessed based on four negative binomial (NB) regressions conducted on four WZ areas. The concept of hierarchical modeling under Bayesian inference was followed, and the WZW delivery distance was considered a random factor to reveal the possible dependency among the individual number of critical TTCs within different clusters of this factor at level-2 of design matrices due to unobserved heterogeneity relating to drivers' behaviors [72]. Equations 6.9 to 6.11 present the details:

$$y_{ij} \sim NB(\lambda_{ij}, \alpha) \quad \text{Equation 6.9}$$

$$E(y_{ij}) = \lambda_{ij} \quad \text{Equation 6.10}$$

$$\ln(\lambda_{ij}) = (\beta_{00} + w_{0j}) + \boldsymbol{\beta X} + e_{ij} \quad \text{Equation 6.11}$$

Where y_{ij} denote the number of observed critical TTCs for i^{th} observation in j^{th} cluster of the provided WZW delivery distance, λ_{ij} and $\alpha > 0$ are, respectively, the location (i.e., expectation) and dispersion parameter (also called shape parameter), β_{00} is the fixed intercept, X is the vector of fixed independent variables (i.e., CV-MPR and traffic volume), $\boldsymbol{\beta}$ is the vector of regression coefficients, and e_{ij} is the individual-level error term. In equation 6, the total of $(\beta_{00} + w_{0j})$ is the random intercept varying across different WZW delivery distances, where for $j \neq j'$, $cov(w_{0j}, w_{0j'}) = 0$ [42], [73]. w_{0j} is the random effect of different WZW delivery distances on β_{00} that follows $\boldsymbol{N}(0, \sigma_0^2)$ and σ_0 is the standard deviation of cluster specified random intercepts to be estimated under Bayesian inference.

6.6.3.1 Results of Statistical Inference

Table 6.4 presents the results of four hierarchical NB regressions corresponding to four WZ areas. Relating to the effect of traffic volume on WZW safety, except for one case, all other estimated coefficients are positive. Although these coefficients are not always statistically significant, their positive signs indicate that an increase in traffic volume can increase the number of critical TTCs because the low traffic volume was considered the reference class for this categorical variable. Model-1, conducted on the advance warning area, shows that when traffic volume is high, critical TTCs would be less than the low traffic volume conditions, which is in line with the literature. Previous studies apprised that, on rural freeways, an increase in traffic volume would decrease the standard deviation of speed [74]. In this situation, speed in the advance warning area would be more homogenous than the low traffic conditions area, leading to a fewer number of conflicts.

Table 6.4 Results of Hierarchical Negative Binomial Regressions Conducted on the Four Work Zone Areas

	Advance Warning Area				Transition Area				Activity Area				Termination Area			
	Model-1				Model-2				Model-3				Model-4			
	Est.	5% CI [a]	95% CI	Sig.	Est.	5% CI	95% CI	Sig.	Est.	5% CI	95% CI	Sig.	Est.	5% CI	95% CI	Sig.
Intercept	3.01	2.46	3.59	*	2.58	1.89	3.34	*	5.64	4.79	6.6	*	2.41	1.74	3.02	*
Traffic-High [b]	-0.40	-0.82	-0.02	*	0.79	0.32	1.27	*	1.32	1.05	1.6	*	0.06	-0.37	0.48	
Traffic-Moderate	0.52	0.13	0.91	*	0.27	-0.20	0.73		0.53	0.25	0.82	*	0.43	0.01	0.82	*
CV-MPR-20% [c]	0.13	-0.36	0.64		0.31	-0.22	0.86		-0.64	-0.97	-0.27	*	0.07	-0.41	0.56	
CV-MPR-40%	-0.13	-0.67	0.4		-0.10	-0.68	0.49		-1.85	-2.19	-1.47	*	-0.01	-0.51	0.49	
CV-MPR-60%	-0.27	-0.80	0.24		-0.19	-0.74	0.37		-2.91	-3.25	-2.52	*	-0.19	-0.70	0.31	
CV-MPR-80%	-0.55	-1.10	-0.02	*	-1.79	-2.42	-1.17	*	-4.62	-5.05	-4.19	*	-1.55	-2.16	-0.99	*
WAIC	**478.1 (11.2) [d]**				**435.6 (15.2)**				**598.8 (28.3)**				**387.1 (13.1)**			
ICC	**0.25**				**0.44**				**0.68**				**0.13**			

* Statistically significant at 95% Confidence Interval

[a] Credible Interval

[b] The reference level for traffic volume is the low traffic volume

[c] The reference level for Market Penetration Rate (MPR) of Connected Vehicles (CVs) is 0%

[d] Values in parentheses indicate standard error of estimation

Furthermore, it can be seen that the activity area is the most sensitive WZ area to any change in CV-MPR, where a significant reduction in the number of critical TTCs can be obtained even at 20% CV-MPR. When CV-MPR increases to 40%, an indication of safety enhancement can be gained because the estimated coefficients corresponding to CV-MPR obtained negative values compared with the reference category (i.e., CV-MPR=0%). However, except for the activity area, these pieces of evidence are not strong enough to result in statistically significant safety enhancements in three other areas. Thus, to attain a statistically significant decrease in the number of critical TTCs in all WZ areas, 80% CV-MPR is needed.

6.6.3.2 Optimal Roadside Unit Placement for WZW Application in WYDOT CV Pilot

According to Table 6.4, the estimated ICCs are greater than zero for all developed NB regressions. Hence, individual observations (i.e., the number of critical TTCs) at level-1 are nested within varying WZW delivery distances at level-2 of the data hierarchy across all WZ areas. As narrated by Figure 6.5, WZW delivery distance could affect the CV-WZW safety performance. However, no fixed coefficient was estimated for this random factor because, in hierarchical modeling, a random factor can randomly affect the response variable at level-2 of the design matrix [40]. In fact, the effect of WZW delivery distance has been embedded into the random intercepts based on Equation 6.11 (i.e., $\beta_{00} + w_{0j}$).

Table 6.5 presents all the estimated random intercepts for four models under varying clusters of WZW delivery distances. According to Equation 6.11, a reduction in the random intercept would reduce the natural logarithm of the estimated number of critical TTCs, enhancing the CV-WZW safety performance. For all WZ areas, when WZW delivery distance increases, the random intercepts would decrease, enhancing the CV-WZW safety performance regardless of CV-MPRs and traffic volume. The mentioned conclusion is reasonable because when WZW delivery distance increases, CV drivers would be given more time and distance to safely execute the received CV notification, augmenting the WZW application efficiency. Another remarkable point can be unveiled from Table 6.5 by comparing the estimated random intercepts corresponding to WZW delivery distances of 4500 feet and 6000 feet that behave very similarly. This similarity means that

an increase in the WZW delivery distance of more than 4500 feet does not have an extra advantage on CVs' safety performance. Therefore, the optimal delivery distance required for the best safety performance of the CV-WZW application is determined to be 4500 feet.

Table 6.5 Random Intercepts under Varying WZW Delivery Distances in Four WZ Areas

WZW [a] Delivery Distance	Advance Warning Area	Transition Area	Activity Area	Termination Area
	Model-1	**Model-2**	**Model-3**	**Model-4**
1500 ft	3.24 (0.26) [b]	2.99 (0.29)	6.39 (0.18)	2.69 (0.26)
3000 ft	3.03 (0.24)	2.62 (0.28)	5.88 (0.18)	2.49 (0.24)
4500 ft	2.88 (0.25)	2.35 (0.28)	5.15 (0.17)	2.22 (0.25)
6000 ft	2.88 (0.25)	2.35 (0.28)	5.16 (0.17)	2.22 (0.25)

[a] Work Zone Warning

[b] Values in parentheses indicate standard error of estimation

7. MITIGATING THE RISK OF HORIZONTAL CURVE-RELATED CRASHES IN CV ENVIRONMENTS

7.1 Introduction

Of rural highway fatalities, 25% occur on horizontal curves, where the crash frequency is three times greater than tangent segment counterparts [75]. Large truck crashes comprise 57% of fatal crashes in rural areas [76]. With a 150% increase in heavy truck traffic volume over the last three decades, serious traffic safety concerns on I-80 in Wyoming are expected [77], [78]. In this regard, one important safety advantage of the WYDOT CV Pilot is providing drivers with advisory messages and real-time information regarding unforeseeable roadway conditions, such as road surface conditions, weather conditions, and roadway geometry. The increased driver situational awareness due to the received CV notifications would potentially reduce the contribution of driver errors to crash occurrences on horizontal curves. Accordingly, this part of the document explains the effectiveness of the WYDOT CV Pilot mitigating the risk of horizontal curve-related crashes utilizing trajectory-level analysis from the high-fidelity driving simulator with the main focus on two critical conditions: horizontal curves with slippery road surface conditions and horizontal curves under foggy weather conditions.

7.2 Horizontal Curves with Slippery Road Surface Conditions

A slippery road surface condition is important because, although drivers can adjust their driving behaviors in adverse weather conditions, they might not be aware of the appropriate speed selection according to road surface conditions [79]–[83]. Therefore, the combination of slippery road surface conditions with challenging roadway geometry, such as horizontal curves, might result in drivers' inability to select the speed appropriately, leading to a remarkable rate of run-off-road, sideswipe, head-on, and rollover crashes. To assess the safety performance of CV TIMs in mitigating the mentioned traffic critical safety events, the research team conducted a with/without analysis at the WyoSafeSim lab to assess the effect of three CV applications, SWIW, CV-VSL, and SA, on truck driver behavior.

7.2.1 Driving Simulator Experimental Design

According to Figure 7.1, the driving scenario was designed based on a divided two-way four-lane freeway segment with a 75 mph speed limit. The scenario included two horizontal curves (HCs) with the same geometric characteristic (radius ≃ 550 meters, length ≃ 800 meters, central angle ≃ 90o, superelevation = 8%, longitudinal grade = 0.01%, and the presence of vertical curve: NO). The two HCs were preceded by independent tangent segments enabling drivers to reach their desired speed [81] before entering the second HC.

The entire scenario was designed under snowy weather conditions. On Curve-A, which is referred to as the HC with regular pavement conditions, it was assumed that the pavement had a minor wet surface, where the friction was not remarkably decreased. Curve-B is located after 600-meter from the beginning of the scenario, which is referred to as the HC with slippery pavement conditions. On Curve-B, the pavement friction was considerably reduced to simulate the patchy ice/black conditions. The combination of reduced pavement side friction, curvature, and superelevation in Curve-B does not provide sufficient stability for those vehicles whose speeds are more than 55

mph. Consequently, Curve-A serves as the baseline, and Curve-B serves as the main focus of the study.

In the CV scenario, the CV notifications included an initial regulatory speed of 75 mph. A variable speed limit (VSL) sign of 65 mph was displayed on the HMI just before Curve-A. A slippery road surface and adverse weather condition notifications were initiated with an advisory speed of 55 mph, which was disseminated before Curve-B. All the CV notifications were accompanied by a simple "beep" sound that would best draw a driver's attention [52]. For the non-CV scenarios, however, CV notifications were muted, and the HMI was kept off. Four study cases were conducted to investigate driver behavior based on vehicle trajectories considering two HCs in the CV and non-CV scenarios.

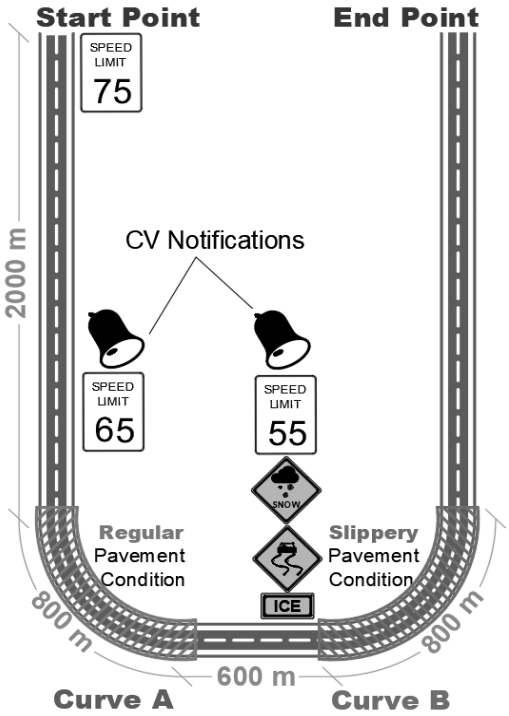


Figure 7.1 Driving Simulator Testbed for Analysis of Run-off-Road Crashes on Slippery Horizontal Curves

7.2.2 Statistical Inference

The research team obtained vehicle trajectories from the DS experiment for 24 drivers and across four studies cases. Driver behaviors were quantified by characterizing several K-SMoS on the trajectory observations, with the main emphasis on drivers' deviation from the pathway (DFP). In addition to investigating vehicles' DFP, the assessment was conducted on other K-SMoS that had the highest correlation with DFP on Curve-A, where the K-SMoS were measured without being affected by the slippery road surface conditions. Accordingly, three other K-SMoS, lateral speed, instantaneous acceleration, and steering angle, were selected for further investigation.

The change in truck drivers' behavior due to CV advisory/warning messages was discerned by analyzing and comparing K-SMoS distribution characteristics on slippery and regular HCs in CV and non-CV environments. In this regard, the non-parametric statistical approach was followed as none of the K-SMoS distributions followed the assumption of normality and homogeneity of variance. The K-SMoS distributions of CVs and non-CVs were compared considering their central tendency and dispersion.

7.2.2.1 Wilcoxon Signed-Rank Test in Central Tendency Analysis of K-SMoS

Measures of central tendency indicate how the locations of a variable can vary systematically under the effect of a specific treatment. Considering this, the Wilcoxon Signed-Rank Test (WSRT) was employed to compare the central tendency of aggregated K-SMoS across four study cases based on their associated medians. WSRT is a nonparametric alternative to the parametric student t-test without requiring normality assumption that evaluates whether the central tendency (i.e.,

population median) of samples being compared are different or not [84]. WSRT is a popular test for a repeated measure design (i.e., with/without analysis) where the same variable is being assessed under the effect of specific treatment [85]. Here, the treatment is the effect of CV advisory/warning messages on driver behaviors to be measured using aggregated K-SMoS. The aggregation of K-SMoS per meter of each HC led to providing paired samples with the same sizes, as is a WSRT requirement with the null hypothesis (H_0) of the samples being compared to come from the same population [86], [87].

