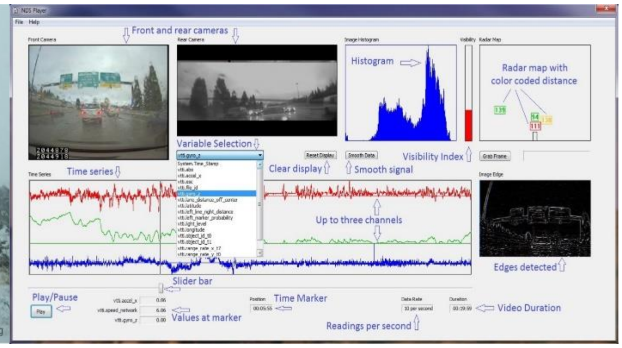
FINAL REPORT
WY-18/05F

SHRP2 Implementation Assistance Program (IAP)—Round 4
Concept to Countermeasures—Research to Deployment Using the SHRP2 Safety Data



DRIVER PERFORMANCE AND BEHAVIOR IN ADVERSE WEATHER CONDITIONS: AN INVESTIGATION USING THE SHRP2 NATURALISTIC DRIVING STUDY DATA—PHASE 2

Final Report
March 2018

Principal Investigator
Mohamed M. Ahmed, Ph.D., P.E.
Assistant Professor

Authors
Mohamed M. Ahmed, Ph.D., P.E.
Ali Ghasemzadeh, Ph.D., Post-doctoral Research Associate
Britton Hammit, Ph.D. Candidate
Nasim Khan, MS Student
Anik Das, MS Student
Elhashemi Ali, Ph.D. Candidate
Rhonda K. Young, Ph.D., P.E.
Hesham Eldeeb, Ph.D. (HIS)

Sponsored by
Wyoming Department of Transportation
(Grant No: RS07216)
From
Department of Civil and Architectural Engineering
University of Wyoming
1000 E. University Avenue, Department 3295
Laramie, Wyoming 82071
Telephone: (307) 766-5550
Fax: (307) 766-2221
E-mail: mahmed@uwyo.edu

U.S. Department of Transportation
Federal Highway Administration

American Association of State Highway and Transportation Officials
AASHTO
The Voice of Transportation

Transportation Research Board
of the National Academies
TRB

FOREWORD

The objective of the second phase of the Second Strategic Highway Research Program (SHRP2) Implementation Assistance Program (IAP) was to conduct a thorough analysis using a larger set of Naturalistic Driving Study trips to extract behavioral trends specific to a wide variety of weather conditions on freeway. These weather conditions included rain, snow, and fog from a diverse driver population from each of the six SHRP2 data collection sites. The objective of Phase 3 of this study is to interpret these findings such that they can be used to inform the development of Wyoming-based safety and reliability countermeasures.

Phase 3 of begins in early 2018 and will conclude in 2019. The solid foundation generated in the first two project phases will be used to enhance the existing weather-dependent VSL system operated by WYDOT. Specifically, the speed selection models will be validated using available data from Wyoming interstates to develop a suitable algorithm for VSL operation. The car-following, acceleration, lane-changing, lane-wandering, and safety critical event analyses will be used to develop weather-related microsimulation model guidance that could be used to evaluate future countermeasures. This report will be available online.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. WY-18/05F	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Driver Performance and Behavior in Adverse Weather Conditions: An Investigation Using the SHRP2 Naturalistic Driving Study Data—Phase 2		5. Report Date February, 2018 6. Performing Organization Code
7. Author(s) Mohamed Ahmed, Ph.D., P.E. (orcid.org/0000-0002-1921-0724) Ali Ghasemzadeh, Ph.D. (orcid.org/0000-0003-1232-251X) Britton Hammit, Ph.D. Candidate (orcid.org/0000-0002-8859-9075) Nasim Khan, MS Student (orcid.org/0000-0001-5996-091X) Anik Das, MS Student (orcid.org/0000-0003-4674-5334) Elhashemi Ali, Ph.D. Candidate (orcid.org/0000-0001-7643-7143) Rhonda K. Young, Ph.D., P.E. (orcid.org/0000-0001-6745-5008) Hesham Eldeeb, Ph.D. (HIS) (orcid.org/0000-0001-9669-5991)		8. Performing Organization Report No.
9. Performing Organization Name and Address Department of Civil & Architectural Engineering University of Wyoming 1000 E. University Avenue, Dept. 3295 Laramie, Wyoming 82071		10. Work Unit No. (TRAIS) 11. Contract or Grant No. RS07216
12. Sponsoring Organization Name and Address Wyoming Department of Transportation 5300 Bishop Blvd, Bldg 6100 Cheyenne, WY 82009-3340		13. Type of Report and Period Covered Final Report February 2018 14. Sponsoring Agency Code FHWA – SHRP2 - WYDOT
15. Supplementary Notes		
16. Abstract The impact of adverse weather conditions on transportation operation and safety is the focus of many studies; however, comprehensive research detailing the differences in driving behavior and performance during adverse conditions is limited. Many previous studies utilized aggregate traffic and weather data (e.g., average speed, headway, and global weather information) to formulate conclusions about the weather impact on network operation and safety; therefore, research into specific factors associated with driver performance and behavior are notably absent. A novel approach, presented in this report, can fill the gap in previous studies by considering disaggregate trajectory-level data available through the SHRP2 Naturalistic Driving Study and Roadway Information Database. Parametric and non-parametric models were utilized to better understand different behavioral factors including speed selection, car following and lane keeping in adverse weather conditions. The purpose of this study was to gather insights into driver behavior and performance in different weather conditions, such that efficient logic can be implemented to introduce a realistic Variable Speed Limit system, aimed at maximizing speed compliance and reducing speed variations. This study provides valuable information related to drivers' interaction with real-time changes in roadway and weather conditions, leading to a better understanding of the effectiveness of operational countermeasures.		
17. Key Words Driver Behavior and Performance, Variable Speed Limit, Adverse Weather Conditions Speed Selection, Lane Keeping, Naturalistic Driving Study, Ordinal Logistic Regression, Classification Tree, CART, SHRP2, Wyoming		18. Distribution Statement This document is available through the National Transportation Library and the Wyoming State Library.
19. Security Classification (of this report) Non-Classified	20. Security Classification (of this page) Non-Classified	21. No. of Pages 123 — 22. Price

Form DOT F 1700.7 (8-72) Reproduction of form and completed page is authorized

METRIC CONVERSION FACTORS

SI* (MODERN METRIC) CONVERSION FACTORS				
APPROXIMATE CONVERSIONS TO SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa
APPROXIMATE CONVERSIONS FROM SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.
(Revised March 2003)

EXECUTIVE SUMMARY

Inclement weather conditions such as fog, snow, ground blizzards, slush, rain, and strong winds negatively affect pavement condition, vehicle performance, visibility, and driver behavior and performance. Driver behavior exhibits high variability and is difficult to quantify, particularly in inclement weather conditions. Driver behavior and performance are imperative to understand when describing the influence of adverse weather conditions on roadways' safety and mobility. Adverse weather inhibits driver's ability to perceive their environment, and visibility reductions – caused by adverse weather events – is known to increase the likelihood of crashes. The effects of adverse weather on safe and efficient operations of transportation networks have been extensively researched; however, specific considerations of driver behavior and performance are noticeably absent from these studies.

The Second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) and Roadway Information Database (RID) provide an unprecedented opportunity for researchers to better understand driving behavior. Having identified the potential benefits and value of this unique dataset, this report aimed to evaluate methods by which the data could be used to study driver behavior and performance characteristics during adverse weather conditions.

This report addresses different gaps in the knowledge by presenting innovative methods to identify and analyze weather-related naturalistic driving data to better understand driver behavior and performance in adverse weather conditions. An innovative methodology to effectively identify weather-related trips in real-time using vehicle wiper status and other complementary methodologies helped to identify naturalistic driving weather-related trips using external weather data sources. In addition, a semi-automated data reduction procedure was developed to process raw trip data files into a format that further analyses and modeling techniques could be easily applied. The novel approaches developed in this report for NDS trip acquisition and reduction could be extended to other naturalistic driving studies worldwide.

In addition to the contributions in data extraction and reduction, preliminary analysis as well as advanced modeling techniques were utilized in this study. These analyses were used to explain the relationship between different levels of speed selection and lane keeping behaviors and a set of contributing factors including roadway characteristics, environmental and traffic conditions and driver demographics on a trajectory level. These modeling techniques ranged from common parametric approaches such as binary logistic regression and ordinal logistic/probit regression models to a more advanced non-parametric/data mining modeling techniques such as Classification and Regression Trees (CART) and Multivariate Adaptive Regression Splines (MARS).

The results from this study suggest that both parametric and non-parametric modeling approaches are important to analyze driver behavior and performance. In fact, this study attempted to maximize the benefits of the advantages of parametric models, such as the ability of interpreting the marginal effects of various risk factors, as well as the advantages of using non-parametric models, including but not limited to the ability of providing high prediction accuracy, handling of missing values automatically, and their capability of handling large number of explanatory variables in a timely manner, which might be extremely beneficial specifically for assessing traffic operations and safety in real-time considering weather and traffic data to be

directly fed into the model. The specific data mining modeling techniques used in this study have an additional advantage compared to most of the data mining and machine learning techniques. Unlike many other non-parametric models, CART and MARS models are interpretable and transparent; therefore, these models do not have the “black box” problem known for most other machine learning techniques.

The results of the developed speed selection models revealed that among various adverse weather conditions, drivers were more likely to reduce their speed in snowy weather conditions compared to other adverse weather conditions. Specifically, the odds of drivers reducing their speeds were 9.29 times higher in snowy weather conditions, followed by rain and fog with 1.55 and 1.29 times compared to clear conditions, respectively. In addition, variable importance analysis using CART method revealed that weather conditions, traffic conditions, and posted speed limits are the three most important variables affecting driver speed selection behavior. Moreover, the results of the developed lane-keeping models revealed that drivers in heavy rain conditions were more likely to have a worse lane-keeping performance compared to clear weather conditions.

The developed speed selection model is a key example of a derived mechanism by which the SHRP2 database can be leveraged to improve Weather Responsive Traffic Management (WRTM) strategies directly. Moreover, the results may shed some light on driver lane keeping behavior at a trajectory level. A better understanding of driver lane-keeping behavior might help in developing better Lane Departure Warning (LDW) systems. Evaluating driver behavior and performance under the influence of reduced visibility due to adverse weather conditions is extremely important to develop safe driving strategies, including Variable Speed Limits (VSL). Many roadways across the U.S. currently have weather-based VSL systems to ensure safe driving environments during adverse weather. Current VSL systems mainly collect traffic information from external sources, including inductive loop detector, overhead radars and Closed Circuit Television (CCTV). However, human factors especially driver behavior and performance such as selection of speed and acceleration/ deceleration behaviors during adverse weather are neglected due to the lack of appropriate driver data. The findings from this study indicated that the SHRP2 NDS data could be effectively utilized to identify trips in adverse weather conditions and to assess the impacts of adverse weather on driver behavior and performance. With the evolution of Connected Vehicles, Machine Vision and other real-time weather social crowd sources such as WeatherCloud®, more accurate real-time data similar to the NDS data will be available in the near future. This study provided early insights into using similar data collected from NDS.

Table of Contents

List of Figures

List of Tables

List of Equations

LIST OF ACRONYMS/ABBREVIATIONS

Abbreviation	Description
AADT	Average Annual Daily Traffic
AASHTO	American Association of State Highway and Transportation Officials
ADAS	Advanced Driver Assistance Systems
ANN	Artificial Neural Network
ATIS	Advanced Traveler Information Systems
BFs	Basis Functions
BNs	Bayesian Networks
CAN	Controller Area Network
CART	Classification and Regression Tree
CAV	Connected and Automated Vehicles
CCTV	Closed Circuit Television
CDO	Climate Data Online
CNDD	Commercially collected NDD
CV	Connected Vehicle
DAS	Data Acquisition Systems
DTS	Dynamic Traffic Status
FARS	Fatality Analysis Reporting System
FFS	Free Flow Speed
FHWA	Federal Highway Administration
FOT	Field Operational Test
GIS	Geographic Information System
GPS	Global Positioning System
HCM	Highway Capacity Manual
IAP	Implementation Assistance Program
iLD	inductive Loop Detectors
ISS	Intelligent Safety Systems
ITS	Intelligent Transportation System
IV	In-vehicle
K-NN	K-Nearest Neighbours
LCC	Latent Class Cluster
LDW	Lane Departure Warning
LR	Logistic Regression
MARS	Multivariate Adaptive Regression Splines
ML	Maximum Likelihood
MOP	Measure of Performance
MUTCD	Manual on Uniform Traffic Control Device
NCDC	National Climatic Data Center
NDD	Naturalistic driving data
NDS	Naturalistic Driving Study
NFOT	Naturalistic field operational test

NHTSA	(US) National Highway Traffic Safety Administration
NOAA	National Oceanic and Atmospheric Administration
NTDS	Naturalistic Teenage Driving Study
OBU	On-Board Units
OPM	Ordered Probit Regression
PDO	Property Damage Only
PII	Personally Identifiable Information
RC	Road Surface Condition
RID	Roadway Information Database
RMSE	Root Mean Square Error
RTMS	Remote Traffic Microwave Sensors
RWIS	Road Weather Information System
SCE	Safety-critical Event
SDLP	Standard Deviation of Lane Position
SHRP2	Second Strategic Highway Research Program
SMoS	Surrogate Measures of Safety
SMOS	Surrogate Measures of Safety
TTC	Time to Collision
VDOT	Virginia Department of Transportation
VI	Visibility Index
VIF	Variance Inflation Factor
VIM	Variable Importance Measure
VL	Visibility Level
VMT	Vehicle Miles Travelled
VRT	Visualization and Reduction Tool
VSL	Variable Speed Limit
VTTI	Virginia Tech Transportation Institute
WAFs	Weather Adjustment Factors
WC	Weather Condition
WRTM	Weather Responsive Traffic Management
WYDOT	Wyoming Department of Transportation

CHAPTER 1-INTRODUCTION

Transportation safety, mobility, and reliability are heavily dependent on weather and roadway conditions. Adverse weather conditions affect the transportation network by reducing average travel speeds, increasing the frequency of crashes, and demanding a substantial portion of agency budgets required for weather-related maintenance. The network-wide impacts of adverse weather conditions are critical for the planning and operation of an agency's network during winter seasons; however, the low resolution data used in these assessments cannot directly detect specific adjustments in individual driver behaviors caused by adverse weather conditions. The underlying cause for decreased mobility during inclement weather events is adjustments of driving behavior in response to reduced visibility, poor pavement conditions, and limited vehicle's performance. Similarly, crash frequency increases during winter weather events because drivers may not adequately shift their behaviors to match current weather conditions *(1)*.

According to the Fatality Analysis Reporting System (FARS), inclement weather of rain, snow and fog/smoke resulted in 5,897 fatal crashes between 2005 and 2014. The NHTSA reported that weather contributed to over 22 percent of the total crashes between 2005 and 2014 *(2)*. In Canada and the UK, such crashes account for approximately 30 percent and 20 percent, respectively *(3, 4)*. The financial burden of weather-related crashes on the U.S. roadways is approximately 42 billion U.S. dollars *(5)*. The impact of vision obstruction caused by adverse weather conditions on traffic safety and operations has been investigated in previous studies. Brodsky and Hakkert, (1988) investigated the risk of traffic crashes in rainy weather conditions *(6)*. The study found that the risk of injury crashes in rainy weather conditions could be significantly 2 to 3 times greater than in clear weather conditions. Two main risk factors were associated with the added risk caused by adverse weather conditions in their study: slippery road surface and reduction in visibility. The study concluded that slippery surface conditions caused by adverse weather negatively affect driver performance, specifically on curves. In addition, the reduction in visibility during adverse weather increased the risk of crashes, which might be exacerbated at night due to distraction and glare produced by shining wet surfaces. A study by Andrey et al. (2003) investigated the impact of adverse weather on the severity of crashes in several Canadian cities *(7)*. They found a 75 percent increase in total crashes and a 45 percent increase in injury crashes in comparison with clear weather conditions. Other studies by Ahmed et al. (2012) reported that an additional one-inch increase in precipitation elevated the risk of a crash by 169 percent *(1)*. In addition, they stated that the added risk of crashes could be doubled in snowy seasons due to the interaction between geometrical characteristics such as steep grades and slippery surface conditions. Although there are some variations in estimated risk of crashes, most of the previous studies confirmed that risk of crashes could be easily elevated during adverse weather, due to affected visibility and surface conditions. Rahman and Lownes, (2012) found that drivers might reduce their speed, maintain a larger headway and drive more cautiously in adverse weather to compensate for reduced visibility and slippery road conditions *(8)*.

Drivers make decisions based on their risk perception. Drivers' decisions to compensate for crash risk can be categorized into three hierarchical levels including strategic, tactical, and operational *(9)*. The strategic level can be defined as those decisions that usually are made off-road and called "off-road decisions". Decisions at this level can be travel mode alternation, route change, trip timing change, etc. Tactical and operational decisions are made on the road. Particularly, these decisions are made in high-risk circumstances, including but not limited to speed adaptation, lane changing, headway selection, gap acceptance, and evasive maneuvers *(10)*. Each of the mentioned decisions stages might be affected by environmental conditions; in particular, adverse weather conditions. Driver attentiveness and control behavior are two important factors that might be negatively affected by adverse weather conditions *(11)*.

As mentioned earlier, previous research aimed to characterize the impacts of adverse weather conditions on the transportation network; however, few studies focused directly on the fundamental cause of network-wide impacts, driving behavior. Comprehensive analysis of driving behavior in adverse weather conditions requires high resolution data collected in a variety of adverse weather events by a multitude of drivers. Collection of this type of data is expensive and out of scope for most research queries.

In efforts to advance the state-of-research and provide data to researchers seeking to understand driving behavior for improving traffic safety and operations, the second Strategic Highway Research Program (SHRP2) generated the Naturalistic Driving Study (NDS) database. The SHRP2 NDS database is comprised of more than 5 Million trips from 3,400 drivers in six geographic regions across the United States *(12)*. In addition, a second database—Roadway Information Database (RID)—was constructed to provide context for NDS trips. The RID contains roadway, traffic operations, environmental, and other information corresponding to the most-travelled roadways traversed by the NDS participants *(13)*. The creation of these SHRP2 databases presents researchers an unprecedented opportunity to advance current understandings of driving behavior.

As part of the SHRP2 Implementation Assistance Program (IAP), the Wyoming Department of Transportation (WYDOT) established a project to investigate the impact of adverse weather conditions on driving behavior for the purpose of establishing practical countermeasures to improve the safety, efficiency, and reliability of the Wyoming transportation network during harsh winter seasons. This project is expected to produce an updated variable speed limit (VSL) algorithm for the existing weather-dependent VSL corridors in Wyoming. In addition, the increased understanding of driving behavior in adverse weather conditions is expected to improve the accuracy of weather-related microsimulation modeling.

The SHRP2 IAP is divided into three phases; this report presents the findings from the second project phase and introduces ongoing efforts that will be completed in the third project phase. The following sections provide an overview of the Wyoming IAP objectives and research questions, summarizes the previous findings from Phase 1, and outlines the remainder of the report.

Project Objectives

The primary goal of the Wyoming SHRP2 IAP project is to leverage the SHRP2 NDS and RID databases to enhance the understanding of how drivers respond to adverse weather and road conditions. The ultimate objective of this research is to use the findings to develop feasible countermeasures that WYDOT can implement on state interstates and highways to improve the reliability of the transportation network VSL systems during adverse weather conditions.

The original project proposal presented the following research questions:

1. Can NDS trips occurring in inclement weather be identified efficiently and effectively using available NDS and RID data?
2. Can driver behavior (e.g., speed selection, car-following, and lane wandering) during inclement weather conditions be characterized efficiently from the NDS data?
3. What are the best surrogate measures for weather-related crashes that can be identified using the NDS data?
4. What type of analysis can be performed and conclusions drawn from the resulting dataset?

After completion of the first project phase—proof of concept—the first research was proven possible and the remaining research questions were deemed feasible. These findings from Phase 1 initiated the start of Phase 2, which was intended to build on the concepts developed in Phase 1 and complete a full analysis on the SHRP2 NDS data. Before the start of Phase 2, the original research questions were fine-tuned and detailed with the knowledge derived from Phase 1. Specifically, additional questions were raised aiming to introduce additional methods for acquiring NDS trips that not only leveraged the SHRP2 data, but also used external weather data sources. In addition, specific statistical methods and the calibration of common driving behavior models were explored for the purpose of characterizing driver behavior differences in adverse weather conditions.

The objective of the second phase was to conduct a thorough analysis using a larger set of NDS trips to extract behavioral trends specific to a wide variety of weather conditions (i.e., rain, snow, and fog). The trips would represent a diverse driver population from each of the six SHRP2 data collection sites. Using experience from Phase 1, efficient data reduction procedures were implemented to process the trip data. Once processed, these data were used in the development and calibration of driver behavior models related to speed selection, car-following, and lane wandering. In addition, a larger number of crashes and near-crashes were collected and evaluated. Ultimately, the final research question—identification of possible analyses and feasible conclusions—was refined to target one specific countermeasure that will be the focus for Phase 3:

- Improvement to existing weather-dependent variable speed limit (VSL) control algorithm used by WYDOT for their interstate VSL systems.

In addition, the Wyoming research team identified parallels with the WYDOT Connected Vehicle (CV) Pilot project, which intends to establish protocols for and test CV applications relevant to rural locations with severe weather conditions and heavy freight traffic *(14)*. As part of the CV Pilot, the research team will use microsimulation modeling as a tool to evaluate the effectiveness of each CV application; therefore, a reliable base model representing driving behavior in adverse weather conditions is required. The SHRP2 IAP project team recognized the potential for collaboration and identified two additional Phase 3 objectives:

- Improved guidance related to microsimulation modeling of adverse weather conditions, and generation of a “base model” to represent driving behavior in adverse weather conditions for use in the Wyoming CV Pilot project impact assessments.

- Evaluation of SHRP2 NDS weather-related vehicle dynamics to support the development of real-time CV applications requiring weather and roadway condition input data.

Phase 1 Overview

Phase 1 of the Wyoming IAP project consisted of a proof-of-concept review of the SHRP2 NDS and RID data that was intended to evaluate the feasibility of answering the original research questions. Therefore, a small sample of NDS trips were queried from the SHRP2 NDS database, aiming to identify trips related to adverse weather conditions and matching trips occurring in clear weather conditions. Phase 1 data acquisition focused on precipitation events and queried trips from only two of the six SHRP2 sites: Washington and Florida. Trips occurring in precipitation were identified by tracking the windshield wiper status of the vehicles and extracting trips with active windshield wipers. In addition, a matching protocol was established in this phase to identify additional trips taken by the same driver on the same route in clear conditions.

Manual data reduction was performed in the first project phase to gain familiarity with the SHRP2 NDS data and suggest procedures for automating various elements of the process. The most time-consuming process involved manual video observation to classify weather conditions for each trip; therefore, the Wyoming research team began the development of the Wyoming NDS Visualization and Reduction software. This software provided an effective platform for viewing NDS data using a convenient graphical user interface and initialized efforts to detect visibility levels from the front video camera.

The preliminary analysis of driver behavior focused on selected speeds, acceleration, headways, and lateral lane position. As part of this analysis, behavior distributions in different classifications of adverse weather and traffic flow conditions were shown to be different. For example, in free flow conditions and heavy rainfall, driver speeds followed a Weibull distribution, while in free flow and clear weather conditions, driver speeds followed a Normal distribution. Additional findings suggested that speed reductions were statistically significant in heavy precipitation, and an increase in the speed variability during precipitation events was detected. Aggressive braking and acceleration events were evaluated and findings suggested that average deceleration was higher in clear weather conditions, when compared to matching adverse weather trips. In addition, lateral lane position was studied by evaluating drivers' ability to maintain their position in their lane and their tendencies to make lane changes. Results showed that the frequency of lane changes is higher in clear weather conditions, and in adverse conditions, the amount of lane-wandering (i.e., lateral movement within the travel lane) was increased. Lastly, driver headways were larger in heavy precipitation—compared to clear weather conditions—and the variability of headways decreased in heavy precipitation.

Another avenue of research aimed to maintain the continuity of a single driver, on a single day, during various weather conditions. Detailed evaluations of specific trips were conducted to analyze the behavior changes of an individual driver on a trip that contained series of weather conditions (e.g., the trip's weather condition was initially classified as light rain, in the middle changed to heavy rain, and at the end was classified as clear conditions). The findings from this analysis indicated the importance of segmenting each trip by weather condition to ensure accurate results.

In addition to evaluating NDS trips using summary statistics, preliminary modeling efforts were conducted to identify speed selection tendencies in different weather conditions. An ordered probit logit model was used to classify speed behavior as a function of traffic, speed limits, surface condition, and weather. Results from this model indicated that weather, speed limits, and traffic conditions were significant, while weather and traffic conditions played the largest role on determining drivers' speed selection. These analyses provided promising preliminary results, and introduced a series of new questions that were later evaluated in Phase 2 of this project.

Finally, a small sample of weather-related crash and near-crash events were analyzed to identify crash surrogate measures. Two vehicle dynamics variables were used as indicators for identifying a potential crash: acceleration/deceleration and yaw rate. Thresholds for these variables were identified from a review of the available crash and near-crash events. The findings from this analysis were used as a baseline for requesting a substantially larger set of weather-related crash and near-crash events from the SHRP2 NDS database for analysis in Phase 2. More information about the Phase 1 findings can be found in Phase 1 Final Report *(15)*.

Report Outline

The remainder of this report presents the findings from Phase 2 of the Wyoming IAP project. The report is organized as follows:

> **Chapter 2 NDS Trip Acquisition and Reduction** provides a detailed overview of the improved data acquisition and reduction protocols established for Phase 2.
>
> **Chapter 3 Research Findings** is divided into sections related to the different areas of driving behavior explored in Phase 2, and specific background information, additional data processing requirements, and findings are reported for each area.
>
> **Chapter 4 Conclusions and Plans for Phase 3** provides a summary of the findings presented in "Research Findings" and relates these findings to practical applications in Wyoming.

CHAPTER 2 - NDS TRIP ACQUISITION AND REDUCTION

The identification and acquisition of weather-related NDS trips was a pivotal objective in this project. As part of Phase 1, precipitation events were identified using the windshield wiper status; however, in order to gather a wider variety of weather conditions, new methodologies for querying the NDS database were required. Once new methods were established to acquire a substantially larger quantity of NDS trips, automated data reduction procedures were needed for efficient analysis.

Data Acquisition section describes the three complementary methodologies the project team developed to identify weather-related NDS trips, and Data Reduction section describes the semi-automated data reduction procedures developed for efficient processing of the acquired NDS trips. Finally, “Wyoming NDS Visualization and Reduction Tool” section discusses the improvements made to the Wyoming NDS Visualization and Reduction tool.

Data Acquisition

Acquisition of weather-related NDS trips is a complex procedure that requires the use of NDS and RID data elements, as well as a variety of additional data sources. In order to compile a comprehensive set of trips occurring in a variety of different weather conditions, three novel complementary methodologies were developed in parallel to extract a sufficient number of trips for detailed analytics. Each method is shown in Figure 1 and will be discussed in detail in the following sections. In addition, subsequent sections provide a summary of the acquired trips and an overview of the crash and near-crash events collected.

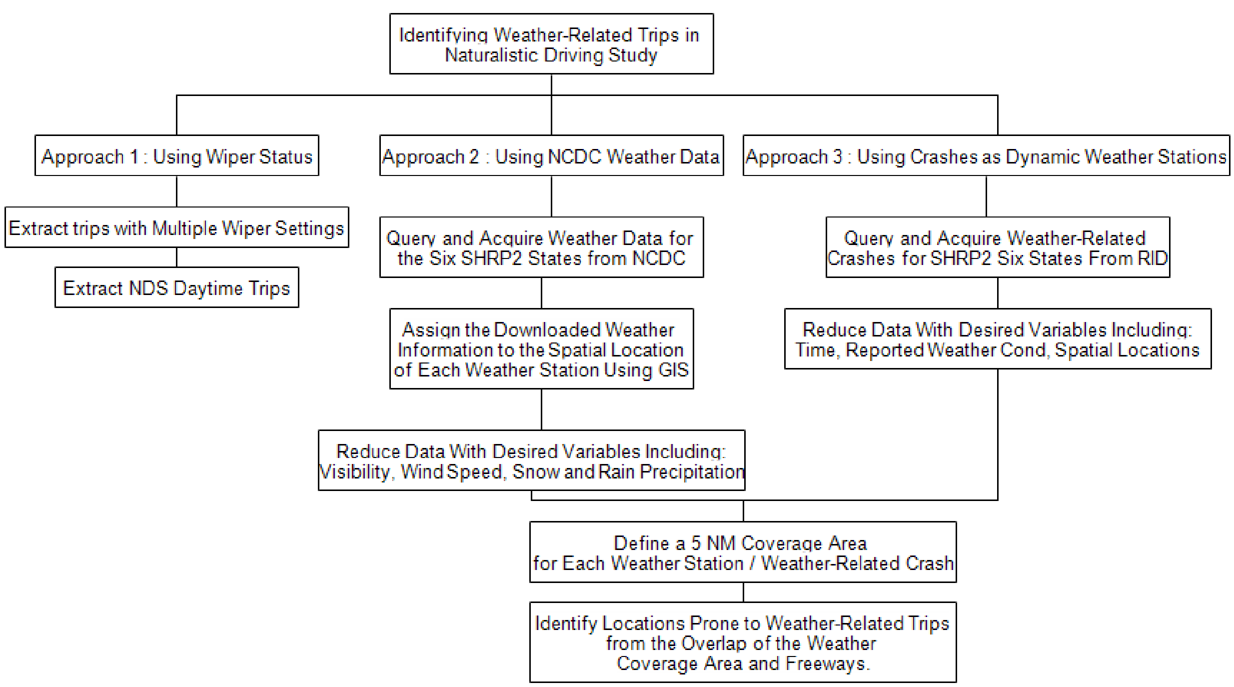


Figure 1 Complementary Methodologies for Identifying Weather-Related NDS Trips *(16)*

Method 1: Wiper Status

The first data acquisition method was developed as part of the first project phase and leveraged the internal vehicle sensor data capturing the windshield wiper status. In this case, the vehicle's front windshield wiper setting was captured to identify events in which precipitation occurred. The time-series variable capturing the windshield wiper status provides an indication of the windshield wiper switch, rather than directly reporting the blade speed. Using the guidance from Phase 1, the following procedures were defined to identify NDS trips in rainfall, without introducing bias:

1. Vehicles with multiple wiper settings were targeted; vehicle data without the full spectrum of values for the wiper status (0, 1, 2, and 3) were filtered out because vehicles with only on/off wiper settings cannot provide an indication of rain intensity.
2. Months with heavy precipitation in all of the SHRP2 data collection regions were targeted.
3. Trips in non-daylight conditions and on non-freeways were removed; freeways were considered based on the project scope and nighttime trips were eliminated due to low video resolution.
4. Extracted trips were tagged with the percentage of the trip in which different wiper settings (0, 1, 2, and 3) were active, which was used to identify trips of interest in various levels of precipitation.

While this method produced a sufficient number of trips for Phase 1, limitations are prevalent when evaluating precipitation rates using the reported wiper setting. For example, all motorists drive differently and have unique tolerance thresholds to different rates of precipitation. In addition, a review of data from Phase 1 indicated that in Honda Civic vehicles, the front camera was positioned above the reach of the wiper blades; therefore, during all rain events, the camera image was blurred and deemed unusable *(15)*. To overcome these limitations, two additional methods were proposed. These complementary methods not only captured additional precipitation events from vehicles without complete wiper status readings, but also enabled the collection of data in different weather conditions (i.e., conditions that did not require windshield wiper activation).

Method 2: NCDC Weather Stations

The second data acquisition method leveraged external weather data sources. For this procedure, weather conditions were queried from the National Climatic Data Center (NCDC), which is a database available through Climate Data Online (CDO). The NCDC archives weather data from various weather stations nationwide, including radar, satellites, airport weather stations, and military weather stations. Among these data sources, the airport weather stations proved to be the most beneficial to identifying adverse weather events. Over 5 GB of weather data from more than 250 weather stations in the six NDS states (between 2010 and 2013) were collected from the National Oceanic and Atmospheric Administration (NOAA) - National Climatic Data Center (NCDC) website.

Airports' automated weather stations monitor weather conditions continuously and record the weather parameters according to predefined changes in their values; for that reason, the data do not follow a specific time pattern, but report weather conditions relative to real time weather changes. The weather parameters collected include visibility, temperature, humidity, wind speed and direction, and precipitation. Among these parameters, visibility is considered one of the most critical factors affecting driver behavior. Visibility can generally be described as the maximum distance that an object can be clearly perceived against the background sky; visibility impairment can be a result of both natural (e.g., fog, mist, haze, snow, rain, windblown dust, etc.) and human induced activities (transportation, agricultural activities, fuel combustion, etc.). The automated weather stations cannot directly measure the visibility, but rather calculate it from a measurement of light extinction, which includes the scattering and absorption of light by particles and gases.

Previous studies concluded that airport weather stations can provide spatial-temporal weather conditions for adjacent roadways within five nautical miles and within a two hour time period at 60 percent to 80 percent accuracy *(17)*. In this study, daily weather data were acquired and NDS trips were requested based on the daily weather information to identify all trips impacted by adverse weather events (such as those conducted on ice or slush road surfaces), not only those occurring during active precipitation or fog. Therefore, the date and time for every weather event was superimposed on the NDS trips for freeways within five nautical miles.

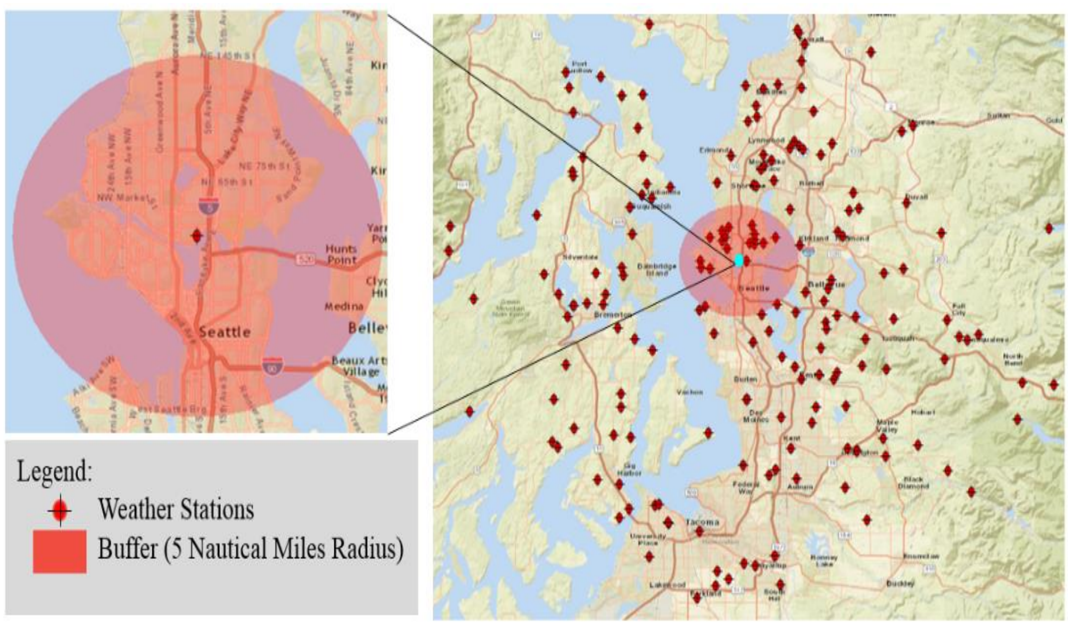


Figure 2 Weather stations and a representation of the five nautical mile coverage area for extrapolating weather conditions, example from Washington State *(16)*

Figure 2 shows weather stations used to identify the snow-related trips in Washington and depicts the concept of a five nautical mile coverage area used in the data acquisition process. In total, 24 GIS-shape files were provided representing rain, snow, and fog conditions for the six NDS states for the extraction of NDS trips.

Method 3: Weather-Related Crashes

The third data acquisition methodology has a similar concept to method 2. Instead of using weather-stations to extrapolate weather conditions onto surrounding freeways, method 3 uses weather-related crashes to identify surrounding NDS trips that were impacted by the same adverse weather conditions. Crash databases were queried and times and locations of weather-related crashes occurring in the NDS data collection regions were extracted. Using the same procedures described in “Method 2: NCDC Weather Stations”, the spatial and temporal data from each crash were uploaded to GIS, and shape files containing these data were provided to Virginia Tech Transportation Institute (VTTI) for NDS trip extraction. Figure 3 illustrates the weather related crashes from the state of Washington and the concept of the five nautical mile coverage area.

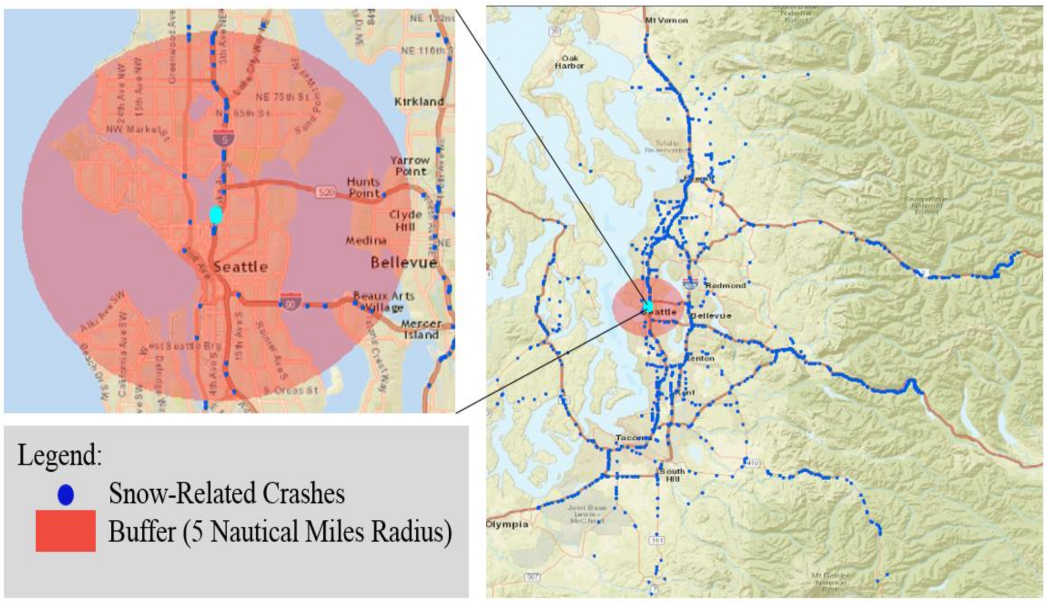


Figure 3 Weather-related crashes and a representation of the five nautical mile coverage area for extrapolating weather conditions, example from Washington State *(16)*

Summary of Acquired NDS Trips

Using these complementary methodologies, the research team received 11,164 trips flagged as being weather-related and 22,328 matched trips in clear weather conditions (i.e., same driver, same route). In total, this resulted in 11,205 hours of driving data. The identified NDS trips involved 1,523 drivers between 16 and 99 years of age with the majority of the drivers in the young age group, 16 to 29 years old. Gender representation was balanced in most age groups, with the exception of a slight overrepresentation of female drivers between 20 and 24 years old. The age and gender distributions for the collected trips and the entire SHRP2 database are shown in Figure 4.

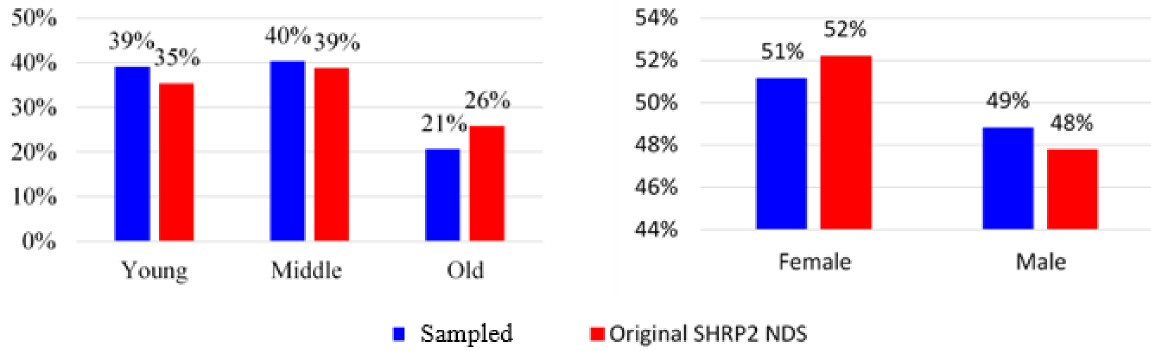


Figure 4 Age and gender distribution for collected vs. original SHRP2 NDS data *(16)*

Data reduction was critical to identifying false-positives, or trips flagged as being weather-related but actually occurred in clear or dry conditions. After extensive data reduction, which is described in detail in Data Reduction section, a total of 4,094 freeway trips (~37 percent) were verified as being adverse weather trips. Of these adverse trips, 3013 occurred in rain, 234 in fog, 320 in snow, 317 in clear conditions with wet pavement, and 210 in clear conditions with snow-covered pavement. A summary of the trips collected from each methodology are shown in Figure 5.

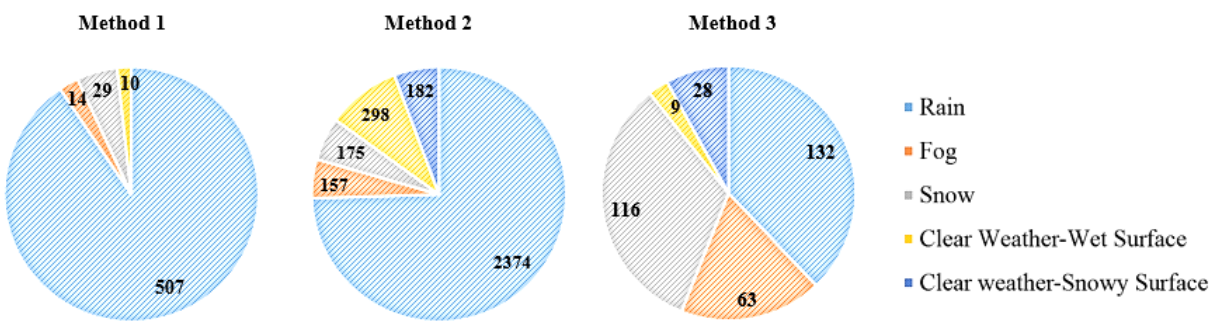


Figure 5 Summary of received trips of each weather classification for the three complementary methodologies *(16)*

Crash and Near Crash Events

In addition to collect full NDS trips, the NDS crash and near-crash databases were queried for weather-related events. This query produced 37 crashes, 266 near crashes, 606 matched non safety-critical events in all represented weather conditions, and 1,176 baseline trips. Manual video verification revealed that only 16 crashes occurred on freeways (7 in rain and 9 in clear weather), and 213 events contained near-crash events (33 in rain or sleet, and 182 in clear weather).

Data Reduction

Once NDS trips were collected, efficient data reduction procedures were required to sift through the 33,000 trip files. Using the experience gained from the first project phase, the research team developed the Wyoming NDS Data Analysis Tool (DAS). The DAS is a python-based analytic tool that ingests the time-series trip data, performs a variety of reduction functions, and produces summary observation templates, summary statistics, and graphical representations of the data. The main functionalities of the DAS are described in the following sections.

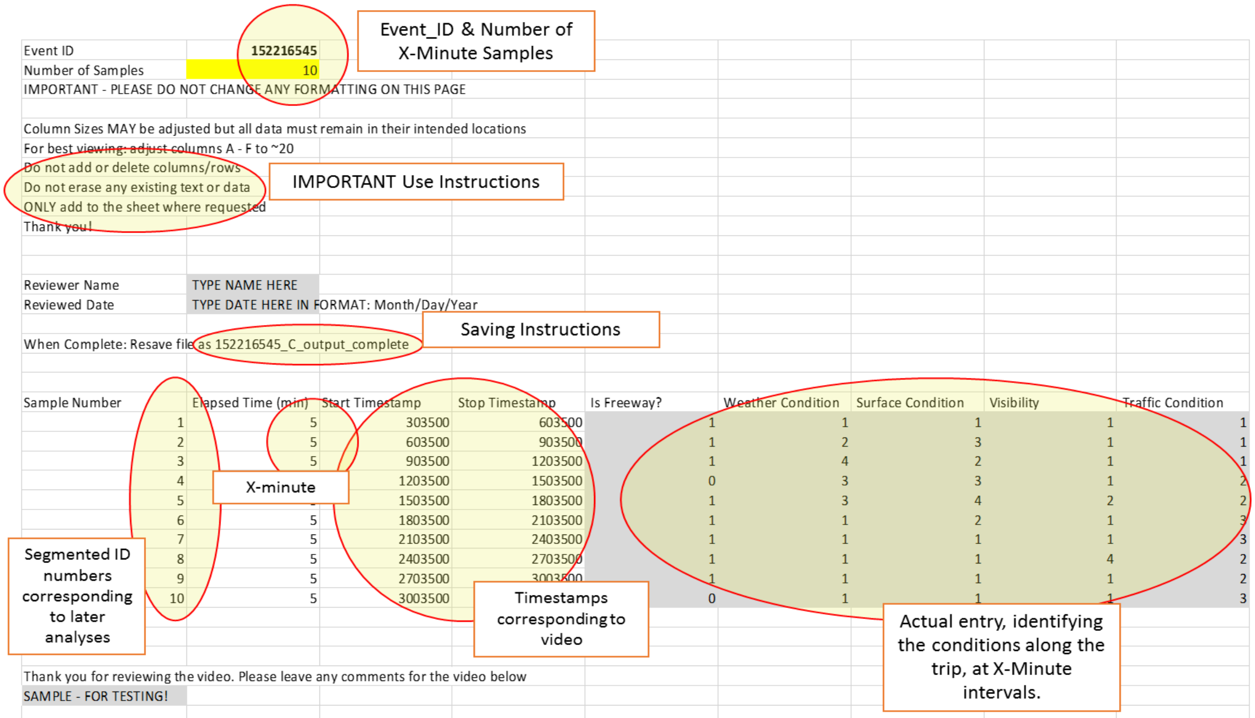
Dimensionality Reduction & Time Chunking

Upon ingesting an NDS trip, the DAS first reduces the data dimensionality by extracting time-series variables identified in the first phase as being relevant for evaluation of driving behavior in inclement weather. Additional variables can be analyzed with minor adjustments to the tool; however, the intention of this step is to reduce the complexity of the data analytics to improve processing speed. The next step included a time-series data processing technique called "time chunking". Analyses in Phase 1 revealed high variability in weather conditions within a single trip; therefore, the concept of time chunking was used to segment out equal sections for analysis. Each trip was divided into one-minute segments to create homogeneous "chunks" with similar environmental and traffic conditions. This process was introduced in Phase 2 to improve the computational efficiency and analysis output quality by improving driving environment (i.e., the driving environment defined by weather, traffic, and roadway conditions) homogeneity.

Video Observation Template

The DAS enables semi-automatic data reduction. Full automation of these procedures is not feasible because identification of the driving environment (i.e., weather, traffic, and roadway conditions) cannot be automatically generated with the current state of vision learning algorithms. A detailed discussion of automated detection efforts using the Wyoming SHRP2 NDS data is found in "Wyoming NDS Visualization and Reduction Tool" section.

In order to enhance and expedite the required manual video observation, the DAS generates observation templates for each trip. The observation template identifies each X-minute segment within the trip by its corresponding start and stop timestamps and leaves data fields for video reviewers to report observed environmental and traffic conditions prevalent in each segment. Figure 6 shows a sample video observation template generated by the DAS using five-minute time chunks. However, most analyses utilized one-minute time chunks, and the trips were reduced with one-minute time chunks.


Event ID 152216545
Number of Samples 10
IMPORTANT - PLEASE DO NOT CHANGE ANY FORMATTING ON THIS PAGE

Column Sizes MAY be adjusted but all data must remain in their intended locations
For best viewing: adjust columns A - F to ~20
Do not add or delete columns/rows
Do not erase any existing text or data
ONLY add to the sheet where requested
Thank you!

Reviewer Name TYPE NAME HERE
Reviewed Date TYPE DATE HERE IN FORMAT: Month/Day/Year

When Complete: Resave file as 152216545_C_output_complete

Sample Number	Elapsed Time (min)	Start Timestamp	Stop Timestamp	Is Freeway?	Weather Condition	Surface Condition	Visibility	Traffic Condition
1	5	303500	603500	1	1	1	1	1
2	5	603500	903500	1	2	3	1	1
3	5	903500	1203500	1	4	2	1	1
4		1203500	1503500	0	3	3	1	2
5		1503500	1803500	1	3	4	2	2
6	5	1803500	2103500	1	1	2	1	3
7	5	2103500	2403500	1	1	1	1	3
8	5	2403500	2703500	1	1	1	4	2
9	5	2703500	3003500	1	1	1	1	2
10	5	3003500		0	1	1	1	3

Thank you for reviewing the video. Please leave any comments for the video below
SAMPLE - FOR TESTING!

Figure 6 Sample video observation template produced by the DAS

Manual Video Observation and Annotation

The most time-intensive component of the data reduction process is the actual manual video observation and annotation. Advanced machine learning algorithms are being explored to identify novel methods to automate elements of this manual process; however, the quality of the NDS video data and variations in the camera angle in different vehicles are challenging obstacles to overcome. More discussions of this effort are provided in "Wyoming NDS Visualization and Reduction Tool" section. Therefore, a systematic procedure was used to gather the driving environment data from the video footage manually. This also served as a ground truth data for the machine vision effort.

This procedure involves the classification of roadway, weather, surface, visibility, and traffic conditions in explicitly defined categories. To eliminate subjectivity and any potential bias in identifying weather, traffic, or roadway conditions, comprehensive training and a detailed description of each condition was provided to video viewers before the manual observation began. The manual observation categories defined for this procedure are shown in Table 1.

Table 1 Manual Observation Categories

Roadway Type		**Surface Conditions**		**Traffic Conditions**	
Freeway	0	Dry	1	LOS A	1
Non-Freeway	1	Wet	2	LOS B	2
Weather Conditions		Snow-covered	3	LOS C	3
Clear	1	Ice-covered	4	LOS D	4
Light Rain	2	**Visibility**		LOS E	5
Heavy Rain	3	Low	1	LOS F	6
Snow	4	Moderate	2		
Fog	5	High	3		
Sleet	6				

Data Aggregation & Summary File Generation

Once manual observations of trips are complete, the DAS ingests the completed files and aggregates the information based on user input to produce summary files, data graphics, and inputs for analytic models. An example of a user input could be requesting all driving data occurring on a freeway, in heavy rain, with wet surface conditions, low visibility, and a level of service (LOS) below C. The DAS could then generate summary statistics related to average speeds, headways, accelerations, etc. for all data available for trips meeting those criteria. In addition, the research team can add another element to the query to focus on driver behaviors from drivers with specific demographics, personally-defined risk taking and perception surveys, or a series of cognitive tests.

Wyoming NDS Visualization and Reduction Tool

The Wyoming NDS Visualization and Reduction Tool (VRT) was initially developed in Phase 1; however, improvements to the functionality and usability were conducted during Phase 2. The VRT was used for manual video observation because it conveniently synchronizes the front and rear video feeds with user-defined time-series variables (e.g., speed, acceleration) and interprets radar data. A screen shot of the VRT graphical user interface is shown in Figure 7. In addition to the data visualization benefits, the research team developed VRT's image processing algorithms using various visibility estimation and machine learning techniques. Ultimately, the goal is to enable automatic video observation, so to further expedite the data reduction procedures. The following sections contain details related to each of these efforts.

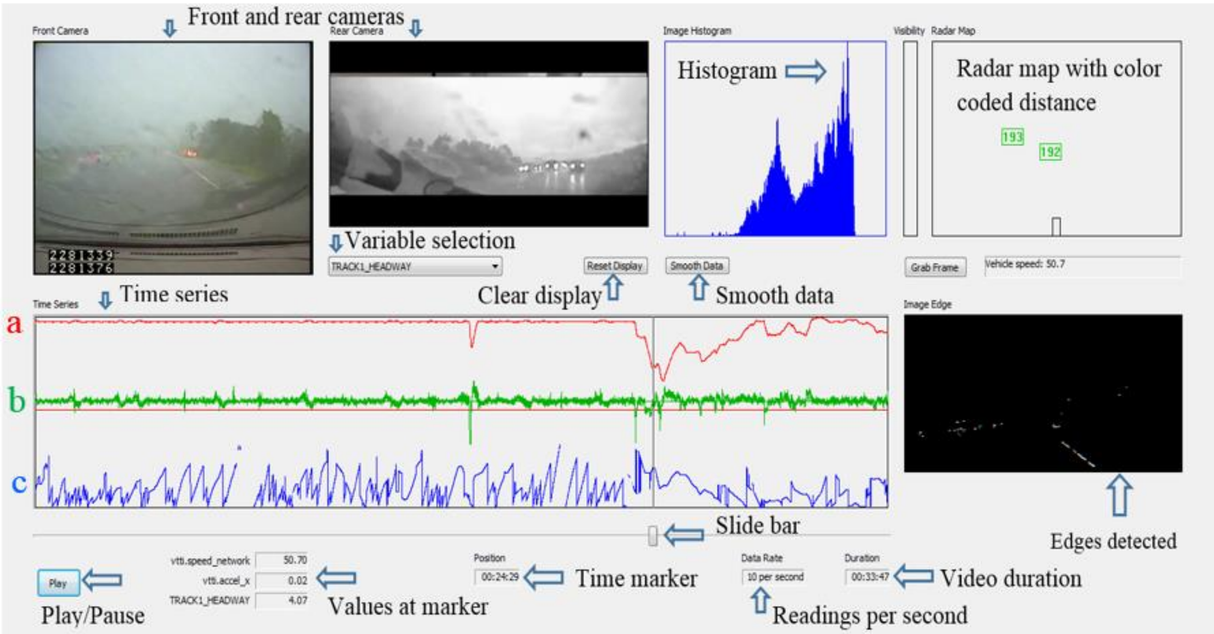


Figure 7 Wyoming Visualization and Reduction Tool *(16)*

Visibility Estimation Algorithm

The visibility estimation algorithm applies methods that look for object boundaries and edges as a way to assess the existence of objects and their clarity in the image. This technique assumes that an image of an adverse weather conditions is, generally, blurrier than that of a clearer weather. The algorithm calculates the Laplacian filter of the image and estimates the visibility level based on its magnitude. The algorithm is heuristic; therefore, accurate results for all input images are not guaranteed. The accuracy of the visibility estimation algorithm depends on numerous factors including: the training data set, the filter magnitude interpreted, cutoff limits used for different weather conditions, and the input image quality. The Visibility Index (VI) is the resulting value given to an image to describe its visibility level. The VI is expressed as a percentage and classified in one of three levels: low, medium, or high. More work is in progress to improve the methods for deriving the VI values, as well as to produce representative ranges of VI values for each classification (i.e., low, medium, high). As the visibility estimation algorithm is still under development and refinement, only experimental accuracies can be defined. Preliminary testing of the algorithm using 19 video files suggests a 79 percent accuracy (14 trips yielding results consistent with human observation; 2 trips yielding partially consistent results, and 3 trips yielding inconsistent results).

CHAPTER 3 - RESEARCH FINDINGS

Once initial data acquisition and reduction were completed, specific research questions related to different aspects of driving behavior were considered by the research team. These research questions were divided into four distinctive research areas: 1) speed selection, 2) car-following behavior, 3) lane-keeping behavior, and 4) safety critical events. Using a variety of analytic modeling techniques, the research team evaluated each research area using a sample of the Wyoming SHRP2 NDS trips. Due to the unique nature of each research area, different trip samples were used for each individual analysis.

One analytic strategy used in many of the research areas was introduced in chapter 2 and is related to the collection of adverse weather trips and matching clear trips. In order to identify distinct behavioral adjustments made during adverse weather events, the research team needed a baseline for comparison. Therefore, clear trips were collected alongside each adverse trip, matching the same driver traversing the same route. As part of Phase 2 data acquisition, two clear trips were collected for each adverse trip, and these matching trip sets are used in various ways in each research area.

The following sections describe the analyses conducted for each specific element of driving behavior studied as part of the Wyoming Phase 2 IAP. Each area of research is prefaced with a brief literature review describing the specific work, an overview of the data preparation procedures required in addition to the initial data reduction, the analysis methodology, analytic results, and a summary of next steps for a continuation of the research.

Speed Selection in Rain and Snow

Literature Review

The impact of adverse weather conditions including rain and snow on driver speed selection has been investigated in a previous study *(18)*. Weather and traffic data were collected from a freeway (Queen Elizabeth Way Mississauga) in Canada. They found a significant speed reduction during adverse weather conditions, including light and heavy rain, light snow, and snow storms. In addition, analysis of traffic operation parameters showed 1.24 to 1.86 mph (2 and 3 kph) reduction in speed caused by light rain and light snow respectively. Moreover, 3.1 to 6.2 mph (5-10 km/h and 38-50 km/h) speed reduction were observed in heavy rain and heavy snow, respectively.

Stern et al. (2003) analyzed the negative impact of adverse weather on traffic flow using eighteen freeway segments and fifteen arterial roadway segments, in total representing 239 miles (384.633 km) *(19)*. Weather data were extracted from three airports in Washington DC, including Dulles International (IAD), Washington National (DCA) and Baltimore/Washington International (BWI). Travel time data were extracted from the "SmarTraveller", a web-based service that provides information including travel time, crash locations, work zones, etc. Different weather conditions were considered including rain, snow, wind, visibility and surface conditions. Rain and snow were categorized into three levels including no rain/snow, light rain/snow, heavy rain and heavy snow/sleet. The wind was categorized into two levels including below 30 mph (48 km) and above 30 mph (48 km). Visibility distance was classified into below 0.25 miles (0.4 km) and above 0.25 miles (0.4 km). Finally, the surface condition was categorized into four levels including dry, snowy, wet, and icy. Their results indicated an increase in travel time during adverse weather conditions. Specifically, an average of 11 and 13 percent increase in travel time during on-peak and off-peak periods were observed under precipitation conditions, respectively.

A study by Hawkins (1988) investigated weather conditions impact on vehicle speed using the loop detector data *(20)*. Weather conditions were categorized into nine categories. Surface conditions and visibility were also considered. The findings showed 25-30 percent speed reduction within 328 feet visibility limit. The speed reduction was found to be higher in snow and slush, 18.6 and 24.9 mph (29.9 and 39.3 kph), respectively. In addition, 2.5 and 3.7 mph (4.02 and 5.95 kph) speed reduction were observed for light and heavy rain respectively.

The impacts of adverse weather on traffic demand, traffic safety, and traffic flow were investigated in a previous study *(21)*. It was found that both weather type and precipitation intensity play important roles in the abovementioned factors. More specifically, the analysis results showed less than 5 percent reduction in traffic volume during heavy rain and 5-80 percent during snowstorms. In addition, 13 percent and 25 percent increase in crash rate were observed during moderate snow and heavy snow respectively. They also found that more than 0.25 in/h (0.635 cm/h) rainfall and 0.5 in/h (1.27 cm/h) snowfall caused 14 and 22 percent reduction in capacity of freeways, respectively.

Another study mentioned that free-flow speed on rural interstate freeways might be impacted by poor surface conditions, visibility, and severe wind speeds *(22)*. They also mentioned that adverse weather conditions should be taken into account in analyzing capacity and level-of-service for a number of the U.S. cities that severe weather conditions occur frequent enough.

Pisano and Goodwin, 2002 investigated the negative effects of different weather conditions on roadways and traffic operation parameters *(23)*. Table 2 shows the impact of weather conditions on roadway environment and operations.

Table 2 Weather Impacts on Roadway Environments *(23)*

Weather Conditions	Roadway Environment Impacts	Transportation System Impacts
Rain, Snow, Sleet, Hail & Flooding	• Visibility Reduction • Reduction in Pavement friction • Lane obstruction • Lane submersion • Reduction in vehicle stability • Reduction in vehicle maneuverability • Increased chemical and abrasive use for snow and ice control • Infrastructure damage	• Reduced roadway capacity • Reduced speeds & increased delay • Increased speed variability • Increased accident risk • Road/bridge restrictions & closures • Loss of communications/power services • Increased maintenance & operations costs
High Winds	• Visibility Reduction (blowing snow or dust) • Lane obstruction (drifting snow) • vehicle stability & maneuverability reduction (Large vehicle tip over)	• Increased delay • Reduced traffic speeds • Road/bridge restrictions & closures
Fog, Smog, Smoke & Glare	• Visibility Reduction	• Reduced speeds & increased delay • Increased speed variability • Increased accident risk • Road/bridge restrictions & closures
Extreme Temperatures & Lightning	• Increased wildfire risk • Infrastructure damage	• Traffic control device failure • Loss of communications & power services • Increased maintenance & operations costs

A previous study showed an inverse proportional relation between traffic demand and inclement weather conditions. In fact, they found that by increasing the severity of adverse weather conditions, the traffic demand decreases. Analyzing the traffic demand on snowy days revealed 20 to 80 percent reduction in demand considering low and high wind respectively *(24)*.

The recently published 2016 Highway Capacity Manual (HCM) has discussed the impact of inclement weather conditions on traffic operation and has provided Weather Adjustment Factors (WAFs) based on (1) weather type and intensity and (2) facility free-flow speed (FFS). As an example, considering a freeway corridor with a 65 mph FFS, freeway capacity reductions are predicted to be 8 percent and 14 percent for medium rain and heavy rain, respectively *(25)*.

Selecting the right speed for the condition is considered as one of the most important driving tasks on high-speed facilities. In fact, one of the key factors to achieve an efficient Intelligent Transportation System (ITS) is collecting high-quality microscopic traffic data. Vehicle speed is one of the main indicators to estimate traffic conditions on freeways. Speed data for this purpose generally collected using inductive Loop Detectors (iLD) and Remote Traffic Microwave Sensors (RTMS) in the literature *(26)*. However, utilizing in-vehicle (IV) data for speed prediction is becoming more common in recent years *(27)*.

In the early 80's, VSL was introduced as a groundbreaking approach that could regulate traffic congestion problems. According to the Federal Highway Administration (FHWA), Connected Vehicle (CV), Variable Speed Limits (VSL), and Advanced Traveler Information Systems (ATIS) are considered the next step in tackling U.S. freeway congestion and safety problems. Hence, VSL systems have been widely implemented in the U.S.

The main benefit of VSL is in its immediate response to real-time conditions including actual traffic conditions, weather conditions, and any other real-time circumstances. In fact, variable speed limit advantages can be classified into three main categories : (1) instantaneous intervention that can influence dynamics of the traffic in any abnormal conditions, (2) traffic management while avoiding diverting or restricting access for unrelated streams of traffic, and (3) speed harmonization that can reduce disturbances of individual driver behavior at microscopic and macroscopic levels as well as improve traffic performance and reduce congestion.

Even though the most common applications of VSLs are incident management, and used mostly near temporary work-zone sites, static speed limits are still the most common way of notifying drivers about the maximum allocated speed on the road *(28)*. In fact, there are technical and policy related challenges in the widespread adoption of VSL systems. The first challenge ahead of researchers is the absence of comprehensive strategies that are independent of specific infrastructure necessities and traffic scenarios. The second important restraining factor is the absence of a comprehensive approach that can connect the VSLs to other existing control systems to gain the maximum efficiency and performance. For instance, the independent design of traffic control systems such as ramp metering and VSL might introduce issues when deploying them simultaneously, since they are not coordinated *(28)*. Finally, the most important factor that received less attention in previous studies is considering driver behavior in VSL scenarios. More specifically, current VSLs are mostly designed to consider weather or traffic conditions. However, driver behavior such as speed selection according to the condition and compliance to traffic control devices could play an important role in determining appropriate speed for VSL based on real-time driver performance on roadways.

The negative impacts of inclement weather conditions on traffic flow have been investigated in many studies; however, there is a lack of studies that have examined the underlying complexity of driver speed selection behavior during adverse weather conditions using trajectory-level data. Therefore, this study utilized NDS trajectory-level weather-related data to provide a better understanding of driver speed selection behavior in adverse weather conditions, which can be used to provide more realistic VSLs. Chapter 3 and 4 provided detailed information about the developed speed selection models and discussed the potential countermeasures.

Data Preparation

In this study, a total of 212 trips in adverse weather conditions (22 trips in fog; 102 trips in rain; and 88 trips in snow – plus 424 matching clear weather trips) were randomly selected from the extensive Wyoming NDS database. The selected NDS trips involve 145 drivers between 16 and 89 years old, with the majority of the drivers in the young age group (16 to 29 years old). Gender was mainly balanced among age groups, except for a slight overrepresentation of female drivers between 20 and 24, which follows the same distribution that is reported by VTTI for all SHRP2 NDS trips.

A total of 14,923 one-minute segments – equivalent to nearly 249 hours and 11,466 miles (18,453 km) equivalent to: Rain: 2,225.7 miles (3,582 km), Snow: 1,003 miles (1,615 km), and Fog: 592.79 miles (954 km) of driving, plus their matching trips in clear weather conditions – were processed. The speed limit data provided in the RID was used to merge speed limits with each one-minute segment. Once non-freeway segments were removed, 10,606 one-minute segments were used to model driver speed selection.

Methodology

In order to identify the impact of weather conditions on driver speed selection, two models using both parametric and non-parametric methods were developed. Parametric models, such as probit and logistic regression models, provide the relationship between a response variable and predictors. However, parametric models have some limitations. More specifically, they cannot provide a high level of prediction accuracy because there are many embedded assumptions *(29)*. Another complication in using parametric models are their inability to automatically handle missing values *(30)*. These shortcomings cannot be addressed using common parametric models such as logistic and probit models *(29, 31)*. Despite their limitations, parametric logistic/probit models are effective in interpreting the marginal effects of various risk factors *(32, 33)*. On the other hand, there are several key advantages of using non-parametric models, including the ability of providing high prediction accuracy, handling of missing values automatically, and their capability of handling large number of explanatory variables in a timely manner, which might be extremely beneficial specifically for assessing traffic operation and safety in real-time considering weather and traffic data to be directly fed into the model *(30)*. However, the trade-off is that their classification results cannot be explicitly interpreted *(29)*.

In this study both ordinal logistic regression (parametric) and classification and regression tree (non-parametric) methods are used for analyzing the impact of different contributing factors (focusing on weather and roadway conditions) on driver speed selection.

Ordinal Logistic Regression (OLR)

Logistic regression is a commonly used model in traffic safety and operation studies. Logistic regression allows the formulation of predictive models on a probabilistic basis. Similar to other regression analyses, it predicts the value of a dependent variable from one or more explanatory variable(s). Logistic regression can be applied to a binary, nominal, or ordinal dependent variable. Logistic regression (Equation 1) can also be utilized to rank the relative importance of the response variables *(34)*.

$$Logit[P(x)] = \log\left(\frac{P(x)}{1-P(x)}\right) = \alpha + \beta x$$ **Equation 1**

Equation 1 shows a logistic regression model with *x* representing the independent variable, and P(*x*) indicating the probability of success for a binary response variable y, considering explanatory variable *x*. α represents the response probability when explanatory variables are at the reference level (or when x=0) ; β represents the regression coefficients. As mentioned earlier, logistic regression can be conducted using an ordinal response variable. The ordinal logistic regression equation is shown in Equation 2.

$$Ln(\Pi_j) = \alpha_j - (\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + ...) \qquad \textbf{Equation 2}$$

An ordinal logic regression (ordered logit) model was applied for this analysis due to the ordinal nature of speed selection that would not be accounted for in multinomial analyses.

Classification and Regression Tree Model (CART)
Decision tree modeling can be used for both continuous and nominal dependent variables. Utilizing a decision tree to classify a nominal dependent variable is called a classification tree *(35, 36)*. Classification can be defined as a procedure for predicting the "class" of an object – considering the object's features *(37)*. Classification models are built from a training dataset in which trends of explanatory and response variables are identified and used to predict the value of the target variable for different datasets *(38)*. The two main components of decision trees are the "root node" and the "leaf node". The "root node" is the node located at the top of the tree, which contains all ingested data and the "leaf node" refers to the termination node, which has the lowest impurity.

The root node is divided into two child nodes, based on the independent variable (splitter) that creates the best homogeneity. This procedure of partitioning the target variable recursively is repeated until all of the data in each node reach their highest homogeneity. At that point, tree growth stops, and the node(s) that do not have any branches become the "leaf node(s)". Each path from the top of the tree (root node) to the bottom/termination of the tree (leaf node) can be considered a rule. Following this sequence, the data in each child node is purer (more homogenous) than the data in the upper parent node *(39)*.

In order to identify possible splits among all variables, a splitting criterion is generated. The splitting criterion is the main design component of a decision tree *(40)*. In a decision-tree learning algorithm, the splitting criterion's role is to measure the quality of each possible split among all variables. Two common tests used to generate splitting criteria are: 1) chi-square and 2) Gini reduction. The Gini splitting criterion is used to select the variable and split pattern to best partition the node. Gini impurity indicates the data purity; in other words, it shows the probability of incorrect classification for a randomly chosen record from the specific node in the data subset.

Variable Importance Measure (VIM) is one of the main outputs of the classification tree model, showing the most important factors affecting target variable *(41)*. In this section, most significant factors affecting driver speed selection considering adverse weather conditions were identified using the VIM.

Analysis

Preliminary Analysis

Table 3 indicates that speed reduction is more likely to occur in adverse weather (snow, rain, fog) conditions in comparison with the matched trips in clear weather conditions. The odds ratios of driving below the speed limit, were 7, 2.7, and 2 times more likely to be in snow, rain, and fog, respectively, than their matching trips in clear weather conditions.

Table 3 Odds Ratios for Speed Behavior in Snow, Rain, and Fog

	Driving below Speed Limit	Driving above Speed Limit	Odds Ratio	Confidence Interval	Significance level
Snow	773	441	6.93	5.89 to 8.14	P < 0.0001
Matched Clear of Snow	386	1525			
Rain	220	251	2.67	2.13 to 3.35	P < 0.0001
Matched Clear of Rain	268	816			
Fog	91	191	2.1	1.53 to 2.89	P < 0.0001
Matched Clear of Fog	119	531			

Additional analyses were conducted to compare speed reduction in each NDS state during free flow speed conditions. Findings indicate that the speed reduction was not similar in each state. For instance, in New York one-minute segments in snow had a speed reduction of about 18 percent being the highest among all NDS states. Important sample size considerations note that 74 percent of the identified snow segments were from New York. Whereas in Pennsylvania with only 7 percent of snow-related segments, the speed reduction was about 9 percent being the lowest. In addition, in rainy weather conditions, trips in Indiana had the highest speed reduction of about 33 percent. Among the considered segments in rain, 1 percent was travelled in Indiana while Washington had 44 percent with the lowest speed reduction of about 3 percent. In fog, the highest speed reduction was 6 percent in Florida with 37 percent of fog-related trips and the lowest speed reduction was in Washington with 14 percent of fog-related segments, where the NDS drivers reduced their speed by nearly 2 percent. These differences are certainly a function of the distribution and sample size of snow, rain, and fog events in each state; nonetheless, the finding indicates that driving behavior in adverse weather conditions must be calibrated based on local driver populations and their familiarity with the weather condition.