Suppose each of the two samples being compared has N number of observations. For $i = 1, 2, \dots, N$, $x_{2,i}$ and $x_{1,i}$ are the i^{th} observation in the second and first sample, comprising N paired observations (i.e., $[x_{2,i}, x_{1,i}]$). First, the absolute value of difference for each pair is calculated (i.e., $|d_i| = |x_{2,i} - x_{1,i}|$) and the pairs should be reordered from the smallest to largest $|d_i|$. Secondly, those pairs with $d_i = 0$ should be removed from the ordered pairs. Thus, the remaining number of pairs would be N_r, where $N_r \leq N$. The remaining ordered pairs should be ranked from one to N_r and the ranks are denoted by R_i. The WSRT statistic (W) would be calculated based on equation 2 to 5, where and $I(\rho)$ is the indicator function and $d_i = (x_{2,i} - x_{1,i})$ [86].

$$I(\rho) = \begin{cases} 1, & if\ \rho\ is\ true \\ 0, & if\ \rho\ is\ false \end{cases} \quad \text{Equation 7.1}$$

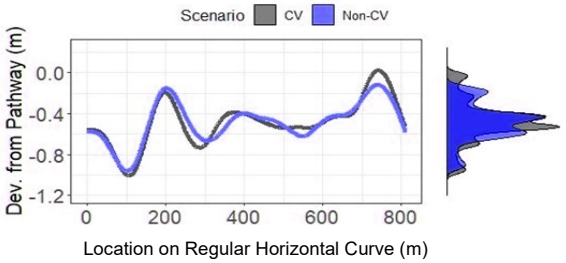
$$W^+ = \sum_{i=1}^{N_r} R_i I(d_i > 0) \quad \text{Equation 7.2}$$

$$W^- = \sum_{i=1}^{N_r} R_i I(d_i < 0) \quad \text{Equation 7.3}$$

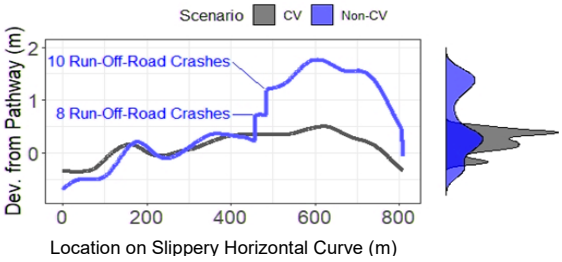
$$W = \min(W^+, W^-) \quad \text{Equation 7.4}$$

It can be shown that under H_0, when N_r increases, the sampling distribution of W would converge to a normal distribution whose mean (μ) and standard deviation (σ) are $\frac{N_r(N_r+1)}{4}$ and $\sqrt{\frac{N_r(N_r+1)(2N_r+1)}{24}}$, respectively [86]. Hence, the significance of WSRT could be performed by comparing standardized W (i.e., $\frac{W-\mu}{\sigma} = Z$) to critical Z-value ($Z_{Critical}$) under a simple two-tailed Z-test whose p-value depends on the considered significance level and N_r [86]. For $|Z| > Z_{Critical}$ (i.e., p-value less than the considered significance level), the H_0 is rejected, indicating there is not enough evidence to show the samples being compared come from the same population.

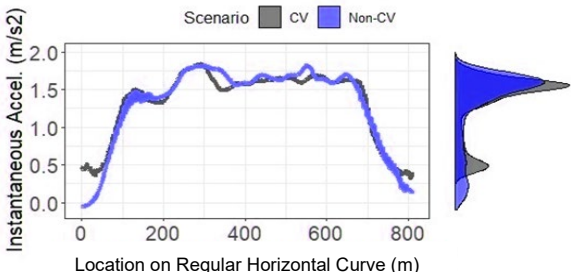
Figure 7.2 narrates truck driver behavior negotiating slippery and regular HCs under the CV and non-CV environments based on vehicle trajectories and using K-SMoS. Table 7.1 presents nonparametric comparisons of central tendency and dispersion of four K-SMoS under CV and non-CV environments negotiating HCs with regular and slippery pavement conditions.



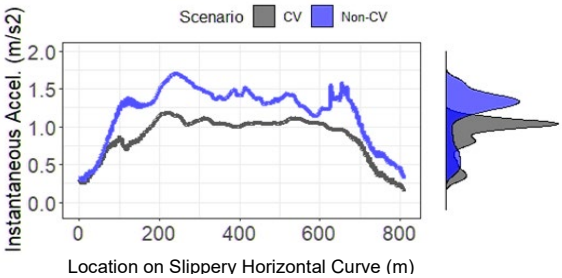
A- Deviation from the Pathway on Regular Horizontal Curve



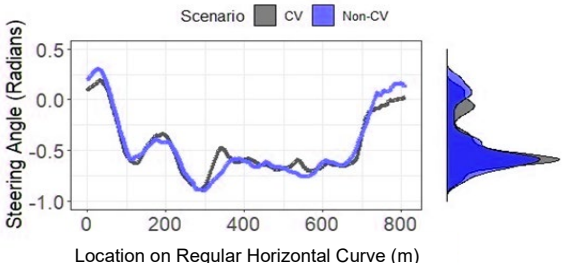
B- Deviation from the Pathway on Slippery Horizontal Curve



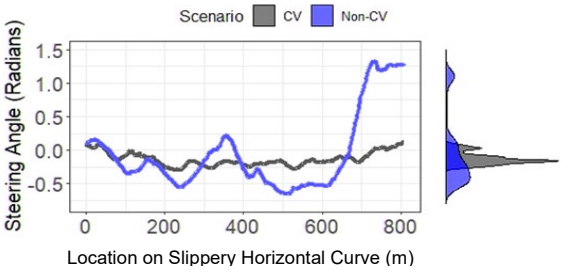
C- Instantaneous Acceleration on Regular Horizontal Curve



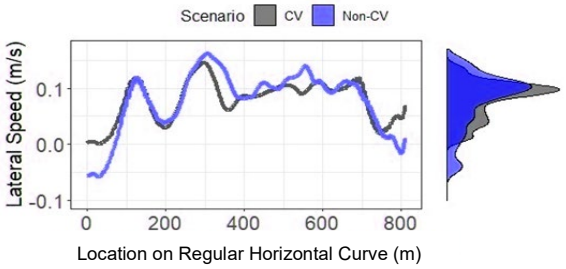
D- Instantaneous Acceleration on Slippery Horizontal Curve



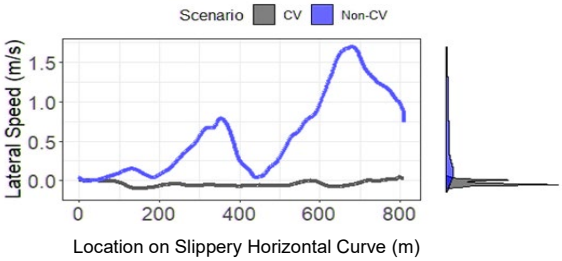
E- Steering Angle on Regular Horizontal Curve



F- Steering Angle on Regular Slippery Curve



G- Lateral Speed on Regular Horizontal Curve



H- Lateral Speed on Regular Slippery Curve

Figure 7.2 Aggregated CVs and Non-CVs K-SMoS on Horizontal Curves

According to Table 7.1, on Curve-A with regular pavement friction, there is no significant difference between the central tendency of four K-SMoS associated with CVs and non-CVs based on the WSRT results. Considering the estimated p-values, it is inferable that on curve-A, CVs and non-CVs trajectories are very similar. These similarities have been visualized in Figures 7.2A, 7.2C, 7.2E, and 7.2G. Considering vehicle trajectories on Curve-B with slippery pavement

conditions, presented in Figure 7.2, the safety benefits of CV advisory/warning messages are evident. According to Figure 7.2B, most non-CVs could tolerate the slippery pavement condition until the middle of the HC. However, after the middle of the curve, 18 non-CVs (i.e., 75% of all participants) could not maintain the appropriate lane position and had ROR crashes. According to Table 7.1, all K-SMoS distributions' central tendencies associated with CVs are statistically different from non-CVs, coincided by shifting the medians of K-SMoS associated with CVs, compared with non-CVs, toward zero, which varies from 23% up to 99% across four K-SMoS.

Table 7.1 Central Tendency Analysis of Kinematic-Based SMoS

Curve-A (Regular Pavement Condition)						
	Median			**Central Tendency Analysis Wilcoxon Signed-Rank Test**		
Kinematic-Based Surrogate Measure of Safety	Non-CVs	CVs	% of Enhancement [a]	Test Stats. (W [b])	p-value	Sig.
Deviation from the Pathway	-0.47	-0.46	2.01	164388	0.971	
Steering Angle	-0.59	-0.58	1.59	164435	0.976	
Instantaneous Acceleration	1.59	1.57	0.60	163211	0.831	
Lateral Speed	0.09	0.08	6.58	163336	0.846	
Curve-B (Slippery Pavement Condition)						
	Median			**Central Tendency Analysis Wilcoxon Signed-Rank Test**		
Kinematic-Based Surrogate Measure of Safety	Non-CVs	CVs	% of Enhancement	Test Stats. (W)	p-value	Sig.
Deviation from the Pathway	0.33	0.16	51.97	71830	<0.001	*
Steering Angle	-0.25	-0.17	31.90	144015	0.002	*
Instantaneous Acceleration	1.32	1.01	23.30	40	<0.001	*
Lateral Speed	0.42	0.01	98.97	416	<0.001	*

[a] Enhancement is defined as the relative proportion of shifting in the central tendency of SMoS toward the zero.
[b] Wilcoxon Signed-Rank Test (WSRT)
* Representative of statistically significant differences under 95% Confidence Interval CI).

7.2.2.2 Median Absolute Deviation in K-SMoS Dispersion Analysis

Dispersion determines the margin of error or uncertainty in making inferences regarding central tendency. Less uncertainty makes the question of causality easier to uncover [88]. Among several statistical dispersion measures, the median absolute deviation (MAD) is considered because of the following reasons: First, the K-SMoS were not normally distributed. Second, the central tendency considerations were performed based on the medians of K-SMoS distributions rather than the mean. MAD is immune to the sample size and outliers compared with other dispersion measures [89]. Equation 7.5 parameterizes MAD calculation to scale dispersion of a set of observation x_i for a random variable X [90], [91].

$$MAD = 1.4826 \times median(|x_i - median(x_i)|) \quad \text{Equation 7.5}$$

The results of MAD, presented in Table 7.2, illustrate the effectiveness of CV advisory/warning messages on both regular and slippery HCs. In both cases, the reduction in variability of K-SMoS in the CV environment compared to non-CV conditions reveals that more certainty of driver behavior is expected to be seen under the CV environment. However, this certainty and reduction of variability are much more obvious on slippery HCs. On slippery HCs, the CV advisory/warning could reduce the dispersion of four K-SMoS from 54% up to 95%, which is more remarkable for lateral speed.

Table 7.2 Dispersion Analysis of Kinematic-Based SMoS

Curve-A (Regular Pavement Condition)			
	Dispersion Analysis Median Absolute Deviation		
Kinematic-Based Surrogate Measure of Safety	Non-CVs	CVs	% of Reduction
Deviation from the Pathway	0.18	0.14	22.40
Steering Angle	0.21	0.14	30.46
Instantaneous Acceleration	0.26	0.17	34.57
Lateral Speed	0.04	0.03	16.56
Curve-B (Slippery Pavement Condition)			
	Dispersion Analysis Median Absolute Deviation		
Kinematic-Based Surrogate Measure of Safety	Non-CVs	CVs	% of Reduction
Deviation from the Pathway	0.95	0.28	70.69
Steering Angle	0.39	0.09	77.65
Instantaneous Acceleration	0.21	0.09	53.99
Lateral Speed	0.52	0.03	95.19

[a] Enhancement is defined as the relative proportion of shifting in the central tendency of SMoS toward the zero.

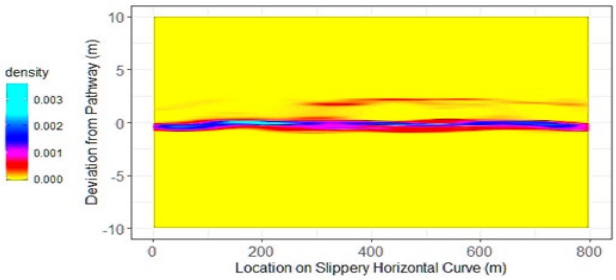
* Representative of statistically significant differences under 95% Confidence Interval (CI).

Because the effect of CV advisory/warning messages on driver behavior is more substantial on a slippery HC, the individual vehicle trajectories negotiating this curve were scrutinized. To this aim, the density distributions of disaggregated vehicle trajectories on the slippery HC have been visualized in Figure 7.3 based on the four K-SMoS. Generally, Figures 7.3A to 7.3H reveal that informing drivers regarding the slippery pavement condition and advising the appropriate speed limit can minimize the SMoS variability. Figures 7.3A and 7.3B show that the CV advisory/warning messages could centralize the deviation from the pathway (DFP) toward zero and minimize its variation on the slippery HC. The maximum DFP for the CVs was 2 meters, which is remarkably less than 5.2 meters for non-CVs. In addition to minimizing the risk of ROR crashes, this reduction can decrease the rate of sideswipe crashes due to skidding on rural I-80 in Wyoming.

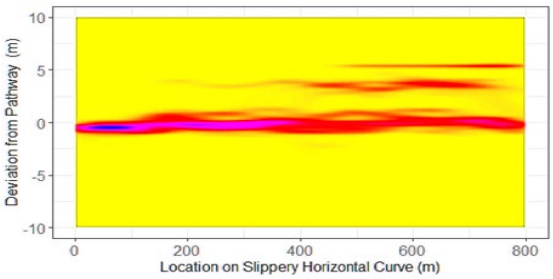
The reduction in variability of steering angle for CVs compared with non-CVs (i.e., 77% according to table 30) is clear from Figures 7.3E and 7.3F. This reduction is one of the main causes of higher stability of CVs versus non-CVs, alleviating ROR crash risk on the HC due to increased driver situational awareness before entering the slippery HC. Mazzae et al. (1999) showed that excessive steering angles are associated with crash-prone conditions. The literature has pointed out that more variation and oversteering maneuvers are associated with more aggressive driving behavior, increasing ROR crash risks [93], [94].