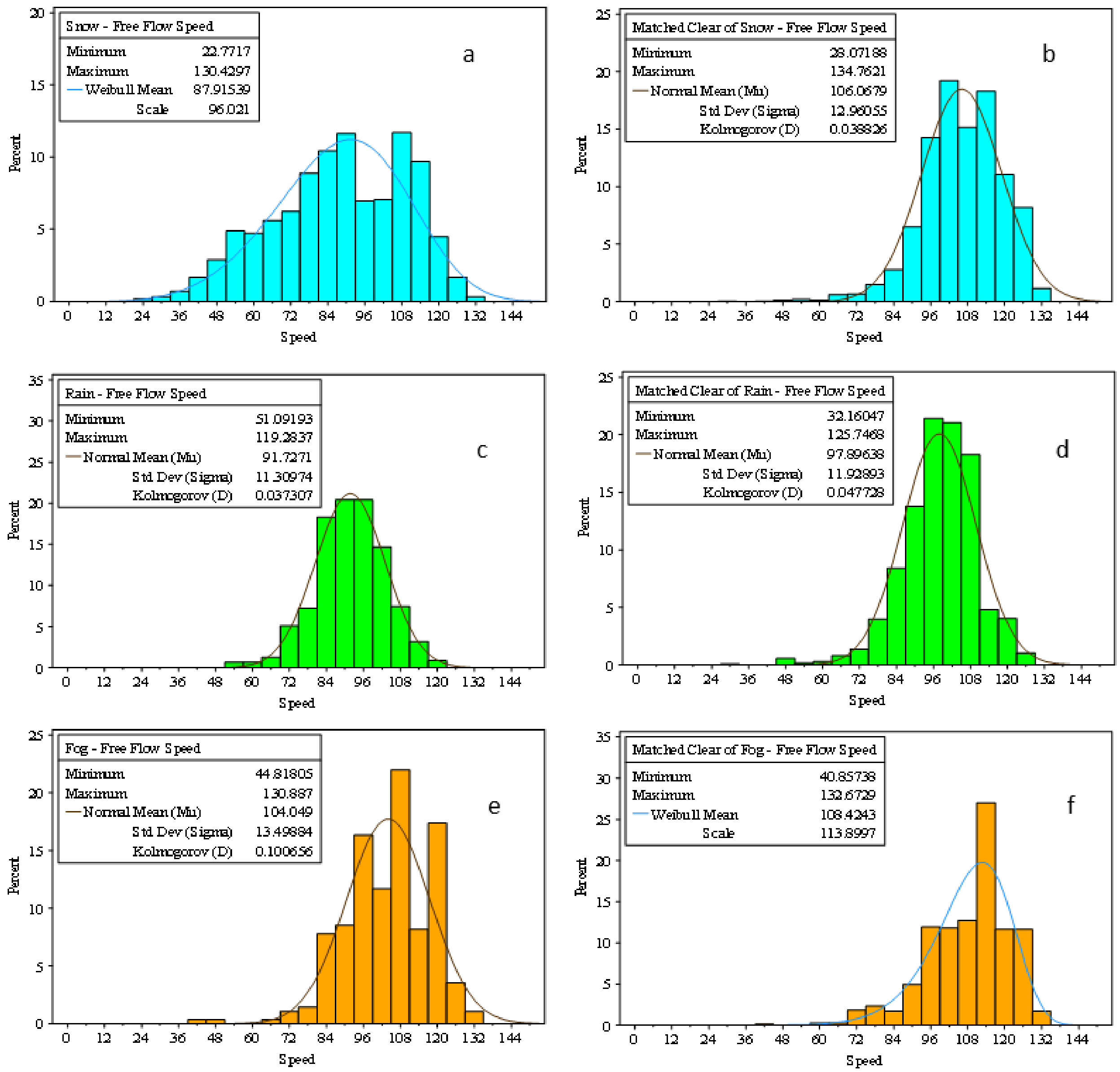


Figure 8 Observed and Fitted Distributions for Speeds during Adverse and Clear Weather under Free-Flow Traffic *(42)*

Direct comparisons between clear weather in free flow speed and driving in adverse weather/ traffic conditions are imperative to identify critical traffic and environmental conditions. GIS was used to illustrate driver speed behavior under various weather conditions.

Trip ID: 13910595
Visibility: Fog (NCDC) – whiteout condition visual observation
Trip Location: New York (NDS TS)
Surface: Snow (Video Observation)
Vehicle Average Speed: 39.6 mph (NDS TS)
Standard Deviation of Speed: 11.86
Wind Speed: 33 mph (NCDC)
Speed Limit: (RID Reduced data)

Trip ID: 13904014
Visibility: Clear (NCDC)
Trip Location: New York (NDS TS)
Surface: Dry (Video Observation)
Vehicle Average Speed: 62 mph (NDS TS)
Standard Deviation of Speed: 12.73
Speed Limit: (RID Reduced data)

Figure 9 Illustration of a Trip in Fog and Whiteout Condition (I-290 New York)

Figure 9 illustrate a significant speed reduction due to whiteout condition on a traversed route in New York, using data from a sample trip and a matching trip. The average speed during the whiteout condition was about 22 mph (35 kph) less than the matched trip in clear weather conditions.

Figure 10 represents the driver performance on roadways considering the risk of crashes. Two heat-maps representing crash-prone zones using the three years of crashes (2011-2013) on I-190 and I-290 were developed. Speed reduction percentages along the travelled routes are shown in a range of colors from green (low crash rate) to red (high crash rate). In addition, a separate color scale along the interstate route represents the speed reduction (compared to the posted speed limit) realized for the trip in whiteout conditions and the matching clear. These maps indicate speed reduction was much greater in whiteout conditions compared to the matched clear weather conditions. The potential benefit of visualizing continuous driver performance data (here speed reduction percentage) and crash-prone locations heat-map is in VSL/VMS application. This information can be utilized in updating VSL/VMS in real-time. More clearly, using this representing GIS maps can highlight not only the crash hotspots but also the possible driver role in crash occurrence. This work will be expanded using more NDS drivers in different weather conditions. In addition, the same concept could be implemented on Phase 3 I-80 VSL corridor.

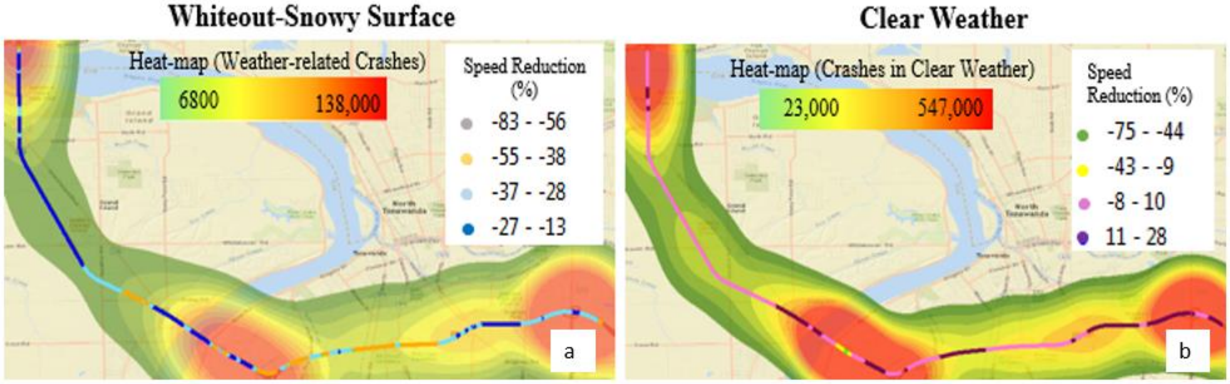


Figure 10 Speed behavior-GIS Representation *(43)*

Direct comparisons between clear weather in free flow speed and driving in adverse weather/ traffic conditions are imperative to identify critical traffic and environmental conditions. GIS was used to illustrate driver speed behavior under various weather conditions. Figure 10 illustrate a significant speed reduction due to whiteout condition on a traversed route in New York, using data from a sample trip and a matching trip. The average speed during the whiteout condition was about 22 mph (35 kph) less than the matched trip in clear weather conditions. Figure 11 shows the drivers' speed and lane keeping behavior in both the clear and whiteout conditions, indicating lower travel speeds and difficulty in maintaining his/her lane in whiteout condition.

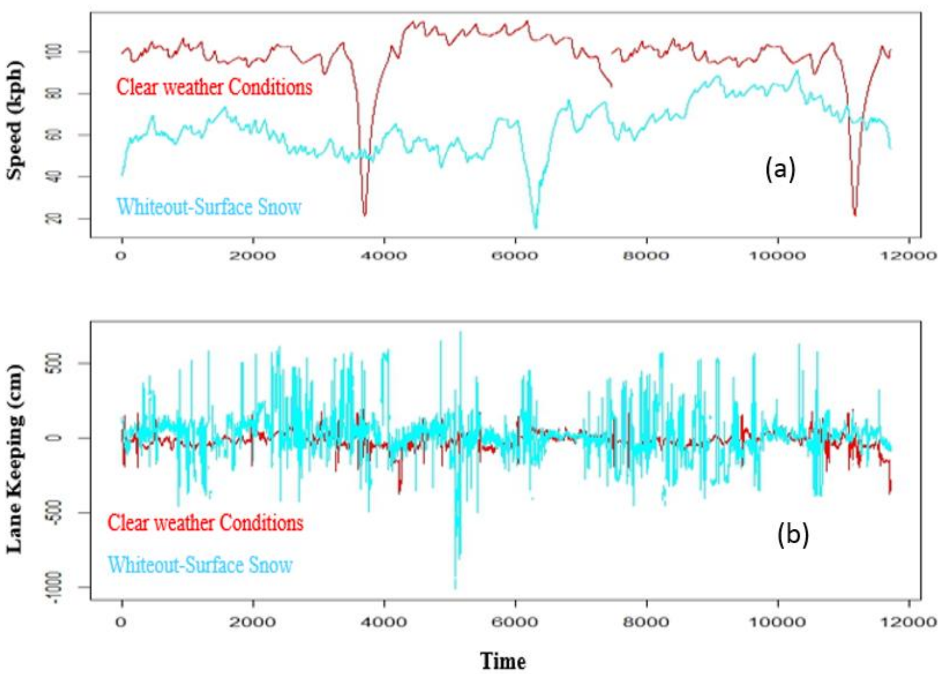


Figure 11 Time-Series Speed and Lane keeping Performance in Clear and Whiteout-Snowy Surface Condition*(43)*

Modeling Speed Selection: Ordered Logit Model

The ordered logit model was calibrated using all available data at the time of the preliminary report; representing a dataset of 10,606 one-minute segments of drivers' speed selections occurring in various weather and traffic conditions (matching is not required). The model was developed for four speed intervals based on the median of the Percent Speed Reduction above or below the speed limit ($\frac{(Speed - Speed\ Limit)}{Speed\ Limit}$) *(42)*. The four-quantile intervals were defined as: 1) more than 14 percent Speed reduction percentage, 2) Speed reduction percentage between 0 to 14 percent, 3) 0-10 percent increase in speed, and 4) more than 10 percent increase in speed. These intervals were used to achieve a sufficient sample size in each category of speed reduction. The remaining variables are exploratory variables, consisting of information extracted from questionnaires including driver demographics (age, marital status, gender, education) and driver experience, roadway factors, and observed environmental and traffic conditions.

To confirm the suitability and fitness of the model, the Log Likelihood Ratio (LR) was used. Table 4 shows the results of the model. The Multi-collinearity was assessed by calculating the Variance Inflation Factor (VIF) for each predictor, which indicates how much the variance of an estimated regression coefficient increases if the predictors are correlated. A VIF between 5 and 10 shows high correlation between predictors and VIF greater than 10 indicates that the regression coefficients are poorly estimated due to multi-collinearity *(44)*. The explanatory variables introduced to the model produced VIF values between 1.03 and 1.40, excluding any concerning multi-collinearity. Only statistically significant variables were retained in the final models.

Table 4 Estimation of Ordered Logit Model for Speed Selection in Different Weather Conditions*(42)*

Analysis of Maximum Likelihood Estimates									
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Odds Ratio	Confidence Interval	
Intercept	4	1	-2.57	0.09	800.41	<.0001	-	-	-
Intercept	3	1	-1.3	0.09	218.23	<.0001	-	-	-
Intercept	2	1	0.32	0.09	13.93	0.0002	-	-	-
Weather Cond.	Rain	1	0.44	0.09	25.35	<.0001	1.55	1.31	1.83
Weather Cond.	Snow	1	2.23	0.06	1612.52	<.0001	9.29	8.33	10.36
Weather Cond.	Fog	1	0.26	0.09	7.61	0.0058	1.29	1.08	1.55
Visibility	Affected	1	0.56	0.09	35.24	<.0001	1.75	1.45	2.1
Traffic Cond.	C-F	1	1.28	0.04	995.02	<.0001	3.6	3.32	3.89
Gender	Female	1	0.09	0.04	5	0.0254	1.09	1.01	1.18
Age	>40	1	0.2	0.05	18.24	<.0001	1.23	1.12	1.35
Marital Status	Divorced	1	0.81	0.09	86.57	<.0001	2.25	1.9	2.67
Marital Status	Widow(er)	1	1.2	0.11	121.33	<.0001	3.31	2.68	4.1
Marital Status	Unmrid-partnrs	1	-0.94	0.1	88.74	<.0001	0.39	0.32	0.48
Marital Status	Married	1	0.34	0.05	45.09	<.0001	1.4	1.27	1.55
Mileage Last Year	10,000 to 20,000	1	-0.5	0.05	122.3	<.0001	0.61	0.56	0.66
Mileage Last Year	>20,000	1	-0.58	0.06	92.33	<.0001	0.56	0.5	0.63

Adverse weather conditions – snow, rain, and fog – were found to have a significant effect on speed selection. Results showed that the odds of a driver reducing their speed were 9.29, 1.55, and 1.29 times higher for drivers travelling in snow, rain, and fog conditions, respectively, in comparison with drivers who were driving in clear weather conditions. Findings related to visibility indicated that the odds of a driver reducing their speed were 1.75 times greater for drivers who were driving in affected visibility conditions versus those driving in good visibility conditions. As expected, traffic conditions indicated a significant negative effect in speed selection. More clearly, the odds of having more speed reduction percentage were 3.6 times greater for drivers who were driving in higher traffic density compared to free flow speed (level of service A and B). Considering drivers' gender, findings indicated that the odds of a female driver reducing her speeds greater than male drivers was 1.09.

Modeling Speed Selection: Classification Tree Model

Classification can be defined as a procedure for predicting the class of an object – considering the object's features *(37)*. Classification models are built from a training dataset in which trends of predictor and target variables are identified and used to predict the value of the target variable for a new datasets *(38)*.

Figure 12 shows the decision tree diagram for the drivers' speed selection in different weather conditions based on the training data described in the previous section. In the node boxes, the node number and the percentage of data in each category are provided.

One beneficial characteristic of a decision tree, compared to other modeling methods, is that it gives decision makers rules to address ''if-then'' questions efficiently. The dataset introduced to this model included 10,606 one-minute segments with time-series vehicle data, weather conditions, driver demographics, and roadway characteristics data. The dataset contains four categories of drivers' speed selection behavior as mentioned before. Of the 10,606 one-minute segments, 60 percent were considered for training dataset, 20 percent were considered for validation, and the remaining 20 percent were used to test the model.

The misclassification rate, based on the training and validation datasets, indicated that the best tree could be obtained with 15 terminal nodes. More clearly, with 15 terminal nodes, the misclassification rate for the model reaches a minimum value of 0.42 and remains fairly steady. Node 3, on the right side, shows the data related to driving in snowy conditions. On the right branch of the tree, there are four terminal nodes (nodes 7, 13, 24, and 25). In three of these terminal nodes, the drivers were predicted to reduce their speed more than 14 percent (Class Label 1), which implies that, if a driver is travelling in snowy conditions, he/she will more likely reduce their speed, regardless of any other variables.

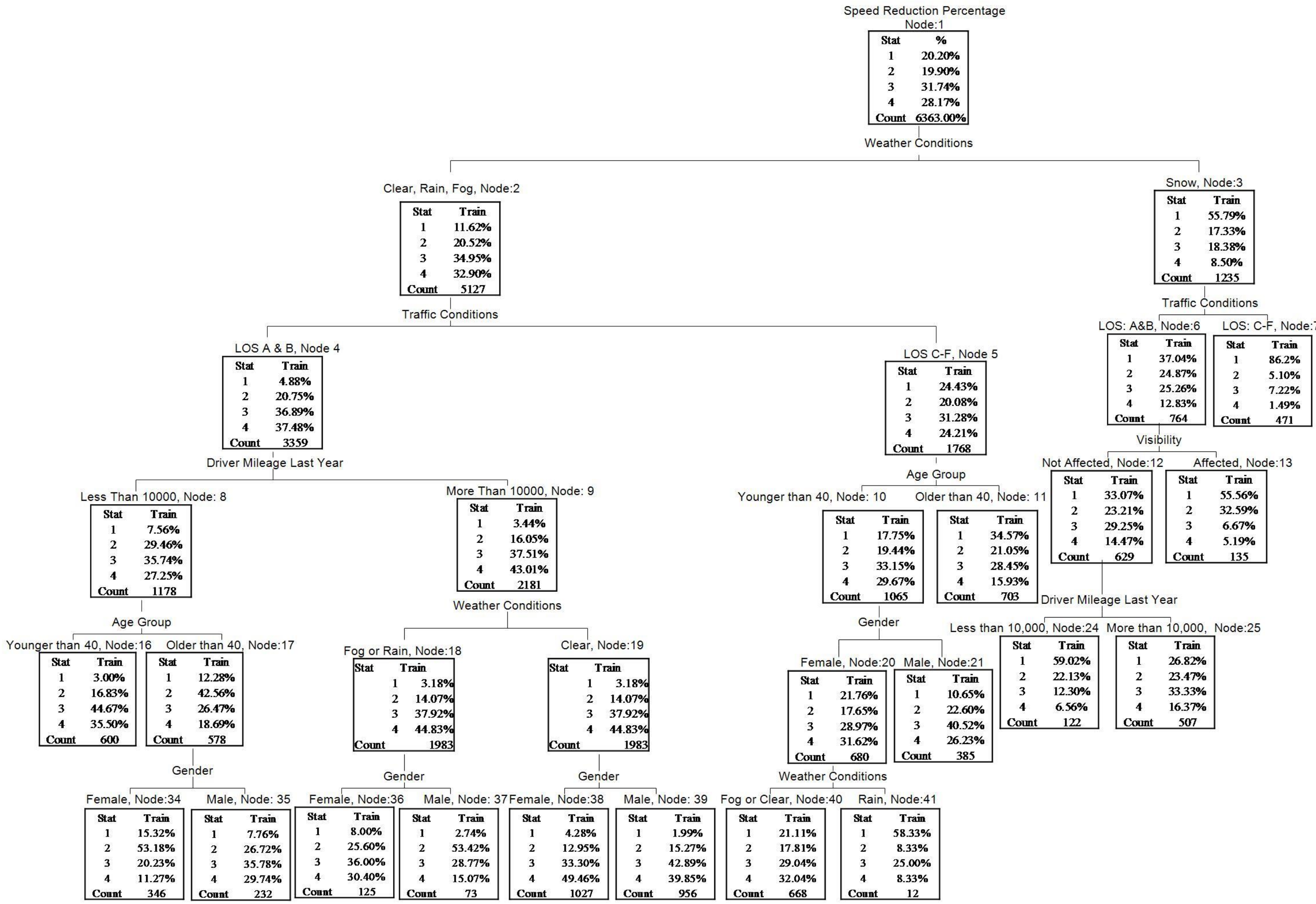


Figure 12 Classification tree diagram for Speed Selection Model*(42)*

As a function of traffic conditions, node 3 is split into node 6 and terminal node 7; terminal node 7 shows that when a driver travels in any level of traffic congestions (not a free flow speed) during snow covered road surface conditions, there is an 86 percent probability that the driver will reduces their speed more than 14 percent. Node 6 is further split into node 12 and terminal node 13 based on visibility conditions. Node 13 shows that drivers are 56 percent likely to reduce their speed more than 14 percent in snowy surface conditions, free flow traffic, and reduced visibility. Node 12 is split into node 24 and 25, based on driver mileage last year. Lastly, node 24 shows that if a driver, drove less than 20,000 miles last year, they were 59 percent more likely to reduce their speed more than 14 percent.

Summary & Next Steps

Both parametric logistic regression and non-parametric classification tree models were developed to better understand driver speed selection in different weather conditions, i.e., snow, rain, and fog. Each modeling technique has its advantages and disadvantages. While, the classification tree model can easily explain the complex interactions between several explanatory variables, it is difficult to fully describe the complicated effects of contributing factors due to non-linearity and the interaction effects in the logistic regression. On the other hand, using parametric logistic regression is beneficial in interpreting the marginal effect of risk factors. Therefore, it is justified to use both models to take the advantage of the benefits and compensate for the shortcomings of each method. Combined, the use of these parametric and non-parametric speed selection models provide a deeper understanding of speed selection behavior in adverse weather conditions. The focus of this analysis was not to show one model is superior to the other one, but it attempts to show how the two proposed complementary parametric and non-parametric approaches can help researchers provide better insights into the factors, which may affect drivers speed selection in adverse weather conditions.

The speed selection models revealed that among various adverse weather conditions, drivers are more likely to reduce their speed in snowy weather conditions. Specifically, the odds of drivers reducing their speed were 9.29 times higher in snowy weather conditions, followed by rain and fog with 1.55 and 1.29 times, respectively (compared to clear conditions). In addition, variable importance analysis using the classification tree method revealed that weather conditions, traffic conditions, and the posted-speed-limit are the three most important variables affecting driver speed selection behavior.

Selecting the appropriate driving speed for prevailing conditions is considered one of the most important driving tasks on high-speed facilities. Due to the previously limited understanding of the interaction between driver behavior/performance and weather conditions, the continuation of this research aims to establish a Connected Human-in-the-Loop VSL system, which is aligned with the SHRP2 Task Force's focus areas. An important component of the driver-weather interaction is the characterization of traffic flow because heterogeneity in driver behavior exists between adverse weather conditions and traffic flow conditions, meaning that driving behavior is different for different levels of congestion and weather conditions. Modeling variation in driver behavior with adverse weather conditions and traffic flow states is crucial to assigning effective VSLs, as these algorithms must consider the impact of both weather and traffic conditions when suggesting the safest and most efficient speed. An additional benefit from these developed models may be introduced in Connected Vehicle (CV) applications, where the VSL system could be expanded to incorporate mobile vehicle data as an input and to export VSL data to on-board units (OBU). The OBUs could then provide speed advisories, regulatory speeds, or other related advisories to the driver. Messages such as, “turn off cruise control”, could be sent in real-time to more effectively regulate driving speed and preserve a safe flow of traffic. If unusual traffic patterns are detected or inclement weather events are forecasted or experienced, these geospatial locations could be flagged for implementation of an appropriate and timely mitigation strategy to reduce the impact of the adverse weather condition.

Speed Selection in Fog

The negative effect of reduced visibility on driver performance has been recognized as one of the main causes of motor vehicle crashes in fog. Although many studies have concentrated on driver behavior during foggy weather in a simulated environment, there is a lack of studies that have addressed the impact of fog on driver behavior and performance in naturalistic settings.

Literature Review

According to the Federal Highway Administration (FHWA), roughly 15 percent of fatal crashes, 19 percent of injury crashes, and 23 percent of Property-Damage-Only (PDO) crashes occur in the presence of adverse weather, which results in approximately 5,100 fatal crashes, 304,800 injury crashes and 922,200 PDO crashes every year *(45).* Many studies have explored the impact of adverse weather on crashes and found that crash rate increases during inclement weather.

Driving in foggy weather is challenging due to reduced visibility, limited contrast, and distorted perception, which causes many accidents every year. From a visual perspective, fog can be described as a reduction in contrast in the visual field *(46)*. In fog, drivers face difficulty in perceiving speed, headway as well as road signs and markings, which are crucial for safe driving. Fog-related crashes tend to involve multiple vehicles and usually have more fatalities compared to crashes under clear weather conditions *(47).* A previous study showed that foggy conditions may increase crashes specifically in lack of road lighting *(48).* Moore and Copper found that drivers usually considered the leading vehicle as a mean of guidance and drove at a speed similar to the leading vehicle while driving in foggy weather. They stated this tendency to be the main cause of rear-end crashes in foggy weather *(49).*

Yan et al. investigated the effect of foggy weather conditions on driver speed control at different risk levels and found that at the high-risk level, driver speed compensation due to fog did not reduce their crash-involvement risk, though it effectively lowered the crash severity *(50).* A study conducted in a driving simulator environment investigated driver speed perception in foggy weather conditions. They found that drivers perceived their speed slower than the actual speed in foggy weather conditions *(50).* However, a number of studies using more sophisticated driving simulators contradict these results. For instance, Owens et al. showed that drivers tended to overestimate their speed and drove at a speed lower than instructed *(51).* In another study, based on 566 surveyed drivers in Florida, Hassan et al. concluded that drivers usually take shorter headways from lead vehicles during limited visibility conditions due to near fog. The study also mentioned that the VSL signs during fog cannot reflect the accurate safe speed limit due to the frequent fluctuation of fog thickness *(52).*

As mentioned in the previous section, the majority of the studies were mainly performed in a simulated environment or utilizing survey questionnaires *(50, 52–54)*.

There is a lack of studies that have examined the impact of fog on driver behavior and performance under naturalistic settings. The data used in this study was collected from the Second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS), which is the largest study on naturalistic driving behavior to date in the US. In addition, Roadway Information Database (RID) was also utilized. The RID was developed mainly to link roadway information to the NDS database.

This section will help in gaining insights into driver dynamics of adapting speed. It will also provide valuable information about how drivers interact with roadway and foggy weather conditions, which can be used for effective countermeasures.

The overall objective of this study is to assess driver behavior and performance during clear and fog weather conditions by utilizing the NDS and RID database. This will be attained by developing an appropriate NDS and RID data acquisition technique, then conducting a preliminary analysis between clear and foggy weather conditions, and finally developing a speed selection model to identify the major contributing factors that might influence driver speed selection under foggy weather conditions.

Data Preparation

A total of 124 trips in fog and 248 matched trips in clear weather conditions have been fully processed using the described data reduction procedure, which resulted in a total of 7,147 one-minute segments (2,549 in fog, 4,598 in clear weather). As mentioned earlier, this section is trying to provide better insights into driver speed selection under foggy weather conditions, which can lead to provide more realistic speed limits for the VSLs on freeways. Therefore, data were requested for freeways only. However, preliminary analysis showed that some trips have some non-freeway segments at the beginning and at the end of the trips (entering and exiting freeways). Therefore, non-freeway segments were removed from the start and end of these trips. Once the non-freeway segments were removed, the number of one-minute segments was reduced to 5,587 (i.e., 1,912 in fog and 3,675 in clear weather). A total of 62 drivers between 16 to 79 years of age participated in the selected NDS trips with more than 80 percent of drivers aged below 34 years. The quantity of male and female drivers were almost equal with a marginally higher percentage (56 percent) of male drivers.

Classification of fog

It is worth mentioning that the classification of fog is not consistent in the literature. The National Oceanic and Atmospheric Administration (NOAA) classified fog into two categories back in 1949 *(55)*. They classified fog as near if the visibility distance falls below 0.25 mile and light if the visibility distance is between 0.3 mile to 6 miles (0.48 to 9.7 km). In 1992, South Carolina Department of Transportation (SCDOT) developed a low visibility warning system, where they defined fog as dense if the visibility falls below 300 ft. (91 meters) and light if the visibility ranges between 300 ft. to 900 ft.(91 to 274 meters) *(56).* Another visibility warning system in Utah used a threshold of 200 ft. (61 meters) to define dense fog *(57).*

However, for this study, we classified fog into two categories including near fog and distant fog, using qualitative-based measures extracted from the NDS videos. Fog was classified based on the visibility of road markings, readability of road signs, roadside surroundings (delineators, guardrail, New Jersey barriers, etc.) and the horizon. The fog was reported as near fog if the video observers:

- Can only see few road markings in front of the NDS vehicle.
- Cannot read the information on the road signs.
- Cannot see the roadside surroundings and traffic ahead clearly.
- Cannot see the horizon.

On the other hand, the fog was classified as a distant fog if the video observers:

- Can see the road markings and read the information on the road signs.
- Can see the roadside surroundings and traffic ahead.
- Cannot see the horizon.

It is worth mentioning that video observers were trained comprehensively and provided with sample pictures of near fog and distant fog to minimize subjectivity. Figures below show some sample pictures of near fog and distant fog.

a b

Figure 13 Near Fog: a) Only one road marking is visible, the sign is unreadable, surroundings and the horizon cannot be seen properly. b) Few road markings are visible, surroundings, traffic, and the horizon cannot be seen clearly.

a b

Figure 14 Distant Fog: a) Road markings are visible, signs are readable, surroundings and traffic can be seen properly to some extent, the horizon cannot be seen clearly. b) Road markings are visible, the speed limit sign is readable, surroundings and traffic can be seen, horizon cannot be seen clearly.

Methodology

To better understand the factors affecting driver speed selection in different weather conditions an ordered logit model was developed. The model was calibrated utilizing a dataset of 5,587 one-minute segments occurring in various weather and traffic conditions. A Logit model has various advantages over other models. For instance, predictors in the logit model can be continuous, categorical, or a mixture of both. In addition, independent variables do not have to be normally distributed or have equal variance in each group *(58)*. Table 5 shows the summary of different variables used in the speed selection model. The response variable of the model is speed selection, which is classified into four levels, based on the median Percent Speed Reduction above or below the speed limit $\frac{(Speed\ limit - speed)}{Speed\ limit}$. The four-quartile intervals were defined as: 1) More than 10 percent increase in speed, 2) 0-10 percent increase in speed, 3) 0-28 percent reduction in speed, and 4) more than 28 percent reduction in speed. These intervals were selected based on quartile values in order to ensure sufficient sample size in each category. The remaining variables are explanatory variables including environmental variables, traffic conditions, driver demographics and roadway factors.

Table 5 Data Descriptions for Speed Selection Model *(59)*

Variable	Description	Type	Levels
Response Variable			
Speed Selection	Percent speed reduction above or below the speed limit	Ordinal	1 = More than 10 percent increase in speed
			2 = 0-10 percent increase in speed
			3 = 0-28 percent reduction in speed
			4 = More than 28 percent reduction in speed
Explanatory Variables			
Environmental Variables			
Weather Conditions	Predominant weather conditions in 1-min video observation	Categorical	1 = Clear
			2 = Distant fog
			3 = Near Fog
Visibility	Predominant visibility conditions in 1-min video observation	Categorical	1 = Not Affected
			2 = Affected
Surface Conditions	Predominant surface conditions in 1-min video observation	Binary	1 = Dry
			2 = Wet
Traffic Variables			
Traffic Condition	Predominant traffic conditions in 1-min video observation	Binary	1 = Free Flow
			2 = Mixed Flow
Speed Limit	Predominant speed limit conditions in 1-min segment	Categorical	1 =<55
			2 = 55-60
			3 = 65-70
Driver Demographics			
Gender	The gender the participant	Binary	1 = Male
			2 = Female
Age	The age group corresponding to the driver's birthdate	Categorical	1 = Less than 40 years
			2 = Greater than 40 years
Education	The highest completed level of education of the participant	Categorical	1 = High school diploma or G.E.D.
			2 = Some education beyond high school but no degree and College degree
			3 = Some graduate or professional school, but no advanced degree and Advanced degree (e.g., J.D.S., M.S. or Ph.D.)
Marital Status	The participant's marital status	Categorical	1 = Single
			2 = Married
			3 = Other (Divorced, Widow, Unmarried Partners)
Driver Mileage Last Year	The approximate number of miles the participant drove last year	Categorical	1 = Less than 10,000 miles
			2 = Between 10,000 to 20,000 miles
			3 = Greater than 20,000 miles
Driving Experience	Number of years driving experience	Categorical	1 = Less than 10 years
			2 = Greater than 10 years
Roadway Factors			
Bridge	Presence of bridge	Binary	1 = No bridge
			2 = On bridge
Curve	Presence of curve	Binary	1 = No curve
			2 = On curve
Superelevation	Average superelevation of the road in 1-min segment	Continuous	-
Curve Length	Average length of curve of the road in 1-min segment	Continuous	-

Analysis

Preliminary Analysis: Summary Statistics

Traffic conditions were categorized into free-flow (Level of service A and B) and mixed traffic condition (Level of service C to F). According to Highway Capacity Manual (HCM), free-flow is defined as low volume roadway conditions, where drivers are free to drive at desired speed and not constrained by the presence of other vehicles *(25)*. In this study, a trip was considered as a free-flow speed when the NDS driver has no leading traffic in any lanes or when a leading vehicle is present at least in one lane, but the NDS driver is still not affected by other vehicles. Other conditions where NDS drivers were affected by other vehicles were considered as mixed traffic conditions.

Each trip in fog was matched with two trips in clear weather considering the same driver, same vehicle, and same location. Matched trips were requested from the VTTI. As mentioned in a previous study, weather conditions may not be consistent within a trip *(15)*. Therefore, considering the possible variations in weather conditions, exact match of the one-minute segments in fog and clear weather was conducted by importing the longitude and latitude of the trips and eliminating non-matching segments in the ArcGIS software.

Removing the non-matching segments resulted in 5,398 one-minute matching segments (1,867 in fog, 3,531 in clear weather), which was equivalent to nearly 90 hours of driving time and 5179 miles (8335 kilometers) of travelled routes. The summary statistics of these 5,398 one-minute segments are provided in Table 6.

Table 6 Summary Statistics of NDS Trips Considered in this Section *(59)*

	Weather Condition	Near Fog	Matched Clear	Distant Fog	Matched Clear	Total
Free-Flow Condition LOS A & B	**Number of One-minute Segments**	241	539	717	1467	2964
	Total Duration (Hour)	4.02	8.98	11.95	24.45	49.40
	Total Length (km)	418.59	973.72	1260.29	2637.26	5289.86
Congested Traffic Condition LOS C to F	**Number of One-minute Segments**	271	405	638	1120	2434
	Total Duration (Hour)	4.52	6.75	10.63	18.67	40.57
	Total Length (km)	340.35	542.85	750.54	1411.11	3044.85
Total Number of One-minute Segments		512	944	1355	2587	5398
Total Duration (Hour)		8.53	15.73	22.58	43.12	89.97
Total Length (km)		758.94	1516.57	2010.83	4048.37	8334.71

Speed Distribution
This study investigated the distribution of speeds between clear and foggy weather conditions in various traffic states. From the NDS sample data, it was concluded that the speeds have a Weibull distribution in near fog under free-flow conditions while the speeds were normally distributed in clear weather for the matching dataset. A similar trend was also found for distant fog as shown in Figure 15. Speed in free-flow conditions is important for VSL application since speed selection here is not affected by the interaction with traffic *(60)*. Other speed distributions for other scenarios were also examined. Speed distribution during near fog in mixed traffic conditions did not fit a specific distribution. However, speeds during distant fog as well as their matched clear weather conditions in mixed traffic fitted a bimodal distribution, which is common during congestion on freeways *(61).* Figure 15 shows the speed distribution for trips in near fog, distant fog and matched trips in clear weather under free-flow and congested (i.e., mixed/near traffic) traffic conditions.

Descriptive Analysis
Driver speed behavior including selection of speeds and accelerations under free-flow conditions were investigated in clear and foggy weather in order to have a better understanding of driver behavior in different weather conditions. Various statistical tests, including t-test, F-test, and Z-test were used to compare driver behavior between foggy and clear weather as shown in Table 7. A t-test indicated that the average speed in near fog, as well as in distant fog, was significantly lower than in clear conditions under free-flow traffic. However, speed reduction was more in near fog compared to distant fog. Speed in near fog and distant fog was found to be 6.8 mph (11 kph) (10 percent reduction) and 1.8 mph (3 kph) (2.8 percent reduction) lower than the speeds in their matching clear weather conditions, respectively. Previous studies have also concluded similar results of speed reduction. For instance, Liang et al. found 3.6-6.2 mph (6 to 10 kph) speed reduction due to the poor visibility caused by fog *(62).* It was also found that speeds had a higher variability during distant fog compared to clear weather under free-flow traffic, which could be an indication of increased safety risk *(30, 63).* However, an opposite trend was found for near fog where speeds in clear weather conditions had more variability.

In addition, the acceleration/deceleration variability was also examined, and ± 0.3g acceleration/deceleration rates were set as a threshold to identify aggressive braking/acceleration events *(64).* However, no acceleration and deceleration were found to be higher or lower than ± 0.3g, indicating the occurrence of zero aggressive events as shown in Table 7.

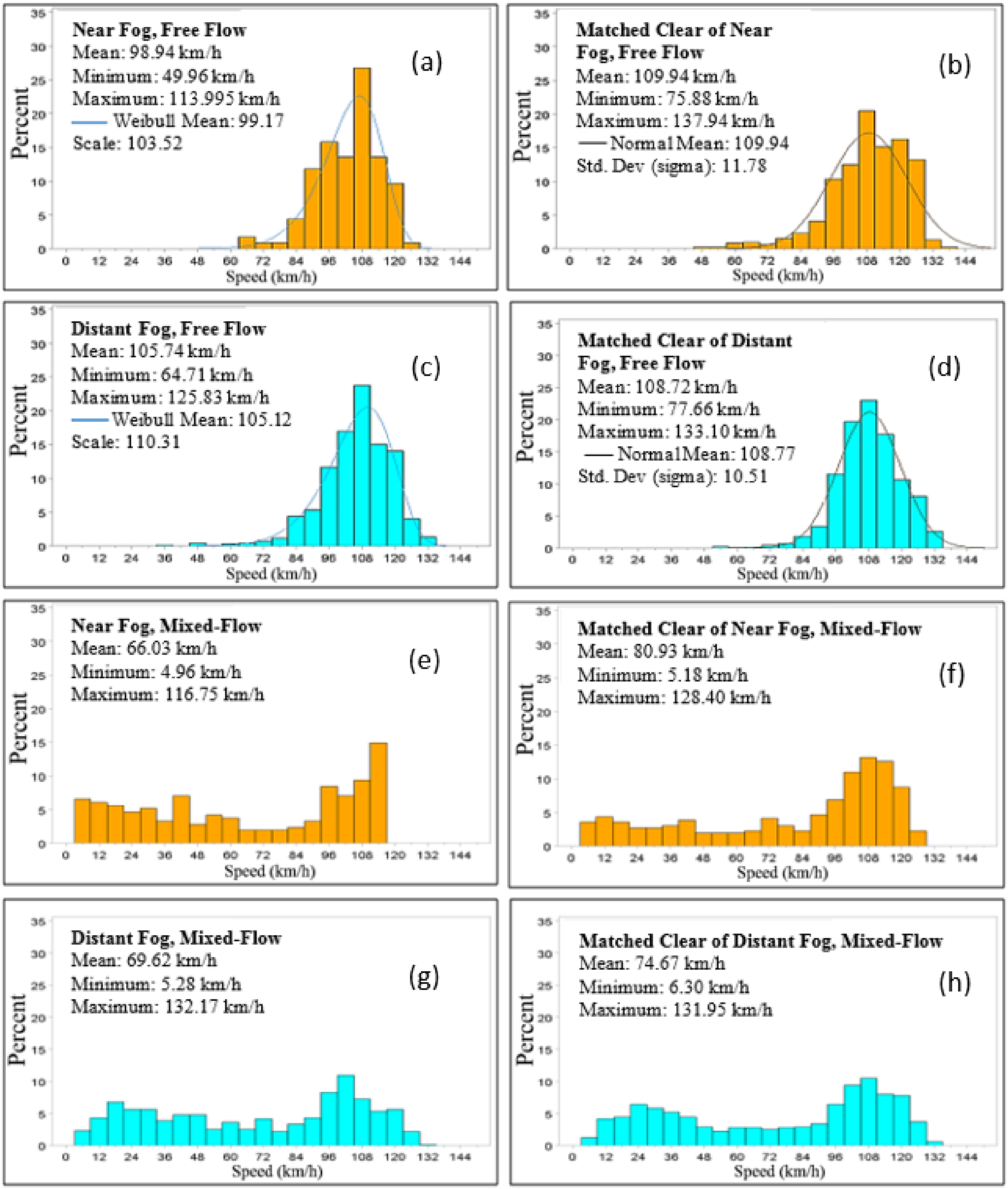


Figure 15 Observed and Fitted Distributions for Speeds during Fog and Clear Weather under Free-Flow and Mixed Traffic Conditions *(59)*

Table 7 Descriptive Statistics for NDS Instrumented Vehicles in Fog *(59)*

	Statistical Test	Free-Flow Traffic							
		Near Fog		Matched Clear		Distant fog		Matched Clear	
		Speed (km/hr)	% Speed Difference from Speed Limit	Speed (km/hr)	% Speed Difference from Speed Limit	Speed (km/hr)	% Speed Difference from Speed Limit	Speed (km/hr)	% Speed Difference from Speed Limit
Speed (km/hr)	Average	98.944	-1.556	109.944	-8.420	105.739	-4.213	108.722	-7.803
	SD	10.436	10.141	11.785	9.831	11.087	10.070	10.511	8.558
	Min.	49.957	-21.360	75.877	-42.627	64.707	-33.525	77.660	-42.104
	Max.	113.995	43.548	137.942	27.530	125.832	33.385	133.097	26.683
	Median	100.313	-2.754	110.195	-8.100	107.018	-5.005	108.887	-7.847
	t-test	Average speed is significantly higher in clear weather. $t(317) = -11.15$, $P<0.05$ Effect size (Cohen's d) = -0.92				Average speed is significantly higher in clear weather. $t(2009) = -5.84$, $P<0.05$ Effect size (Cohen's d) = - 0.29			
	F-test	Speed variability is significantly higher in clear weather. $F_{523,201} = 1.32$, $P<0.05$				Speed variability is significantly higher in distant fog. $F_{648,1361} = 1.11$, $P<0.05$			
	Z-test	Proportion of violation ≥ 10 km/hr above the speed limit is significantly higher in clear weather. $Z = -5.73$, $P<0.05$				No significant difference between the proportion of speeding ≥ 10 km/hr in distant fog and clear weather. $Z = 0.69$, $P > 0.05$			
Acceleration/ Deceleration (g)		Acc. (g)	Dec. (g)	Acc. (g)	Dec. (g)	Acc. (g)	Dec. (g)	Acc. (g)	Dec. (g)
	Average	0.017	-0.013	0.016	-0.012	0.016	-0.019	0.017	-0.016
	SD	0.017	0.014	0.017	0.0140	0.014	0.016	0.013	0.015
	Min.	0.000	-0.061	0.000	-0.067	0.000	-0.070	0.000	-0.067
	Max.	0.082	0.000	0.079	-0.000	0.065	0.000	0.072	0.000
	Median	0.011	-0.009	0.011	-0.006	0.011	-0.012	0.013	-0.011
	t-test	Average acceleration is significantly higher in clear weather. $t(453) = 0.13$, $P<0.05$ Effect size (Cohen's d) = -0.02 No significant difference between the average deceleration in near fog and clear weather. $t(218) = -0.80$, $P>0.05$ Effect size (Cohen's d) = -0.12				No significant difference between the average acceleration in distant fog and clear weather. $t(1191) = -1.14$, $P>0.05$ Effect size (Cohen's d) = -0.10 No significant difference between the average deceleration in distant fog and clear weather. $t(816) = -1.60$, $P>0.05$ Effect size (Cohen's d) = -0.12			
	F-test	No significant difference between the acceleration variability in near fog and clear weather. $F_{1011,362} = 1.09$, $P>0.05$ No significant difference between the deceleration variability in near fog and clear weather. $F_{86,144} = 1.03$, $P>0.05$				No significant difference between the acceleration variability in distant fog and clear weather. $F_{384,807} = 1.06$, $P>0.05$ No significant difference between the deceleration variability in distant fog and clear weather. $F_{263,553} = 1.14$, $P>0.05$			
	Z-test	No acceleration/ deceleration were found higher/lower than ± 0.3g				No acceleration/ deceleration were found higher/lower than ± 0.3g			

Speed Selection

Speed above the speed limit in fog and respective matching trips in clear weather were examined to determine driver compliance to the speed limit in different weather conditions. It was found that NDS drivers drove consistently above the speed limit in all conditions including near fog. For instance, NDS drivers drove with a speed 1.6 percent above the speed limit in near fog; whereas in clear weather NDS drivers drove with a speed 8.4 percent above the speed limit. Similar results were also found for distant fog. A Z-test as shown in Table 7Table 8 indicates that the violation of speed greater than 6.2 mph (10 kph) was significantly lower in near fog compared to matching trips in clear weather. However, no significant difference was found between the violation of speed greater than 6.2 mph (10 kph) in distant fog and corresponding matched clear trips, indicating no effect of distant fog on speeding behavior.

According to Table 8 the majority of the drivers drove with a speed above the limit. For instance, speeds of about 66 percent of the trips in near fog and almost 83 percent of the trips in matching clear weather were above the speed limit as shown in Table 8. Similarly, about 71 percent of the trips in distant fog and 82 percent of the trips in respective clear weather were driven at speeds more than the speed limit. Table 8 also indicates that speed reduction was more likely to occur in foggy weather conditions in comparison with the matched trips in clear weather conditions. The odds ratios of driving below the speed limit, in general, were 2.4 and 1.9 times more likely to be in near fog and distant fog respectively, than matching trips in clear weather conditions.

Table 8 Odds Ratio for Speed Behavior*(59)*

Weather Condition	Driving below Speed Limit	Driving above Speed Limit	Odds Ratio	Confidence Interval	Significance level
Near Fog	77 (33.7 %)	151 (66.3%)	2.42	1.69 to 3.45	P < 0.0001
Matched Clear of Near Fog	386 (17.4%)	1525 (82.6%)			
Distant fog	196 (29%)	480 (71%)	1.90	1.54 to 2.36	P < 0.0001
Matched Clear of Distant fog	253 (17.7%)	1180 (82.3%)			

Speed Selection Model Results

The log likelihood ratio (LR) was used to confirm the fitness of the model. The LR test statistic as shown in Table 9 falls into the rejection region with a p-value < 0.05, which indicates the overall explanatory variables of the model have significant effect on the response at a statically significant level of 95 percent. To check the possible presence of multicollinearity, Variance Inflation Factor (VIF) was calculated for each predictor. The VIF measures how much the variance of an estimated regression coefficient increases if predictors are correlated. A VIF between 5 and 10 shows a high correlation between predictors and a VIF greater than 10 indicates that the regression coefficients are poorly estimated due to multicollinearity *(65).* However, the VIF value of all the predictors in the speed selection model fell below 2.5, indicating no multicollinearity problem. Only the statically significant variables were retained in the final model. Table 9 shows the results of the speed selection model.

Table 9 Estimation of Ordered Logit Model for Speed Selection *(59)*

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	P-value	Odds Ratio	95% confidence Interval	
Intercept	**4**	1	-1.4876	0.2516	34.9552	<.0001	0.226	-	-
Intercept	**3**	1	0.00174	0.2508	0.0000	0.9945	1.002	-	-
Intercept	**2**	1	1.4051	0.2509	31.3594	<.0001	4.076	-	-
Weather	**Clear**	1	-	-	-	-	-	-	-
	Distant Fog	1	0.2528	0.0741	11.6346	0.0006	1.288	1.113	1.490
	Near Fog	1	0.2726	0.1234	4.8808	0.0272	1.313	1.031	1.673
Visibility	**Not Affected**	1	-	-	-	-	-	-	-
	Affected	1	0.4953	0.1377	12.9452	0.0003	1.641	1.253	2.149
Surface Condition	**Dry**	1	-	-	-	-	-	-	-
	Wet	1	0.7642	0.1878	16.5508	<.0001	2.147	1.486	3.103
Traffic Condition	**Free-Flow (A-B)**	1	-	-	-	-	-	-	-
	Mixed-Flow (C - F)	1	1.9217	0.0585	1080.4932	<.0001	6.833	6.093	7.662

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	P-value	Odds Ratio	95% confidence Interval	
Speed Limit	**< 55 mph**	1	-	-	-	-	-	-	-
	55-60 mph	1	-1.4557	0.2215	43.1710	<.0001	0.233	0.151	0.360
	65-70 mph	1	-1.8661	0.2238	69.5241	<.0001	0.155	0.100	0.240
Gender	**Male**	1	-	-	-	-	-	-	-
	Female	1	-0.2529	0.0580	18.9851	<.0001	0.777	0.693	0.87
Age	**< 40 years**	1	-	-	-	-	-	-	-
	> 40 years	1	0.1628	0.0801	4.1344	0.0420	1.177	1.006	1.377
Education	**High School**	1	-	-	-	-	-	-	-
	Beyond High School	1	0.3776	0.1107	11.6391	0.0006	1.459	1.174	1.812
	Advance degree	1	0.9529	0.1267	56.5368	<.0001	2.593	2.023	3.324
Marital Status	**Single**	1	-	-	-	-	-	-	-
	Married	1	0.2804	0.0794	12.4663	0.0004	1.324	1.133	1.547
	Others	1	-0.3879	0.0936	17.1765	<.0001	0.678	0.437	0.741
Driver' s Mileage Last Year	**< 10000 miles**	1	-	-	-	-	-	-	-
	10,000 – 20,000 miles	1	-0.5999	0.0731	67.4055	<.0001	0.549	0.475	0.633
	> 20,000 miles	1	-0.3879	0.0936	17.1765	<.0001	0.678	0.565	0.815
Drivers Experience	**< 10 years**	1	-	-	-	-	-	-	-
	> 10 years	1	0.5890	0.0934	39.7381	<.0001	1.802	1.500	2.165
Bridge	**No**	1	-	-	-	-	-	-	-
	Yes	1	0.7457	0.2241	11.0670	0.0009	2.108	1.358	3.27
Super elevation	-	1	0.0258	0.0113	5.2036	0.0225	1.026	1.004	1.049
Curve length	-	1	- 0.00113	0.000244	21.4193	<.0001	0.999	0.998	1.000
Speed Limit ×Curve Length	**55-60 mph, Curve Length**	1	0.000983	0.000247	15.8866	<.0001	0.457 At, Avg. Curve length = 683.46 m	0.327	0.638
	65-70 mph Curve Length	1	0.00112	0.000246	20.9590	<.0001	0.334 At, Avg. Curve length = 683.46 m	0.237	0.469
Drivers Experience ×Visibility	**> 10 years, Affected**	1	-0.6644	0.1646	16.2971	<.0001	0.927 At, affected visibility	0.664	1.295
Curve × Weather	**Curve × Near Fog**		0.3835	0.1897	4.0838	0.0433	1.467 At Near Fog	1.012	2.128

Fit Statistics:

Likelihood Ratio Test: χ^2 = 2284.36, Df = 24, P-value < 0.0001
Score Test for the Proportional Odds Assumption: χ^2 = 1313.20, Df = 48, P-value < 0.0001
Akaike Information Criterion (AIC) = 13254.61
-2 Log L = 13200.614

Discussion of Key Factors

Fourteen variables and three interaction terms were found to be significant in the speed selection model. As expected, fog had a significant effect on speed selection. Results showed that drivers were likely to travel at significantly lower speeds during foggy weather conditions; more specifically the odds of drivers reducing their speeds were 1.31 and 1.28 times higher for drivers traveling in near fog and distant fog respectively, in comparison with drivers who were driving in clear weather conditions. Driving over the speed limit could be hazardous especially during inclement weather conditions including fog; because drivers might not have enough time to respond to or mitigate an unexpected event *(66).* This study showed that drivers reduced their speeds to compensate for the negative effect of fog on the primary driving tasks.

Findings related to visibility indicated that the odds of drivers reducing their speeds were 1.64 times greater for drivers who were driving in affected visibility versus those driving in good visibility conditions. A similar result was also found for surface conditions. Wet surface was found to have a significant impact on speed reduction. More clearly, the odds of drivers reducing speeds on wet surfaces were 2.15 times higher compared to dry surfaces.

Traffic conditions had a positive coefficient as expected. Controlling for all other variables, drivers were 6.83 times more likely to reduce their speeds in mixed traffic conditions (level of service C to F) compared to free-flow conditions (level of service A and B). It was found that female drivers were less likely to reduce their speed compared to male drivers. More clearly, the odds of female drivers reducing their speeds were 1.29 times less compared to male drivers (OR = 0.777). As expected older drivers had more speed reduction compared to young drivers. More specifically, the odds of having more speed reduction percentage were 1.18 times higher for drivers older than 40 years compared to drivers of 40 years of age or younger. Education level also came out to be significant in the model. It was found that with the increase of education level, drivers became more compliant with the speed limit. For instance, the odds of a driver with an advanced degree were 2.59 times more likely to reduce speed compared to a driver who is a high school graduate. Marital status was also found to be significant with a usual trend of married drivers being the safest compared to single drivers *(67)*. More clearly, married drivers were 1.32 times more likely to reduce speed compared to single drivers.

Several factors related to the roadway, including the presence of bridge, superelevation and curve length were found to have a significant effect on driver speed selection. Considering interaction terms, it was found that at an affected visibility conditions, experienced drivers (driving experience > 10 years) were 10 percent less likely to reduce speed compared to less experienced drivers (driving experience < 10 years), which indicates experienced drivers are usually more confident in reduced visibility compared to less experienced drivers *(68)*. Similarly, the interaction between weather conditions and curves indicated that the drivers were 1.47 times more likely to reduce their speed on curves compared to their speed on tangents during near fog.

Summary

The main focus of this study was to attain better insights into driver behavior in general and speed selection in particular during clear and foggy weather conditions using the SHRP2 NDS dataset. The preliminary analysis showed a Weibull speed distribution in near fog under free-flow conditions while the speeds were normally distributed in clear weather for the matching dataset (i.e., same vehicle, driver, route, and traffic state). Descriptive analysis indicated about 10 percent reduction in speed during near fog and about 3 percent reduction in speed during distant fog. The results from the ordered logit model revealed that weather-related factors including the presence of fog, visibility, and surface conditions have a significant impact on driver speed selection behavior. For instance, results showed that drivers were more likely to select significantly lower speeds during foggy weather conditions. More specifically, the odds of drivers reducing their speeds from the posted speed limit were 1.31 and 1.28 times higher for drivers traveling in near fog and distant fog respectively compared to drivers who were driving in clear weather conditions. As mentioned before, the majority of the participants in the SHRP2 NDS were young (39 percent of the NDS drivers were below 25 years old). Considering the fact that the data used in this study is representing the age distribution of the actual SHRP2 NDS data, a more representative sample of age groups might provide different results.

Foggy weather conditions can negatively affect driver speed perception and ability to see objects on the roadway, which is one of the main causes of rear-end and lane departure crashes on freeways. Advanced Driver Assistance Systems (ADAS) such as Adaptive Cruise Control (ACC), Collision Avoidance Systems (CAS), Collison Warning System (CWS), Dynamic Brake Support (DBS), Autonomous Emergency Braking (AES), Lane Departure Warning (LDW), etc. are currently being used to improve roadway safety in different adverse weather conditions including fog. The main focus of these technologies is to prevent crashes by detecting a conflict, alerting drivers and aiding in taking appropriate and timely actions. For instance, the ACC can assist drivers to keep a safe distance from a lead vehicle while maintaining a chosen speed *(69)*. However, most of the ADAS are based on Machine Vision Techniques, which might be inefficient during adverse weather conditions due to the difficulty of detecting objects in adverse weather. In comparison, sensor based technologies are usually more effective as they are less susceptible to adverse weather. As the ACC works based on ultrasonic, laser, or LiDAR sensors, it is more effective in adverse weather conditions including fog compared to other machine based ADAS.

Evaluating driver behavior and performance under the influence of reduced visibility due to foggy weather conditions is extremely important to developing safe driving strategies, including Variable Speed Limits (VSL). Many roadways across the US currently have weather-based VSL systems to ensure safe driving environments during adverse weather. Current VSL systems mainly collect traffic information from external sources, including inductive loop detector, overhead radar and Closed Circuit Television (CCTV). However, human factors especially driver behavior and performance such as selection of speed and acceleration during adverse weather are neglected due to the lack of appropriate driver data. The SHRP2 NDS database has huge potential in becoming a good source for driver data. The findings from this study indicated that the NDS data could be effectively utilized to identify trips in foggy weather conditions and to assess the impacts of fog on driver behavior and performance.

The results from this section provided insights into incorporating naturalistic speed selection behavior in Variable Speed Limit systems. While the vast majority of VSL systems are based on Road Weather Information System (RWIS) data, previous studies noted many limitations of these systems. Utilization of 1-minute real-time weather and surface conditions, and visibility limits may improve VSL logics significantly. With the evolution of connected vehicles, Machine Vision and other real-time weather social networks such as WeatherCloud, more accurate real-time data similar to the NDS data will be available in the near future. This study provided early insights into using similar data collected from NDS.

Car-Following Behavior

Car-following is a crucial element of driving behavior as it describes the interaction between vehicles in a shared lane. Car-following defines the longitudinal motion of a following vehicle, as it perceives a lead vehicle in its path. As opposed to free-flow conditions, where a driver selects his/her acceleration and deceleration to match their desired speed, car-following conditions entail the adjustment of a driver's behavior in recognition of another road-user. Car-following behavior is a key process in microscopic simulation models—as well as in the fundamental understanding of traffic flow theory—which attempts to bridge the gap between driver-level behavior and macroscopic network-level outputs *(70)*. Car-following models have evolved from basic understandings to sophisticated and complicated models over the past half-century, resulting in hundreds of deviations of different car-following models *(71)*.

Building from the state-of-research, the Wyoming SHRP2 NDS dataset is used to calibrate the Gipps car-following model in clear and adverse conditions using the matching trip sets described in the introduction of research finding section. The following sections provide a brief literature review, describe the data preparation procedures required to identify car-following behavior from the SHRP2 NDS, summarize the methodology used to understand deviations in car-following behavior for adverse and clear weather conditions, and present the analytic results. Finally, the implications of the car-following research conducted as part of the Wyoming IAP Phase 2 is discussed and next steps for implementation in the third phase are described.

Literature Review

A variety of car-following models have been used to investigate the impact of adverse weather conditions on driving behavior. Hoogendoorn et al. examined the impact of fog conditions on driving behavior using a psychophysical plane (i.e., the evaluation of driving behavior relative to the following distance and relative velocity between vehicles) *(72)*. The study utilized a driving simulator that induced fog conditions to compare the location of "action points" (i.e., instances defined by relative speed and following distance where the driver reacted to a lead vehicle) in clear and adverse conditions. Results indicated that in reduced visibility, the location of the action points were significantly more dispersed than those in clear conditions. In addition, drivers appeared to be less perceptive to small changes in relative velocity in adverse conditions *(72)*.

In a report prepared for the FHWA, Rakha et al. discussed the findings from an extensive study aiming to measure the traffic impact of inclement weather conditions using microscopic (i.e., trajectory-level) data *(73)*. As part of this study, car-following models were compared and the Gipps car-following model was identified as being highly flexible, capable of capturing driver behavior under multiple regimes and across multiple road types; therefore, it was suggested for use in weather-related car-following model calibration. Additional studies indicate similar results using NDS data from the 100-Car Study *(74)* and trajectory-level data from an instrumented research vehicle (IRV) *(75)*.

Data Preparation

Capturing car-following behavior using NDS trajectory-level data requires the use of the forward facing radar, vehicle velocity, and vehicle acceleration. In addition, forward-facing video was used to confirm expected conditions. For this analysis, segments of continuous car-following behavior (car-following events) from a single follower-leader pair were extracted based on the following criteria:

- The car-following event lasted a minimum of 20 seconds;
- The following distance did not surpass 197 ft. (60 meters);
- The minimum speed of the subject (following) vehicle was greater than 3.2 ft/hr (1 m/s).

A minimum car-following event length was selected based on a review of literature which showed a range from 5 seconds to 30 seconds by which research teams extracted following behavior *(74, 76)*. For this research, a length of 20 seconds was selected to ensure behavioral continuity without severely limiting the number of car-following events. The maximum threshold for following distance was identified from the VTTI technical guidance, which indicated that radar quality decreased at measured distances greater than 60 meters. Lastly, a minimum speed was identified to ensure events occurred during expected freeway conditions, in which a vehicle never comes to a complete stop.

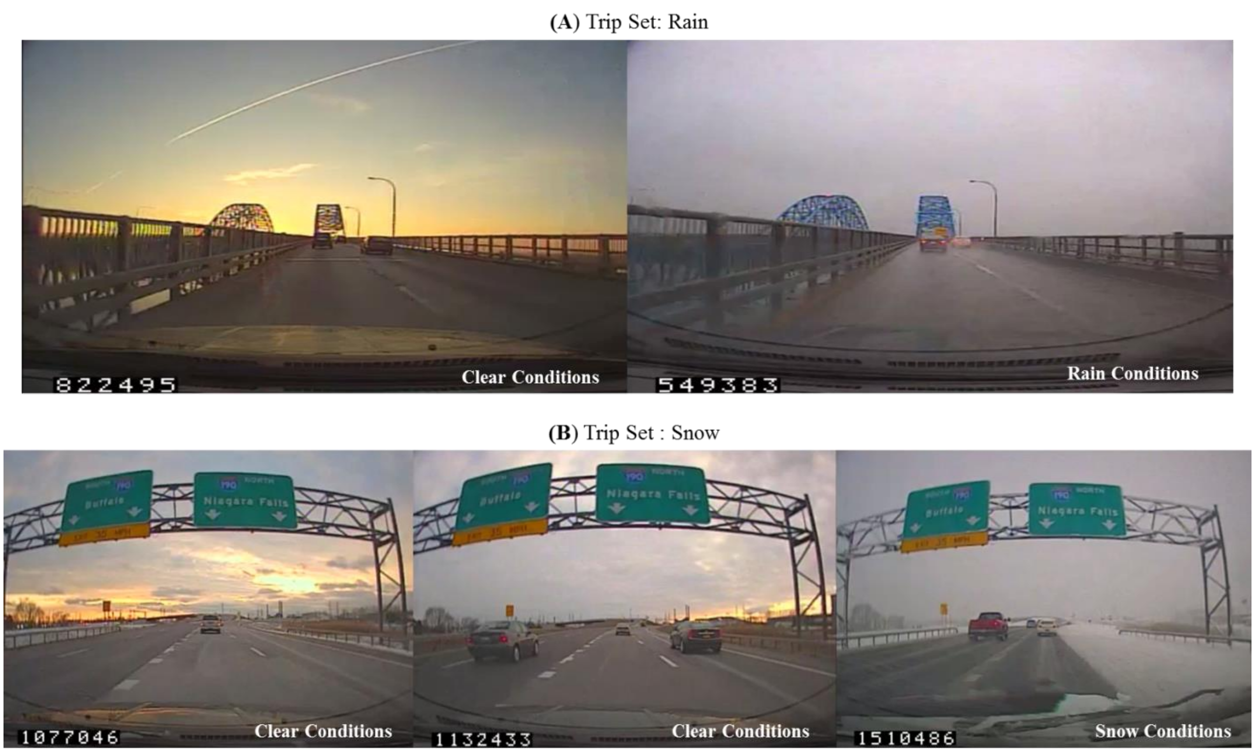


Figure 16 (A) Trip set that includes one trip in clear conditions (A, left) and one trip in rain conditions (A, right). (B) Trip set that includes two trips in clear conditions (B, left and middle) and one trip in snow conditions (B, right) *(77)*

As described in previous sections, the car-following analysis also leveraged the trip sets, which include matching sets of trips with a single driver in differing weather conditions (one trip in adverse conditions and two trips in clear conditions). Evaluation on the basis of trip sets enabled the research team control over inter-driver heterogeneity (or differences between different drivers), while focusing on weather induced intra-driver heterogeneity (or differences between driving behavior in different environmental conditions). When evaluating driver heterogeneity, many external and internal factors impact driver behavior; however, with the use of matching trip sets with the same drivers, same routes, and often the same time of day, isolation of behavioral changes due to weather conditions is highlighted.

As mentioned, identification of car-following events requires the use of the vehicle network speed (from CAN-Bus), vehicle acceleration (from CAN-Bus), following distance (from radar unit), and relative velocity from lead vehicle (from radar unit). In efforts to increase the usability of the radar data, VTTI processed the radar data to sync the radar fields with the time-series fields reported from the CAN-Bus and GPS units, as well as address other challenges related to inherently noisy radar data. Once received and evaluated, Wyoming's research team conducted one additional processing step using a moving average filter to smooth the reported relative velocity values (as the relative velocity readings were significantly more noisy than those of the following distance) *(77)*.

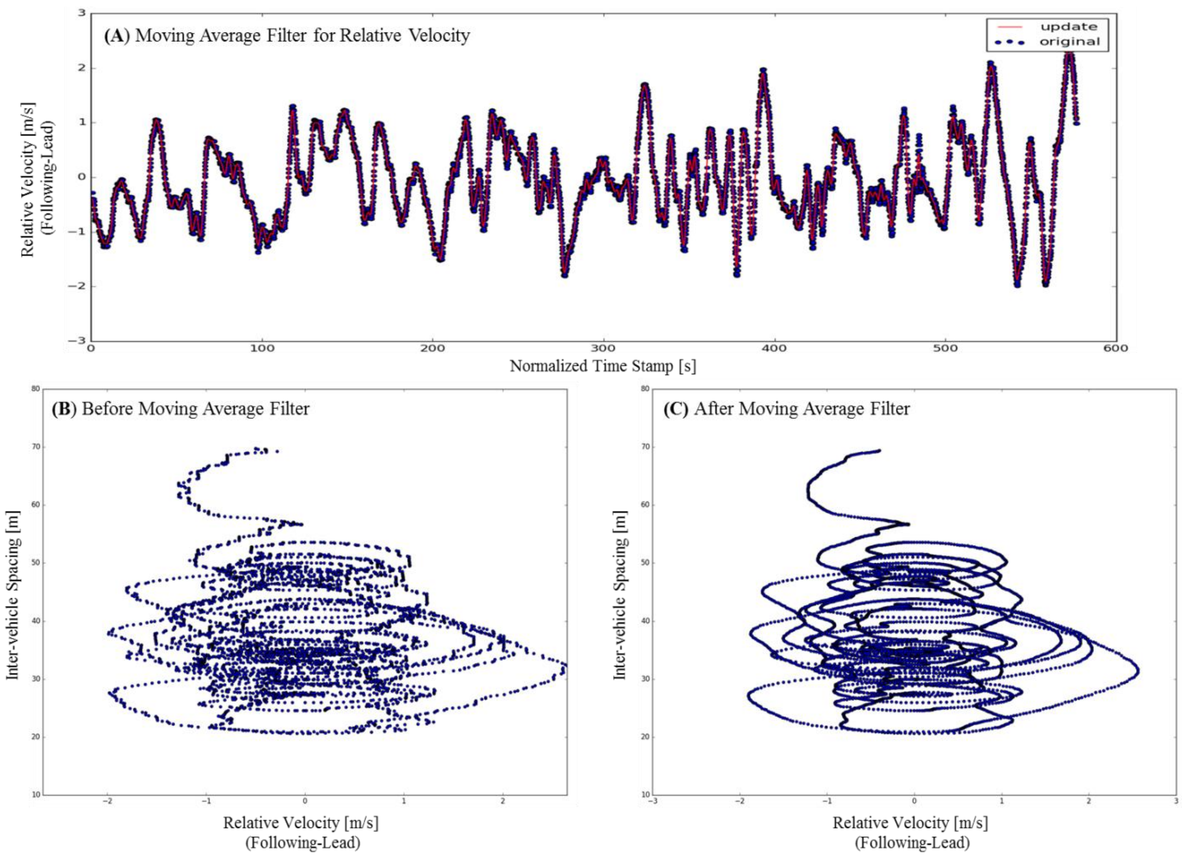


Figure 17 Effect of Moving Average Filter. (A) Plot A shows the relative velocity in m/s over time; the blue dots indicate the VTTI processed range-rate, while the red line shows the updated moving average filter. (B) Plot B shows the relative velocity in the x-axis and the inter-vehicle spacing in meters between lead and following vehicles in the y-axis using the VTTI processed radar data. (C) Plot C shows the relative velocity in the x-axis and the inter-vehicle spacing in the y-axis using the smoothed relative velocity data from the moving average filter *(77)*

Once the data were processed, an efficient procedure was developed and implemented to automatically identify car-following events within each trip. This routine ingests the vehicle speed, acceleration, following distance, and relative velocity, and exports the time segments by which car-following events meeting the designated criteria were identified. Automation of the car-following event detection process proved to be successful in multiple studies *(74, 76)* and is required for processing the quantity of trips collected for this study. Validation of the automated procedure was conducted in order to provide additional confidence for the methodology.

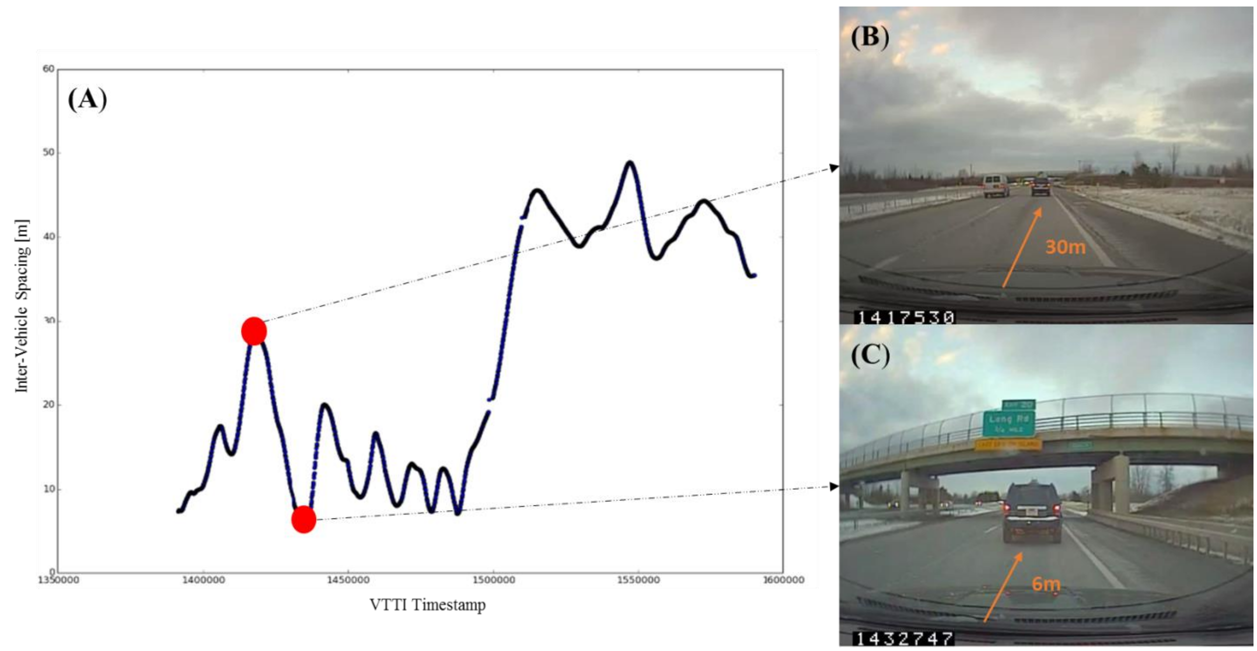


Figure 18 Verification of Automatic Identification of Car-Following Events. (A) Plot A shows the inter-vehicle spacing (y-axis) in meters with the corresponding VTTI timestamp for the selected car-following event (x-axis). The two red dots indicate points during the car-following event, which correspond to the images provided in B and C. (B) Image B corresponds to a point within the car-following event at VTTI Timestamp 1417530 (which can be cross-referenced in Plot A) where the inter-vehicle spacing was reported to be approximately 30m. (C) Image C corresponds to a point within the car-following event at VTTI Timestamp 1432747 – 15 seconds after the first image – where the inter-vehicle spacing is approximately 6m *(77)*

Once car-following events were extracted for each trip, summary statistics related to each individual car-following event and the complete trip were aggregated. In order to analyze a trip set for car-following behavior, a sufficient amount of car-following behavior must be prevalent within each trip. After a review of a selection of trips, a minimum of 4.6 minutes of car-following behavior was required from each trip, as well as a minimum of 2 distinct car-following events. This procedure eliminated individual trips that did not contain enough instances of car-following for analysis; therefore, trip sets were evaluated to ensure at least one adverse trip and one clear trip remained. Next, in order to maintain the correlation within trip sets, a maximum of 50 percent difference in car-following time between each clear trip and adverse trip within a trip set was required. A 50 percent difference was selected as a threshold to ensure that comparisons are only conducted between trips with similar amounts of car-following behavior because the quantity of car-following data impacts the calibration procedure. The magnitude of the threshold value was selected by balancing these comparison requirements with the sample size requirements. Once eliminating all trip sets that did not meet these conditions, the remaining trip sets (containing at a minimum one adverse trip and one clear trip) were used in the analysis.

Methodology

Car-following model calibration is a crucial factor to this study, as successful methods may enable practical application in agency microsimulation models; therefore, the Gipps car-following model—a common model used in the AIMSUN microsimulation software—was selected for analysis.