As mentioned earlier and based on Table 7.2, the lateral speed was the most affected K-SMoS in the CV scenario. This K-SMoS showed a 99% shifting in the central tendency toward zero and a 95% reduction in variability under the CV environment compared with non-CVs. These changes are extremely important in terms of enhancing drivers' lane-keeping behavior negotiating slippery HCs [95], which, in turn, might minimize the likelihood of run-off-road, sideswipe, and head-on crashes. Figures 7.3G and 7.3H depict that the maximum lateral speeds for CVs, either to the left or right sides, are half of the non-CVs. This reduction in lateral speed is more critical for high traffic volume corridors where the average of time headways is usually less than low volume corridors. In such a situation, and considering the limited human perception reaction time, faster lateral movement (i.e., higher lateral speed) of vehicles can be associated with higher crash

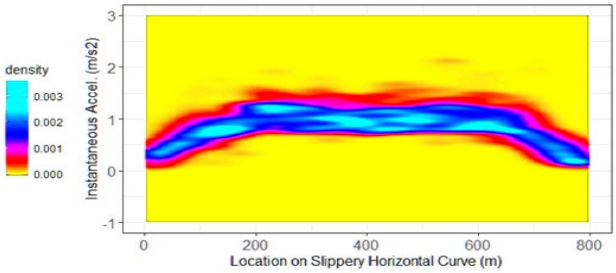
probability. Therefore, CV advisory/warning messages preceding slippery horizontal curves can alleviate this crash-prone condition, more specifically on high traffic volume highways.



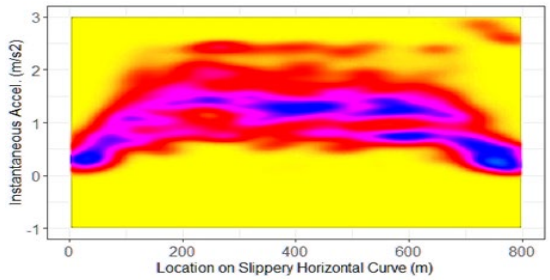
A- Deviation from the Pathway for Connected Vehicles



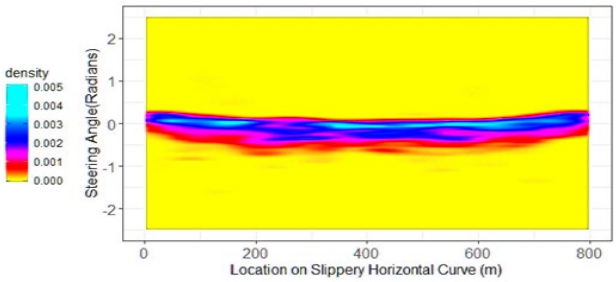
B- Deviation from the Pathway for Non-Connected Vehicles



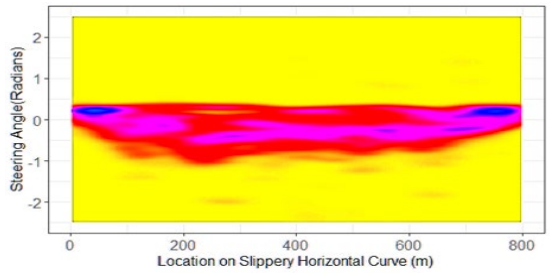
C- Instantaneous Acceleration for Connected Vehicles



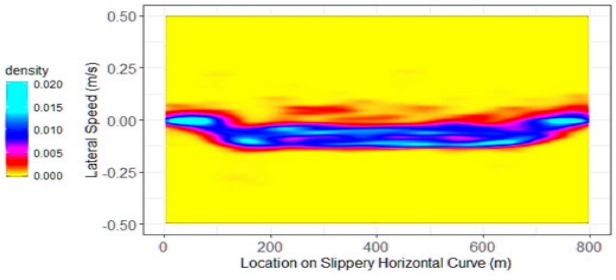
D- Instantaneous Acceleration for Non-Connected Vehicles



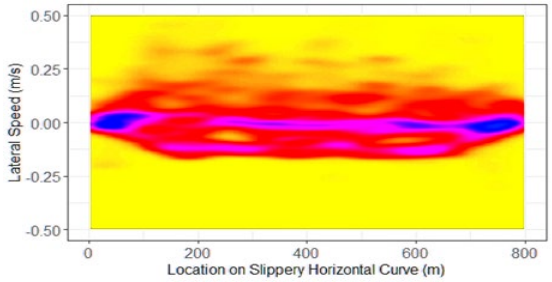
E- Steering Angle for Connected Vehicles



F- Steering Angle for Non-Connected Vehicles



G- Lateral Speed for Connected Vehicles



H- Lateral Speed for Non-Connected Vehicles

Figure 7.3 Density Distributions of CV and Non-CV Trajectories on the Slippery Horizontal Curve

7.3 Horizontal Curves under Foggy Weather Conditions

In addition to the safety performance of CVs on HCs with a slippery road surface, the research team analyzed the safety impact of CV technology on HCs located in a dense fog with reduced visibility. Previous studies have explained the negative impact of reduced visibility and complex roadway geometry on traffic safety separately. More crucially, it is known that foggy conditions will increase the crash likelihood at road alignment transition locations [96], particularly at HCs [95], [97]. Accordingly, to show truck drivers' behavioral alterations due to CV notifications in the aforementioned critical situation, a with/without approach to vehicle trajectories from the driving simulator (DS) experiment was analyzed.

7.3.1 Driving Simulator Experimental Design

Figure 7.4 narrates the DS testbed, consisting of 5500 meters of a divided two-way four-lane freeway with a 75 mph initial speed limit, replicating the primary traffic operation performance on I-80. The scenario includes two horizontal curves, Curve-A and Curve-B, with the same geometric characteristic (longitudinal grade = 0.5%, radius ≃ 550 meters, length ≃ 800 meters, central angle ≃ 90o, superelevation = 8%, and the presence of vertical curve: NO). The entire scenario was designed based on a dry pavement condition with a regular friction coefficient. The weather condition was clear throughout the first 3700 meters of the testbed. However, just before Curve-B, it switched to dense foggy weather conditions with a visibility of 110 meters (360 feet), which remained 500 meters after the end of Curve-B. This visibility distance was considered because, according to the Virginia Department of Transportation, visibility of lower than 360 feet notably reduces traffic safety performance [98]. It should be noted that in the non-CV scenario, where HMI was off, no information regarding the upcoming weather condition and the associated advisory speed was given to drivers. However, in the CV scenario, two notifications regarding the forthcoming foggy weather condition and the advisory speed of 45 mph were given to the drivers via HMI, accompanied by a "beep" sound to best draw drivers' attention [52].

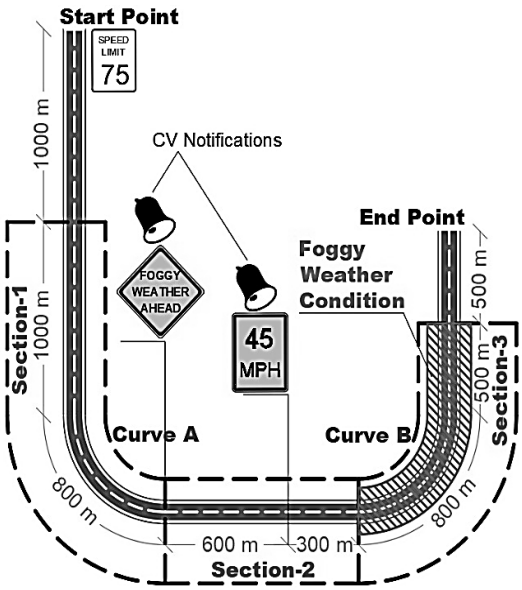


Figure 7.4 Driving Simulator Testbed for Analysis of Run-off-Road and Rear-End Crashes on Horizontal Curves under Foggy Conditions

7.3.2 Data Analysis

The effect of the CV notifications on driver behavior was unveiled by analyzing vehicle trajectories in CV and non-CV environments. According to Figure 7.5, the trajectory-level observations were divided into three sections (Section-1, 2, and 3) for further investigations. In Section-1, no TIMs were given to CV drivers. Thus, it is expected to see very similar behavior between CVs and non-CVs, if not the same. Throughout Section-2, CV drivers received two notifications, which will be analyzed to reveal immediate CV driver reactions and compliance to the received TIMs. Section-3, however, is the study's main focus, where CV safety performance under combined foggy weather with horizontal curves was assessed by characterizing K-SMoS.

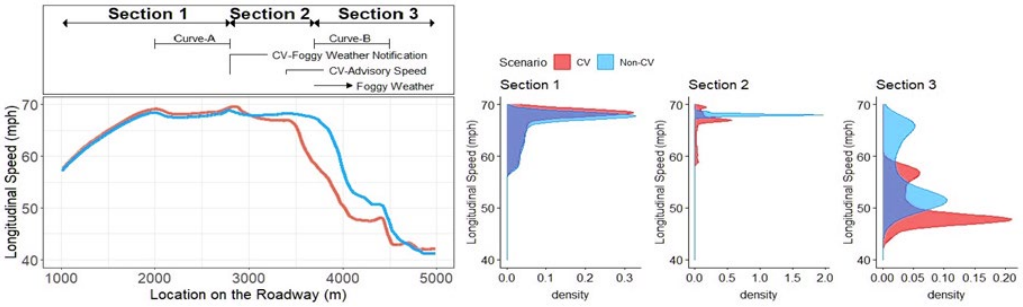
The alteration of driver behavior due to CV notifications was analyzed longitudinally to account for the high risk of rear-end crashes in foggy weather conditions [9], and laterally due to the presence of horizontal curve, as an influential factor to increasing the risk of run-off-road, sideswipe, and rollover crashes [7], [59], [81]. Therefore, longitudinal speed, lateral speed, and steering, directly measured in the DS, were considered for further investigation. For the sake of dispersion analysis and assessing uncertainty in driver behavior, spatial standard deviations (SSD) of these K-SMoS were computed across vehicle trajectories and throughout the simulated roadway for CVs and non-CVs. The coefficient of variation (CoV) in longitudinal speed was also regarded because this variable is one of the most important real-time speed-related factors contributing to crash probability [9].

The safety performance of CV notifications was evaluated by comparing K-SMoS distributions in CV and non-CV environments (with/without analysis) using data visualization and statistical analysis. Similar to the previous analysis on an HC with a slippery road surface, the statistical analysis emphasizes two characteristics of K-SMoS distributions: central tendency and dispersion.

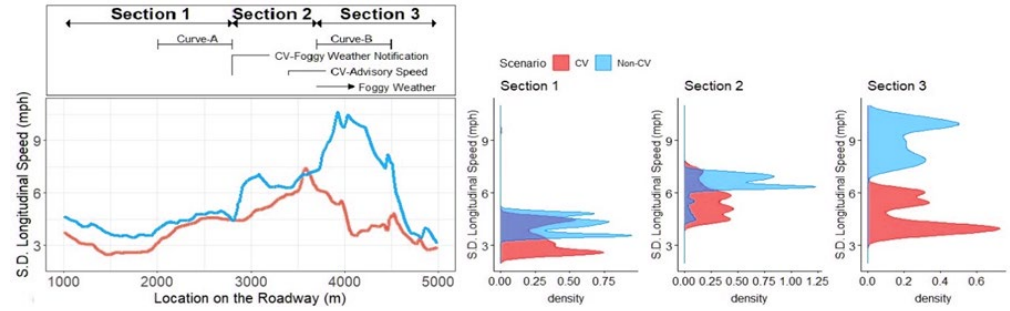
7.3.2.1 Visual Inspection of Trajectory-Level Observations

Figures 7.5 and 7.6 present the trajectory-level investigations for all K-SMoS across the 23 participants in CV and non-CV scenarios in three sections. Based on Figure 7.5, a systematic reduction in K-SMoS relating to CVs' longitudinal driving behaviors can be seen, gradually increasing from Section-1 to Section-3. In Section-3, almost the entire values of CoV and SSD in CVs' longitudinal speed are less than the non-CVs, implying the safety enhancement due to drivers' increased situational awareness [9]. Figure 7.5A shows that CV drivers smoothly reduced their longitudinal speed once they received advisory speed, implying speed compliance in the CV environment. This smooth reduction is not comparable with non-CVs' abrupt speed reduction at the beginning of foggy conditions that would increase traffic backward shock waves and rear-end crash probability in non-CV environments.

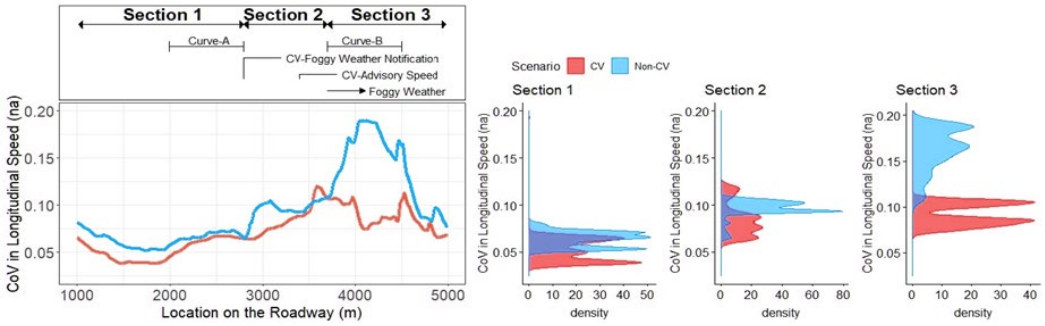
Switching to the effect of CV notifications on truck drivers' lateral driving behaviors, Figure 7.6 narrates the similarity of truck drivers' lateral behavior in Section-1 and even in Section-2. However, Section-3 does not indicate systematic changes or a remarkable difference for lateral speed and steering between CV and non-CV trajectories. Nonetheless, delving into the SSD of lateral speed and steering would reveal a systematic reduction in the spatial variation of K-SMoS for CVs compared with non-CVs in Section-3, specifically on Curve-B (see Figure 7.6B and 7.6D). These reductions in SSDs illustrate more certainty in CVs' lateral behavior than non-CVs, implying the effectiveness of CV notifications on altering drivers' lateral behavior.



A- The Effect of CV Notifications on the Longitudinal Speed



B- The Effect of CV Notifications on the Spatial Standard Deviation (SSD) of Longitudinal Speed

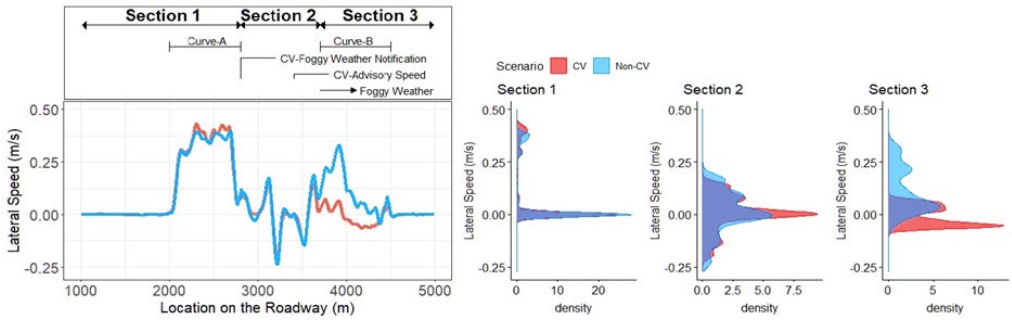


C- The Effect of CV Notifications on the Coefficient of Variation (CoV) in Longitudinal Speed

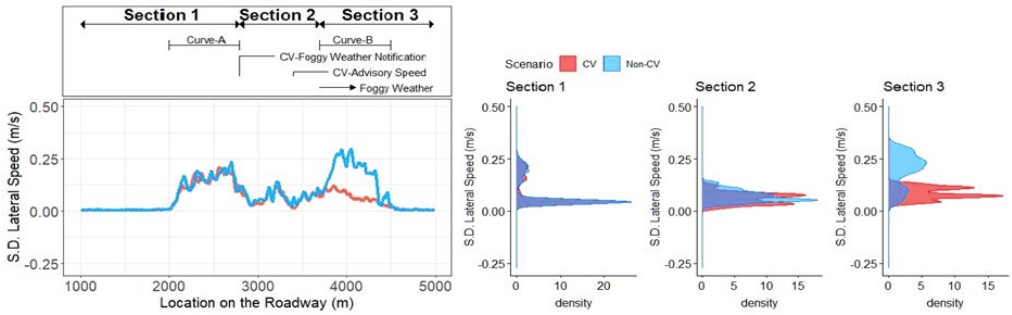
Figure 7.5 Alteration of Longitudinal K-SMoS Distributions due to CV Notifications

7.3.2.2 Statistical Inference

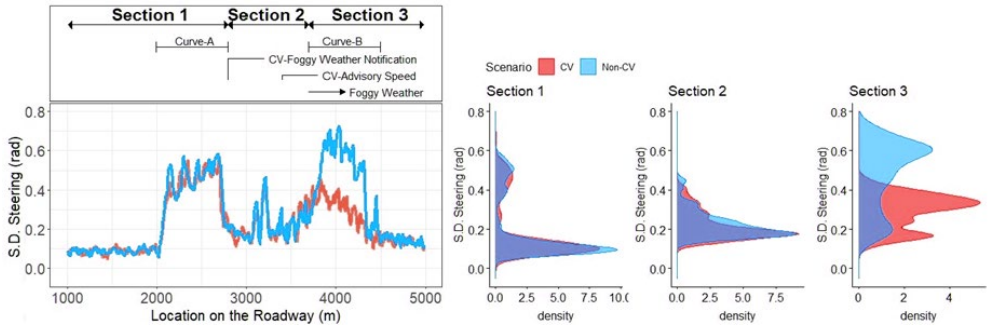
According to Table 32, Shapiro-Wilk normality tests led to statistically significant results for all K-SMoS associated with CVs and non-CVs across three sections, confirming the appropriateness of the nonparametric approach for the central tendency analysis. Considering the central tendency analysis of K-SMoS utilizing WSRT, it can be seen that none of the WSRT at Section-1 were statistically significant. This inference affirms the previous visual inspection, indicating similarity in CV and non-CV behavior, which was expected because no notifications were given to CVs in this section. In Section-2, the WSRTs indicate statistically significant differences in the central tendency for longitudinal speed and its SSD and the coefficient of variation, all associated with longitudinal driving behavior and not drivers' lateral behavior. This inference implies truck drivers' immediate compliance to CV notifications, which is very apparent in their longitudinal and speed adaptation behavior, consistent with the literature [99]. These differences in central tendencies coincided with reductions of these K-SMoS for CVs compared with non-CVs, promoting speed harmonization and traffic safety levels [9], [39], [100].



A- The Effect of Connected Vehicle Notifications on the Lateral Speed



B- The Effect of Connected Vehicle Notifications on the Spatial Standard Deviation (SSD) of Lateral Speed



C- The Effect of Connected Vehicle Notifications on Steering

Figure 7.6 Alteration of Lateral K-SMoS Distributions due to CV Notifications

In Section-3, with a horizontal curve located in foggy weather, the situation is totally different. Without exception, all the WSRT indicated statistically significant differences in longitudinal and lateral behavior of CV compared with non-CV drivers. In terms of the longitudinal speed adaptation, the central tendencies of the SSD of speed and coefficient of variation in speed have been reduced by almost 55% and 48% in the CV scenario, respectively. According to the literature, these two variables are among the most influential real-time crash contributing factors, directly increasing the risk of crashes [9], [39]. The SSD of speed is associated with spatial dispersion of longitudinal driving behavior. The mentioned reductions indicating more certainty in driver behavior negotiating HCs in foggy conditions are expected to be observed more from CVs than non-CVs. Therefore, the results affirm the safety advantage and the effectiveness of CV technology by altering and minimizing uncertainty in drivers' longitudinal behavior, alleviating the causal effect of these factors contributing to crash occurrences and promoting speed harmonization in Section-3.

Table 7.3 Shapiro-Wilk Test of Normality and Central Tendency Analysis of K-SMoS

Section-1

Kinematic Based Surrogate Measure of Safety	Scenario	Test of Normality: Shapiro-Wilk Test			Central Tendency Analysis: Wilcoxon Signed-Rank Test				
		Test Stats.	p-value	Sig.	Test Stats.	p-value	Sig.	Median	% of Reduction [a]
Longitudinal Speed (mph)	CV	0.79	<0.001	*	1629595	0.825		68.064	-0.89 [b]
	Non-CV	0.79	<0.001	*				67.462	
SSD[c] of Longitudinal Speed (mph)	CV	0.84	<0.001	*	1574371	0.122		3.321	21.12
	Non-CV	0.86	<0.001	*				4.210	
CoV [d] Longitudinal Speed (mph)	CV	0.74	<0.001	*	1566099	0.070		0.051	22.21
	Non-CV	0.80	<0.001	*				0.065	
Lateral Speed (m/s)	CV	0.60	<0.001	*	3919265	0.899		0.012	7.69 [e]
	Non-CV	0.58	<0.001	*				0.013	
SSD of Lateral Speed (m/s)	CV	0.64	<0.001	*	3918882	0.904		0.005	11.03
	Non-CV	0.63	<0.001	*				0.005	
Steering (rad)	CV	0.64	<0.001	*	3884658	0.656		-0.015	31.81 [e]
	Non-CV	0.61	<0.001	*				-0.022	
SSD of Steering (rad)	CV	0.51	<0.001	*	3927610	0.791		0.118	-4.48
	Non-CV	0.70	<0.001	*				0.113	

Section-2

Kinematic Based Surrogate Measure of Safety	Scenario	Test of Normality: Shapiro-Wilk Test			Central Tendency Analysis: Wilcoxon Signed-Rank Test				
		Test Stats.	p-value	Sig.	Test Stats.	p-value	Sig.	Median	% of Reduction
Longitudinal Speed (mph)	CV	0.80	<0.001	*	185675	<0.001	*	66.913	1.52
	Non-CV	0.97	<0.001	*				67.945	
SSD of Longitudinal Speed (mph)	CV	0.93	<0.001	*	805504	<0.001	*	5.437	17.48
	Non-CV	0.93	<0.001	*				6.589	
CoV Longitudinal Speed (mph)	CV	0.90	<0.001	*	201091	<0.001	*	0.081	16.13
	Non-CV	0.82	<0.001	*				0.097	
Lateral Speed (m/s)	CV	0.94	<0.001	*	401174	0.886		0.004	47.70 [e]
	Non-CV	0.97	<0.001	*				0.008	
SSD of Lateral Speed (m/s)	CV	0.96	<0.001	*	399423	0.762		0.065	2.19
	Non-CV	0.95	<0.001	*				0.066	
Steering (rad)	CV	0.96	<0.001	*	425301	0.085		0.016	23.81 [e]
	Non-CV	0.92	<0.001	*				-0.021	
SSD of Steering (rad)	CV	0.89	<0.001	*	400728	0.854		0.191	5.87
	Non-CV	0.87	<0.001	*				0.203	

Section-3

K-SMoS	Scenario	Test of Normality: Shapiro-Wilk Test			Central Tendency Analysis: Wilcoxon Signed-Rank Test				
		Test Stats.	p-value	Sig.	Test Stats.	p-value	Sig.	Median	% of Reduction
Longitudinal Speed (mph)	CV	0.84	<0.001	*	134512	<0.001	*	48.135	9.53
	Non-CV	0.86	<0.001	*				53.204	
SSD of Longitudinal Speed (mph)	CV	0.85	<0.001	*	471492	<0.001	*	4.162	54.81
	Non-CV	0.91	<0.001	*				9.211	
CoV Longitudinal Speed (mph)	CV	0.88	<0.001	*	13332	<0.001	*	0.087	47.64
	Non-CV	0.91	<0.001	*				0.166	
Lateral Speed (m/s)	CV	0.89	<0.001	*	90759	<0.001	*	-0.023	65.15 [e]
	Non-CV	0.91	<0.001	*				0.066	
SSD of Lateral Speed (m/s)	CV	0.95	<0.001	*	87483	<0.001	*	0.062	67.35
	Non-CV	0.93	<0.001	*				0.191	
Steering (rad)	CV	0.94	<0.001	*	477803	<0.001	*	-0.311	41.65 [e]
	Non-CV	0.96	<0.001	*				-0.533	
SSD of Steering (rad)	CV	0.95	<0.001	*	132643	<0.001	*	0.308	41.72
	Non-CV	0.90	<0.001	*				0.529	

* Representative of statistically significant under 95% Confidence Interval (CI).
a The percentage of reduction is the decrease in the median of K-SMoS in CVs compared to non-CVs relative to non-CVs.
b Negative signs for the percentage of reduction showed the central tendency of K-SMoS of CVs is more than non-CVs
c Spatial Standard Deviation (SSD).
d Coefficient of Variation (CoV).
e The percentage of reduction was calculated based on the absolute values of medians to make this measure directional free.

Specifically, for Curve-B located in Section-3 under foggy conditions, steering and lateral speed have been selected as K-SMoS to unveil changes in drivers' lateral behavior due to the effect of CV notifications. Although no statistically significant differences were found regarding the central tendency of these two factors in Section-1 and Section-2 between CVs and non-CVs, the findings in Section-3 indicate another inference. In Section-3, the central tendencies of steering and lateral speed have been reduced by almost 42% and 65%, respectively, in favor of CVs. Note that the reduction percentages for these two K-SMoS were computed based on the absolute values of medians to make these reductions directional-free. According to the friction ellipse theory [59], [101], a reduction in lateral speed would minimize lateral skidding, mitigating the risk of rollover and sideswipe crashes due to CV notifications on Curve-B. It is known that steering can be effectively used to examine driver performance in controlling vehicles' lateral positions in adverse traffic conditions [7]. A larger absolute value of steering angle is highly correlated with high-risk drivers and a small safety margin [101]. Hence, reducing the central tendency of steering angle in the CV scenario would remarkably enhance traffic safety performance.

In addition to analysis of central tendency, it is essential to assess the spatial variability of steering and lateral and longitudinal speed across CVs and non-CVs throughout the simulated roadway. The reason for this necessity is that a specific treatment might result in a very similar central tendency for two samples being compared but with totally different levels of variability, where the higher variability in a sample depicts a higher uncertainty regarding the effect of treatment (e.g., CV notifications) [88]. Therefore, to assess the uncertainty level in CV and non-CV behavior, the dispersion analysis was also applied by delving into the SSD of the mentioned K-SMoS.

It was possible to simply consider the standard deviation of these K-SMoS. However, this approach would result in one value per section per K-SMoS for CVs and non-CVs, impeding trajectory-level dispersion comparison in K-SMoS between CV and non-CV drivers throughout three sections. In fact, it was necessary to reveal how CV notifications might affect variability and certainty in CV drivers' behavior compared with non-CVs at every single location of the roadway. Accordingly, it was decided to investigate the SSD of K-SMoS across all drivers in CV and non-CV scenarios as a separate variable.

Considering the lateral speed and steering in Section-3, the central tendencies of SSD for these K-SMoS have been reduced by 67% and 42%, respectively, in the CV scenario. This inference indicates more certainty and a higher homogeneity in lateral driving behavior that could be observed by providing drivers with real-time advisory/warning notifications in the CV environment, minimizing the risk of horizontal curve-related crashes, and augmenting the overall traffic safety performance.

8. CRASH DETECTION IN CV ENVIRONMENTS

8.1 Introduction

In the crash detection domain, the purpose is to accurately and rapidly detect crash time and location [102], leading to reducing delay and inconvenience imposed by crashes, swiftly dispatching emergency services [103]–[107], mitigating the risk of secondary crashes [108], and even transmitting distress notifications in a CV environment [109]. Regardless of the variety of methodology used in the literature, many previous analyses only rely on traditional traffic observations for traffic crash detection in the absence of detailed individual driving data. Therefore, this part of the document presents the application of CV technology in detecting traffic crashes, specifically for I-80 in Wyoming, based on the generated large-scale integrated trajectory-level observations embedded in BSMs from driving simulator experiments.

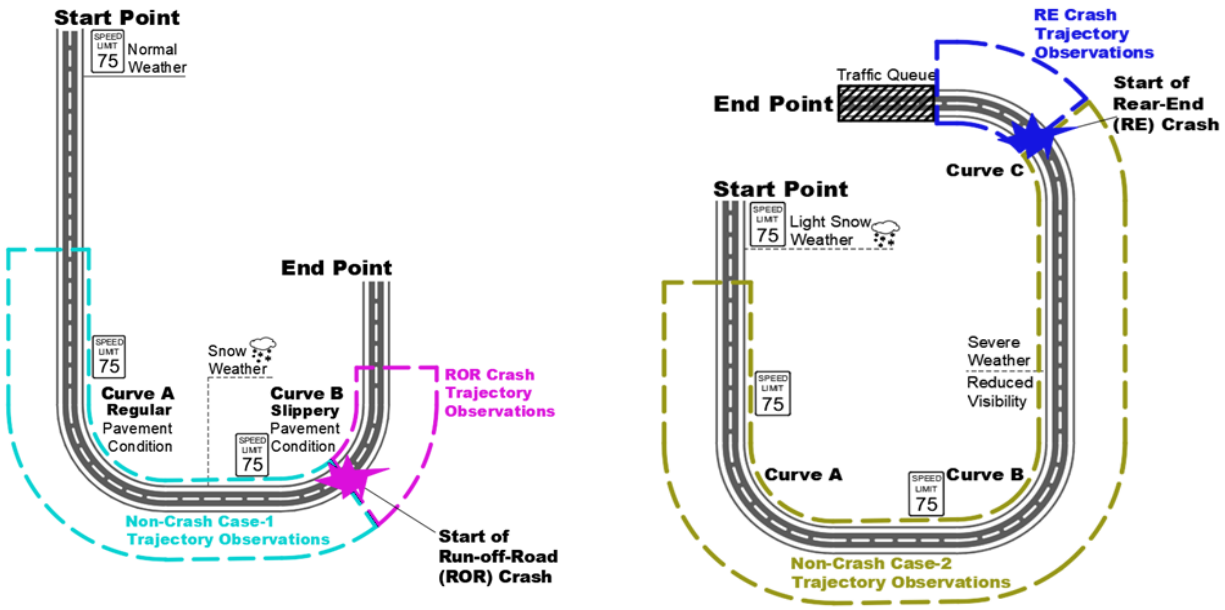
The core of BSMs includes CVs' dynamic and kinematic information relating to location, speed, acceleration/deceleration, steering, etc. [110]. Accordingly, there is continuous and real-time access to CV trajectories. Conversion of these trajectories into a well-characterized K-SMoS might quantify CV driver styles and volatilities under different traffic and environmental conditions [7]. On the other side, it has been shown that extreme driver volatility is highly correlated with critical safety events [111], implying the suitability of extreme value theory for analyzing extreme driver behaviors. Accordingly, the research team conducted driving simulator scenarios to form conditions similar to run-off-road (ROR) and rear-end (RE) crashes. The crash detection matter for each crash type was formulated as a binary classification of two continuous K-SMoS generalized extreme value (GEV) distributions under safe and unsafe traffic conditions, wherein the optimal thresholds have been set to detect crashes.