Equation 3 illustrates the Gipps car-following model, which is a safety-distance car-following model that predicts the following vehicle's speed in such a way to maintain a safe following distance and avoid a collision. Model details can be found in the original model presentation *(78)* and in recent studies evaluating the model *(79, 80)*.

$$v_f(t+\tau) = min\left\{\begin{matrix} v_f(t) + 2.5a_f\tau * \left(1 - \frac{v_f(t)}{V_f}\right) * \sqrt{0.025 + \frac{v_f(t)}{V_f}} \\ b_f\tau + \sqrt{{b_f}^2\tau^2 - b_f * (2\left[x_l(t) - s_l - x_f(t)\right] - v_f(t)\tau - \frac{v_l(t)^2}{\widehat{b_l}})} \end{matrix}\right\} \quad \textbf{Equation 3}$$

where

$_f$	subscript indicating following vehicle
$_l$	subscript indicating lead vehicle
V	desired speed [m/s]
a	maximum acceleration desired [m/s^2]
τ	true reaction time [s]
b	most severe braking desired [m/s^2]
$\hat{b}$	estimated desired severe braking [m/s^2]
s	vehicle size + minimum following distance at stop (speed=0) [m]
$x(t)$	location of front of vehicle at time (t)
$v(t)$	speed of vehicle at time (t) [m/s]

Car-following events within an NDS trip were extracted and the Gipps model was calibrated using common calibration procedures identified through a thorough literature search:

- The Root Mean Square Error (RMSE) was selected as the objective function to compare modeled and actual driving behavior.
- The following distance was selected as the Measure of Performance (MOP), which is the input variable to the objective function.
- A genetic algorithm was constructed as the search mechanism used to minimize the objective function.
- The search space for each of the six calibratable parameters was defined referencing the original Gipps model formulation and additional studies that used the Gipps model.

Once the Gipps model calibration was completed for each individual trip, the trips were aggregated into their trip sets. Next, the trip sets were separated into categories based on the adverse weather condition: {fog, very light rain, light rain, moderate rain, heavy rain, and snow}. Preliminary evaluation of calibration scores and parameter sets within these categories showed few statistically significant differences among different weather conditions. These results were expected, as previous studies identified correlation between parameters that makes it difficult to draw conclusions from parameter values (even when those parameter values are expected to correspond to understandable driving characteristics; e.g., reaction time and desired deceleration) *(81)*. Therefore, the research team developed a novel trajectory validation technique which enabled the evaluation of modeled car-following behavior in differing adverse weather conditions. Once expected characteristics of car-following behavior (i.e., following distance, relative velocity, and acceleration behaviors) were shown to deviate more significantly in increasing weather intensities (compared with clear weather conditions), the modeled behavioral changes were compared with the observed behavioral changes from the actual car-following events. For example, the modeled reduction of headway in snow conditions is compared with the actual reduction of headway in snow conditions.

The second phase of this IAP is intended to enable a full-scale analysis in preparation for the third phase, which includes actual countermeasure implementation. To this end, substantial effort was placed on automating each of the procedures described above and implementing them into a comprehensive tool for wide-spread analysis.

Analysis

For this analysis, 389 trip sets (i.e., representing 1165 trips) were selected and processed through data preparation procedures. Of these, 111 trip sets (i.e., representing 270 trips) passed the criteria ensuring sufficient existence of car-following behavior and trip set matching. Table 10 presents summary statistics describing the trips and car-following behavior available for each adverse weather condition.

Table 10 Data used for analysis *(77)*

Weather Conditions	Count		Average					
	Trip Sets	Trips	Trip Length [min]	Distance Traveled [km]	%Time in car-following	Time in car-following [min]	No. of car-following events	Mean trip velocity [m/s]
All	111	270	26.4	11.4	47.7%	11.5	6.8	25.9
Fog	2	5	24.2	11.2	59.8%	14.7	6.6	28.3
Very Light Rain	24	60	26.4	11.2	51.3%	12.5	7.5	25.6
Light Rain	59	146	25.9	11.2	46.9%	11.3	6.7	25.9
Moderate Rain	17	40	31.1	13.6	41.8%	11.0	6.6	25.4
Heavy Rain	3	7	19.1	9.0	53.3%	10.3	6.4	28.4
Snow	4	10	27.5	13.2	38.1%	8.1	4.1	27.1

As described in the methodology, the calibration procedure produced Gipps model parameters and a calibration score (i.e., the metric describing how well the modeled trajectory from the resulting parameter values matches the actual trajectory). Table 11 shows the average calibration scores of all trip sets in each weather condition. These scores are divided between the clear trips within each trip set and the adverse trips within the trip set. A comparison between the calibration scores of the clear and adverse trip sets were conducted to evaluate if the Gipps model can better represent driving behavior in clear or adverse conditions; however, a statistically significant difference was only found to exist for rain at a moderate intensity.

Table 11 Calibration scores, differentials, and t-test evaluation *(77)*

Weather Conditions	Average Clear Score (RMSE)	Average Adverse Score (RMSE)	% Score Difference (+, clear is higher)	T-test, P-value, 2side, 1paired
Fog	6.19	5.35	14.6%	0.698
Very Light Rain	5.79	5.44	6.3%	0.224
Light Rain	5.23	5.29	-1.1%	0.773
Moderate Rain	6.00	4.51	28.3%	0.005*
Heavy Rain	4.36	4.24	2.7%	0.837
Snow	3.20	3.28	-2.5%	0.927

** Statistically significant at a 95% Confidence Level*

Table 12 shows the average calibrated Gipps parameter values for each weather condition. The first noticeable trend is that the average value for reaction time during clear weather conditions is lower than the reaction time for adverse weather conditions, which is an intuitive behavioral shift as drivers may have more difficulty perceiving their environment during inclement weather events. A similar observation is shown for the minimum following distance at a stop. For trips with clear conditions, the average parameter value is lower than the corresponding value in each adverse weather condition. This finding is also supported by the notion that drivers are likely to increase their following distance during adverse weather conditions.

Table 12 Average calibrated Gipps parameter values for each adverse weather condition *(77)*

Weather Conditions	Average						
	τ	V	a	b	$\hat{b}$	$s^{\wedge}$	$b/\hat{b}$
Clear	0.8	33.1	1.6	-2.5	-2.4	2.9	1.1
Fog	1.1	32.2	1.0	-2.9	-2.7	4.4	1.1
Very Light Rain	1.1	33.2	1.6	-2.5	-2.3	3.1	1.1
Light Rain	1.0	31.1	1.7	-2.5	-2.2	3.2	1.1
Moderate Rain	1.0	30.2	1.5	-2.5	-2.3	3.6	1.1
Heavy Rain	1.7	34.3	1.3	-2.2	-1.7	3.8	1.2
Snow	1.2	31.8	1.5	-2.0	-2.0	3.5	1.0

**Due to the nature of the NDS data, the equations were adjusted to consider parameter "s^" as the distance between the front bumper of the following vehicle and the rear bumper of the lead vehicle at a stop.*

Discernable trends for the remaining Gipps parameter values are not captured in Table 12. While each of the Gipps parameters are intended to correspond to easily understandable elements of driving behavior, taking the average of each parameter introduces an analytic error caused by parameter correlation *(81)*. An evaluation of parameter correlation showed significant correlation between the reaction time and predicted lead vehicle deceleration, as well as between the desired deceleration and predicted lead vehicle deceleration.

For this reason, a new validation procedure was introduced to identify behavioral shifts captured by the calibrated Gipps model for each adverse weather condition. The validation procedure is called <u>trajectory validation</u> and entails the comparison of following vehicle behavior in response to a single lead vehicle trajectory for each calibrated Gipps parameter set. Using this method, driving behavior is normalized and can be averaged for comparison among weather conditions.

Results from the trajectory validation procedure are shown in Table 13 and Table 14. Table 13 provides the RMSE between driver behaviors in clear and adverse weather conditions. Driver behavior is captured using intuitive metrics: following distance, relative velocity, and acceleration. The magnitude of the RMSE value indicates the difference between the clear and adverse trips within all trip sets for each weather condition. Therefore, a smaller RMSE value indicates greater similarity between the driving behavior in adverse and clear weather conditions, and a larger RMSE value indicates greater difference in driving behavior. Results from each driving behavior show that as weather intensity increases, the corresponding difference between modeled behavior in clear and adverse conditions also increases.

Table 13 Average difference for following distance, relative velocity, and acceleration *(77)*

Weather Conditions	Average RMSE		
	Following Distance	Relative Velocity	Acceleration
Fog	5.705	0.494	0.240
Very Light Rain	5.640	0.304	0.211
Light Rain	5.377	0.323	0.217
Moderate Rain	6.827	0.368	0.244
Heavy Rain	12.374	0.532	0.371
Snow	13.738	0.504	0.304

Similarly, Table 14 presents the average R correlation coefficient values comparing driver behavior in adverse and clear conditions for all trip sets within each weather condition. While RMSE describes the arithmetic difference between the modeled clear and adverse trajectories within each trip set, it does not capture *in what way* the trajectories are different. Therefore, the R correlation coefficient was calculated to measure the linear dependency between the trajectories. An R value of +1 represents an exact positive linear correlation and a value of 0 represents no linear correlation. Similar trends are found when comparing correlation between the clear and adverse trajectories within each trip set.

Table 14 Average correlation errors for following distance, relative velocity, and acceleration *(77)*

Weather Conditions	Average R Coefficient		
	Following Distance	Relative Velocity	Acceleration
Fog	0.938	0.784	0.862
Very Light Rain	0.931	0.883	0.887
Light Rain	0.912	0.863	0.876
Moderate Rain	0.792	0.848	0.837
Heavy Rain	0.896	0.815	0.684
Snow	0.874	0.791	0.751

These results support the understanding that drivers deviate from their normal behavior more significantly in severe adverse weather conditions compared to less severe weather conditions. Since this finding was derived from calibrated Gipps parameter sets, it supports the notion that calibration of the Gipps car-following model can capture deviations in driver behavior inflicted by inclement weather events.

While these findings show calibrating the Gipps model from car-following events identified in adverse weather conditions and clear weather conditions produce increasingly different driver behaviors as weather intensity increases, it doesn't explicitly indicate how the driver behavior is changing (i.e., a driver maintained an X-second larger headway in snow conditions compared to clear conditions). This question is conducted using the driver behavior derived from the trajectory validation mechanism, and the accuracy of the results are compared with the actual driving behavior from the original car-following events.

This analysis was conducted for the following time gap (i.e., time headway measured from the front bumper of the following vehicle to the back bumper of the lead vehicle) and the relative velocity between the lead and following vehicles. The results of these comparisons are shown in Table 15 and Table 16.

Table 15 compares the observed and calibrated time gaps maintained by drivers in different weather conditions. A negative value indicates the value for adverse conditions are greater than for clear conditions. For clarity, an example of result interpretation is provided:

> In observed conditions, the mean time gap during *very light rain* is 0.11 seconds greater than matching clear conditions. Comparatively, the model predicted a mean time gap of 0.09 seconds greater in *very light rain* conditions.

The results indicate that in very light rain, light rain, and moderate rain conditions, the maintained time gap is translated accurately from the observed behaviors to the calibrated model. The remaining weather conditions show less correlation in time gap; however, this is likely due to the trip sample size available for these conditions as part of this analysis.

Table 15 Observed and calibrated time gap differences between clear and adverse weather conditions *(77)*

Weather Conditions	Average Difference in					
	Actual Trajectories			**Validation Trajectories**		
	Mean Time Gap [s]	Percentile 85 Time Gap [s]	Time Gap Standard Deviation [s]	Mean Time Gap [s]	Percentile 85 Time Gap [s]	Time Gap Standard Deviation [s]
Fog	-0.07	-0.05	0.11	-0.40	-0.59	-0.17
Very Light Rain	-0.11	-0.10	0.02	-0.09	-0.10	-0.02
Light Rain	-0.13	-0.17	-0.04	-0.11	-0.16	-0.05
Moderate Rain	-0.24	-0.28	-0.09	-0.14	-0.20	-0.04
Heavy Rain	-0.15	-0.25	-0.10	-1.05	-1.55	-0.31
Snow	-1.64	-2.25	-0.58	-0.89	-1.20	-0.22

Table 16 compares the observed and calibrated relative velocity maintained by drivers in different weather conditions. The relative velocity is defined as the following vehicle speed minus the lead vehicle speed. A negative value indicates the value calculated for adverse conditions are greater than for clear conditions. For clarity, an example of result interpretation is provided:

In observed conditions, the maximum relative velocity during *very light rain* is 0.08 seconds <u>less than</u> matching clear conditions. Comparatively, the model predicted a maximum relative velocity of 0.27 seconds <u>greater</u> in *very light rain* conditions.

These results show less correlation than those for time gap in Table 15. This is likely due to the unique nature of relative velocity for each car-following event scenario. A comparison of maximum relative velocity (i.e., the maximum speed differential when the following vehicle is traveling faster than the lead vehicle) between actual and modeled validation trajectories show little correlation due to the stochasticity of the different car-following events represented; however, focusing on the modeled validation trajectory data, a slight positive trend is detected as weather intensity increases. This trend is likely associated with the increased reaction time for more severe weather conditions.

Table 16 Observed and calibrated relative velocity differences between clear and adverse weather conditions *(77)*

Weather Conditions	Average Difference in			
	Actual Trajectories		**Validation Trajectories**	
	Maximum Relative Velocity [m/s]	Relative Velocity Standard Deviation [m/s]	Maximum Relative Velocity [m/s]	Relative Velocity Standard Deviation [m/s]
Fog	0.54	0.09	-0.20	-0.09
Very Light Rain	0.08	-0.04	-0.27	-0.08
Light Rain	-0.09	-0.04	-0.19	-0.06
Moderate Rain	-0.02	0.01	-0.24	-0.07
Heavy Rain	-0.73	-0.19	-1.45	-0.40
Snow	0.15	0.16	-0.88	-0.30

These results produce evidence of the ability of the Gipps car-following model to replicate driving behavior in differing weather conditions. In addition, the calibration procedures and derived parameter sets lay the foundation for deriving weather-specific microsimulation guidance, which could be used to evaluate various strategies (e.g., VSLs) based on drivers behavior in differing weather conditions.

Summary

The car-following analysis in Phase 1 of the Wyoming IAP focused on developing the scope by which the project team would analyze car-following behavior. Phase 2 focused on transforming the scope into concrete analytic procedures, and then automating those procedures to enable analysis on a large number of NDS trips in a wide range of adverse weather conditions. The intricate car-following behavior analysis completed in this project phase will be expanded in Phase 3 to include a greater sample size that will better represent each weather condition, as well as the calibration of other common car-following models. After, the findings will be analyzed to ensure that calibrated models adequately represent driving behavior. The aim of Phase 3 is to conduct a more comprehensive analysis with more NDS trips, as well as use the findings to provide tangible guidance for microsimulation modeling of driver behavior in adverse weather conditions needed for the Wyoming Connected Vehicle Pilot Program.

Lane-keeping Behavior: Preliminary Analysis

Literature Review

According to the Federal Highway Administration (FHWA), 90 percent of crashes are related to driver behavior, and human error is identified as the primary factor contributing to over 60 percent of crashes *(82)*. Many studies in the literature have analyzed drivers' lane-keeping ability from distraction perspective *(83–86)*. While these studies are important to understand how different forms of distracted driving affect lane-keeping ability, the impact of heavy rain on lane-keeping ability has not been researched in naturalistic settings before. Adverse weather conditions such as fog, snow, ground blizzard, slush, rain, and strong wind have been recognized to have significant effects on traffic flow dynamic, drivers' performance and severity of crashes *(87, 88)*. Previous studies showed that the probability of rear-end crashes increases during adverse weather conditions *(89, 90)*. According to the FHWA, weather contributed to over 24 percent of the total crashes between 1995 and 2008. In Canada and the UK, such crashes account for approximately 30 percent and 20 percent respectively *(3, 4)*.

Several studies concluded that crashes increase due to vision obstruction during rainfall by 100 percent or more *(6, 91)*, while others found more moderate (but still statistically significant) increases *(92, 93)*. Sudden reduction in visibility was found to increase the severity level of crashes and tend to involve more vehicles. While these studies provided insights into the impacts of adverse weather conditions on traffic safety, they failed to provide comprehensive understanding of the underlying causes of weather-related crashes due to lack of driver behavior data.

Drivers' lane-keeping performance is one of the vital factors that can affect run-off-road events. Deterioration of lane-keeping ability might be exacerbated by adverse weather conditions due to reduction in visibility and slippery surface conditions *(94, 95)*.

Understanding drivers' responses, when the visibility falls below a certain threshold, might be helpful not only in reducing the lane-departure related crashes in heavy rain, but also in finding a new efficient threshold for Lane Departure Warning (LDW) systems in adverse weather conditions *(96)*. Although, many studies have been conducted in analyzing driver behavior, there are not many researches studies that have focused on the effects of heavy rain on driver performance on a microscopic scale *(17, 97)*. In the last few years, naturalistic driving studies (NDS) have made it possible to obtain more information about driver behavior and performance in different conditions. The NDS data will allow for better understanding of how drivers adjust their behaviors to compensate for increased risk due to reduction in visibility.

The main goal of this section is to investigate the feasibility of using the Second Strategic Highway Research Program (SHRP2) NDS data to analyze drivers' lane-keeping ability in heavy rain and slippery road conditions. This was conducted by compiling a sample dataset from the SHRP2 NDS data, then extracting and reducing data for heavy rain trips and their matching clear weather condition trips on freeways.

Data Preparation

Data extraction and reduction are crucial steps in this study. As mentioned earlier, a subset of data reduced from the SHRP2 NDS were requested to examine driver response in rain/heavy rain in the States of Florida and Washington. In particular, 50 NDS trips during rain/heavy rain on freeway segments were targeted. The provided NDS data included forward-facing and rear-facing videos, basic trip characteristics, and selected vehicle time-series variables. The RID as well as visual inspection of aerial and street view images from Google maps were also utilized. To address the first research question of identifying appropriate trips in rainy conditions, a preliminary criterion for data extraction was developed.

An additional 100 matching NDS trips during clear weather on the same segments and subjects in Florida and Washington States were extracted. A total of 147 valid trips with requested characteristics in rain/heavy rain and their matching clear weather trips were considered in this study. Although most of the trips in heavy rain were matched with two trips in clear weather conditions, only a matching rate of 1:1 was achieved in this study due to data limitation; some of the provided trips in rain did not have matching trips in clear weather and thus were excluded from the analysis. Matching is important to control for sundry factors such as driver population, and roadway geometry.

During the manual verification of the trips, some trips were found to be driven in both free flow and heavy traffic conditions. These trips were considered as mixed traffic. Real-time traffic data are not available in the NDS data. To isolate the impact of heavy rain on driver behavior, trips in free-flow traffic were identified. Traffic conditions were characterized and categorized into two groups including heavy traffic and free flow conditions. Traffic density was determined based on the number of vehicles present in the NDS driver's travel lane, the ability of selecting speed and the ability of maneuvering between lanes. A trip was considered as a free flow speed when the NDS driver has no leading traffic in any lanes or when a leading vehicle is present at least in one lane, but NDS driver is still not affected by other vehicles. Other conditions where NDS drivers were affected by other vehicles were considered as other traffic conditions. All the NDS trips were manually checked to identify the accurate traffic conditions. Travel times were also used to identify trips in free-flow/light traffic. More clearly, if a trip was travelled within the speed limit range, trip was considered as a free-flow condition, otherwise, the trip was considered as other traffic conditions (mixed/heavy traffic). As mentioned earlier, roadway characteristics including speed limit information are provided in the RID.

For automatic identification of trips in rain, other basic trip characteristics such as number of brake activations, high variability in headway times and distances, Electronic Stability Control (ESC), roadway departures, number of Anti-Lock Braking System (ABS) activations, and number of traction control activations were examined in this study. A preliminary analysis on trips in rain/heavy rain indicated that there were no ABS, traction control, or electronic stability control activations in any of the trips. This could be explained due to the fact that the activation of these safety features is not common in rain on freeway segments; moreover, these variables are not available in the NDS data for all vehicles. As mentioned earlier, 147 NDS total trips were acquired, but only 56 were considered for the preliminary analysis when matching is needed. The total 147 acquired trips were utilized in developing the lane-keeping logistic regression model.

Methodology

Logistic regression has been used to develop the lane-keeping model and investigate the factors that affect drivers' lane-keeping ability in different weather conditions. The dependent variable in the model is Standard Deviation of Lane Position (SDLP), and the explanatory variables are the factors which may have significant influence on the lane-keeping ability. The SDLP is a binary variable which defined as two levels including SDLP less than 19.6 in (50 cm) and SDLP greater than 50 cm *(98)*. More specifically, if the average SDLP is maintained within 19.6 in (50 cm) during the trip, lane-keeping performance can be considered within an acceptable reliability level and vice versa. It is worth mentioning that 19.6 in (50 cm) threshold was selected with considering Lane Departure Warning (LDW) system application. Average of SDLP (11.85 in [30.1cm] during heavy rain) was not used as a threshold in this study as the driver would be still safe within this range given that it is not due to distraction. Standard Deviation of Lane Position (SDLP) has been widely used in examining lane-keeping ability. Previous studies used SDLP for assessing drivers' lane-keeping ability *(99, 100)*. SDLP can be considered as a surrogate for overall driving safety due to the fact that an increase in SDLP is associated with an increase in the probability of lane departure events (i.e., when the outside edge of the vehicle tires crosses the lane marking), a precursor of run-off-road crashes *(101)*.

In order to consider the SDLP as a crash surrogate measure, SDLP was calculated for each NDS trip. Weather conditions was used as explanatory variable in this section. Weather conditions were considered in 3 levels: clear weather, light rain, and heavy rain. The model also accounted for traffic conditions, posted speed limit, and speed behavior. In case of speed behavior, the 5 kph (3 mph) interval was considered based on the Variable Speed Limit application (variable speed limits are adjusted at 5 kph/mph increments). Also the median of the speed limits was considered as the threshold. Driver demographics and vehicle characteristics (make, model, and year) data were not provided and hence, only environmental and traffic variables were considered. The Table 17 below is a summary of the different variables used in the lane-keeping model.

Table 17 Data Description *(102)*

Variable	Description	Type	Levels
Response Variable			
SDLP	Standard Deviation of Lane Position	Binary	SDLP<=50
			SDLP>50
Explanatory Variables			
Traffic	Traffic Condition	Binary	0= Free-flow
			1= Traffic
Speed Limit	Posted Speed Limit	Categorical	0= below 90 km/hr
			1= above 90 km/hr
Speed Behavior	Speed selection in various weather conditions	Categorical	More than 5 km/hr below the speed limit
			0–5 km/hr below the speed limit
			0–5 km/hr above the speed limit
			More than 5 km/hr above the speed limit
Weather	Type of weather condition	Categorical	Clear
			Light Rain
			Heavy Rain

A lane-keeping model was developed using logistic regression to better understand factors affecting drivers' lane-keeping ability in different weather conditions. Logit models have been utilized in previous studies*(103, 104)*. One of the advantages of logistic regression in comparison with ordinary least-squares regressions is that independent variables do not have to be normally distributed, or have equal variance in each group. Also, predictors in the logistic regression can be continuous, categorical, or a mixture of both continuous and categorical. Equation 4 shows logistic regression model with x as an independent variable, P(x) as a probability of having success for a binary response variable y considering explanatory variable x, and α is the probability of response when explanatory variables are the reference level (or when x=0) *(34)*. Also the conditional probability of positive outcome can be determined by equation 5.

$$\text{Logit}[P(x)]=\log\left(\frac{P(x)}{1-P(x)}\right)=\alpha+\beta x$$ **Equation 4**

$$P(x)=\frac{\exp(\alpha+\beta x)}{1+\exp(\alpha+\beta x)}=\frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$$ **Equation 5**

The maximum likelihood (ML) method was used to measure the associations by constructing the likelihood function as follows. For more discussion regarding ML method refer to (*35*).

$$l(\beta)=\prod_{i=1}^{n}P(x_i)^{y_i}(1-P(x_i))^{1-y_i}$$ **Equation 6**

In Equation 6, y_i represents the ith observed outcome, with the value of either 0 or 1, and i=1, 2, 3,…, n, where n is the number of observations. The best estimate of β could be obtained by maximizing the log likelihood expression as:

$$LL(\beta)=\ln(l(\beta))=\sum_{i=1}^{n}\left\{y_i\ln(P(x_i))+(1-y_i)\ln(1-P(x_i))\right\}$$ **Equation 7**

Odds ratio is used in many studies to interpret the logistic regression results (*36*). By exponentiating the coefficient (β), odds ratio could be obtained in a logistic regression model (*35*).

$$OR=\exp(\beta_j)$$ **Equation 8**

Analysis

Descriptive Statistics

The NDS video data were manually analyzed to verify and validate results. Classifying the NDS data into two different traffic states (free-flow and mixed traffic) resulted in a total of 56 trips that were considered for the preliminary analysis.

Table **18** shows a summary of the statistics for the number of trips, length of routes, total travel times, and percentages of wiper use at different settings along with their matching clear weather trips. All corresponding RID data were identified and linked to the provided NDS data. The 56 NDS trips constituted a total of about 1,103 miles (1,775 interstate kilometers) traveled over 21.94 hours on six interstate routes in the states of Florida and Washington. These trips occurred mostly on I-4, I-75, and I-275 in Florida, and on I-5, I-90, and I-405 in Washington.

Table 18 Summary Statistics of NDS Trips Considered in this Section *(102)*

	Weather Condition		Heavy Rain	Matched Clear	Light Rain	Matched Clear	Total
Free-Flow Condition	Number of Trips		7	7	9	9	32 trips
	Total Duration (hr)		3.26	2.80	1.42	1.37	8.85 hr
	Total Length (km)		308.67	308.67	172.76	172.76	962.86 km
	% Wiper Setting	0	6.1%	99.5%	0.0%	96.6%	
		1	0.0%	0.0%	60%	3.4%	
		2	0.0%	0.0%	22%	0.0%	
		3	93.9%	0.5%	18%	0.0%	
Heavy/ Mixed Traffic	Number of Trips		3	3	9	9	24 trips
	Total Duration (hr)		1.34	1.64	5.44	4.67	13.09 hr
	Total Length (km)		95.3	95.3	309.64	312.05	812.29 km
	% Wiper Setting	0	0.0%	99.9%	6%	91.2%	
		1	10%	0.0%	50%	8.8%	
		2	14%	0.0%	26%	0.0%	
		3	75.2%	0.1%	18%	0.0%	
Total Number of Trips			10	10	18	18	56

Analysis of wipers status as well as visual inspections of all NDS videos were utilized to identify heavy/light rain and clear weather condition trips.

Table **18** provides a breakdown of the percentage of the time that the wipers were engaged at each level. If the wipers were engaged at level 3 for greater than 75 percent of the whole trip duration, the trip will be considered as a heavy rain trip. Heavy rain trips in free-flow traffic had about 94 percent active wipers at setting 3. Similarly, if the wipers were active at level 1 or level 2 for greater than 75 percent, the trip would be considered as a light rain trip (light rain trips in free-flow conditions had 82 percent active wipers at settings 2 and 3). A trip with inactive wipers (level 0) for more than 91 percent of the time would be marked as a clear weather trip (0 percent for settings 2 and 3). This classification was used to provide a general consensus of the impact of heavy and light rain only on drivers' lane-keeping ability as well as other driving behaviors for the free-flow conditions only.

Table 19 to Table 23 show preliminary analysis and various statistical tests for the main time-series variables of interest for heavy rain/clear weather in the free-flow conditions. In addition, descriptive statistics are shown for trips that included heavy rain and clear weather conditions within the same trips. Cohen's d effect size which is an indication of the magnitude of the difference between heavy rain and clear weather is also provided in Table 19 to Table 23. Cohen's d effect size can be interpreted as d=0.2 small size effect; d=0.50 medium size effect; and d=0.80 large size effect (105).

Table 19 Preliminary Analysis for the NDS Instrumented Vehicles-Speed *(102)*

	Statistical Tests	Free-Flow Traffic (Matched Trips)		Comparison within Trips	
		Heavy Rain	**Matched Clear**	**Heavy Rain**	**Clear Weather**
Speed (km/hr)	Average	85.07	101.39	91.8	106.36
	SD	14.69	11.25	14.65	6.53
	Min.	17.4	70.4	35.09	53
	Max.	109.4	133.5	125.5	125.9
	Median	87.5	101	94.19	106
	t-Test	Avg. Speed is significantly lower in Heavy Rain. t(21021)=-303, P<0.05 Effect size (Cohen's d)=-1.24		Avg. Speed is significantly lower in Heavy Rain. t(3713)=-164.6, P<0.05 Effect size (Cohen's d)=-1.28	
	F-Test	Speed variability is higher in Heavy Rain $F_{1,9969,12454}$=0.990, p<0.05		Speed variability is higher in Heavy Rain $F_{1,30006,46129}$=5.5, p<0.05	
	Z-Test	Proportion of violation ≥ 10 km/h above the speed limit is significantly higher in Clear Weather. Z=206.6731 , P<0.05		Proportion of violation ≥ 10 km/h above the speed limit is significantly higher in Clear Weather. Z=50.47, P<0.05	

Notes: Analysis was performed for one-minute aggregation level and 95% confidence interval.
Matched data have equal trips distance, different travel times are due to lower speed because of weather

Table 20 Preliminary Analysis for the NDS Instrumented Vehicles-Acc/Dec *(102)*

	Statistical Tests	Free-Flow Traffic (Matched Trips)				Comparison within Trips			
		Heavy Rain		**Matched Clear**		**Heavy Rain**		**Matched Clear**	
		Acc	Dec	Acc	Dec	Acc	Dec	Acc	Dec
Acceleration/ Deceleration (g) (Positive columns= Acceleration)	Average	0.0263	-0.0266	0.0253	-0.0276	0.0213	-0.0282	0.0158	-0.0162
	SD	0.0181	0.0214	0.0184	0.0225	0.0157	0.0245	0.0160	0.0185
	Min.	0.0029	-0.3132	0.0015	-0.4321	0.0015	-0.2842	0.0029	-0.2610
	Max.	0.2059	-0.0029	0.1769	-0.0015	0.1769	-0.0015	0.1624	-0.0029
	Median	0.0232	-0.0232	0.0203	-0.0232	0.0174	-0.0218	0.0116	-0.0087
	t-Test	Average Acc. is significantly higher in Heavy Rain and avg. Dec. is higher in Clear Weather Acc: t(11232)=8.64, P<0.05, Effect size (Cohen's d)=0.05 Dec: t(8199)=6.49, P<0.05, Effect size (Cohen's d)=0.04				Average Acc./Dec. is significantly higher in Heavy Rain Acc: t(3223)=33.68, P<0.05, Effect size (Cohen's d)=0.37 Dec: t(2199)=-45.51, P<0.05, Effect size (Cohen's d)=-0.61			
	F-Test	Acc./Dec. variability is higher in Clear Weather Acc: $F_{1,7251,5258}$=0.97, p<0.05 Dec: $F_{1,4256,4031}$=0.90, p<0.05				Acc./Dec. variability is higher in Clear Weather Acc: $F_{1,1507,2520}$=0.95, p<0.05 Dec: $F_{1,1228,1633}$=1.75, p<0.05			
	Z-Test	Proportions of Dec. lower than -0.3g is significantly greater in Clear Weather. No Acc. were found higher than +0.3g Dec: Z=-4.2732, P<0.05				No Acc./ Dec. were found higher/lower than ±0.3g			

Notes: Analysis was performed for one-minute aggregation level and 95% confidence interval.
Matched data have equal trips distance, different travel times are due to lower speed because of weather

Table 21 Preliminary Analysis for the NDS Instrumented Vehicles-Yaw Rate *(102)*

	Statistical Tests	Free-Flow Traffic (Matched Trips)				Comparison within Trips			
		Heavy Rain		**Matched Clear**		**Heavy Rain**		**Matched Clear**	
		Acc	Dec	Acc	Dec	Acc	Dec	Acc	Dec
Yaw Rate (deg/s) (negative sign=left rotation)	Average	0.84	-0.97	0.89	-0.8	1.01	-0.97	0.64	-0.61
	SD	0.73	0.65	0.71	0.59	0.88	0.86	0.41	0.46
	Min.	0.33	-8.78	0.33	-3.9	0.16	-8.78	0.16	-4.55
	Max.	6.83	-0.33	5.85	-0.33	10.08	-0.16	3.25	-0.16
	Median	0.65	-0.65	0.65	-0.65	0.65	-0.65	0.49	-0.33
	t-Test	Yaw rate (right rotation) is significantly higher in Clear Weather—no significant difference in left rotation. Right rotation: t(2515)=-6.4, P <0.05 Effect size (Cohen's d)=-0.08 Left rotation: t(3022)=0.3, P>0.05 Effect size (Cohen's d)=0.003				Yaw rate is significantly higher in Heavy Rain Right rotation: t(1010)=34.62, P <0.05, Effect size (Cohen's d)=0.69 Left rotation: t(1793)=-41.62, P<0.05 Effect size (Cohen's d)=-0.62			
	F-Test	Yaw rate variability is higher in Heavy Rain Right rotation: $F_{1,2704,1258}$=1.05, p<0.05 Left rotation: $F_{1,4504,1586}$=1.2, p<0.05				Yaw rate variability is higher in Heavy Rain Right rotation: $F_{1,755,958}$=4.64, p<0.05 Left rotation: $F_{1,1229,1505}$=3.48, p<0.05			

Notes: Analysis was performed for one-minute aggregation level and 95% confidence interval.
Matched data have equal trips distance, different travel times are due to lower speed because of weather

Table 22 Preliminary Analysis for the NDS Instrumented Vehicles-Lane Offset *(102)*

	Statistical Tests	Free-Flow Traffic (Matched Trips)				Comparison within Trips			
		Heavy Rain		**Matched Clear**		**Heavy Rain**		**Matched Clear**	
		Acc	Dec	Acc	Dec	Acc	Dec	Acc	Dec
Lane Offset (cm)	Average	24.4	-23.04	62.26	-71.92	39.55	-45.99	34.56	-43.39
	SD	22.55	26.87	130.79	135.39	76.44	83.33	65.58	75.06
	Max	964.95	0	999.86	-0.01	838.83	-0.01	955.04	-999.59
	Min	0	-590.8	0.05	-999.12	0.05	-998.61	0.05	-0.04
	Median	20.32	-17.02	18.66	-29.08	16.85	-26.94	15.54	-26.88
	t-Test	Avg. lane offset to the right and left from the lane center is significantly higher in Clear Weather Right: t(1450)=-34.23, P<0.05 Effect size (Cohen's d)=-0.57 Left: t(4113)=66.80, P<0.05 Effect size (Cohen's d)=0.66				Avg. lane offset to the right and left from the lane center is significantly higher in Heavy Rain Right: t(1493)=4.91, P<0.05 Effect size (Cohen's d)=0.08 Left: t(4200)=-3.78, P<0.05 Effect size (Cohen's d)=-0.03			
	F-Test	Lane offset to the right and left variability is higher in Clear Weather Right: $F_{1,3424,1415}$=0.02, p<0.05 Left: $F_{1,2494,3649}$=0.03, p<0.05				Lane offset variability is higher in Heavy Rain Right: $F_{1,810,1392}$=1.36, p<0.05 Left: $F_{1,2174,3650}$=1.23, p<0.05			

Notes: Analysis was performed for one-minute aggregation level and 95% confidence interval.
Matched data have equal trips distance, different travel times are due to lower speed because of weather

Table 23 Preliminary Analysis for the NDS Instrumented Vehicles-Headway *(102)*

	Statistical Tests	Free-Flow Traffic (Matched Trips)		Comparison within Trips	
		Heavy Rain	**Matched Clear**	**Heavy Rain**	**Matched Clear**
Headway(sec)	Average	2.17	2.01	1.98	2.02
	SD	1.00	1.12	1.16	1.14
	Max	7.84	6.65	7.58	6.68
	Min	0.16	0.08	0.12	0.15
	Median	2.10	1.99	1.83	1.81
	t-Test	Headway is significantly higher in Heavy Rain $t(8268)=-21.93$, $P<0.05$ Effect size (Cohen's d)=-0.15		No significant difference	
	F-Test	Headway variability is higher in Clear Weather. $F_{1,4030,4303}=1.04$, $p<0.05$		No significant difference	

Notes: Analysis was performed for one-minute aggregation level and 95% confidence interval.
Matched data have equal trips distance, different travel times are due to lower speed because of weather

As can be seen in Table 19, a t-test indicated that the average speed in heavy rain under the free-flow traffic conditions was significantly 10.14 mph (16.32km/hr) lower than in clear weather and free-flow traffic conditions. Speed in free-flow conditions is important for variable speed limit (VSL) application because the speed choice here is not affected by the interaction with traffic. It was also found that speeds have higher variability during heavy rain in comparison with clear conditions under free-flow traffic, which could be an indication of increased safety risk *(30)*.