8.2 Driving Simulator Scenario Development

Two experiments were developed based on a four-lane freeway segment with a 75 mph speed limit to replicate the actual traffic operation performance of I-80. At the end of each scenario, one specific crash condition, ROR and RE crashes, was simulated. The rest of each scenario was considered non-crash conditions. Thus, four cases of study based on two crash conditions and two corresponding non-crash conditions were conducted. For each case, vehicle trajectories were obtained to analyze and compare driving behaviors under crash conditions and normal traffic conditions.

8.2.1 Run-off-Road (ROR) Crash Scenario

According to Figure 8.1A, this scenario consists of two horizontal curves (HCs) with the same geometric characteristic (radius = 509.3 m, length= 800 m, central angle= 90º, superelevation= 6%, longitudinal grade= 0%, presence of vertical curve: No). The scenario starts with a normal weather condition until it passes the first curve (Curve-A). After Curve-A, dynamic adverse weather is simulated within the tangent segment between the first and the second curve (Curve-B). Unlike Curve-A, with sufficient pavement friction to stabilize vehicles negotiating the curve, Curve-B's pavement friction was remarkably reduced due to light snow conditions with a very low temperature, simulating ROR crash conditions.



A- Run-off-Road Crash Scenario

B- Rear-End Crash Scenario

Figure 8.1 Driving Simulator Scenarios for the Crash Detection

8.2.2 Rear-End (RE) Crash Scenario

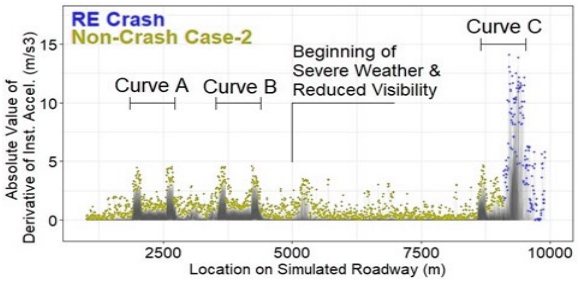
According to Figure 8.1B, in this scenario, the roadway geometry consisting of three HCs with the same geometric as the first scenario was designed. The pavement surface was designed to simulate a very slight snowy surface, whose friction coefficient was close to dry road surface conditions throughout the entire scenario. The scenario started with very light snow weather and clear visibility until the second HC. After passing the second curve, a sudden severe weather condition and low visibility due to heavy snow conditions were simulated. At the end of the scenario, a traffic queue resulting from a presumptive primary crash was simulated. Due to the reduced visibility on Curve-C, drivers could recognize the traffic queue ahead very late, forcing them to react within a limited time to collision, simulating a rear-end crash.

8.3 Data Analysis

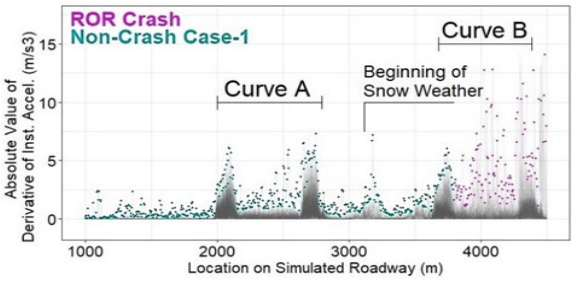
Vehicle trajectories under four study cases, including ROR crash, RE crash, and their corresponding non-crash cases, were converted into two K-SMoS: absolute value of derivative of instantaneous acceleration (ADInstAccel) and absolute value of derivative of steering (ADSteering). Accordingly, eight GEV distributions under Bayesian inference were fitted for two K-SMoS under four study cases.

8.3.1 Block Maxima Approach in Extreme Value Theory (EVT)

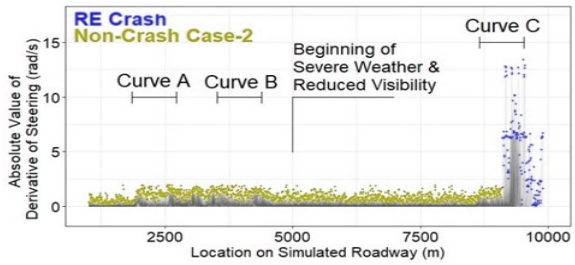
Extreme value theory (EVT) establishes the probabilistic and statistical tools to quantify those events that are more extreme than what often has been seen. This statistical area is advantageous for crash detection based on extreme changes in driver behaviors quantified by K-SMoS. The extreme values of K-SMoS were modeled using the block maxima (BM) approach (Figure 8.2) and estimating their corresponding GEV distributions [112]–[115] under crash and non-crash conditions.



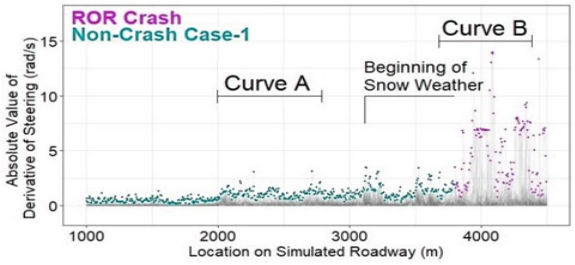
A- Block Maxima of Absolute Value of Derivative of Instantaneous Acceleration in the RE Scenario



B- Block Maxima of Absolute Value of Derivative of Instantaneous Acceleration in the ROR Scenario



C- Block Maxima of Absolute Value of Derivative of Steering in the RE Scenario



D- Block Maxima of Absolute Value of Derivative of Steering in the ROR Scenario

Figure 8.2 Block Maxima Approach in Extreme Value Analysis

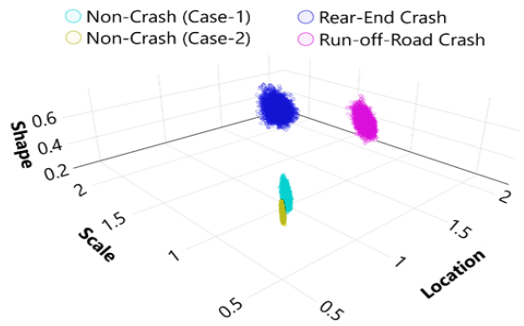
Based on the BM approach, the maxima of K-SMoS were identified across all vehicle trajectories with a resolution of 60 observations per second (60 Hz) in every five-meter spatial block under crash and non-crash cases, as shown in Figure 8.2. Using the captured maxima, probability density functions (PDF) of GEV distributions for two K-SMoS in four cases of the study were fitted according to Equation 8.1 [116].

$$H(\xi,\mu,\sigma;x)=\begin{cases} exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} & if\ \xi\neq 0\ and\ \ x\colon 1+\xi\left(\frac{x-\mu}{\sigma}\right)>0 \\ exp\left\{\left[-\exp\left(-\left(\frac{x-\mu}{\sigma}\right)\right)\right]\right\} & if\ \ \xi=0\ and\ x\in\mathbb{R} \end{cases} \quad \text{Equation 8.1}$$

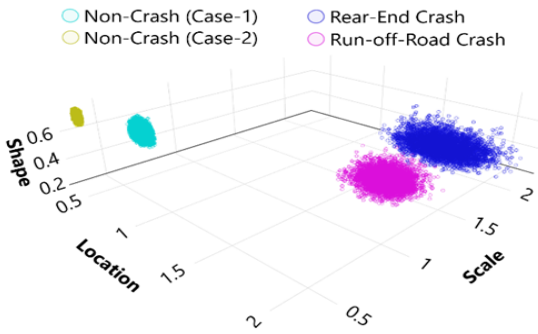
In Equation 8.1, μ is the location parameter akin to the mean and measures the central tendency of maxima/minima with a larger value localizing the distribution of the maxima at higher values. The scale parameter $\sigma > 0$ is an extent of dispersion and variance of maximal whose higher values indicate a more dispersed distribution of the maxima. Shape parameter ξ, also called the tail index, is a measure of skewness and tail behavior with higher values corresponding to heavier tails.

8.3.2 Bayesian Inference

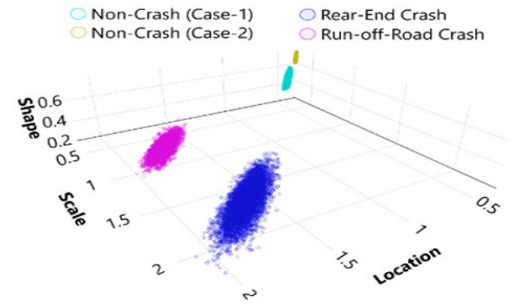
The estimation of GEV parameters was conducted under Bayesian inference using Markov chain Monte Carlo (MCMC) simulation techniques based on non-informative priors. Bayesian inference was followed because it is concerned with the posterior probability distribution lying within regions of that space [117], helping in uncertainty analysis in crash detection. The visual assessment of K-SMoS GEV parameters' joint posterior density distributions could qualitatively reveal these parameters' uncertainties [118]. Figure 8.3 illustrates distributions of GEV parameters in three-dimensional perspectives for two K-SMoS to demonstrate how these distributions are similar/different to/from each other by considering uncertainty.



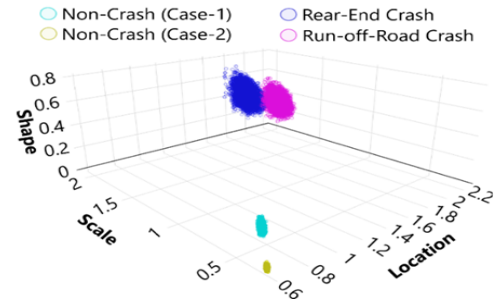
A- Joint Posterior Probability of GEV Distributions for ADInstAccel **(View-1)**



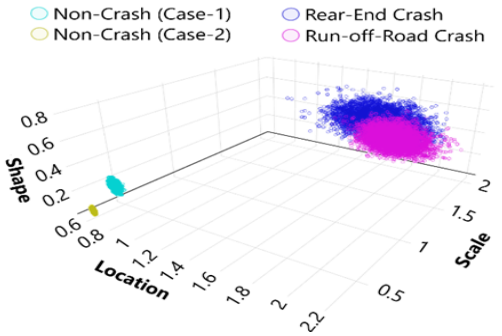
B- Joint Posterior Probability of GEV Distributions for ADInstAccel **(View-2)**



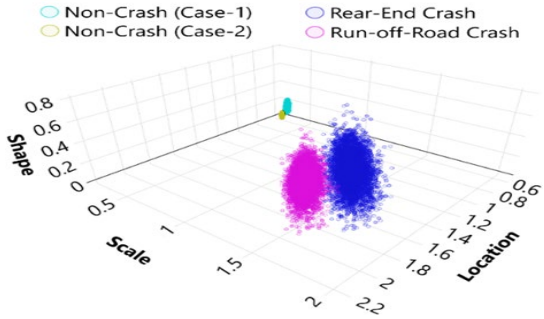
C- Joint Posterior Probability of GEV Distributions for ADInstAccel **(View-3)**



D- Joint Posterior Probability of GEV Distributions for ADSteering **(View-1)**



E- Joint Posterior Probability of GEV Distributions for ADSteering **(View-2)**



F- Joint Posterior Probability of GEV Distributions for ADSteering **(View-3)**

Figure 8.3 Joint Posterior Probability Distributions of Eight K-SMoS GEV Parameters for Crash and Non-crash Cases in a Three-Dimensional Space

In Figure 8.3, every single point represents one unique GEV distribution, and each of the point clouds corresponds to one study case. Having 5000 samples generated by MCMC after warming up, each point cloud consists of 5000 GEV distribution representations. The presented samples in Figure 8.3 are generated by random sampling of MCMC from the joint posterior distributions [119] of three parameters (i.e., location, shape, and scale) for each GEV, corresponding to K-SMoS, fitted under crash and non-crash cases. For each K-SMoS, it is clear that the corresponding point clouds under non-crash cases are remarkably denser than the crash counterparts, indicating higher certainty in the GEV parameter estimations. This visual inference indicates that extreme driving styles under normal traffic conditions are much more similar than under critical conditions.

In addition to Figure 8.3, Table 8.1 presents eight fitted GEV distribution parameter estimations under four study cases using two K-SMoS. Again, the narrower credible intervals (CIs) associated with GEV parameter estimations in non-crash cases show more certainty in extreme driving styles in normal traffic than crash conditions. Based on Table 8.1, the location parameters of K-SMoS GEV distributions under crash conditions are notably higher than their corresponding values under non-crash conditions. The fitted GEVs for extreme K-SMoS have systematically shifted away from zero in crash conditions. Depending on the driving simulator scenario and the type of K-SMoS, this shifting is between two to four times the reference levels (i.e., estimated location parameter in non-crash cases). Hence, the central tendency of extreme driver behaviors captured by K-SMoS would take remarkably higher values in crash compared with non-crash conditions.