The acceleration/deceleration variable was examined (Table 20), and ±0.3g acceleration/deceleration rates were set as a threshold to identify aggressive braking/acceleration events *(106)*. The preliminary analysis showed that while heavy rain has a wider range of acceleration and statistically has a higher average, the average deceleration was found to be statistically higher in the matching clear weather conditions. The variability of acceleration and deceleration and the proportions of deceleration that were lower than -0.3g were found to be greater in clear weather conditions. These findings coupled with the observed reduction in speed during heavy rain indicate that drivers compensate for the slippery surface conditions by not decelerating by rates greater than -0.3g.

The lane offset variable in the NDS data is estimated using machine vision techniques. Lane offset is an indication of either a lane change or a deviation from the lane. Lane change is defined as an intended and substantial lateral shift of a vehicle *(107)*. Lane change could be modeled using multiple variables: turn signal, steering angle, yaw rate, and machine vision lane offset. Although lane change is not the main focus of this section, distinguishing lane change from lane wandering is important to understand driver behavior in heavy rain condition. Utilizing time-series and video data, lane changes were separated from lane wandering.

A criterion for lane offset values within ±0.3 meters was set to flag lane wandering events (Table 22), especially when these events varied to the right and left over a short duration of time. Continuous and steady lane offset within a threshold greater than ±0.3 meters to ±9.5 meters in one direction was considered as a full lane change. A past NDS study indicated that using a threshold of ±0.1 meters resulted in a higher than expected number of lane departures *(108)*.

Preliminary analysis indicated that the number of lane changes is higher in clear weather conditions while lane wandering was found to be significantly higher in heavy rain.

Yaw rate and steering angle are additional variables that could be used to analyze lane maintenance. Unfortunately, steering wheel position was only available for a fraction of vehicles (only two trips included steering angle data). Yaw rates were analyzed (table Table 21) for events with lane offset within ±0.3 meters where there were no lane changes. Yaw rates were significantly higher in heavy rain, which, as mentioned earlier, might indicate frequent evasive maneuvers to mitigate an increased risk. On the one hand, average headways (Table 23) were found to be significantly higher in heavy rain compared to clear weather conditions under free-flow traffic. On the other hand, the variability of headways was found to be significantly higher in clear conditions. This could be explained by the fact that drivers tend to compensate for the increased risk due to the limitation in visibility by maintaining longer headway times.

Analyzing the NDS time-series data in conjunction with video data revealed that the estimated NDS machine vision lane offset is noisy but still reliable in heavy rain weather condition.The min/max values for the lane offset also revealed a very interesting finding: drivers tend to change multiple lanes (2–3 lanes) during clear weather condition versus a single lane change in heavy rain conditions. Controlling for entry and exit of the freeway maneuvers, lane changes that occurred in heavy rain were mostly evasive maneuvers to mitigate an increased risk. From video observations, it was found that drivers opted out of speed reduction behind a slower vehicle more often than changing lanes.

Additional analyses were conducted on an individual (no matching) seven NDS trips that were identified to have both clear and heavy rain conditions within the same trip. All seven trips were in the free-flow traffic conditions. There was an agreement across the seven trips that speeds were reduced significantly with a higher standard deviation in heavy rain than in clear conditions. Also, the acceleration/deceleration and lane change/maintenance were affected. The number of braking, decelerations, and accelerations were significantly higher in heavy rain than in the clear portion of the trips.

There were 44 and 22 braking events in heavy rain and clear weather conditions, respectively. High variability in yaw rate might indicate either too many lane changes or poor lane maintenance. Although the number of lane changes was very limited in heavy rain compared to clear conditions, the high variability in yaw rate during heavy rain suggested worse lane maintenance capabilities than in the clear condition.

Lane-Keeping Model Results and Discussion

To confirm the suitability and fitness of the model, the log likelihood ratio and the pseudo R^2 were used. Table 24 shows the results of the model; the Likelihood Ratio (LR) test statistic falls into the rejection area (p-value < 0.05), which means that the overall explanatory variables of the model have significant influence on the response at a statistical significance level of 95 percent. Only statistically significant variables were retained in the final models.

Table 24 Logistic Regression Model for lane-keeping ability in Different Weather Conditions *(102)*

Analysis of Maximum Likelihood Estimates									
Parameter		**DF**	**Estimate**	**Standard Error**	**Wald Chi-Square**	**Odds Ratio**	**P value**	**95% Confidence Limits**	
Intercept		1	-0.4630	0.4621	1.0037	-	0.3164	-	-
Weather	**Clear**		-	-	-	-	-	-	-
	Light Rain	1	-0.7671	0.8352	0.8435	0.464	0.3584	0.090	2.387
	Heavy Rain	1	1.3389	0.5554	5.8117	3.815	0.0159	1.284	11.331
Speed Limit	**Below 90 km/hr**		-	-	-	-	-	-	-
	Above 90 km/hr	1	-2.7258	1.1092	6.0395	0.065	0.0140	0.007	0.576
Traffic	**Free-Flow**		-	-	-	-	-	-	-
	Traffic	1	-1.5778	0.5387	8.5792	0.206	0.0034	0.072	0.593

As can be seen in Table 24, heavy rain has a statistically positive effect on SDLP. It means that standard deviation of lane position is more likely to be higher in heavy rain condition. Particularly, driver lane-keeping ability would be reduced (SDLP would be increased) by increasing the precipitation intensity. This may be attributed to the shorter sight distance and also low visibility of lane marking in heavy rain condition. This finding is in agreement with other previous studies, showing the negative effect of adverse weather on drivers' performance *(46, 109)*. More clearly, drivers in heavy rain condition are 3.8 times more likely than clear weather to have higher SDLP (OR=3.8). It is also shown that driving in light rain condition does not have any effect on lane-keeping ability.

Interestingly, maximum posted speed limit was found to be significant with a negative coefficient in the developed lane-keeping model. This might be due to the fact that drivers pay more attention to the road ahead considering the higher speed. It is worth mentioning that road segments with higher speed limits might have better geometry design and sight distance in comparison with segments with lower speed limit. Obtained negative association between lane-keeping and posted speed limit could be because of the mentioned advantages of segments with higher speed limits that can compensate for the negative effects of rainy weather condition to some extent. Driving in a segment with higher speed limit does not necessarily mean that the driver has higher speed. More specifically, drivers who are driving in road segments with posted speed limit less than 50 mph (90 km/hr) are 15 times more likely to have higher SDLP in comparison with those who are driving in segments with posted speed limit above 50 mph (90 km/hr) (OR=0.065). It is known that drivers reduce their speed during adverse weather conditions *(110)*. Lower speed can enhance drivers' performance especially at the start of the rain as the surfaces are most slippery because of the oil and dust that have not washed away mix with the moisture. Moreover, lower speed can increase the headway spaces providing more time to prepare for the appropriate maneuver as driving becomes risky with low visibility.

Traffic conditions were found to be statistically significant as expected. The negative sign depicts the fact that by increasing traffic congestion, drivers have less ability to swerve, change lane, and generally are forced to have better lane-keeping. More clearly, drivers who drive in a free flow condition are 4.8 times more likely to have higher SDLP in comparison with those who are driving in traffic congestion condition (OR=0.206).

Summary

Descriptive statistics were used to understand the difference between drivers' behavior in clear and heavy rain weather conditions, and logistic regression was utilized to identify the main contributing factors affecting drivers' lane-keeping ability in different weather conditions.

Based on the obtained results from the performed descriptive analysis, heavy rain had a wider range and a higher average of acceleration; however, average deceleration was found to be higher in matching trips in clear weather condition. The number of lane changes is higher in clear weather; however, lane wandering is higher in heavy rain conditions. Yaw rates and average headways were found to be statistically higher in heavy rain in comparison with clear weather conditions. Acceleration, deceleration, speed, headway, and lane-keeping can be used as indicators of safety. Weather, speed limit, and traffic conditions were found to be significant contributing factors in the developed lane-keeping model.

Analyzing drivers' behavior at a microscopic level has become an important topic for different tasks in transportation engineering. The Naturalistic Driving Study (NDS) data in particular may help in developing driving models that could be applied to different areas *(111–113)*: i) performing safety analyses based on individual driver data, ii) calibration of driving behavior models to update microscopic models for traffic simulation, specifically in various traffic and weather conditions, iii) developing control logics for Advanced Driving Assistance Systems (ADAS), and Connected and Automated Vehicles (CAV). While the results from this analysis may improve our understanding about lane-keeping behavior in heavy rain at a microscopic individual level, the results may also help in developing better Lane Departure Warning (LDW) systems. The NDS data may address limitations of these systems during adverse weather conditions. Individual drivers' data may provide more insights into drivers' behavior and performance in different traffic and weather conditions than the commonly used macroscopic level of speed, volume and occupancy; the understanding gained from these data may help in updating microsimulation models.

Lane-keeping Behavior Considering Driver Demographics and Roadway Characteristics: Using Non-parametric MARS Modeling Technique

One of the most unpredictable factors in the driver, vehicle and roadway triangle is the driver behavior *(114, 115)*. In fact, driver behavior might be changed while driving considering the physical conditions as well as distractions provided by in-vehicle technologies etc. Lane-keeping behavior has been identified as one of the principal behavioral-performance factors with a broader implication of driving task. The effect of distracted driving on lane keeping ability has been investigated in previous studies *(15, 83)*. Even though the results from these studies are extremely important and show the effect of different distractions on the lane-keeping ability, the weather impact on lane-keeping ability has not been studied exclusively in previous studies. Considering the increase in use of the Naturalistic Driving studies (NDS) in recent years, researchers have a better opportunity to study driver performance and behavior at a microscopic level. The main goal of this section is to investigate driver lane-keeping ability using more NDS data and at a higher resolution (1-min chunks), as well as advanced non-parametric modeling technique to better understand driver lane-keeping ability in heavy rain and slippery road conditions.

Data Preparation

Of the received 2,881 trips in rain, 196 trips were randomly selected for further analysis in this chapter. In addition, 392 matching trips in clear weather conditions (2:1 matching ratio) have been fully processed in this chapter. The selected NDS trips involved 141 drivers between 19 and 89 years of age with the majority of the drivers in the young age group (19 to 29 years old). A total of 12,320 one-minute segments – equivalent to nearly 205 hours of driving have been fully processed for this chapter *(116)*.

Methodology

Two lane-keeping models were developed using the logistic regression and Multivariate Adaptive Regression Splines (MARS) to better understand factors affecting driver lane-keeping ability in clear and rainy weather conditions. Advantages of using MARS include the capacity to intake continuous response variables, promising predictive power, and overcoming the black-box limitations *(116)*.

Multivariate Adaptive Regression Splines (MARS)

MARS can be defined as a piecewise, multivariate regression that can consider the complex relationships among variables. This model was introduced by Friedman (1991) *(117)*. In the MARS model, the space of predictors is divided into multiple knots, and a spline function is fitted between those knots *(118)*. Basis functions (BFs) are those elements that can be used to fit a MARS model. Each basis function can be a main function or an interaction of different variables. Equation 9 shows a general form of the MARS model *(119, 120)*.

$$\hat{y} = \alpha_0 + \sum_{m=1}^{M} \alpha_m \beta_m(x)$$ **Equation 9**

Where $\hat{y}$ can be defined as the predicted response variable, α_0 is the coefficient of the constant basis function, α_m is the coefficient of the mth basis function, $\beta_m(\mathrm{X})$ is the mth basis function, and M is the total number of basis functions in the developed MARS model.

There are two main steps to fit a MARS model. These two steps can be summarized as first, forward-stepwise regression selection and second, backward-stepwise elimination procedure *(121)*. In the constructive phase, the initial model starts with just a constant, then the model searches for a possible variable-knot combination, and the improvement of the model is measured. The process would be repeated to identify the best variable-knot combination. The search process will continue until reaching the maximum number of basis functions. In the elimination phase, MARS identifies a BF to drop based on residual sum of squares criteria. After refitting the model, another BF is selected to drop based on the same criteria. The process is repeated until all the BFs have been deleted. Finally, the result of the backward-stepwise elimination procedure is a distinctive series of candidate models *(121)*. The final selection of the model would be based on the generalized cross-validation (GCV) criterion as shown in Equation 10 *(119–121)*:

$$\mathrm{GCV} = \frac{\sum_{i=1}^{N}\left(y_i - \hat{y}\right)^2}{N\left(1-\frac{C(M)}{N}\right)^2}$$

Equation 10

where N is the number of observations; y_i is the response for observation *i*; and $C(M)$ is a complexity penalty function. To be specific, in Equation 10, the numerator represents the lack of fit on the model with M basis function and the denominator contains a term to consider a penalty for model complexity, which is $C(M)$. This term is related to the number of estimated parameters in the model *(121)*.

Logistic Regression

Logistic regression is a commonly used model in traffic safety and operation studies. For more information about the logistic regression please see the previous section or *(122)*.

Analysis

Matching trips in rain and clear weather were required, specifically for comparative analysis. As mentioned earlier, weather conditions may not be consistent within a trip. Therefore, considering the entire clear trip as matched to a particular rain trip would not provide appropriate results. Therefore, the data were further reduced by importing the coordinates of the rain and clear trips into the ArcGIS software and eliminating non-matching segments. Removing the non-freeway, non-matching segments resulted in 4,434 one-minute matching segments (486 segments in light rain, 1090 in heavy rain and 2,858 segments in clear weather) in free flow speed conditions, which was equivalent to nearly 74 hours of driving time and 7,334 kilometers. The summary statistics and various statistical tests for the variables of interest including speed, acceleration/deceleration, yaw rate, and lane offset considering the matched trips in clear and rainy weather conditions are presented in Table 26 and Table 27. Preliminary analysis indicated that the number of lane changes is higher in clear weather conditions while lane wandering was found to be significantly higher in heavy rain. Table 25 is a summary of the different variables used to set lane-keeping models. In addition Table 26 and Table 27 provide the comparison of key variables in light rain, heavy rain and their matching trips in clear weather conditions.

Table 25 Data Description *(123)*

Variable	Description	Type	Source	Definition	Assigned Code	Reference Level
Response Variable:						
SDLP	Standard deviation of lane position offset Distance to the left or right of the center of the lane based on machine vision	Categorical	Naturalistic driving time series data	SDLPO<=20 cm	1	
				SDLPO>20 cm	2	*
Explanatory Variables:						
Environmental Factors						
Weather Conditions	Predominant weather conditions in 1-min video observation	Categorical	Video Observation	Clear	1	*
				Rain	2	
				Heavy Rain	3	
Speed Limit	Predominant Posted Speed limit in 1-min video observation	Binary	Roadway Information Database (RID)	<=60 mph (median of speed limit: 60)	0	
				>60 mph	1	
Traffic Conditions	Predominant Traffic conditions (LOS) in 1-min video observation	Binary	Video Observation	A & B	1	
				C-F	2	*
Demographics						
Gender	The gender the participant identifies with	Binary	Electronic online questionnaire administered during participant in-processing	Male	1	*
				Female	2	
Age	The age group corresponding to the driver's birthdate.	Categorical	Electronic online questionnaire administered during participant in-processing	Young<25	1	*
				Middle(25-44)	2	
				Old>44	3	
Education	The participant highest completed level of education	Categorical	Electronic online questionnaire administered during participant in-processing	Below High School Diploma	1	*
				Above High School Diploma	2	
Driver Mileage Last Year Details	The approximate number of miles the participant drove last year	Categorical	Electronic online questionnaire administered during participant in-processing	Less than <=12,000	1	*
				>12000	2	
Driving Experience	Number of years driving experience	Categorical	Electronic online questionnaire administered during participant in-processing	Less than 3 years.	1	
				More than 3 years	2	*
Roadway Characteristics						
Curve	Whether the majority of 1-min driving was driven on tangent or curve	Binary	RID	Tangent	1	*
				Curve	2	
Number of lanes	Number of lanes that the majority of the 1-min driving was travelled on	Count	RID	-	-	-

Table 26. Preliminary Analysis using the Matched Trips in Light Rain

	Statistical Test	Free-Flow Traffic			
		Light Rain		Clear	
Speed (km/hr.)		**Speed**	**% Speed Reduction from Speed Limit**	**Speed**	**% Speed Reduction from Speed Limit**
	Average	98.199	- 1.846	102.881	- 6.213
	SD	13.322	12.673	12.271	10.840
	Min.	40.653	- 37.346	50.270	- 46.888
	Max.	130.744	55.646	137.448	47.595
	Median	99.302	- 3.339	103.077	- 6.164
	t-test	Average speed is significantly higher in matched clear. t(1964) = -9.55, $P<0.05$ Effect size (Cohen's d) = -0.37			
	F-test	Speed variability is significantly higher in light rain. $F_{1053,2094} = 1.18$, $P<0.05$			
	Z-test	No significant difference between the proportion of speeding ≥ 10 km/h in light rain and clear weather. Z = -1.56, $P>0.05$			
Acceleration/ Deceleration(g)		**Acceleration**	**Deceleration**	**Acceleration**	**Deceleration**
	Average	0.014	-0.015	0.015	-0.020
	SD	0.013	0.016	0.014	0.044
	Min.	0.000	-0.111	0.000	-0.458
	Max.	0.097	0.000	0.090	0.000
	Median	0.011	-0.011	0.012	-0.012
	t-test	Average acceleration is significantly higher in clear weather. t(1656) = -1.64, $P<0.05$ Effect size (Cohen's d) = -0.08 Average deceleration is significantly higher in clear weather. t(1448) = 3.5, $P<0.05$ Effect size (Cohen's d) = 0.14			
	F-test	No significant difference between average acceleration variability in light rain and clear weather. $F_{1073,583} = 1.06$, $P>0.05$ Deceleration variability is significantly higher in clear weather. $F_{1035,505} = 7.71$, $P<0.05$			
	z-test	No acceleration/ deceleration were found higher/lower than ± 0.3g			
Yaw Rate, negative sign = left rotation (deg./s)		**Positive**	**Negative**	**Positive**	**Negative**
	Average	0.345	-0.719	0.512	-0.687
	SD	0.345	2.075	0.895	2.123
	Min.	0.001	-21.377	0.001	-22.231
	Max.	2.400	0.000	7.058	-0.002
	Median	0.251	-0.318	0.267	-0.267
	t-test	Average right rotation is significantly higher in clear weather. t(926) = -4.13, $P<0.05$ Effect size (Cohen's d) = -0.22 No significant difference between average left rotation in light rain and clear weather. t(2106) = -0.34, $P>0.05$ Effect size (Cohen's d) = -0.015			
	F-test	Right rotation variability is significantly higher in clear weather. $F_{650,296} = 6.74$, $P<0.05$ No significant difference between left rotation variability in light rain and clear weather. $F_{1332,774} = 1.05$, $P>0.05$			
Lane Offset (cm)		**Positive**	**Negative**	**Positive**	**Negative**
	Average	20.790	-17.672	18.401	-21.713
	SD	21.390	19.120	16.316	22.070
	Min.	0.462	-276.155	0.017	-266.983
	Max.	136.296	-0.102	121.291	-0.001
	Median	16.299	-11.165	14.787	-16.199
	t-test	Average lane offset to the right is significantly higher in light rain. t(666) = 1.93, $P<0.05$ Effect size (Cohen's d) = 0.13 Average lane offset to the left is significantly higher in clear weather. t(1709) = 4.31, $P<0.05$ Effect size (Cohen's d) = 0.19			
	F-test	Lane offset to the right variability is significantly higher in light rain. $F_{398,694} = 1.72$, $P<0.05$ Lane offset to the left variability is significantly higher in clear weather. $F_{1280,733} = 1.33$, $P<0.05$			

Table 27. Preliminary Analysis using the Matched Trips in Heavy Rain

	Statistical Test	Free-Flow Traffic			
		Heavy Rain		Clear	
Speed (km/hr.)		**Speed**	**% Speed Reduction from Speed Limit**	**Speed**	**% Speed Reduction from Speed Limit**
	Average	93.711	- 0.089	101.553	- 8.939
	SD	11.023	12.795	9.186	9.697
	Min.	51.092	- 34.673	72.384	- 35.280
	Max.	128.542	49.458	130.786	21.260
	Median	93.668	- 0.740	100.966	- 9.739
	t-test	Average speed is significantly higher in matched clear. t(899) = -12.99, P<0.05 Effect size (Cohen's d) = - 0.79			
	F-test	Speed variability is significantly higher in heavy rain. $F_{483,745}$ = 1.44, P<0.05			
	Z-test	No significant difference between the proportion of speeding ≥ 10 km/h in heavy rain and clear weather. Z = - 0.31, P>0.05			
Acceleration/ Deceleration(g)		**Acceleration**	**Deceleration**	**Acceleration**	**Deceleration**
	Average	0.019	-0.014	0.021	-0.020
	SD	0.017	0.015	0.020	0.019
	Min.	0.000	-0.072	0.000	-0.105
	Max.	0.121	0.000	0.093	0.000
	Median	0.014	-0.010	0.014	-0.015
	t-test	No significant difference between average acceleration in heavy rain and clear weather. t(648) = -1.43, P>0.05 Effect size (Cohen's d) = - 0.11 Average deceleration is significantly higher in clear weather. t(509) = 4.01, P<0.05 Effect size (Cohen's d) = 0.33			
	F-test	Acceleration variability is significantly higher in clear weather. $F_{380,281}$ = 1.37, P<0.05 Deceleration variability is significantly higher in clear weather. $F_{366,203}$ = 1.65, P<0.05			
	z-test	No acceleration/ deceleration were found higher/lower than ± 0.3g			
Yaw Rate, negative sign = left rotation (deg./s)		**Positive**	**Negative**	**Positive**	**Negative**
	Average	0.424	-1.591	0.425	-0.501
	SD	0.381	2.269	0.424	0.458
	Min.	0.004	-7.631	0.001	-2.578
	Max.	2.215	-0.001	1.883	-0.001
	Median	0.301	-0.479	0.269	-0.368
	t-test	No significant difference between average left rotation in heavy rain and clear weather. t(330) = -0.03, P>0.05 Effect size (Cohen's d) = -0.003 Average left rotation is significantly higher in heavy rain. t(348) = -8.60, P<0.05 Effect size (Cohen's d) = -0.72			
	F-test	No significant difference between Left rotation variability in heavy rain and clear weather. $F_{234,144}$ = 1.24, P>0.05 Left rotation variability is significantly higher in heavy rain. $F_{329,459}$ = 24.53, P<0.05			
Lane Offset (cm)		**Positive**	**Negative**	**Positive**	**Negative**
	Average	15.377	-16.747	17.721	-17.631
	SD	11.651	21.887	20.305	17.513
	Min.	0.095	-224.663	0.191	-139.816
	Max.	74.906	-0.308	227.831	-0.049
	Median	13.568	-11.480	14.136	-11.963
	t-test	Average lane offset to the right is significantly higher in clear weather. t(521) = -1.7, P<0.05 Effect size (Cohen's d) = -0.136 Average lane offset to the left is significantly higher in clear weather. t(412) = 0.52, P<0.05 Effect size (Cohen's d) = 0.046			
	F-test	Lane offset to the right variability is significantly higher in clear weather. $F_{317,224}$ = 3.04, P<0.05 Lane offset to the left variability is significantly higher in heavy rain. $F_{232,376}$ = 1.56, P<0.05			

Lane-keeping Model Results

Previous studies showed that increasing the number of interactions in MARS may increase the model complexity; therefore, the applicability of the model and interpretability of the results might be decreased *(119, 120)*. Hence, the maximum order of interactions was defined as two in this study. Table 28 and Table 29 present the developed MARS and logistic regression models for driver lane-keeping ability.

Table 28 Driver Lane Keeping Model Using MARS *(123)*

BF	Basis function	Basis function description	Coefficient	Standard Error	P-value
Intercept	Intercept	Constant	0.103	0.017	0.00000
BF1	(Weather in (3))	Main effect	0.028	0.039	0.0001
BF2	(Weather in (2, 1))	Not sig.	Not sig.	Not sig.	Not sig.
BF3	max(0, Number of lanes- 2) * BF1	Interaction	0.157	0.022	0.00000
BF5	(TRAFFIC_CAT in (1)) * BF2;	Interaction	0.064	0.013	0.00000
BF7	(Age in (1)) * BF2	Interaction	-0.073	0.014	0.00000
BF9	(Education level in (2)) * BF1	Interaction	0.146	0.033	0.00001
BF11	max(0, Number of lanes- 2) * BF2	Interaction	0.063	0.015	0.00001
BF13	(Speed limit in (below 60 mph)) * BF1	Interaction	0.132	0.033	0.00005
BF15	(Education level in (2)) * BF2	Interaction	-0.032	0.013	0.01401
BF19	max(0, Number of lanes- 3)	Not sig.	Not sig.	Not sig.	Not sig.
BF20	max(0, 3- Number of lanes)	Not sig.	Not sig.	Not sig.	Not sig.
BF21	(Driving experience in (2)) * BF19	Interaction	-0.096	0.020	0.00000
BF27	(Speed limit in (below 60 mph)) * BF20	Interaction	0.045	0.011	0.00006

Analyzing the relative importance of variables from the MARS model revealed that weather condition was the most important variable affecting lane-keeping ability, such that it is 3 times more important than the second variable. It indicates that weather conditions play a key role in driver lane-keeping ability *(123)*. This finding is consistent with the previous study that demonstrated the effect of weather conditions on driver behavior in general and lane maintaining in specific *(102)*. The second important variable is speed limit and the third one is traffic conditions. In addition, Table 4 reveals that age, driving experience, and number of lanes are other important factors affecting drive lane-keeping ability.

Table 29 Estimation of Logistic Regression Model for Lane Keeping Behavior *(116)*

Parameter	Description		Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Odds ratio
Intercept	-	-	-1.58	0.31	25.88	<.0001	0.206
Weather	Rain	-	-0.86	0.22	14.53	0.0001	0.425
Weather	Heavy Rain	-	0.79	0.25	9.82	0.0017	2.203
Speed Limit	>60 mph	-	-0.27	0.11	5.62	0.0177	0.764
Age	Middle Age	-	0.31	0.17	3.58	0.0584	1.37
Gender	Female		0.58	0.27	4.72	0.0298	1.785
Traffic Cond.	LOS A-B	-	0.25	0.06	20.95	<.0001	1.288
Driver Mileage Last Year	>12000 miles	-	-0.12	0.06	4.40	0.0358	0.885
Interaction between Weather cond. and Curve	Light Rain	Curves	0.26	0.15	3.00	0.083	1.298
Interaction between Weather Cond. and Traffic Cond.	Light Rain	LOS A-B	0.32	0.08	15.96	<.0001	1.373
Interaction between Weather Cond. and Traffic Cond.	Heavy Rain	LOS A-B	-0.29	0.09	11.15	0.0008	0.747
Interaction between Weather cond. and Age	Heavy Rain	Older Drivers	-0.29	0.16	3.42	0.0645	0.747
Interaction between Weather cond. and Driver Mileage Last Year	Light Rain	>12000 miles	-0.19	0.08	6.63	0.01	0.824
Interaction between Weather cond. and Gender	Heavy Rain	Female	0.13	0.08	2.74	0.098	1.141
Interaction between Weather cond. and Number of Lanes	Heavy Rain	-	0.21	0.07	10.51	0.0012	1.238
Interaction between Weather cond. and Speed Limits	Light Rain	>60 mph	0.31	0.17	3.40	0.065	1.367
Interaction between Weather cond. and Speed Limits	Heavy Rain	>60 mph	-0.29	0.16	3.10	0.0783	0.749
Interaction between Gender and Driving Experience	Female	Less than 3 years.	0.60	0.27	5.12	0.0236	1.827
Interaction between Gender and Driver Mileage Last Year	Female	>12000 miles	0.20	0.05	18.64	<.0001	1.219

Model Fit Statistics:	
AIC	3905.016
SC	4118.405
Log-likelihood at convergence	-4236.771
Number of observations	4753

Hosmer and Lemeshow Goodness-of-Fit Test Chi square = 10.1067,8, p = 0.258.

Logistic regression results indicated that heavy rain has a significant effect on driver lane-keeping ability, which could be due to the shorter sight distance and low visibility of lane marking in heavy rain conditions. This finding is in agreement with previous studies, showing the negative effect of adverse weather on driver performance in general and lane-keeping ability in *(46, 60, 109)*. The effect of maximum-posted speed limits on driver lane-keeping ability was found to be significant. This could be due to drivers more attention to the road ahead in higher speeds and better geometry design and sight distance *(102)*. In addition, driver age was a significant factor in the lane-keeping model. To be specific, middle-aged drivers were 1.4 times more likely to have worse lane-keeping ability in comparison with young drivers.

Traffic conditions were found to have a significant effect on lane-keeping ability. This is not surprising as by increasing traffic congestion, drivers are limited and do not have enough space for maneuvers including swerve and lane change. In other words, drivers are forced to have a better lane keeping in congested traffic. The results provided in Table 5 reveals that drivers who drove in a free flow condition were 1.3 times more likely to have higher SDLP (worse lane keeping) in comparison with those who were driving in congested traffic.

Driver mileage last year was found to be a statistically significant factor that affect driver lane-keeping ability. More specifically, those drivers that drove more than 12,000 miles last year were more likely to have a better lane-keeping in comparison with drivers that drove less than 12,000 miles. This variable can be an index for driving experience; therefore, the result shows the significant effect of driving experience on lane-keeping ability.

Summary

The results from the MARS model revealed that weather conditions were not linearly associated with the SDLP. In addition, among all the explanatory variables considered in the model, weather conditions turned out to be the most important variable affecting lane-keeping performance.

As the importance of analyzing driver behavior in real-time at a trajectory level is becoming more important for various tasks in transportation engineering, the NDS data may help in not only providing a reliable source of trajectory-level driver-behavioral and vehicle information, but also in developing driving models that could be applied to different areas, including but not limited to safety analysis considering microscopic individual driver data; calibrating driver behavior models, specifically in different weather and traffic conditions; and developing control logics for Advanced Driving Assistance Systems (ADAS), and Connected and Automated Vehicles (CAV).

Safety Critical Events (Crashes /Near Crashes)

Literature Review

It is an essential objective in many studies to address causal factors that may lead to an increase in crash risk. One of the efforts that researchers did is to utilize naturalistic driving dataset as the 100-car *(124, 125)*, and SHRP2 NDS dataset *(126)*. In many studies, vehicle kinematics were used to detect crash or near-crash event through having a well-defined parameters signature for vehicle collisions, critical jerks, evasive maneuver, etc. *(127–129)*. VTTI performed a specific methodology to address the factors that may cause an increase of risky events on roadways *(130, 131)*. Firstly, they selected essential vehicle kinematics and their thresholds "triggers" that would help in identifying Safety Critical Events (SCEs). Secondly, a video inspection step was executed on the candidate events to eliminate any false non-SCEs. According to the literature, the first effort done for utilizing NDS dataset was the 100-car study *(125)*. It aimed at providing a primary threshold for vehicle kinematics that can be used to define SCEs. Then many studies tried to modify and adjust these thresholds to capture any risky driving patterns. Table 30 shows the kinematics parameters thresholds specified by the 100-car study and the Naturalistic Teenage Driving Study (NTDS) *(132)*. In addition, a study in Germany 2012, used 102,000 trips collected over 2 years and with total driving miles equal 500,000 miles. This study added new kinematic thresholds for other parameters such as; time headway should be less than or equal to 0.5 second, time to lane crossing should be less than one second, and when the longitudinal deceleration is less than $2 m/s^2$, the jerk should be less than $2 m/s^3$ *(133)*. Another study in the US 2013, utilized NDS dataset collected form 204 drivers over 31 days to study the impact of cell phone distraction by comparing three types of phone answering/usage *(134)*. This study added the type of roadway to the deceleration threshold by specifying aggressive deceleration that is less than -0.3g while the vehicle was traveling above 64 km/h.

Table 30. Different Studies Thresholds for Vehicle Kinematics

Kinematic Parameter	Previous Studies Thresholds	
	100-Car NDS *(125)*	**NTDS** *(132)*
Lateral Acceleration	≥ 0.7g	≥ 0.75
Longitudinal Acceleration	> 0.6g	> 0.65
Longitudinal Deceleration	<- 0.6g	< -0.65
Forward Time-to-Collision (TTC)	≤ 4sec under a condition of having an acceleration ≥ 0.5g or deceleration ≤ -0.5g and with a corresponding forward range ≤ 100 ft	≤ 4sec under a condition of having an acceleration ≥ 0.5g or deceleration ≤ -0.5g and with a corresponding forward range ≤ 100 ft and ≤ 4sec under a condition of having an acceleration ≥ 0.60g or deceleration ≤ -0.60g
Yaw Rate	Vehicle swerve from ±4 deg/sec to ±4 deg/sec within 3sec	Vehicle swerve from ±4 deg/sec to ±4 deg/sec within 3sec

In 2017, a study focused on the performance of kinematic thresholds in detecting crashes and near-crash events using NDS dataset and Canada NDS (CNDS) *(131)*. Sensitivity and specificity were the approaches used to reduce the effort done in validating risky events by developing and validating vehicle kinematics thresholds that can be used for detecting crash risk events. In addition, the study recommended improving the methods used to define crash risk by using advanced statistical techniques and artificial intelligence methods. Statistical methods used to analyze crash data are mainly classified into two types: parametric and non-parametric models. Parametric models were commonly used in previous studies to identify factors contributed to SCE such as negative binomial regression *(135, 136)*, and Multivariate Poisson Lognormal *(137, 138)*. However, non-parametric statistical models were used in the literature such as Latent Class Cluster (LCC) *(139, 140)*, Hierarchical and k-mean Clustering *(140, 141)*, and Bayesian Networks (BNs) *(139)*. It worth to mention that the data used in these studies were just crash frequencies without involving any naturalistic driving datasets in the input data. Therefore, this section is focusing on analyzing NDS time-series data of crash events.

Data Preparation

All near-crash events occurred on freeways reported in the SHRP2 NDS data were verified using visual inspection and by observing any changes in driving patterns in vehicle kinematics time-series datasets. This step classified if a time interval of a trip should be flagged as a risky driving or normal driving. In addition, video records helped in flagging any unexpected behavior from a driver such as sudden lane changes, driving on the shoulder, etc. To identify near-crash events and match them with normal driving trips, latitude and longitude data along with video records were used in the Geographic Information System® (GIS) and the Wyoming NDS Visualization and Reduction Software® *(15)*.

Methodology

The research team utilized 30 near-crash events in rainy weather and randomly selected 60 events in clear weather with a ratio equal 1:2. Additionally, normal trips were used as a baseline while comparing vehicle kinematics associated with events to those in normal driving with a matching ratio to near-crash events equal 2:1 (i.e., 2 normal driving trips for each 1 near-crash event). This technique was implemented to account for any confounding factors related to roadway and driver characteristics. The number of total events used in this section was 30 near-crash trips in rainy weather and 58 matching normal driving trips. In addition to 60 near-crash trips in clear weather and 120 matching normal driving trips. After the research team identified all trips of interest, data reduction step was extended to include traffic and environmental conditions such as the Weather Condition (WC), Dynamic Traffic Status (DTS), Visibility Level (VL), and Road Surface Condition (RC). The extraction process was done by inspecting the video records and annotating any changes in these variables. It is worth mentioning that the time-series NDS data were aggregated over 5-second, 10-second, 15-second, and 60-second time windows as shown in Figure 19. In this research, time chunking was assumed and tested to determine appropriate sampling rate to effectively capture changes in vehicle kinematics. Further, vehicle kinematics are compared in different weather conditions in a trajectory-level analysis. Factors contributing to near-crash events are investigated using parametric logistic regression and several non-parametric techniques such as decision trees and k-nearest neighbors algorithm (k-NN).