Table 8.1 GEV Parameter Estimations for Two K-SMoS under Four Study Cases

Absolute value of Derivative of Instantaneous Acceleration (ADInstAccel)						
Parameters of Generalized Extreme Value (GEV) Distribution	Non-crash (Case-1) [a]	Run-off-Road Crash	Relative Percentage Change (%) [b]	Non-crash (Case-2) [a]	Rear-End Crash	Relative Percentage Change (%)
Location	0.38 (0.37, 0.39) [c]	1.99 (1.88, 2.10)	423.68	0.55 (0.52, 0.57)	1.91 (1.72, 2.11)	247.27
Scale	0.36 (0.35, 0.37)	1.02 (0.92, 1.13)	183.33	0.60 (0.57, 0.63)	1.63 (1.45, 1.80)	171.66
Shape	0.70 (0.66, 0.73)	0.58 (0.49, 0.67)	-17.14	0.59 (0.53, 0.65)	0.55 (0.35, 0.60)	-6.77
Absolute value of Derivative of Steering (ADSteering)						
Parameters of Generalized Extreme Value (GEV) Distribution	Non-crash (Case-1)	Run-off-Road Crash	Relative Percentage Change (%)	Non-crash (Case-2)	Rear-End Crash	Relative Percentage Change (%)
Location	0.49 (0.48, 0.50)	2.05 (1.89, 2.21)	318.36	0.60 (0.58, 0.62)	1.79 (1.58, 2.01)	198.33
Scale	0.27 (0.26, 0.28)	1.15 (1.02, 1.31)	325.92	0.39 (0.38, 0.41)	1.41 (1.20, 1.61)	261.53
Shape	0.22 (0.21, 0.24)	0.45 (0.32, 0.57)	104.54	0.19 (0.16, 0.24)	0.50 (0.35, 0.66)	163.15

a The non-crash case-1 and case-2 respectively relate to vehicles' trajectory observations in normal traffic conditions in the run-off-road and rear-end crash scenarios.

b The Relative Percentage Change (%) indicates the change of estimated GEV's parameter from a non-crash to the corresponding crash case divided by the value of GEV's parameter in the non-crash case, as the reference value, expressed in percentage term. The positive values of Relative Percentage Change show GEVs' parameters have been shifted away from zero in crash conditions compare to non-crash counterparts.

c Values in parentheses indicate the lower and upper bound of 95% Credible Intervals.

8.4 Crash Detection

8.4.1 Binary Classification

The purpose is to discriminate crash versus its corresponding non-crash conditions using fitted GEV PDFs associated with K-SMoS. In this regard, it is virtually impossible to see a perfect separation between PDFs of a K-SMoS under crash and non-crash conditions. Hence, PDFs associated with a K-SMoS might have an overlap. Having said that, for any possible hypothetical selected threshold to discriminate between the two PDFs, there might be true positive, true negative, false negative, and false positive diagnoses of the crash and non-crash conditions. Hence, to assess the discriminative power of each K-SMoS to cluster ROR and RE crashes from non-crash cases, the receiver operating characteristic (ROC) was conducted, and the area under the curve (AUC) was calculated. Since the binary classifications are to be conducted between two continuous PDFs of each K-SMoS under crash and non-crash conditions, sensitivity, and specificity were calculated according to Equations 8.2 and 8.3 for a predefined threshold T [120], where $f(x)$ is the PDF of a K-SMoS in the non-crash case and $g(x)$ is its corresponding PDF under crash conditions throughout the range of K-SMoS (x). Sensitivity is defined as the probability of a true positive, whereas specificity is defined as the probability of a true negative.

$$Sensitivity: \int_{T}^{+\infty} g(x)dx = 1 - g(T) \quad \text{Equation 8.2}$$

$$Specificity: \int_{-\infty}^{T} f(x)dx = f(T) \quad \text{Equation 8.3}$$

Therefore, for any predefined T, unique sensitivity and specificity were obtained. To establish ROC and calculate AUC associated with each of the K-SMoS under each scenario, sensitivity and specificity were calculated by altering T with a 0.1-unit increment from the minimum to the maximum value taken by each of the K-SMoS. Figure 8.4 presents the discriminating ability of each K-SMoS for clustering vehicle trajectories from crash versus non-crash conditions based on ROC and associated AUCs. For each of the crash scenarios, ADSteering would result in better accuracy compared with ADInstAccel. For rear-end and ROR crash scenarios, ADSteering resulted in 81% and 79% accuracies, respectively, which are notably more than the 70% and 71% accuracies obtained by ADInstAccel for the mentioned scenarios.

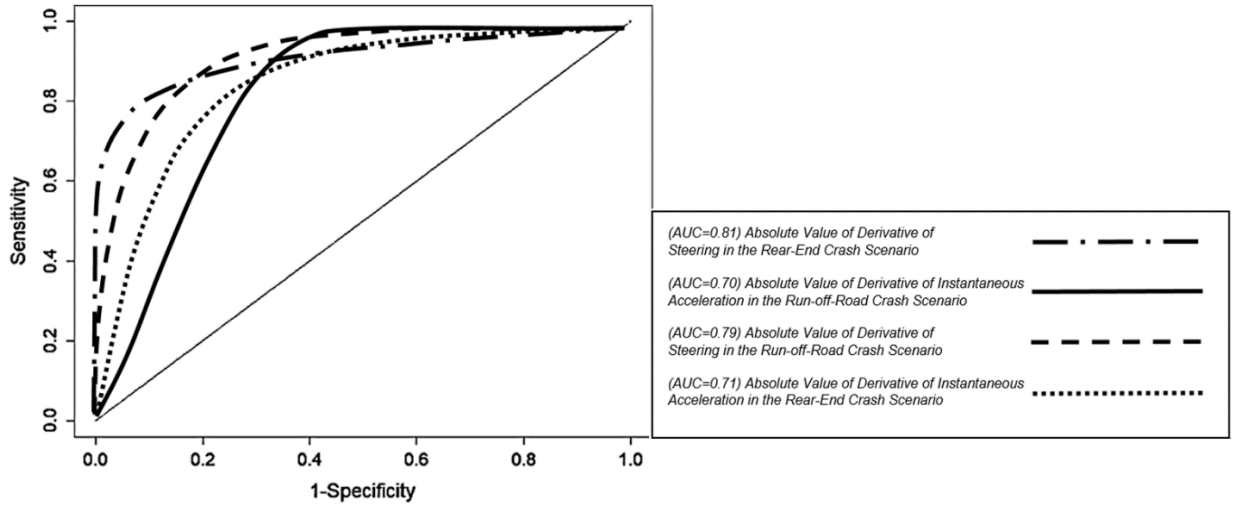


Figure 8.4 Comparison of Discriminating Ability of Two K-SMoS in Clustering Two Crash Types from Their Non-Crash Corresponding Cases

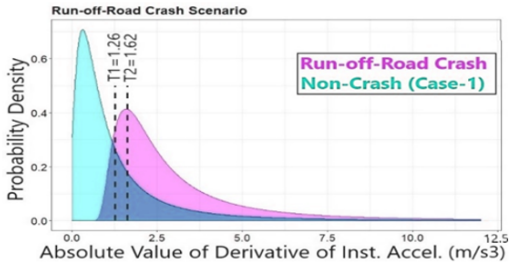
8.4.2 Optimal Threshold Setting for K-SMoS in Crash Detection

The optimal thresholds throughout the range of the K-SMoS were identified based on two distinct and widely used methods [121]. According to Equation 8.4, the first method takes advantage of maximizing the sum of sensitivity and specificity. However, the second method defines it for such a threshold leading to equal sensitivity and specificity in binary classification, presented by Equation 8.5 [121].

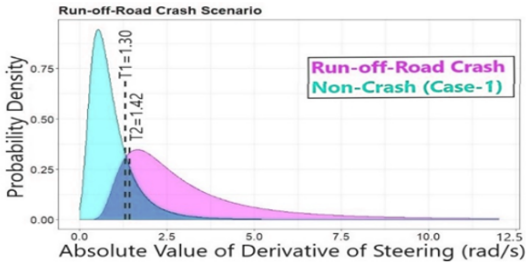
$$Method\ 1\ (T_1): T_1 = argmax_T\ (f(T) + 1 - g(T)) \quad \text{Equation 8.4}$$

$$Method\ 2\ (T_2): T_2 = (\exists T \in \mathbb{R}) \mid (f(T) = 1 - g(T)) \quad \text{Equation 8.5}$$

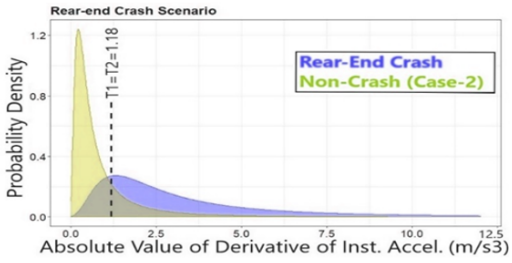
Accordingly, for each of the K-SMoS, four optimal thresholds were defined based on two different driving simulator scenarios and two threshold setting methods. Four defined thresholds were averaged to report the final optimal threshold per K-SMoS. These threshold optimizations were performed by specificity and sensitivity analysis in the binary classification under varying cut-off points throughout a range of each K-SMoS. Figure 8.5 presents the location of the defined optimal thresholds on the fitted GEV distributions for each of the K-SMoS under two crash scenarios.



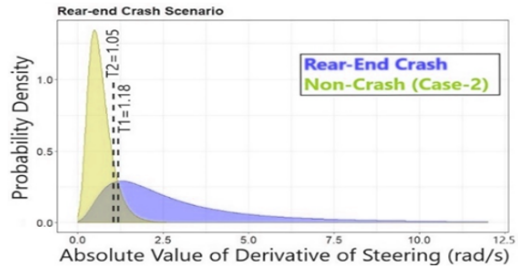
A- GEV Distributions of Absolute Value of Derivative of Instantaneous Acceleration in the Run-off-Road Scenario



B- GEV Distributions of Absolute Value of Derivative of Steering in the Run-off-Road Scenario



C- GEV Distributions of Absolute Value of Derivative of Steering in the Rear-End Scenario



D- GEV Distributions of Absolute Value of Derivative of Instantaneous Acceleration in the Rear-End Scenario

Figure 8.5 K-SMoS GEV Distributions under Crash and Non-Crash Conditions

The final thresholds for ADSteering and ADInstAccel to detect crash conditions in the CV environment were achieved by averaging the four optimal thresholds associated with each of the K-SMoS. Table 8.2 presents the details. The final ADSteering and ADInstAccel thresholds were defined to be 1.24 (rad/s) and 1.31 (m/s3), respectively. The higher reliability of ADSteering in detecting crash conditions has been quantified by comparing its final threshold performance to ADInstAccel final threshold in clustering two crash types from their corresponding non-crash cases. According to Table 8.2, it is clear that ADSteering has notably outperformed ADInstAccel based on all three metrics (i.e., accuracy, sensitivity, and specificity). ADSteering could cluster each of the two crash conditions with an accuracy of almost 85%, which coincided with achieving a high rate of sensitivity and specificity.

Table 8.2 Optimal Threshold Definition for Crash Detection Based on Two K-SMoS

	Run-off-Road Crash Scenario		**Rear-End Crash Scenario**		**Final Threshold**	
	Optimal Thresholds Setting					
K-SMoS	Method-1 [a]	Method-2 [b]	Method-1	Method-2	Average	Standard Deviation
ADSteering	1.3	1.42	1.05	1.18	**1.24**	0.16
ADInstAccel	1.26	1.62	1.18	1.18	**1.31**	0.21

Discrimination Ability of the Final Thresholds ($T_{ADSteering}$= 1.24 , $T_{ADInstAccel}$= 1.31) for Crash Detection

K-SMoS	**Run-off-Road Crash Scenario**			**Rear-End Crash Scenario**		
	Accuracy (AUC)	Sensitivity	Specificity	Accuracy (AUC)	Sensitivity	Specificity
ADSteering	0.84	0.77	0.91	0.86	0.92	0.79
ADInstAccel	0.79	0.69	0.88	0.78	0.81	0.76

Note:
K-SMoS = Kinematic-based Surrogate Measure of Safety; ADSteering = Absolute value of Derivative of Steering; ADInstAccel = Absolute value of Derivative of Instantaneous Acceleration; AUC = Area Under the Curve.
a The maximized sum of Specificity and Sensitivity
b The equalized Specificity and Sensitivity

9. CONCLUSIONS AND RECOMMENDATIONS

This research provided new traffic safety research perspectives to assess the safety performance of the Wyoming Connected Vehicle (CV) Pilot Deployment Program on Interstate-80 (WYDOT CV Pilot). Through the use of advanced statistical modeling and machine learning techniques, the research explained the procedure and the analytical inference for developing baseline and analysis, modeling, and simulation (AMS) framework under two distinct but complementary approaches: before/after analysis to explore real-time traffic-related crash/crash severity causations during CV pre-deployment as a baseline; and with/without analyses to quantify drivers' behavioral alterations under the effect of various CV applications.

The purpose of the first approach was to explore real-time traffic-related crash contributing factors and critical crash causations under the concept of the matched-case control design (MCCD) before deploying CV technology on 402 miles of I-80 in Wyoming. Comparing the causal effect of significant real-time traffic-related variables between CV pre-and post-deployment would reveal how CV technology affects I-80 safety performance under varying CV market penetration rates (MPRs). Accordingly, real-time crash prediction models (CPMs) were conducted, and traffic flow oscillations during crash precursors were analyzed throughout I-80's lengthy corridor with relatively low traffic volume compared with other interstates. The research team combined nonparametric data-driven statistical techniques with the machine learning method to identify the most significant real-time traffic-related factors contributing to the crash likelihood. Upon this identification, the concept of post-hoc interpretation was employed to visualize and interpret the causal effect of crash contributing factors using a variety of visualization tools.

Furthermore, the research team performed comprehensive analyses to identify critical crash causations under hierarchical Bayesian inference, which can account for unobservable factors (i.e., unobserved heterogeneity) affecting the likelihood of critical crashes on I-80 in Wyoming. The results provided WYDOT with the recognition of how real-time traffic-related and environmental factors significantly affect the probability of critical crashes. This recognition served as a baseline for the WYDOT CV Pilot, revealing the effect of CVs on mitigating the risk of these crashes on I-80 in Wyoming during post-deployment.

Switching to the with/without approach, the research team developed a reliable AMS framework, calibrated and validated based on real-time traffic observations preprocessed from the before/after analysis. Using the high-fidelity driving simulator lab at the University of Wyoming (WyoSafeSim Lab), drivers' behavioral alterations due to various CV notifications under the WYDOT CV Pilot were quantified. This quantification was analyzed separately on vehicle trajectories and was conflated with traffic microsimulation modeling to show the effect of the CV technology on the entire mixed traffic stream under varying CV MPRs. In the experiments, various CV applications were tested in a controlled environment, including spot weather impact warning (SWIW), distress notification (DN), situational awareness (SA), CV variable speed limit (CV-VSL), work zone warning (WZW), forward collision warning (FCW), and rerouting applications.

Due to the main focus of the WYDOT CV Pilot on improving safety performance for freight movement and commercial trucks, the analysis was mainly conducted on the truck driving simulator in the WyoSafeSim Lab with a total of six separate studies. In general, the research team inferred the positive impact of CV notifications on increasing drivers' situational awareness and conservativeness level, enhancing future traffic safety performance of I-80 in Wyoming.

The analysis of conflict-based surrogate measure of safety (C-SMoS) for WZW, SWIW, and CV-VSL in work zone areas under foggy weather conditions showed a remarkable safety enhancement in the work zone areas as the CV-MPR increases. The spatiotemporal investigation of speed harmonization depicted that the standard deviation of speed and mean speed in all work zone areas would decrease when the MPR of CVs increases. The research team introduced the WZW delivery distance as another macro-level contributing factor to the safety performance of CVs. It was inferred that the safety performance of CVs under different traffic volumes and in all WZ areas is not only affected by varying CV-MPRs but also by different WZW delivery distance influences. Findings from the traffic microsimulation modeling showed the number of traffic conflicts at the work zone would decrease if either CV MPR or the delivery distance increases. The findings indicated that the same safety level at higher MPRs could be achieved even at lower MPRs only if the WZW delivery distance increases. This inference is practically beneficial for the WYDOT CV Pilot because CV MPR would increase gradually, and its higher rate is not achievable soon. Accordingly, from the system design perspective, it has been shown that if CVs receive WZW 4500 feet before the transition area in work zones, their safety performance would be maximized, regardless of CV-MPR and traffic volume. Hence, the results recommend installing road-side units (RSUs) at the mentioned location in work zones. This finding is specifically essential for I-80, as a monotonous rural corridor, where providing full communication coverage for CVs might not be a cost-effective approach. Hence, WYDOT would achieve the maximum CV safety performance by utilizing mobile RSUs mounted on trailers at 4500 feet upstream of the transition area.

Furthermore, using trajectory-level observations from driving simulator experiments, the effect of SWIW, SA, and CV-VSL on truck drivers under limited visibility with complex roadway geometry was scrutinized. The research team investigated these application effects on longitudinal and lateral driver behaviors by comparing distributions of kinematic-based SMoS (K-SMoS). Results revealed an immediate and high driver compliance level in CV environments compared with non-CVs, which was more notable in drivers' longitudinal behavior and speed adaptation. Notable reductions in K-SMoS central tendencies in the CV compared with the non-CV scenario were observed that would potentially minimize the risk of rear-end crashes in foggy weather conditions and horizontal curve-related crashes under reduced visibility. Afterward, in a with/without analysis, the research switched to the safety assessment of the aforementioned CV applications on another black-spot point on I-80 where horizontal curves coincide with slippery pavement conditions. Data visualization and statistical inference indicated the safety outperformance of CVs compared with non-CVs, where the risk of run-off-road (ROR) and rear-end (RE) collisions was notably reduced, specifically due to CV-VSL and SWIW.

More importantly, the research team unlocked one of the most important potentials of basic safety messages (BSMs) for the WYDOT CV Pilot in a simulated environment for traffic crash detection, leading to swift dispatching of emergency services and minimizing the risk of secondary crashes. It was shown that analysis of K-SMoS characterized on BSMs could quantify extreme driver volatility and detect traffic crashes with a reasonable accuracy according to pre-identified thresholds for K-SMoS. These thresholds would have practical implications in automatically detecting crash-prone conditions in the CV environment. Real-time access to the large-scale trajectory-level datasets embedded in BSMs in a spatiotemporal dimension would enable WYDOT to monitor the entire I-80 corridor to detect critical safety events and automatically dispatch DNs, minimizing the risk of secondary crashes.

10. REFERENCES

[1] J. SAE, "3016: 2014 Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems," *Soc. Automot. Eng.*, 2014.

[2] NHTSA, "Critical reasons for crashes investigated in the national motor vehicle crash causation survey," 2015.

[3] F. Kitchener et al., "Connected Vehicle Pilot Deployment Program Phase 2, Data Management Plan-Wyoming," United States. Dept. of Transportation. ITS Joint Program Office, 2017.

[4] F. Kitchener et al., "Connected Vehicle Pilot Deployment Program Phase 2 Updated Performance Measurement and Evaluation Support Plan – Wyoming," 2018.

[5] AASHTO, *Highway Safety Manual*. Washington, D.C.: American Association of State Highway and Transportation Officials, 2011.

[6] S. M. S., Mahmud, L. Ferreira, M. S. Hoque, and A. Tavassoli, "Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs," *IATSS Res.*, vol. 41, no. 4, pp. 153–163, 2017.

[7] A. Khoda Bakhshi, S. M. Gaweesh, and M. M. Ahmed, "The Safety Performance of Connected Vehicles on Slippery Horizontal Curves through Enhancing Truck Drivers' Situational Awareness: A Driving Simulator Experiment," *Transp. Res. part F traffic Psychol. Behav.*, vol. 79, pp. 118–138, 2021.

[8] A. Khoda Bakhshi and M. M. Ahmed, "Bayesian Extreme Value Analysis of Kinematic-Based Surrogate Measure of Safety to Detect Crash-Prone Conditions in Connected Vehicles Environment: A Driving Simulator Experiment," *Transp. Res. Part C Emerg. Technol. (In Press.*, 2021.

[9] M. Hossain, M. Abdel-Aty, M. A. Quddus, Y. Muromachi, and S. N. Sadeek, "Real-time crash prediction models: State-of-the-art, design pathways and ubiquitous requirements," *Accid. Anal. Prev.*, vol. 124, pp. 66–84, 2019.

[10] M. M. Ahmed and M. A. Abdel-Aty, "The viability of using automatic vehicle identification data for real-time crash prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 459–468, 2011.

[11] M. M. Ahmed, M. Abdel-Aty, and R. Yu, "Assessment of interaction of crash occurrence, mountainous freeway geometry, real-time weather, and traffic data," *Transp. Res. Rec.*, vol. 2280, no. 1, pp. 51–59, 2012.

[12] M. Ahmed and M. Abdel-Aty, "A data fusion framework for real-time risk assessment on freeways," *Transp. Res. Part C Emerg. Technol.*, vol. 26, pp. 203–213, 2013.

[13] G. Louppe, L. Wehenkel, A. Sutera, and P. Geurts, “Understanding variable importances in forests of randomized trees,” in *Advances in neural information processing systems*, 2013, pp. 431–439.

[14] M. N. Wright and A. Ziegler, “Ranger: A fast implementation of random forests for high dimensional data in C++ and R,” *arXiv Prepr. arXiv1508.04409*, 2015.

[15] C. Strobl, A.-L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis, “Conditional variable importance for random forests,” *BMC Bioinformatics*, vol. 9, no. 1, p. 307, 2008.

[16] B. Gregorutti, B. Michel, and P. Saint-Pierre, “Correlation and variable importance in random forests,” *Stat. Comput.*, vol. 27, no. 3, pp. 659–678, 2017.

[17] S. Nembrini, I. R. König, and M. N. Wright, “The revival of the Gini importance?” *Bioinformatics*, vol. 34, no. 21, pp. 3711–3718, 2018.

[18] A. P. White and W. Z. Liu, “Bias in information-based measures in decision tree induction,” *Mach. Learn.*, vol. 15, no. 3, pp. 321–329, 1994.

[19] M. N. Wright, S. Wager, P. Probst, and M. M. N. Wright, “Package ‘ranger,’” 2019.

[20] Y. Lao, G. Zhang, Y. Wang, and J. Milton, “Generalized nonlinear models for rear-end crash risk analysis,” *Accid. Anal. Prev.*, vol. 62, pp. 9–16, 2014.

[21] Y. Xie and Y. Zhang, “Crash frequency analysis with generalized additive models,” *Transp. Res. Rec.*, vol. 2061, no. 1, pp. 39–45, 2008.

[22] X. Li, D. Lord, and Y. Zhang, “Development of accident modification factors for rural frontage road segments in Texas using generalized additive models,” *J. Transp. Eng.*, vol. 137, no. 1, pp. 74–83, 2010.

[23] T. J. Hastie, “Generalized additive models,” in *Statistical Models in S*, Routledge, 2017, pp. 249–307.

[24] K. Jones and N. Wrigley, “Generalized additive models, graphical diagnostics, and logistic regression,” *Geogr. Anal.*, vol. 27, no. 1, pp. 1–18, 1995.

[25] K. Jones and S. Almond, “Moving out of the linear rut: the possibilities of generalized additive models,” *Trans. Inst. Br. Geogr.*, pp. 434–447, 1992.

[26] A. Khoda Bakhshi and M. M. Ahmed, “A Note on Random Forest Visualization Tools in Post-Hoc Interpretation of Nonparametric Real-Time Risk Assessment Models,” in *Transportation Research Board 100th Annual Meeting*, 2021.

[27] A. Khoda Bakhshi and M. M. Ahmed, “Utilizing Black-Box Visualization Tools to Interpret Non-Parametric Real-Time Risk Assessment Models,” *Transp. A Transp. Sci.*, vol. 17, no. 4, pp. 739–765, 2020.

[28] M. Du, N. Liu, and X. Hu, “Techniques for interpretable machine learning,” *Commun. ACM*, vol. 63, no. 1, pp. 68–77, 2019.

[29] F. Yang, M. Du, and X. Hu, “Evaluating explanation without ground truth in interpretable machine learning,” *arXiv Prepr. arXiv1907.06831*, 2019.

[30] J. H. Friedman, “Greedy function approximation: a gradient boosting machine,” *Ann. Stat.*, pp. 1189–1232, 2001.

[31] C. Molnar, *Interpretable Machine Learning*. Leanpib, 2020.

[32] B. M. Greenwell, “pdp: An R Package for Constructing Partial Dependence Plots,” *R J.*, vol. 9, no. 1, pp. 421–436, 2017.

[33] A. Goldstein, A. Kapelner, J. Bleich, and E. Pitkin, “Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation,” *J. Comput. Graph. Stat.*, vol. 24, no. 1, pp. 44–65, 2015.

[34] D. W. Apley, “Visualizing the effects of predictor variables in black box supervised learning models,” *arXiv Prepr. arXiv1612.08468*, 2016.

[35] A. Khoda Bakhshi and M. M. Ahmed, “Application of Cross-Classified Random-Effects Modeling to Account for Unobserved Heterogeneity in Real-Time Clustering Critical versus Non-Critical Crashes,” in *ASCE, International Conference on Transportation & Development*, 2021.

[36] A. Khoda Bakhshi and M. M. Ahmed, “Does Random Slope Hierarchical Modeling Always Outperform Random Intercept Counterpart? Accounting for Unobserved Heterogeneity in a Real-Time Empirical Analysis of Critical Crash Occurrence,” *J. Transp. Saf. Secur. (In Press.*, 2022.

[37] F. L. Mannering and C. R. Bhat, “Analytic methods in accident research: Methodological frontier and future directions,” *Anal. methods Accid. Res.*, vol. 1, pp. 1–22, 2014.

[38] A. Khoda Bakhshi, M. M. Ahmed, and E. Adomah, “Accounting for Human-Related Unobserved Heterogeneity in the Safety Performance of Connected Vehicles Work Zone Warning Application,” *IATSS Res. (In Press.*, 2021.

[39] A. Khoda Bakhshi and M. M. Ahmed, “Practical advantage of crossed random intercepts under Bayesian hierarchical modeling to tackle unobserved heterogeneity in clustering critical versus non-critical crashes,” *Accid. Anal. Prev.*, vol. 149, p. 105855, 2021.

[40] L. M. Collins and L. A. Seitz, *Advances in Data Analysis for Prevention Intervention Research*. US Department of Health and Human Services, Public Health Service, National Institutes of Health, 1994.

[41] A. Khoda Bakhshi and M. M. Ahmed, "Coping with Endogeneity and Unobserved Heterogeneity in Real-Time Clustering Critical Crash Occurrences Nested Within Weather and Road Surface Conditions," *Int. J. Inj. Contr. Saf. Promot.*, vol. 28, no. 2, pp. 208–221, 2021.

[42] L. M. O'Dwyer and C. E. Parker, "A primer for analyzing nested data: multilevel modeling in SPSS using an example from a REL study," 2014.

[43] H. Quené and H. Van den Bergh, "Examples of mixed-effects modeling with crossed random effects and with binomial data," *J. Mem. Lang.*, vol. 59, no. 4, pp. 413–425, 2008.

[44] S.-J. Cho and S. Rabe-Hesketh, "Alternating imputation posterior estimation of models with crossed random effects," *Comput. Stat. Data Anal.*, vol. 55, no. 1, pp. 12–25, 2011.

[45] R. H. Baayen, D. J. Davidson, and D. M. Bates, "Mixed-effects modeling with crossed random effects for subjects and items," *J. Mem. Lang.*, vol. 59, no. 4, pp. 390–412, 2008.

[46] H. Goldstein, *Multilevel Statistical Models*, vol. 922. John Wiley & Sons, 2011.

[47] Y. Shi, W. Leite, and J. Algina, "The impact of omitting the interaction between crossed factors in cross-classified random effects modelling," *Br. J. Math. Stat. Psychol.*, vol. 63, no. 1, pp. 1–15, 2010.

[48] G. D. Garson, "Fundamentals of hierarchical linear and multilevel modeling," *Hierarchical linear Model. Guid. Appl.*, pp. 3–25, 2013.

[49] A. Fielding and H. Goldstein, "Cross-classified and multiple membership structures in multilevel models: An introduction and review," DfES, 2006.

[50] N. Sommet and D. Morselli, "Keep calm and learn multilevel logistic modeling: A simplified three-step procedure using Stata, R, Mplus, and SPSS," *Int. Rev. Soc. Psychol.*, vol. 30, no. 1, 2017.

[51] M. D. Hoffman and A. Gelman, "The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo.," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1593–1623, 2014.

[52] M. M. Ahmed, G. Yang, and S. Gaweesh, "Assessment of Drivers' Perceptions of Connected Vehicle–Human Machine Interface for Driving Under Adverse Weather Conditions: Preliminary Findings From Wyoming," *Front. Psychol.*, vol. 11, 2020.

[53] G. D. Israel, "Determination of sample size," Florida, 1992.

[54] Ahmed, S. Gaweesh, and G. Yang, "A Preliminary Investigation into the Impact of Connected Vehicle Human-Machine Interface on Driving Behavior," *IFAC-PapersOnLine*, vol. 51, no. 34, pp. 227–229, 2019.

[55] M. Ahmed, G. Yang, S. Gaweesh, R. Young, and F. Kitchener, "Performance evaluation framework of Wyoming connected vehicle pilot deployment program: summary of Phase 2 pre-deployment efforts and lessons learned," *J. Intell. Connect. Veh.*, 2019.