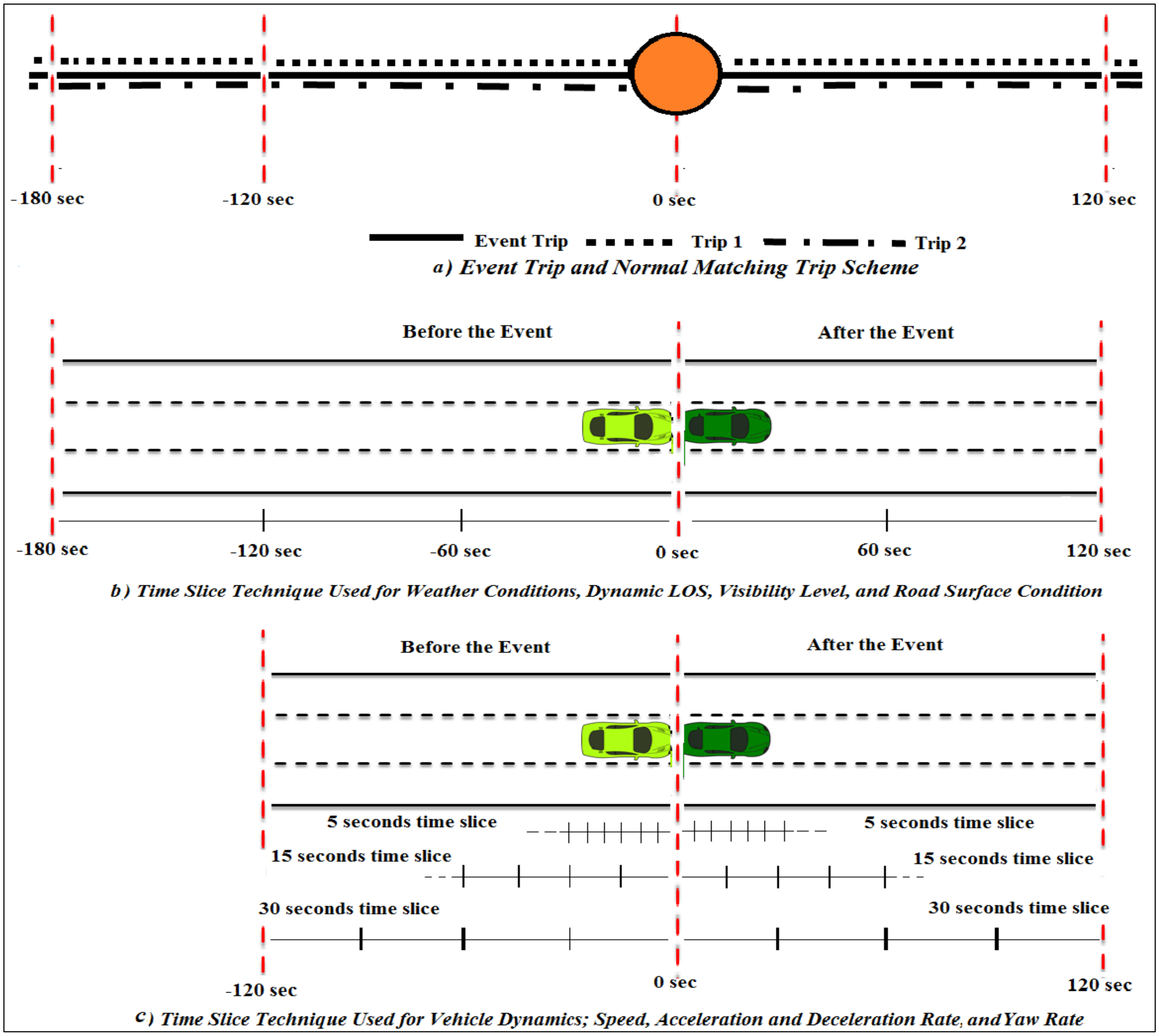


Figure 19 Fixed Length Aggregation levels Technique Used for Data Reduction

Analysis

Vehicle Kinematics for Near-Crash Events in Clear and Rainy Weather

This section is intended to utilize the NDS data in the context of Connected Vehicle (CV) to provide an early understanding of Surrogate Measures of Safety (SMoS) on freeways, and to illustrate the effect of weather conditions on traffic safety by comparing vehicle kinematics of all near-crash events reduced in this report. The zone of interest (i.e., time duration) and parameter extreme values of each vehicle in rainy weather were compared to the corresponding values in clear weather. This visual presentation is considered as a guidance step towards an automated process of extracting different driving patterns that can be used in a CV environment.

Figure 20 shows a comparison between speed, acceleration and deceleration rate, and yaw rate for near-crash events in rainy weather and events in clear weather. The average speeds in clear weather were higher than in rainy weather. However, in clear weather, events had higher standard deviations of speeds than in rain. Figure 20 (b and e) show the difference between acceleration and deceleration rates in rainy and clear weather. The zone of interest of acceleration and deceleration rate in rainy weather started approximately 10-second before an event timestamp while in clear weather it started approximately 5-second before the events. Furthermore, maximum deceleration rates in rain was slightly lower than in clear weather indicating a compensation for wet surface conditions. Yaw rates were lower in rain than their matched clear events. This indicates that drivers in clear weather chose to change lanes more often than in rain.

Figure 21 (a, b, c, d, e, f) summarizes normal trips matching events in clear and rain weather. The reasons for using these visual comparisons was to show how vehicle kinematics look like in normal trips regardless of driver behavior and freeway geometry. It can be observed that wider ranges of acceleration and deceleration rates during events compared to matching normal driving trips. Nevertheless, maximum yaw rates for matching normal driving trips to events in rain were still higher than those trips matched to clear weather condition. This indicated that the change in road geometry might be the reason for increasing yaw rate in rain events and their matching trips. The speed selection range was not affected by having an event or not, but it was affected mostly by weather conditions, traffic conditions, and road geometry.

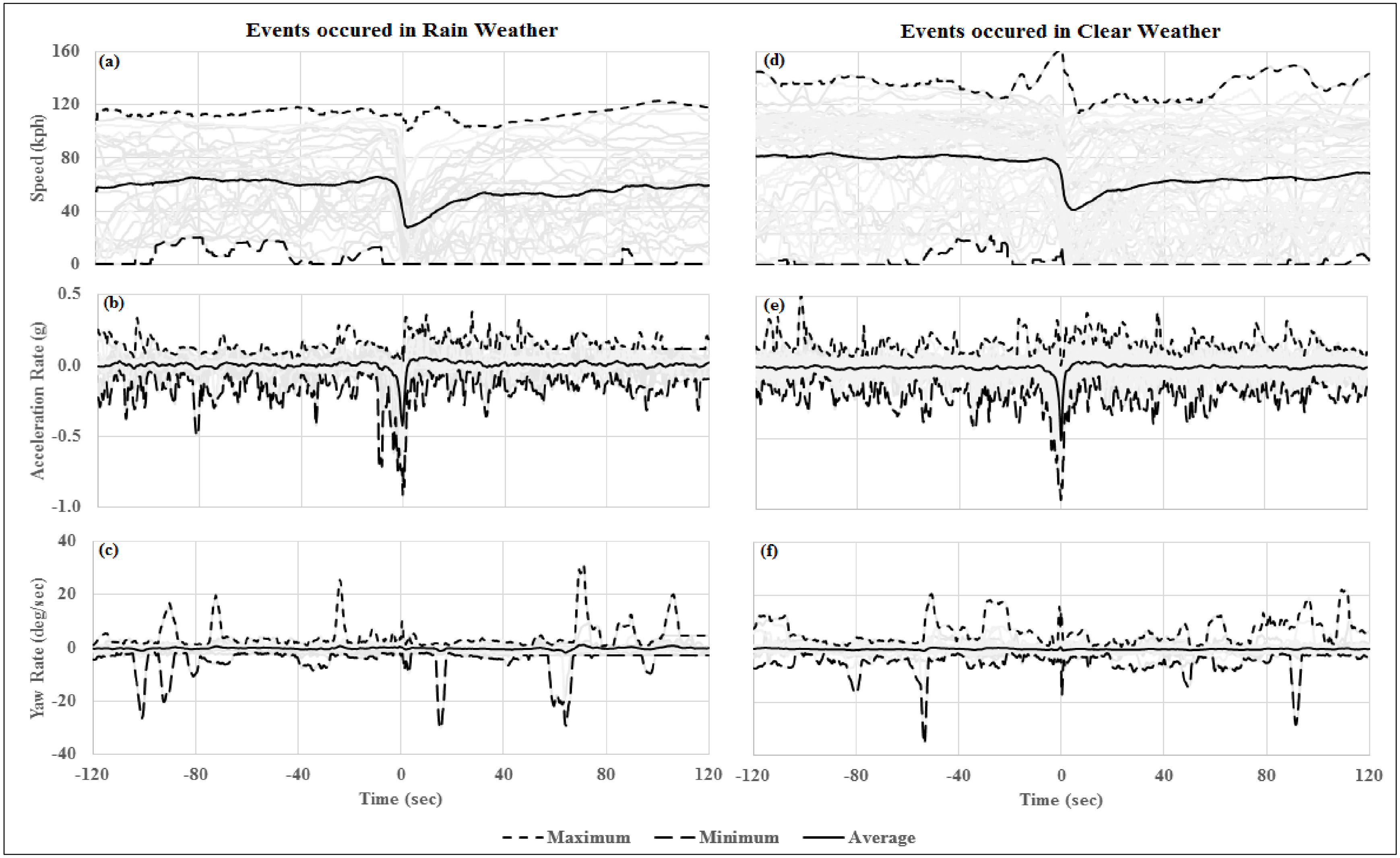


Figure 20 Vehicle Kinematics Trajectories for Events in Rainy and Clear Weather

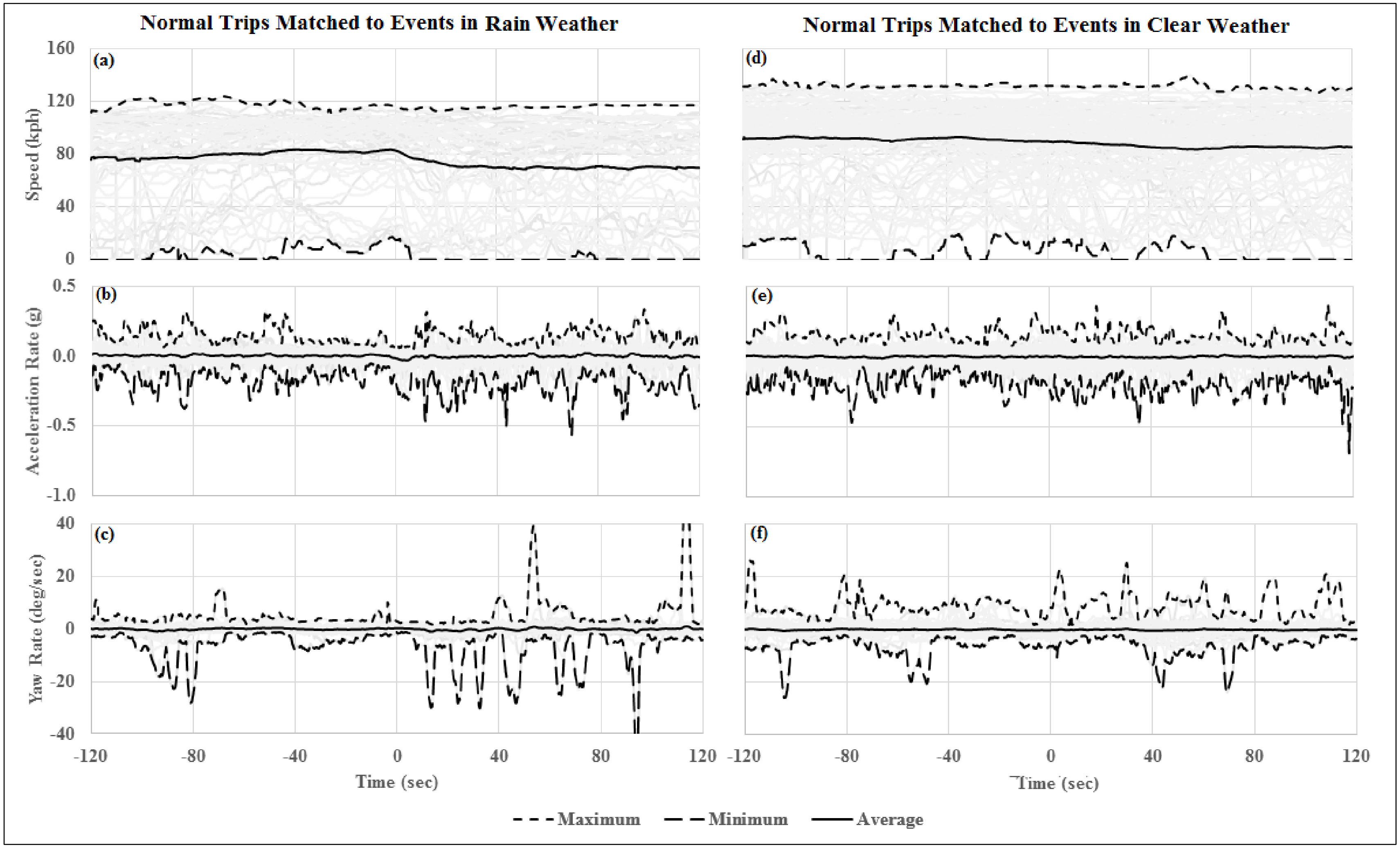


Figure 21: Vehicle Kinematics Trajectories for Matched Normal Trips and Events in Rainy and Clear Weather

Table 31 provides descriptive statistics for near-crash events in rainy and clear weather, and their matching normal trips in clear weather. Average speed, average acceleration and deceleration rate, and average yaw rate were significantly lower for events occurred in rain than clear weather. Standard deviations of acceleration and deceleration rate, and yaw rate were significantly higher in rain than in clear weather. However, the variability of speed was significantly lower in rain compared to clear weather. Furthermore, Table 31 provides comparisons between events in different weather and their normal matching trips. The variability of all vehicle kinematics in rainy and clear weather was higher for the events than for normal trips. Conclusively, weather conditions affected the speed selection, as in clear weather the speed selection had a wider range compared to rainy weather. Additionally, drivers in rainy weather were less likely to change lanes than in clear weather.

Table 31 Descriptive Statistics for the Near-crash Events and Matching Normal Trips

Statistical Tests	**Events in Rainy and Clear Weather**		**Events in Rainy Weather and Matching Normal Trips**		**Events in Clear Weather VS Matching Normal Trips**	
	Rain Condition	**Clear Condition**	**Events in Rain**	**Matching Trips**	**Events in Clear**	**Matching Trips**
Speed (kph)						
Average	56.6	70.5	56.6	75.5	70.5	89.0
Variance	59.2	119.3	59.2	27.1	119.3	9.8
t-test	Avg. Speed was significantly higher in Rain		Avg. Speed was significantly lower in Rain Events		Avg. Speed was significantly lower in Clear Events	
F-test	Speed variability was significantly lower in Rain Events		Speed variability was significantly higher in Rain Events		Speed variability was significantly higher in Clear Events	
Acceleration / Deceleration Rate (g)						
Average	0.0017	-0.0014	0.0017	0.0023	-0.0014	0.0004
Variance	0.0015	0.0014	0.0015	0.0001	0.0014	0.0000
t- Test	Avg. Acc. / Dec. rate was significantly higher in Rain Events		No significant difference in Avg. Acc. / Dec. rate		Avg. Acc. / Dec. rate was significant lower in Clear Events	
F- test	Acc. / Dec. rate variability was higher in Rain Events		Acc. / Dec. rate variability was higher in Rain Events		Acc. / Dec. rate higher in Clear Events	
Yaw Rate (deg/sec)						
Average	-0.2097	-0.0955	-0.2097	-0.0633	-0.0955	-0.1009
Variance	0.1401	0.0471	0.1401	0.1162	0.0471	0.0248
t- Test	Avg. Yaw Rate was significantly lower in Rain Events		Avg. Yaw Rate was significantly lower in Rain Events		No significantly difference between Avg. Yaw Rate	
F- test	Yaw Rate variability was significantly higher in Rain Events		Yaw Rate variability was significantly higher in Rain Events		Yaw Rate variability was significantly higher in Clear Events	

Near-Crash Events Detection on Freeways

The relation between data reduced form video records and SMoS was demonstrated to predict the occurrence of a near-crash event. A parametric model and non-parametric techniques were used in this step. Input data were sampled over 1-second and 5-second aggregation levels. The selecting of the efficient aggregation level was done after assuming five different aggregation levels: 1, 5, 10, 15, and 60-second, and running a binary logistic model. The model was used to validate the aggregation levels using trial and error technique. The results indicated that using the aggregation level of 1-second and 5-second length could show more reliable results, while other aggregation levels failed to provide any significant predictors. Through these models, the importance of SMoS and environment factors was illustrated, and how the probability of having a near-crash can be estimated using SMoS.

Parametric Model
Factors contributing to near-crash occurrence can be identified using both parametric and nonparametric models. The logistic regression model could provide relationship between the probability of near-crash occurrence and significant predictors. One of the advantages of using logistic regression model is the feasibility of interpreting the effects of predictors.

Table 32 presents the results of a binary logistic regression model. The model results for the 1-second aggregation level indicates that significant predictors were standard deviation of acceleration and deceleration rate, coefficient of variation for acceleration and deceleration rate, and the coefficient of variation for yaw rate at six, eleven, and two seconds before the event timestamp, respectively. However, model results for 5-second aggregation level were weather condition, dynamic traffic state, visibility level, and coefficient of variation of yaw rate before event timestamps. The logistic regression model proved no evidence of model poor fitting. Moreover, the models showed a significant difference when different aggregation levels were used (i.e., 1-second and 5-second) and how this would affect model accuracy. The results succeeded to verify the contribution of weather condition, visibility level, and traffic flow status within 5-second before near-crash events. Additionally, the results showed how vehicle kinematics could help in predicting near-crash on freeways in the last 11 seconds before the near-crash depending on time-series data collected by the NDS instrumented vehicle. This finding can be used in CV applications to enhance traffic safety.

Table 32 Logistic Regression Results for Modeling Near-Crash Occurrence on Freeways

A. 1-second Aggregation level

Variable	Coefficient	Standard Error	Wald Chi-Square	Significance
Intercept	-5.51	0.81	45.96	<.0001
Standard Deviation of Acc. and Dec. Rate (T106)*	39.31	16.89	5.41	0.02
Standard Deviation of Acc. and Dec. Rate (T120)*	134.60	21.36	39.70	<.0001
Coefficient of Variance of Acc. and Dec. Rate (T100)*	0.24	0.10	5.84	0.02
Coefficient of Variance of Yaw Rate (T118)*	-0.40	0.20	4.23	0.04

B. 5-second Aggregation level

Variable	Coefficient	Standard Error	Wald Chi-Square	Significance
Intercept	-1.79	0.30	34.88	<.0001
WC	1.37	0.34	16.37	<.0001
DTS	0.90	0.32	7.97	0.00
VL	0.76	0.33	5.29	0.02
Coefficient of Variance of Yaw Rate (T120)*	0.03	0.02	3.20	0.07

*Vehicle Kinematics time slices (T) were numbered as following:

- The data were used from 30 seconds before the event (time slice number 90) till the event timestamp (T120).
- For 1-second aggregation level: T91, T92, T93, …, T118, T119, T120.
- For 5-second aggregation level: T95, T100, T105, …, T110, T115, T120

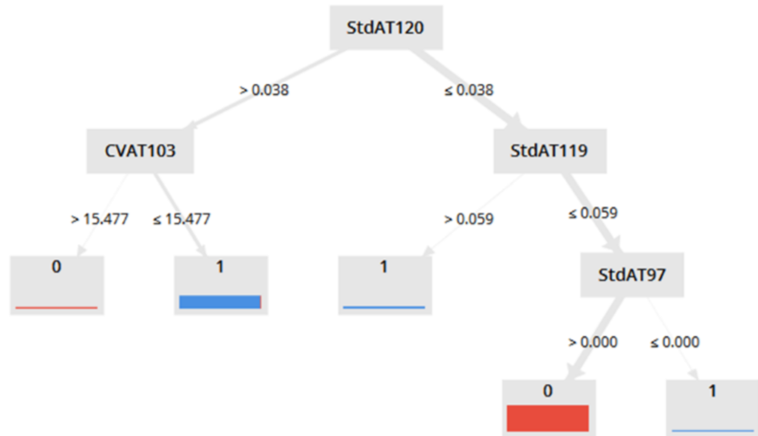
Non-Parametric Models

One of the main advantages using a non-parametric model over a parametric model is that no certain assumptions are needed between dependent and independent variables. In addition, a non-parametric model can handle massive datasets while a parametric model has some limitations on data size *(141, 142)*. This analysis selected three supervised Machine Learning algorithms: Decision Tree, K-Nearest Neighbors (K-NN) Classification, and Deep Learning Artificial Neural Network (ANN), Classification models. These models were used to predict the occurrence of a near-crash event on a freeway in clear and rainy weather. These models were evaluated based on the accuracy of the developed model to predict near-crash events.

Figure 22 shows the results of the decision tree model. Results indicate the contribution of vehicle kinematics in predicting a near-crash event is in a zone of interest approximately 20 seconds time window before the event occurrence. Additionally, weather condition, i.e., rain or clear, was the root node of the decision tree when 5-second aggregation level was used.

a. Time Slice of 1-second Length

StdAT120 > 0.038
| CVAT103 > 15.477: 0 {1=0, 0=2}
| CVAT103 ≤ 15.477: 1 {1=81, 0=1}
StdAT120 ≤ 0.038
| StdAT119 > 0.059: 1 {1=7, 0=0}
| StdAT119 ≤ 0.059
| | StdAT97 > 0.000: 0 {1=0, 0=175}
| | StdAT97 ≤ 0.000: 1 {1=2, 0=0}



b. Time Slice of 5-second Length

WC = 0
| StdY100 > 0.085: 0 {1=50, 0=154}
| StdY100 ≤ 0.085
| | StdA100 > 0.025
| | | Sex = 0: 0 {1=0, 0=4}
| | | Sex = 1: 1 {1=2, 0=0}
| | StdA100 ≤ 0.025: 1 {1=8, 0=0}
WC = 1
| StdA120 > 0.030
| | CvY100 > 2.171: 0 {1=0, 0=3}
| | CvY100 ≤ 2.171
| | | CvS105 > 0.345: 1 {1=29, 0=4}
| | | CvS105 ≤ 0.345: 0 {1=0, 0=2}
| StdA120 ≤ 0.030: 0 {1=1, 0=11}

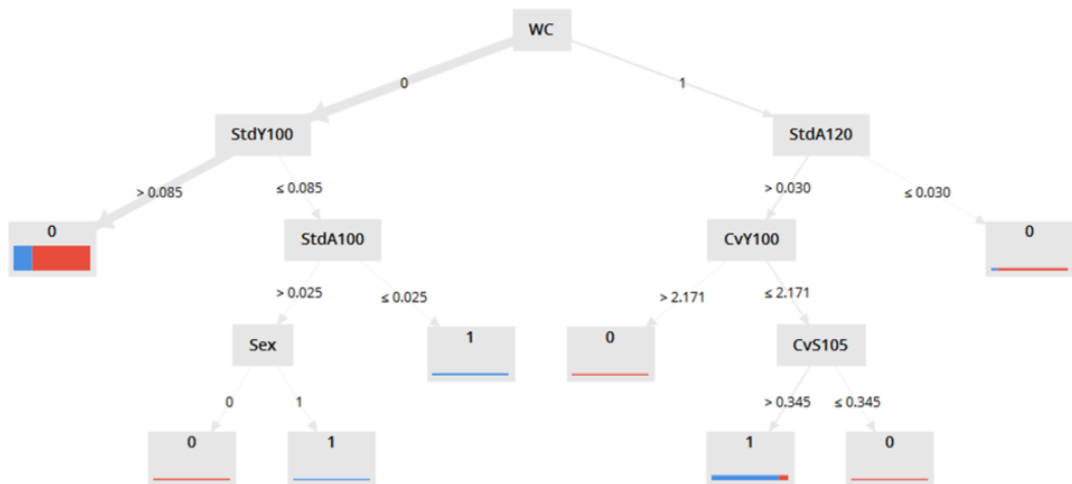


Figure 22 Result of Decision Tree Model Applied on 1-second and 5-second Aggregation Levels

Table 33 shows models accuracy and the confusion matrix associated with 1-second and 5-second aggregation levels. Results show a significant difference in accuracy based on the aggregation level. For example, the results of 1-second show that the Decision Tree model had the highest accuracy of 96 percent, followed by Deep Learning ANN with an accuracy equal 84 percent, then K-NN with an accuracy equal to 81 percent. However, for 5-second, Deep Learning ANN had the highest accuracy of 85 percent, followed by Decision Tree model with an accuracy equal 69 percent, then the K-NN model with an accuracy equal 63 percent.

Table 33: Results of Non-Parametric Models Used for Predicting Near-Crash Events

Non-Parametric Models	Aggregation Level	Overall Classification Rate	True Positive Rate	True Negative Rate	False Positive Rate	False Negative Rate
1. Decision Tree	1- Second	96%	93%	97%	3%	7%
	5- Seconds	69%	13%	97%	3%	87%
2. K-NN	1- Second	81%	49%	98%	2%	51%
	5- Seconds	63%	20%	85%	15%	80%
3. Deep Learning ANN	1- Second	84%	87%	83%	17%	13%
	5- Seconds	85%	93%	81%	19%	7%

Summary & Next Steps

Continuous vehicle kinematics datasets enhanced the understanding of the rainfall effects on increasing the probability of having a near-crash event, through dissimilarity in the driver's behavior, vehicle kinematics, and driving patterns while having a safe or risky driving. The speed, acceleration and deceleration rate, and yaw rate can be utilized as SMoS indicators to differentiate between normal driving and near-crash events. Supervised machine learning techniques proved to enhance the verification process by predicting near-crash occurrence on freeways. Moreover, the time chunking technique attempted in this approach would help future studies in outlining the interest zone of vehicle kinematics used as SMoS for near-crash events.

The effort done in summarizing vehicle kinematics signatures, by adding environmental conditions as a new dimension, will help in extracting different driving pattern through developing an algorithm that can handle CV datasets. Results from parametric and non-parametric models might be used in Advanced Driver Warning Systems (ADAS), as warning messages could be displayed to drivers between 5 and 20 seconds before the occurrence of a risky event. Additionally, future work will be extended to automate the data reduction and analysis for CV applications through utilizing supervised and unsupervised machine learning techniques.

CHAPTER 4 – CONCLUSIONS AND PLANS FOR PHASE 3

This report provides a detailed overview of the findings from the second phase of the Wyoming Department of Transportation (WYDOT) SHRP2 Implementation Assistance Program (IAP) project. The IAP is a FHWA sponsored program that was created to encourage state DOTs to use the SHRP2 Safety Data to conduct research that leads to practical countermeasure development. The program is divided into three stages, starting with a proof-of-concept phase, moving into a full research analysis phase, and ending with countermeasure development and adoption. Due to Wyoming's severe winter seasons, the project team elected to address driver behavior characteristics prevalent during adverse weather and roadway conditions. Therefore, using the data available from the SHRP2 Naturalistic Driving Study (NDS) and Roadway Information Database (RID), the research team investigated the impact of adverse conditions on driver behavior.

The following sections provide an overview of the WYDOT SHRP2 IAP research objectives and how these objectives have been addressed in the first two phases or will be addressed in the third phase. Next, specific contributions from Phase 2 are described and core objectives for Phase 3 are explained. Finally, a summary of lessons learned related to the project team's experience using the SHRP2 Safety Data are provided.

Review of Project Objectives

The project objectives described in chapter 1 are listed below with a description of how the presented research met each objective.

1. Can NDS trips occurring in inclement weather be identified efficiently and effectively using available NDS and RID data?

Findings from the first project phase indicated the feasibility of collecting weather-related trips using the NDS vehicles' windshield wiper status. Phase 2 introduced two additional data acquisition methodologies which enabled the collection of trips from a wide variety of weather conditions. The novel data acquisition procedure developed for Phase 2 uses three complementary methodologies: (i) windshield wiper status, (ii) weather stations' data, and (iii) weather-related crash reports. In order to capture a wide range of roadway conditions—those in active weather events (e.g., during a blizzard) and those not in active weather events (e.g., black ice road conditions)—a 24 hour time window surrounding recorded weather events in methods (ii) and (iii) was used. Using these three methodologies, more than 11,000 NDS trips were extracted as being *potentially* weather-related trips. Through efficient data reduction and processing procedures, these trips were evaluated and 4,094 NDS trips were confirmed to have been influenced by adverse weather conditions. Of these adverse trips, 3,013 occurred in rain, 234 in fog, 320 in snow, 317 in clear conditions with wet pavement, and 210 in clear conditions with snow-covered pavement. In addition to collecting full trips occurring during adverse conditions, weather-related crash and near crash event data were collected. From these collected events, 7 crashes and 33 near crashes occurred in rain or sleet conditions.

2. Can driver behavior (e.g., speed selection, car-following, and lane wandering) during inclement weather conditions be characterized efficiently from the NDS data?

A central focus of Phase 2 was deriving appropriate methods for characterizing different elements of driving behavior such that behavior differences in adverse and clear conditions could be compared. The research team successfully implemented a series of speed selection models which identified desired driving speeds during different weather conditions (i.e., snow, rain, and fog). Parametric ordinal logistic regression and non-parametric classification tree modeling techniques were leveraged to advance the understanding of drivers' selected speeds during different levels of adverse weather. Drivers' car-following behavior was analyzed through the calibration of the Gipps car-following model. The analysis identified whether observed changes in drivers' following distance, following headway, and relative speed kept with the leading vehicle matched the predicted behaviors from the calibrated model. Drivers' lane-keeping behavior during adverse conditions was analyzed using logistic regression and multivariate adaptive regression splines (MARS). The results indicated that weather conditions were a significant factor related to drivers' lane-keeping ability. Finally, safety critical events (i.e., crashes and near crashes) were reviewed. Both parametric and non-parametric models were used to describe vehicle kinematics during these events.

3. What are the best surrogate measures for weather-related crashes that can be identified using the NDS data?

The third project objective was addressed in the analysis of safety critical events. The purpose of generating surrogate measures for weather-related crashes is to inform research aiming to identify near crash events in real time. With the introduction of connected vehicle technology, vehicle kinematics data will be available in real time; therefore, the identification of specific surrogate measures related to crash and near crash events, especially those occurring in adverse conditions, is critical. The findings from Phase 2 indicate that vehicle speed, acceleration, deceleration, and yaw rate can be utilized as surrogate measures of safety to differentiate between normal driving and near crash events.

4. What type of analysis can be performed and conclusions drawn from the resulting dataset?

Analyzing the role of adverse conditions on driving behavior is crucial for developing countermeasures to improve the safety and reliability of the transportation network during and after adverse weather events. The behavior analyses presented in this report take many forms in order to better characterize different elements of the complex driving task. These analyses included parametric and non-parametric modeling techniques, calibration techniques of existing behavioral models, and machine learning techniques. The results of these analyses will directly contribute to the improvement of the Wyoming Variable Speed Limit (VSL) control algorithm, weather-related microsimulation modeling procedures needed for the Wyoming Connected Vehicle Pilot project, and future work aiming to identify weather-related safety critical events in real time.

5. Can the NDS data be extrapolated to provide real-time weather information in the context of the Road Weather Connected Vehicle Applications?

The last project objective was addressed by developing a tool called Wyoming NDS Visualization and Reduction Software to identify the visibility conditions in real-time using AI techniques. The software is still under development to increase the accuracy and precision of estimates. In addition, data mining methods were utilized to detect weather events from vehicle kinematics. As mentioned earlier, considering the similarities between the trajectory-level NDS and the CV data, the results provided an early understanding of Surrogate Measures of Safety (SMoS) on freeways, and illustrated the effect of weather conditions on traffic safety by comparing vehicle kinematics of near-crash events.

Review of Phase 2 Contributions

Chapter 2 describes the data acquisition procedures used to query the SHRP2 NDS database. The three complementary methodologies developed by the project team introduce a novel method for identifying weather-related natural driving trips using both internal vehicle data and external weather information. Using these procedures, trips influenced by weather—both during and after the weather event—were captured and used for analysis. These complementary methodologies not only contribute a method to extract weather-related trips from the SHRP2 NDS, but are extendable to other NDS and vehicle-trajectory databases available worldwide.

Chapter 2 also discusses the data reduction tools developed to efficiently and effectively process the extracted trip files from the SHRP2 NDS. Due to the large quantity and size of trip files, efficient data processing procedures were developed. In order to automate a portion of the data reduction process, the project team created the Wyoming NDS Data Analysis Tool, which is a python-based analytic tool that produces summary statistics, performs time chunking procedures, and creates video observation templates. As part of this tool, summary files can be generated that describe entire trips or specific trip segments, such as individual time chunks or a single car-following event.

The second component to data reduction and processing is video observation in order to identify weather, roadway, and traffic conditions. For this project phase, manual video processing was required to identify the roadway type, weather condition, road surface conditions, visibility, and level of service. While explicit definitions of discrete condition categories were described to all video-reviewers, manual video observation was not preferred due to its time-intensive nature. For this reason, the project team developed the Wyoming Data Visualization and Reduction tool aiming to reduce the burden of manual video reduction. This tool utilizes state-of-the-art techniques for image processing to identify visibility levels based on three thresholds (low, moderate, and high). In addition, the research team is exploring alternate machine learning techniques for more effectively characterizing other aspects of the video image including surrounding traffic and road surface conditions. Due to the complex nature of these methods, this tool is still under development and the work will be continued during the third project phase.

The purpose of Phase 2 of the IAP is to conduct a series of in depth research analyses to answer the project research questions and identify countermeasures from the research experience and findings to improve network safety and reliability. As described, the WYDOT IAP is centered upon advancing the understanding of driver behavior during adverse weather conditions. To this end, a series of analyses related to different elements of the driving task were investigated separately and their results are provided in chapter 3. As part of this in-depth analysis, speed selection, car-following behavior, lane-keeping behavior, and safety critical events were investigated during adverse weather conditions, compared with matching clear conditions. In order to perform these analyses, the project team leveraged a trip matching technique which ensured that all adverse weather trips were matched with at least two clear weather trips taken by the same driver, on the same route. In this way, drivers' reactions to weather conditions could be better isolated and the research team could identify findings with higher confidence.

As the core of this report, chapter 3 presents the behavior analysis for each of the described elements of the driving task, as individual avenues of research focused on specific areas of driver behavior were identified and investigated separately. These areas of driver behavior research include speed selection, car-following, lane-keeping, and safety critical events. While these research areas were presented separately due to their distinctive nature, each is a crucial element in identifying driver behaviors specific to different weather conditions. Improved understanding of weather-induced driver behavior can inform decisions made by transportation agencies through the identification of key roadway characteristics or specific roadway segments in which drivers' behavior is more severely impacted during adverse weather. A summary of the contribution from each of these research avenues are provided below.

- Generation of a series of speed selection models highlighting drivers' desired speeds during various weather conditions. Parametric ordinal logistic regression and non-parametric classification trees modelling were utilized to better understand speed selection behavior in adverse weather conditions. The purpose of this analysis was to gather insights into driver speed preferences in different weather conditions, such that efficient logic can be implemented to introduce a realistic Variable Speed Limit system, aimed at maximizing speed compliance and reducing speed variations. The analysis also provides valuable information related to drivers' interaction with real-time changes in roadway and weather conditions, leading to a better understanding of the effectiveness of operational countermeasures.
- Evaluation of car-following behavior in clear and adverse weather conditions to distinguish the difference in drivers' perception of a leading vehicle. The Gipps car-following model was calibrated using both clear and adverse trips to determine if the subtle changes in car-following behavior could be captured by model calibration. Results are promising and future work in Phase 3 will evaluate the impact of improved car-following model calibration in microsimulation needed for the Wyoming Connected Vehicle Pilot Program.
- Investigation of drivers' lane-keeping ability during adverse weather events. Both parametric logistic regression and non-parametric MARS modelling approaches were utilized to identify contributing factors affecting driver lane keeping ability in different weather conditions. Results from this study may provide insights into automating the activation and deactivation of lane departure warning systems.
- Development of a framework to analyse and identify factors affecting safety critical events in a trajectory level utilizing classical logistic regression and data mining decision

tree and K-NN approaches. The results will help in extracting different driving pattern through developing an algorithm that can handle CV datasets.

As described, the intent of the Wyoming SHRP2 IAP project is to analyze weather-related driving behavior for the purpose of identifying specific countermeasures and applications that WYDOT can adopt and champion to improve transportation safety and reliability. Therefore, the critical contribution of this project phase is the identification of three explicit areas in which the research presented in Chapter 3 can be used in practice:

1. Improvement to existing weather-dependent variable speed limit (VSL) control algorithm used by WYDOT for their interstate VSL systems.