[56] Ahmed, G. Yang, and S. Gaweesh, "Development and assessment of a connected vehicle training program for truck drivers," *Transp. Res. Rec.*, vol. 2673, no. 2, pp. 113–126, 2019.

[57] O. Raddaoui and M. M. Ahmed, "Evaluating the Effects of Connected Vehicle Weather and Work Zone Warnings on Truck Drivers' Workload and Distraction using Eye Glance Behavior," *Transp. Res. Rec.*, vol. 2674, no. 3, pp. 293–304, 2020.

[58] G. Yang, M. M. Ahmed, and B. Subedi, "Distraction of Connected Vehicle Human–Machine Interface for Truck Drivers," *Transp. Res. Rec.*, vol. 2674, no. 9, pp. 438–449, 2020.

[59] S. F. Eftekharzadeh and A. Khodabakhshi, "Safety evaluation of highway geometric design criteria in horizontal curves at downgrades," *Int. J. Civ. Eng.*, vol. 12, no. 3, pp. 326–332, 2014.

[60] NHTSA, "Fatality Analysis Reporting System (FARS) Encyclopedia: 2014-2016," 2019. .

[61] Y. Wu, M. Abdel-Aty, J. Park, and J. Zhu, "Effects of crash warning systems on rear-end crash avoidance behavior under fog conditions," *Transp. Res. part C Emerg. Technol.*, vol. 95, pp. 481–492, 2018.

[62] M. Shakouri, L. H. Ikuma, F. Aghazadeh, K. Punniaraj, and S. Ishak, "Effects of work zone configurations and traffic density on performance variables and subjective workload," *Accid. Anal. Prev.*, vol. 71, pp. 166–176, 2014.

[63] E. Adomah, A. Khoda Bakhshi, and M. M. Ahmed, "Safety Impact of Connected Vehicles on Driver Behavior in Rural Work Zones under Foggy Weather Conditions," *Transp. Res. Rec.*, 2021.

[64] N. L. Jehn and R. E. Turochy, "Calibration of Vissim Models for Rural Freeway Lane Closures: Novel Approach to the Modification of Key Parameters," *Transp. Res. Rec.*, vol. 2673, no. 5, pp. 574–583, Apr. 2019.

[65] W. Zhizhou, S. Jian, and Y. Xiaoguang, "Calibration of VISSIM for shanghai expressway using genetic algorithm," in *Proceedings of the Winter Simulation Conference, 2005.*, 2005, pp. 4-pp.

[66] P. Holm, D. Tomich, J. Sloboden, and C. F. Lowrance, "Traffic Analysis Toolbox Volume IV: Guidelines for Applying Corsim Microsimulation Modeling Software," United States. Department of Transportation. Intelligent Transportation …, 2007.

[67] American Association of State Highway and Transportation Officials (AASHTO), *A Policy on Geometric Design of Highways and Streets*. Washington, D.C., 2004.

[68] A. Khoda Bakhshi, E. Adomah, and M. M. Ahmed, “Accounting for Human-Related Unobserved Heterogeneity to Enhance the Safety Performance of Work Zone Warning Application in Connected Vehicles Environment,” in *ASCE, International Conference on Transportation & Development*, 2021.

[69] J. Liu and A. J. Khattak, “Delivering improved alerts, warnings, and control assistance using basic safety messages transmitted between connected vehicles,” *Transp. Res. part C Emerg. Technol.*, vol. 68, pp. 83–100, 2016.

[70] S. M. Gaweesh, A. Khoda Bakhshi, and M. M. Ahmed, “Safety Performance Assessment of Connected Vehicles in Mitigating the Risk of Secondary Crashes: A Driving Simulator Study,” in *Transportation Research Board 100th Annual Meeting*, 2021.

[71] FHWA, *Manual on Uniform Traffic Control Devices for Streets and Highways*. 2009.

[72] A. Khoda Bakhshi and M. M. Ahmed, “The Effect of Work Zone Warning Delivery Distances on Connected Vehicles Safety Performance: A Conflation of Microsimulation Modeling and Hierarchical Negative Binomial,” in *Transportation Research Board 101st Annual Meeting*, 2022.

[73] G. Shieh, “Choosing the best index for the average score intraclass correlation coefficient,” *Behav. Res. Methods*, vol. 48, no. 3, pp. 994–1003, 2016.

[74] J. L. Crunkleton and P. J. Tarnoff, “Statistical Patterns of Traffic Data and Sample Size Estimation,” 2015.

[75] AASHTO Strategic Highway Safety, “Volume 7: A Guide for Reducing Collisions on Horizontal Curves,” 2004.

[76] Federal Motor Carrier Safety Administration, “Large Truck and Bus Crash Facts 2017,” 2019.

[77] Wyoming Department of Transportation, “Master Plan Implementation Report, I-80 Corridor Study,” 2018.

[78] S. M. Gaweesh and M. M. Ahmed, “Evaluating the safety effectiveness of a weather-based variable speed limit for a rural mountainous freeway in Wyoming,” *J. Transp. Saf. Secur.*, pp. 1–26, 2019.

[79] A. Khoda Bakhshi, S. M. Gaweesh, and M. M. Ahmed, “Mitigating the Risk of Horizontal Curve-Related Crashes via Altering Drivers’ Behaviors in Connected Vehicles Environment,” in *ASCE, International Conference on Transportation & Development*, 2021.

[80] K. Fitzpatrick, P. Carlson, M. Brewer, and M. D. Wooldridge, “Design speed, operating speed, and posted speed limit practices,” in *82nd Annual Meeting of the Transportation Research Board, Washington, DC*, 2003.

[81] R. Lamm, B. Psarianos, and T. Mailaender, *Highway Design and Traffic Safety Engineering Handbook*. 1999.

[82] C. Druta, A. Kassing, R. Gibbons, and V. A. Alden, "Assessing driver behavior using shrp2 adverse weather data," *J. Safety Res.*, vol. 73, pp. 283–295, 2020.

[83] A. Morgan and F. L. Mannering, "The effects of road-surface conditions, age, and gender on driver-injury severities," *Accid. Anal. Prev.*, vol. 43, no. 5, pp. 1852–1863, 2011.

[84] I. C. A. Oyeka and G. U. Ebuh, "Modified Wilcoxon signed-rank test," *Open J. Stat.*, vol. 2, no. 2, pp. 172–176, 2012.

[85] B. Rosner, R. J. Glynn, and M. T. Lee, "The Wilcoxon signed rank test for paired comparisons of clustered data," *Biometrics*, vol. 62, no. 1, pp. 185–192, 2006.

[86] F. Wilcoxon, S. K. Katti, and R. A. Wilcox, "Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test," *Sel. tables Math. Stat.*, vol. 1, pp. 171–259, 1970.

[87] S. M. Taheri and G. Hesamian, "A generalization of the Wilcoxon signed-rank test and its applications," *Stat. Pap.*, vol. 54, no. 2, pp. 457–470, 2013.

[88] C. J. Wild and M. Pfannkuch, "Statistical thinking in empirical enquiry," *Int. Stat. Rev.*, vol. 67, no. 3, pp. 223–248, 1999.

[89] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median," *J. Exp. Soc. Psychol.*, vol. 49, no. 4, pp. 764–766, 2013.

[90] N. Chung et al., "Median absolute deviation to improve hit selection for genome-scale RNAi screens," *J. Biomol. Screen.*, vol. 13, no. 2, pp. 149–158, 2008.

[91] T. Pham-Gia and T. L. Hung, "The mean and median absolute deviations," *Math. Comput. Model.*, vol. 34, no. 7–8, pp. 921–936, 2001.

[92] E. N. Mazzae, G. H. S. Baldwin, and D. V McGehee, "Driver crash avoidance behavior with ABS in an intersection incursion scenario on the Iowa driving simulator," *SAE Trans.*, pp. 2353–2363, 1999.

[93] R. J. Kiefer, D. J. LeBlanc, and C. A. Flannagan, "Developing an inverse time-to-collision crash alert timing approach based on drivers' last-second braking and steering judgments," *Accid. Anal. Prev.*, vol. 37, no. 2, pp. 295–303, 2005.

[94] J. Lee and K. Jang, "A framework for evaluating aggressive driving behaviors based on in-vehicle driving records," *Transp. Res. part F traffic Psychol. Behav.*, vol. 65, pp. 610–619, 2019.

[95] A. Das, M. M. Ahmed, A. Ghasemzadeh, M. N. Khan, A. Ghasemzadeh, and M. M. Ahmed, "Using trajectory-level SHRP2 naturalistic driving data for investigating driver lane-keeping ability in fog: An association rules mining approach," *Accid. Anal. Prev.*, vol. 129, no. 16, pp. 250–262, Jun. 2019.

[96] X. Yan, X. Li, Y. Liu, and J. Zhao, "Effects of foggy conditions on drivers' speed control behaviors at different risk levels," *Saf. Sci.*, vol. 68, pp. 275–287, 2014.

[97] A. Khoda Bakhshi and M. M. Ahmed, "Trajectory-Level Analysis of Truck Drivers' Behavioral Alteration in Connected Vehicles Environment under Fog with Complex Roadway Geometry: A Driving Simulator Study," in *Transportation Research Board 101st Annual Meeting*, 2022.

[98] D. Gonzales and M. D. Fontaine, "Impacts of the I-77 Variable Speed Limit System on Speed and Crash Characteristics During Low Visibility Conditions," Virginia Transportation Research Council, 2018.

[99] G. Yang, M. M. Ahmed, and S. Gaweesh, "Impact of variable speed limit in a connected vehicle environment on truck driver behavior under adverse weather conditions: driving simulator study," *Transp. Res. Rec.*, vol. 2673, no. 7, pp. 132–142, 2019.

[100] A. Khoda Bakhshi and M. M. Ahmed, "Real-Time Crash Prediction for a Long Low-Traffic Volume Corridor Using Corrected-Impurity Importance and Semi-Parametric Generalized Additive Model," *J. Transp. Saf. Secur.*, 2021.

[101] R. N. Jazar, *Vehicle Dynamics: Theory and Application*. Springer, 2017.

[102] M. Motamed, "Developing a real-time freeway incident detection model using machine learning techniques." 2016.

[103] A. Hosseinzadeh and R. Kluger, "Do EMS times associate with injury severity?" *Accid. Anal. Prev.*, vol. 153, p. 106053, 2021.

[104] A. Hosseinzadeh, A. Karimpour, R. Kluger, and R. Orthober, "A Framework to Link Crashes to Emergency Medical Service Runs and Trauma Admissions for Improved Highway Safety Monitoring and Crash Outcome Assessment," in *Transportation Research Board. 99th Annual Meeting Transportation Research Board*, 2020.

[105] A. Hosseinzadeh, M. Haghani, and R. Kluger, "Exploring Influencing Factors on Crash-related Emergency Response Time: A Machine Learning Approach," in *Transportation Research Board 100th Annual MeetingTransportation Research Board*, 2021.

[106] A. Hosseinzadeh and R. Kluger, "Data Linkage for Traffic Safety in Jefferson County, Kentucky," in *International Conference on Transportation and Development*, 2021, pp. 243–250.

[107] A. Khoda Bakhshi and M. M. Ahmed, “Crash Detection in Connected Vehicles Environment: Extreme Value Analysis of Trajectory-Level Observations from Driving Simulator Experiments,” in *Transportation Research Board 101st Annual Meeting*, 2022.

[108] S. M. Gaweesh, A. Khoda Bakhshi, and M. M. Ahmed, “Safety Performance Assessment of Connected Vehicles in Mitigating the Risk of Secondary Crashes: A Driving Simulator Study,” *Transp. Res. Rec.*, 2021.

[109] T. Huang, S. Wang, and A. Sharma, “Highway crash detection and risk estimation using deep learning,” *Accid. Anal. Prev.*, vol. 135, p. 105392, 2020.

[110] D. Henclewood, M. Abramovich, and B. Yelchuru, “Safety pilot model deployment–one day sample data environment data handbook,” *Res. Technol. Innov. Adm. Res. Technol. Innov. Adm. US Dep. Transp. McLean, VA*, 2014.

[111] F. Feng, S. Bao, J. R. Sayer, C. Flannagan, M. Manser, and R. Wunderlich, “Can vehicle longitudinal jerk be used to identify aggressive drivers? An examination using naturalistic driving data,” *Accid. Anal. Prev.*, vol. 104, pp. 125–136, 2017.

[112] E. Mannshardt, P. F. Craigmile, and M. P. Tingley, “Statistical modeling of extreme value behavior in North American tree-ring density series,” *Clim. Change*, vol. 117, no. 4, pp. 843–858, 2013.

[113] L. Zheng, K. Ismail, T. Sayed, and T. Fatema, “Bivariate extreme value modeling for road safety estimation,” *Accid. Anal. Prev.*, vol. 120, pp. 83–91, 2018.

[114] H. Farah and C. L. Azevedo, “Safety analysis of passing maneuvers using extreme value theory,” *IATSS Res.*, vol. 41, no. 1, pp. 12–21, 2017.

[115] P. Songchitruksa and A. P. Tarko, “The extreme value theory approach to safety estimation,” *Accid. Anal. Prev.*, vol. 38, no. 4, pp. 811–822, 2006.

[116] E. Gilleland and R. W. Katz, “extRemes 2.0: An extreme value analysis package in R,” *J. Stat. Softw.*, vol. 72, no. 1, pp. 1–39, 2016.

[117] A. Gelman, J. Hwang, and A. Vehtari, “Understanding predictive information criteria for Bayesian models,” *Stat. Comput.*, vol. 24, no. 6, pp. 997–1016, 2014.

[118] G. Evin, C. Thomas, and N. Eckert, “Has fire policy decreased the return period of the largest wildfire events in France? A Bayesian assessment based on extreme value theory,” *Nat. Hazards Earth Syst. Sci.*, vol. 18, no. 10, pp. 2641–2651, 2018.

[119] M. E. Glickman and D. A. Van Dyk, “Basic Bayesian Methods,” *Top. Biostat.*, pp. 319–338, 2007.

[120] B. C. T. Cabella, M. J. Sturzbecher, W. Tedeschi, O. Baffa Filho, D. B. de Araújo, and U. P. da C. Neves, “A numerical study of the Kullback-Leibler distance in functional magnetic resonance imaging,” *Brazilian J. Phys.*, vol. 38, no. 1, pp. 20–25, 2008.

[121] E. A. Freeman and G. G. Moisen, “A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa,” *Ecol. Modell.*, vol. 217, no. 1–2, pp. 48–58, 2008.