Due to the limited understanding of the interaction between driver behavior/performance and weather conditions, the continuation of this research into Phase 3 aims to establish a Connected Human-in-the-Loop VSL system, which is aligned with the SHRP2 Task Force's focus areas. An important component of the driver-weather interaction is the characterization of traffic flow because heterogeneity driving behavior is different among differing weather conditions and levels of congestion. Modeling variation in driver behavior with adverse weather conditions and traffic flow states is crucial to assign effective VSLs, as these algorithms must consider the impact of both weather and traffic conditions when suggesting the safest and most efficient speed. In fact, the proposed speed selection models in this study will be evaluated for direct integration into the VSL algorithms considering detailed traffic, weather, and driver information.

The updated VSL algorithms could provide speed limits that are better suited to real time weather and traffic conditions, which is expected to improve drivers' speed compliance. For instance, the results of the CART speed selection modeling revealed that 56% of drivers were likely to reduce their speed by more than 14 percent in snowy surface conditions, free-flow traffic, and affected visibility conditions. From these results, behavior adjustment factors for similar weather conditions can be generated and applied to the VSL algorithm to improve speed compliance.

An additional benefit from the developed speed selection models may be introduced in Connected Vehicle (CV) speed harmonization applications, where the VSL system could be expanded to ingest mobile vehicle data as an input and transfer VSL notifications to on-board units (OBU). The OBUs could then provide speed advisories, regulatory speeds, or other related advisories to the driver. Messages such as, "turn off cruise control", could be sent in real-time to more effectively regulate driving speed and preserve a safe flow of traffic. If unusual traffic patterns are detected, or inclement weather events are forecasted or experienced, these geospatial locations could be flagged for implementation of an appropriate and timely mitigation strategy to reduce the impact of the adverse weather condition.

2. Improved guidance related to microsimulation modeling of adverse weather conditions, and generation of a "base model" to represent driving behavior in adverse weather conditions for use in the Wyoming CV Pilot project impact assessments.

The inclusion of microsimulation analyses in agency decision-making procedures is common. The results from these models can guide cost-benefit analyses by projecting the impact of different traffic control strategies. Agencies often leverage the detailed analysis to compare alternative intersections designs, forecast traffic congestion with and without added lanes, and assess cutting-edge traffic control strategies before they're implemented in the field (e.g., diverging diamond interchanges were originally developed and evaluated using microsimulation software). Successful microsimulation base models accurately reflect real conditions—roadway geometry, existing traffic control strategies, travel demand, and driving behavior—on the selected roadway segment. From the base model, different alternatives and projections are completed by adjusting one or more of these factors based on potential construction or travel demand. Once roadway geometry, traffic control strategies, or travel demands are adjusted in a model, the glue holding the model together is the accurate representation of driving behavior. Without calibrated driver behavior data reflecting specific driver tendencies and driving conditions prevalent on the modeled corridor, accurate predictions cannot be achieved.

Compared to the rest of the country, the state of Wyoming has only a few isolated areas of high travel demand and congestion in which the common approach to microsimulation modeling is required. Rather, due to frequent shifts in weather conditions that have severe impacts on the rural transportation network, microsimulation modeling could be used to assess the impact of large weather events on existing roadways or scenarios such as proposed work zones and large events drawing an increased travel demand. Therefore, this research supports the development of a weather-dependent calibration procedure through the evaluation of driving behaviors (i.e., speed selection, car-following, and lane keeping). Unprecedented insight into driver behaviors in adverse weather conditions available through the SHRP2 NDS database enabled the project team to develop procedures to calibrate components of microsimulation models considering weather as a fundamental behavioral factor. These findings will be used to generate a microsimulation "base model" representing the impact of various weather conditions on driving behavior. Future studies will evaluate the transferability of these weather-related behaviors in different geographical regions.

Lastly, the project team's motivation for developing procedures to calibrate a weather-dependent microsimulation model is related to the needs of the Wyoming Connected Vehicle (CV) Pilot project. The Wyoming CV Pilot is developing a rural CV application to communicate standard messages between equipped vehicles and infrastructure devices. These messages are expected to improve safety and roadway efficiency by providing drivers with greater awareness of roadway and traffic conditions, with a specific emphasis on adverse weather events. A pivotal component of this project is assessing the impact the applied applications have on driver behavior and on overall roadway efficiency. To this end, the CV Pilot team will rely on microsimulation modeling to assess the impact of the applied CV applications and will likely leverage the procedures established in this SHRP2 IAP to calibrate their base (non-CV) model in adverse weather conditions.

3. Evaluation of SHRP2 NDS weather-related vehicle dynamics to support the development of real-time CV applications requiring weather and roadway condition input data.

The third practical output from Wyoming's SHRP2 IAP relates to the detailed analysis of safety critical events occurring in clear and adverse weather conditions. The future of connected vehicle (CV) technology relies on successful application of the surplus of data that will be generated by the constant communication between vehicles, infrastructure, and other road users. Basic Safety Messages (BSMs) describing vehicle dynamics will only be pertinent for a brief amount of time; therefore, efficient protocols and algorithms are required to identify and react to events of interest in real time.

The findings from this study identify surrogate measures of safety—or critical vehicle kinematics data elements—that may serve as indicators of near crash or crash events using the SHRP2 NDS data. Due to the lack of available CV data, researchers are required to leverage surrogate data sources to begin developing protocols for handling the influx of CV data. The SHRP2 NDS provides vehicle data very similar to the data available from BSMs; therefore, researchers took advantage of the opportunity to use these data to add to the state of research. The goal of this objective is to identify specific measures and thresholds that might distinguish between weather-related and non-weather-related safety critical events. With the limited sample size available, the project team produced a set of findings identifying the significance of vehicle kinematics variables in predicting safety critical events and will continue to develop recommendations for future real time CV applications.

Plans for Phase 3

As detailed earlier, the objective of the second phase was to conduct a thorough analysis using a larger set of NDS trips to extract behavioral trends specific to a wide variety of weather conditions (i.e., rain, snow, and fog) from a diverse driver population from each of the six SHRP2 data collection sites. The objective of Phase 3 is to interpret these findings such that they can be used to inform the development of Wyoming-based safety and reliability countermeasures.

Phase 3 of the Wyoming SHRP2 IAP project begins in October 2017 and will conclude by the end of September 2019. The solid foundation generated in the first two project phases will be used to enhance the existing weather-dependent VSL system operated by WYDOT. Specifically, the speed selection models will be validated using available data from Wyoming interstates to develop a suitable algorithm for VSL operation. The car-following and lane-keeping findings will be used to develop weather-related microsimulation model guidance that could be used to evaluate future countermeasures. Finally, the analyses from safety critical events in adverse and clear conditions will be used to provide recommendations for surrogate measures of safety that could enable the detection of crash and near crash events in real time using connected vehicle data.

Summary of Lessons Learned

The SHRP2 NDS and RID databases are extremely rich databases that introduce substantial potential to researchers for better understanding driving behavior in a large number of natural driving environments. In order to support the development of new project statements and inform researchers interested in using the SHRP2 data, the project team aggregated a list of lessons learned from their experience in Phase 1 and 2 of the IAP.

- Sole utilization of the wiper setting variable to identify weather-related trips is not sufficient as this variable is missing for significant number of NDS trips—which could be due to an error in the DAS system or the overrepresentation of old vehicles in the SHRP2 dataset—and only enables the collection of trips during active precipitation.
- Use of multiple sources of weather data, such as weather stations and weather-related crashes, while costly and labor intensive, enhanced the capability of the research team to identify sufficient weather-related events.
- Employing an effective data reduction procedure is a critical requirement when using the SHRP2 NDS and RID databases due to their massive dimensionality. Creative reduction methods (such as, one-minute sampling rates) are useful in efficiently analysing a substantial number of trips.
- Manual observation of the trips to identify surface, weather, visibility, and traffic conditions was a successful procedure; however, detailed training of video observers about different levels of each specific variable should be considered to reduce the subjectivity. In addition, considering the time and labor work, more advanced techniques should be considered in future research to automate the process as much as possible.
- Related to the previous point, the research team is currently working on new methods to capture the visibility conditions from the front view camera using AI and Deep Learning techniques. However, video quality—specifically in adverse weather—and variations in physical camera location among different instrumented vehicles are challenges for this effort.
- Assessing the effectiveness of different countermeasures on NDS traversed routes requires the exact time of the trips as well as the external data related to each specific countermeasure. Even though some information about each countermeasure might be acquired from different agencies, the important component of exact time of the trips is considered as Personally Identifiable Information (PII), which is one of the biggest challenges to get the maximum benefits out of SHRP2 NDS dataset. The USDOT FHWA is actively working on this issue by leveraging the SHRP2 NDS data with sanitized PII data and provide them to the NDS time-series data.
- Driver behavior characteristics are often considered site specific, therefore transferability of the obtained results is not guaranteed. However, the diverse data collection sites and the large amount of available data through the SHRP2 NDS enabled the assessment of result transferability and decreased probability of model overfitting. The continuation of this research will compare the trajectory-level speed data collected from the CV pilot deployment project in Wyoming to provide a better insight regarding the transferability of SHRP2 results to I-80 corridor.
- The SHRP2 NDS data has a remarkable similarity to CV data; therefore, this data source can be used as surrogate data source to assess CV applications before CV data are readily available.

Concluding Statements
Adverse weather conditions severely impact the operations and safety of the transportation network. Due to unique challenges of severe adverse weather conditions on remote highways and interstates throughout the state of Wyoming, WYDOT is motivated to identify solutions that will enhance travel safety and reliability. While substantial research has been conducted to identify the impact of adverse weather on the transportation network, few studies have focused on the root cause of those impacts. All network-wide impacts attributed to various adverse weather events are caused by specific driver behavior adjustments in response to those adverse conditions weather conditions. With the generation of the SHRP2 NDS and RID, researchers have a new opportunity to evaluate driver behavior in a multitude of different driving environments, including adverse weather conditions. Recognizing this opportunity WYDOT entered the SHRP2 Implementation Assistance Program (IAP) and was awarded three project phases to research the impact of adverse weather conditions on individual driver behaviors and develop the findings into countermeasures to address these impacts.

This report presents the findings from the second phase of the Wyoming SHRP2 IAP project. The reported findings include novel data acquisition and efficient data reduction strategies aiming to increase the usability of the SHRP2 NDS data. In addition, advanced machine learning and image processing strategies were leveraged to automatically extract valuable data from the NDS video footage. Distinctive areas of research related to driver behavior (i.e. speed selection, car following, lane keeping) were evaluated using a series of analytic and modeling techniques to interpret drivers' decision making and represent the behavior adjustments in mathematics and empirical models for future prediction. The report concludes with a discussion of the findings and an introduction to the third project phase which will build upon these findings to produce tangible countermeasures to improve transportation safety and reliability during adverse weather events.

ACKNOWLEDGEMENTS

This work was conducted under the second Strategic Highway Research Program (SHRP2), which is administrated by the Transportation Research Board (TRB) of the National Academies of Sciences, Engineering, and Medicine, and it was sponsored by the Federal Highway Administration (FHWA) in cooperation with the American Association of State Highway and Transportation Officials (AASHTO).

The authors of this report would also like to extend their gratitude to the Wyoming Department of Transportation (WYDOT) for supporting and funding this research project. All statements and opinions presented in this report are the sole responsibility of the authors, and may not necessarily reflect those of WYDOT.

REFERENCES

1. Ahmed, M., M. Abdel-Aty, and R. Yu. Assessment of interaction of crash occurrence, mountainous freeway geometry, real-time weather, and traffic data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2280, 2012, pp. 51–59.
2. Hamilton, B. A. Ten-year averages from 2005 to 2014 based on NHTSA data. 2015.
3. Hambly, D., J. Andrey, B. Mills, and C. Fletcher. Projected implications of climate change for road safety in Greater Vancouver, Canada. *Climatic Change*, Vol. 116, No. 3–4, 2013, pp. 613–629.
4. Vogt, A., and J. Bared. Accident models for two-lane rural segments and intersections. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1635, 1998, pp. 18–29.
5. Strong, C. K., Z. Ye, and X. Shi. Safety effects of winter weather: the state of knowledge and remaining challenges. *Transport reviews*, Vol. 30, No. 6, 2010, pp. 677–699.
6. Brodsky, H., and A. S. Hakkert. Risk of a road accident in rainy weather. *Accident Analysis & Prevention*, Vol. 20, No. 3, 1988, pp. 161–176.
7. Andrey, J., B. Mills, M. Leahy, and J. Suggett. Weather as a chronic hazard for road transportation in Canadian cities. *Natural Hazards*, Vol. 28, No. 2–3, 2003, pp. 319–343.
8. Rahman, A., and N. E. Lownes. Analysis of rainfall impacts on platooned vehicle spacing and speed. *Transportation research part F: traffic psychology and behaviour*, Vol. 15, No. 4, 2012, pp. 395–403.
9. Fuller, R. Towards a general theory of driver behaviour. *Accident Analysis & Prevention*, Vol. 37, No. 3, 2005, pp. 461–472.
10. Afrin, S. *The influence of winter weather on high-crash days in Southern Ontario*. University of Waterloo, 2013.
11. Andrey, J., and R. Olley. The relationship between weather and road safety: past and future research directions. *Climatological Bulletin*, Vol. 24, No. 3, 1990, pp. 123–127.
12. Virginia Tech Transportation Institute. InSight SHRP2 NDS.
13. Smadi, O., N. Hawkins, Z. Hans, B. A. Bektas, S. Knickerbocker, I. Nlenanya, R. Souleyrette, and S. Hallmark. *Naturalistic Driving Study: Development of the Roadway Information Database, SHRP 2 Report S2-S04A-RW-1*. 2014.
14. WYDOT. Wyoming department of transportation connected vehicle pilot: Improving safety and travel reliability on I-80 in Wyoming.
15. Ahmed, M. M., A. Ghasemzadeh, H. Eldeeb, S. Gaweesh, J. Clapp, K. Ksaibati, and R. Young. *Driver Performance and Behavior in Adverse Weather Conditions: An Investigation Using the SHRP2 Naturalistic Driving Study Data—Phase 1. DOI: 10.13140/RG.2.2.24061.05602*. 2015.
16. Ghasemzadeh, A., B. Hammit, M. M. Ahmed, and H. Eldeeb. Complementary Methodologies to Identify Weather Conditions in Naturalistic Driving Study Trips: Lessons Learned from the SHRP2 Naturalistic Driving Study & Roadway Information Database. 97th Transportation Research Board Annual Meeting. Washington, D.C. 2018.
17. Ahmed, M. M., M. Abdel-Aty, J. Lee, and R. Yu. Real-time assessment of fog-related crashes using airport weather data: A feasibility analysis. *Accident Analysis and Prevention*,

Vol. 72, 2014, pp. 309–317.
18. Ibrahim, A. T., and F. L. Hall. Effect of adverse weather conditions on speed-flow-occupancy relationships. *Transportation Research Record*, No. 1457, 1994, pp. 184–191.
19. Stern, A. D., V. Shah, L. Goodwin, and P. Pisano. Analysis of weather impacts on traffic flow in metropolitan Washington DC. 2003.
20. Hawkins, R. K. Motorway traffic behaviour in reduced visibility conditions. 1988.
21. Agarwal, M., T. H. Maze, and R. Souleyrette. *Impact of Weather On Urban Freeway Traffic Flow Characteristics And Facility Capacity*. 2006.
22. Kyte, M., Z. Khatib, P. Shannon, and F. Kitchener. Effect of Weather on Free-Flow Speed. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1776, No. 1, 2001, pp. 60–68.
23. Pisano, P., and L. C. Goodwin. Surface transportation weather applications, Federal Highway Administration in concert with Mitretek Systems, presented at the 2002 Institute of Transportation Engineers Annual Meeting. 2002.
24. Agarwal, M., T. H. Maze, and R. Souleyrette. Impacts of weather on urban freeway traffic flow characteristics and facility capacity. 2005.
25. Highway Capacity Manual 2016. *Transportation Research Board, National Research Council, Washington, DC*, 2016.
26. Jiang, H., Y. Zou, S. Zhang, J. Tang, and Y. Wang. Short-term speed prediction using remote microwave sensor data: machine learning versus statistical model. *Mathematical Problems in Engineering*, Vol. 2016, 2016.
27. Wang, Y., P. Beullens, H. Liu, D. Brown, T. Thornton, and R. Proud. A practical intelligent navigation system based on travel speed prediction. 2008.
28. Mahajan, N. *Integrated Approach to Variable Speed Limits & Ramp Metering*. Delft University of Technology, 2014.
29. Weng, J., Q. Meng, and D. Z. W. Wang. Tree- Based Logistic Regression Approach for Work Zone Casualty Risk Assessment. *Risk analysis*, Vol. 33, No. 3, 2013, pp. 493–504.
30. Ahmed, M., and M. Abdel-Aty. A data fusion framework for real-time risk assessment on freeways. *Transportation Research Part C: Emerging Technologies*, Vol. 26, 2013, pp. 203–213.
31. Das, S., and X. Sun. Investigating the Pattern of Traffic Crashes under Rainy Weather by Association Rules in Data Mining 2. 2014.
32. Abdel-Aty, M., and H. Abdelwahab. Modeling rear-end collisions including the role of driver's visibility and light truck vehicles using a nested logit structure. *Accident Analysis & Prevention*, Vol. 36, No. 3, 2004, pp. 447–456.
33. Lord, D., and F. Mannering. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice*, Vol. 44, No. 5, 2010, pp. 291–305.
34. Agresti, A. *An Introduction to Categorical Data Analysis: Second Edition*. John Wiley & Sons, Inc., Publication, 2007.
35. Ghasemzadeh, A., and M. M. Ahmed. A Tree-Based Ordered Probit Approach to Identify Factors Affecting Work Zone Weather-Related Crashes Severity in North Carolina Using the Highway Safety Information System Dataset. 96th Transportation Research Board Annual Meeting. Washington, D.C. 2017.
36. Ghasemzadeh, A., and M. M. Ahmed. A Probit-Decision Tree Approach to Analyze Effects

of Adverse Weather Conditions on Work Zone Crash Severity Using Second Strategic Highway Research Program Roadway Information Dataset. 96th Transportation Research Board Annual Meeting. Washington, D.C. 2017.

37. Han, J., and M. Kamber. Data Mining: Concepts and Techniques, University of Illinois at Urbana-Champaign.
38. Kashani, A. T., and A. S. Mohaymany. Analysis of the traffic injury severity on two-lane, two-way rural roads based on classification tree models. *Safety Science,* Vol. 49, No. 10, 2011, pp. 1314–1320.
39. Weng, J., and Q. Meng. Decision tree-based model for estimation of work zone capacity. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2257, 2011, pp. 40–50.
40. Kovalerchuk, B., and E. Vityaev. *Data mining in finance: advances in relational and hybrid methods*. BOOK. Springer Science & Business Media, 2000.
41. Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen. *Classification and regression trees*. BOOK. CRC press, 1984.
42. Ghasemzadeh, A., B. Hammit, M. Ahmed, and R. Young. Using Parametric Ordinal Logistic Regression and Non-Parametric Decision Tree Approaches for Assessing the Impact of Weather Conditions on Driver Speed Selection Using Naturalistic Driving Data. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. In press, 2018.
43. Ahmed, M. M., R. K. Young, A. Ghasemzadeh, B. Hammit, A. Elhashemi, N. Khan, A. Das, and H. Eldeeb. *Implementation of SHRP2 results within the Wyoming connected vehicle variable speed limit system: Phase 2 early findings report and phase 3 proposal*. 2017.
44. Marquaridt, D. W. Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics*, Vol. 12, No. 3, 1970, pp. 591–612.
45. FHWA. How Do Weather Events Impact Roads? - FHWA Road Weather Management.
46. Brooks, J. O., M. C. Crisler, N. Klein, R. Goodenough, R. W. Beeco, C. Guirl, P. J. Tyler, A. Hilpert, Y. Miller, and J. Grygier. Speed choice and driving performance in simulated foggy conditions. *Accident Analysis & Prevention*, Vol. 43, No. 3, 2011, pp. 698–705.
47. Whiffen, B., P. Delannoy, and S. Siok. Fog: Impact on road transportation and mitigation options. *National Highway Visibility Conference*, 2004, pp. 1–12.
48. Wanvik, P. O. Effects of road lighting: an analysis based on Dutch accident statistics 1987–2006. *Accident Analysis & Prevention,* Vol. 41, No. 1, 2009, pp. 123–128.
49. Moore, R. L., and L. COOPER. FOG AND ROAD TRAFFIC. *Transport and Road Research Laboratory, Rept No TRRL-LR-446,* 1972.
50. Yan, X., X. Li, Y. Liu, and J. Zhao. Effects of foggy conditions on drivers' speed control behaviors at different risk levels. *Safety Science,* Vol. 68, 2014, pp. 275–287.
51. Owens, D. A., J. Wood, and T. Carberry. Effects of reduced contrast on the perception and control of speed when driving. *Perception*, Vol. 39, No. 9, 2010, pp. 1199–1215.
52. Hassan, H. M., M. A. Abdel-Aty, K. Choi, and S. A. Algadhi. Driver behavior and preferences for changeable message signs and variable speed limits in reduced visibility conditions. *Journal of Intelligent Transportation Systems*, Vol. 16, No. 3, 2012, pp. 132–146.
53. Snowden, R. J., N. Stimpson, and R. A. Ruddle. Speed perception fogs up as visibility

drops. *Nature*, Vol. 392, No. 6675, 1998, p. 450.
54. Owens, J. M., L. Angell, J. M. Hankey, J. Foley, and K. Ebe. Creation of the Naturalistic Engagement in Secondary Tasks (NEST) distracted driving dataset. *Journal of safety research*, Vol. 54, 2015, pp. 33-e29.
55. Mueller, A. S., and L. M. Trick. Driving in Fog: The Effects of Driving Experience and Visibility on Speed Compensation and Hazard Avoidance. *Accident Analysis and Prevention*, Vol. 48, 2012, pp. 472–479.
56. Pisano, P., and L. C. Goodwin. Best Practices For Road Weather Management, Version 2.0. 2003.
57. Perrin, J. Effects of variable speed limit signs on driver behavior during inclement weather. *Institute of Transportation Engineers*, 2000.
58. Jaeger, T. F. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language*, Vol. 59, No. 4, 2008, pp. 434–446.
59. Khan, N. M., A. Ghasemzadeh, and M. M. Ahmed. Investigating the Impact of Fog on Freeway Speed Selection Using the SHRP2 Naturalistic Driving Study Data. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. In press, 2018.
60. Ahmed, M. M., and A. Ghasemzadeh. Exploring the Impacts of Adverse Weather Conditions on Speed and Headway Behaviors Using the SHRP2 Naturalistic Driving Study Data. 96th Transportation Research Board Annual Meeting, 2017. 2017.
61. Jun, J. Understanding the variability of speed distributions under mixed traffic conditions caused by holiday traffic. *Transportation Research Part C: Emerging Technologies*, Vol. 18, No. 4, 2010, pp. 599–610.
62. Liang, W. L., M. Kyte, F. Kitchener, and P. Shannon. Effect of Enviromental Factors on Driver Speed-A Case study. *Transportation Research Record*, 1998, pp. 155–161.
63. Abdel-Aty, M. A., H. M. Hassan, M. Ahmed, and A. S. Al-Ghamdi. Real-time prediction of visibility related crashes. *Transportation research part C: emerging technologies*, Vol. 24, 2012, pp. 288–298.
64. Bagdadi, O. Assessing safety critical braking events in naturalistic driving studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 16, 2013, pp. 117–126.
65. Marquardt, D. W. Generalized Inverse, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation. *Technometrics*, Vol. 12, No. 1, 1970, pp. 591–612.
66. Ghasemzadeh, A., and M. M. Ahmed. Crash Characteristics and Injury Severity at Work Zones Considering Adverse Weather Conditions in Washington Using SHRP 2 Roadway Information Database. 95th Transportation Research Board Annual Meeting. Washington, D.C. 2016.
67. Tazul Islam, M., L. Thue, and J. Grekul. Understanding Traffic Safety Culture: Implications for Increasing Traffic Safety. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2635, 2017, pp. 79–89.
68. Campbell, K. L. The SHRP 2 naturalistic driving study: Addressing driver performance and behavior in traffic safety. *TR News*, No. 282, 2012.
69. Horrey, W. J., A. L. Alexander, and C. D. Wickens. Does workload modulate the effects of in-vehicle display location on concurrent driving and side task performance. 2003.

70. Krauß, S. Towards a unified view of microscopic traffic flow theories. No. 30, 1997, pp. 901–905.
71. Brackstone, M., and M. Mcdonald. Car-following: A historical review. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 2, No. 4, 1999, pp. 181–196.
72. Hoogendoorn, R. G., S. P. Hoogendoorn, K. A. Brookhuis, and W. Daamen. Longitudinal driving behavior under adverse conditions: A close look at psycho-spacing models. *Procedia - Social and Behavioral Sciences*, Vol. 20, 2011, pp. 536–546.
73. Rakha, H., D. Krechmer, G. Cordahi, I. Zohdy, S. Sadek, and M. Arafeh. *Microscopic analysis of traffic flow in inclement weather*. Blacksburg, VA, 2009.
74. Sangster, J., H. Rakha, and J. Du. Application of naturalistic driving data to modeling of driver car-following Behavior. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2390, 2013, pp. 20–33.
75. Soria, I. S., and L. Elefteriadou. Assessment of car-following models using field data. 2011.
76. Chong, L., and M. Abbas. Driver Car-following Behavior Simulation using Fuzzy Rule-based Neural Network. 2011.
77. Hammit, B., A. Ghasemzadeh, M. M. Ahmed, and R. K. Young. Evaluation of Weather-Related Freeway Car-Following Behavior using the SHRP2 Naturalistic Driving Study. 97th Transportation Research Board Annual Meeting. Washington, D.C. 2018.
78. Gipps, P. G. A behavioural car-following model for computer simulation. *Transportation Research Part B: Methodological*, Vol. 15, No. 2, 1981, pp. 105–111.
79. Ciuffo, B., V. Punzo, and M. Montanino. Thirty years of gipps' car-following model. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2315, 2012, pp. 89–99.
80. Wilson, R. E. An analysis of gipps' car-following model of highway traffic. *Institute of Mathematics and its Applications Journal of Applied Mathematics*, Vol. 66, 2001, pp. 509–537.
81. Kim, J., and H. S. Mahmassani. Correlated parameters in driving behavior models. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2249, 2011, pp. 62–77.
82. Herbel, S., L. Laing, and C. McGovern. *Highway Safety Improvement Program Manual*. BOOK. US Department of Transportation, Federal Highway Administration, Office of Safety, 2010.
83. Reed-Jones, J., L. M. Trick, and M. Matthews. Testing assumptions implicit in the use of the 15-second rule as an early predictor of whether an in-vehicle device produces unacceptable levels of distraction. *Accident Analysis & Prevention*, Vol. 40, No. 2, 2008, pp. 628–634.
84. Harbluk, J. L., Y. I. Noy, and M. Eizenman. *The impact of cognitive distraction on driver visual behaviour and vehicle control*. 2002.
85. Chisholm, S. L., J. K. Caird, and J. Lockhart. The effects of practice with MP3 players on driving performance. *Accident Analysis & Prevention*, Vol. 40, No. 2, 2008, pp. 704–713.
86. Collet, C., A. Guillot, and C. Petit. Phoning while driving I: a review of epidemiological, psychological, behavioural and physiological studies. *Ergonomics*, Vol. 53, No. 5, 2010, pp. 589–601.
87. Shankar, V., F. Mannering, and W. Barfield. Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. *Accident Analysis &*

Prevention, Vol. 27, No. 3, 1995, pp. 371–389.
88. Koetse, M. J., and P. Rietveld. The impact of climate change and weather on transport: An overview of empirical findings. *Transportation Research Part D: Transport and Environment*, Vol. 14, No. 3, 2009, pp. 205–221.
89. Brackstone, M., B. Waterson, and M. McDonald. Determinants of following headway in congested traffic. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 12, No. 2, 2009, pp. 131–142.
90. Van Winsum, W. The human element in car following models. *Transportation research part F: traffic psychology and behaviour*, Vol. 2, No. 4, 1999, pp. 207–211.
91. National Transportation Safety Board. Fatal Highway Accidents on Wet Pavement—The Magnitude, Location, and Characteristics. *National Technical Information Service, Springfield*, Vol. HTSB-HSS-8, 1980.
92. Andreescu, M. P., and D. B. Frost. Weather and traffic accidents in Montreal, Canada. *Climate Research*, Vol. 9, No. 3, 1998, pp. 225–230.
93. Allen, R. W., Z. Parseghian, and A. C. Stein. A driving simulator study of the performance effects of low blood alcohol concentration. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 40, No. 18, 1996, pp. 943–946.
94. Allen, R. W., and Z. Parseghian. A Driving Simulator Study of the Performance Effects of Low Blood Alchohol Concentration. *Proceedings of The Human Factors and Ergonomics Society 40th Annual Meeting*, 1996, pp. 943–946.
95. Young, K., M. Regan, and M. Hammer. Driver distraction: A review of the literature. *Distracted driving*, 2007, pp. 379–405.
96. Johnson, T., R. Chen, R. Sherony, and H. C. Gabler. Characterization of the lateral control performance by human drivers. *SAE International Journal of Transportation Safety*, Vol. 4, No. 2, 2016, pp. 2016-01–1449.
97. Yu, R., M. A. Abdel-Aty, M. M. Ahmed, and X. Wang. Utilizing microscopic traffic and weather data to analyze real-time crash patterns in the context of active traffic management. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 15, No. 1, 2014, pp. 205–213.
98. Ghosh, D., S. L. Jamson, P. D. Baxter, and M. W. Elliott. Continuous measures of driving performance on an advanced office-based driving simulator can be used to predict simulator task failure in patients with obstructive sleep apnoea syndrome. *Thorax*, 2012, p. thoraxjnl-2011.
99. Anttila, V., and J. Luoma. Surrogate in-vehicle information systems and driver behaviour in an urban environment: A field study on the effects of visual and cognitive load. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 8, No. 2, 2005, pp. 121–133.
100. Horrey, W. J., M. F. Lesch, and A. Garabet. Assessing the awareness of performance decrements in distracted drivers. *Accident Analysis & Prevention*, Vol. 40, No. 2, 2008, pp. 675–682.
101. Peng, Y., L. N. Boyle, and S. L. Hallmark. Driver's lane keeping ability with eyes off road: Insights from a naturalistic study. *Accident Analysis & Prevention*, Vol. 50, 2013, pp. 628–634.
102. Ghasemzadeh, A., and M. M. Ahmed. Drivers' Lane-Keeping Ability in Heavy Rain: Preliminary Investigation Using SHRP 2 Naturalistic Driving Study Data. *Transportation*

Research Record: Journal of the Transportation Research Board, No. 2663, 2017, pp. 99–108.

103. Chen, Ping-Ling, Rong-Chang Jou, Wafaa Saleh, and C. P. Accidents involving pedestrians with their backs to traffic or facing traffic: an evaluation of crash characteristics and injuries. *Journal of Advanced Transportation*, 2016.
104. Brown, D. C. Models for Ordered and Unordered Categorical Variables. https://www.utexas.edu/cola/prc/_files/cs/Multinomial_Ordinal_Models.pdf.
105. Cohen, J. Statistical power analysis for the behavioral sciences (revised ed.).
106. Bagdadi, O. Assessing safety critical braking events in naturalistic driving studies. *Transportation research part F: traffic psychology and behaviour*, Vol. 16, 2013, pp. 117–126.
107. Chovan, J. D., L. Tijerina, G. Alexander, and D. L. Hendricks. Examination of Lane Change Crashes and Potential IVHS Countermeasures. *Final Report, National Highway Traffic Safety Administration, U.S. Department of Transportation*, No. March, 1994, p. 60.
108. Hallmark, S., S. Tyner, and D. Mcgehee. Development of a Conceptual Model of Curve Driving for Isolated Rural Two Lane Curves Using SHRP 2 Naturalistic Driving Data. No. 319, 2015, pp. 7–8.
109. Hamdar, S. H., L. Qin, and A. Talebpour. Weather and road geometry impact on longitudinal driving behavior: Exploratory analysis using an empirically supported acceleration modeling framework. *Transportation Research Part C: Emerging Technologies*, Vol. 67, 2016, pp. 193–213.
110. Maze, T., M. Agarwai, and G. Burchett. Whether weather matters to traffic demand, traffic safety, and traffic operations and flow. *Transportation research record: Journal of the transportation research board*, No. 1948, 2006, pp. 170–176.
111. Bifulco, G. N., F. Galante, L. Pariota, M. R. Spena, and P. Del Gais. Data Collection for Traffic and Drivers' Behaviour Studies: a large-scale survey. *Procedia-Social and Behavioral Sciences*, Vol. 111, 2014, pp. 721–730.
112. Goodall, N. J., B. L. Smith, and B. "Brian" Park. Microscopic estimation of freeway vehicle positions from the behavior of connected vehicles. *Journal of Intelligent Transportation Systems*, Vol. 20, No. 1, 2016, pp. 45–54.
113. Birrell, S. A., M. Fowkes, and P. A. Jennings. Effect of using an in-vehicle smart driving aid on real-world driver performance. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 15, No. 4, 2014, pp. 1801–1810.
114. Zhou, J., H. Peng, and T. J. Gordon. Characterization of the lateral control performance by human drivers on highways. *SAE International Journal of Passenger Cars-Mechanical Systems*, Vol. 1, No. 2008-01–0561, 2008, pp. 450–458.
115. Azizi, L., and A. Sheikholeslami. Safety Effect of U-Turn Conversions in Tehran: Empirical Bayes Observational Before-and-After Study and Crash Prediction Models. *Journal of Transportation Engineering*, Vol. 139, No. 1, 2012, pp. 101–108.
116. Ghasemzadeh, A., and M. Ahmed. Multivariate Adaptive Regression Splines and Logistic Regression Models to Identify the Impact of Rainy Weather on Driver Lane-keeping Performance Considering Driver Demographics and Roadway Characteristics Using SHRP2 Naturalistic Driving Data. 2018.
117. Friedman, J. H. Multivariate adaptive regression splines. *The annals of statistics*, 1991, pp. 1–67.

118. Abraham, A., D. Steinberg, and N. S. Philip. Rainfall forecasting using soft computing models and multivariate adaptive regression splines. *IEEE SMC Transactions, Special Issue on Fusion of Soft Computing and Hard Computing in Industrial Applications*, Vol. 1, 2001, pp. 1–6.
119. Haleem, K., A. Gan, and J. Lu. Using multivariate adaptive regression splines (MARS) to develop crash modification factors for urban freeway interchange influence areas. *Accident Analysis & Prevention*, Vol. 55, 2013, pp. 12–21.
120. Park, J., and M. Abdel-Aty. Assessing the safety effects of multiple roadside treatments using parametric and nonparametric approaches. *Accident Analysis & Prevention*, Vol. 83, 2015, pp. 203–213.
121. Osei-Bryson, K.-M., and O. Ngwenyama. *Advances in research methods for information systems research*. BOOK. Springer, 2014.
122. Agresti, a. *An Introduction to Categorical Data Analysis Second Edition*. 2007.
123. Ghasemzadeh, A., and M. M. Ahmed. Utilizing naturalistic driving data for in-depth analysis of driver lane-keeping behavior in rain: Non-parametric MARS and parametric logistic regression modeling approaches. *Transportation Research Part C*, Vol. In press, 2018.
124. Neale, V. L., S. G. Klauer, R. R. Knipling, T. a. Dingus, G. T. Holbrook, and a. Petersen. The 100 Car Naturalistic Driving Study Phase I – Experimental Design. *US DOT, National Highway Traffic Safety Administration*, No. December, 2002.
125. Dingus, T. a., S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. a. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. R. Doerzaph, J. Jermeland, and R. R. Knipling. The 100-Car naturalistic driving study phase II – Results of the 100-Car field experiment. *Dot Hs 810 593*, No. April, 2006, p. No. HS-810 593.
126. Dingus, T. A., J. M. National Research Council (U.S.). Transportation Research Board., J. F. Second Strategic Highway Research Program (U.S.), S. E. Lee, L. Eichelberger, K. E. Stulce, D. McGraw, M. Perez, and L. Stowe. *Naturalistic driving study : technical coordination and quality control*. 2015.
127. ROBERTSON, D. I., M. A. WINNETT, and R. T. Herrod. ACCELERATION SIGNATURES. *Traffic Engineering and Control*, Vol. 33, No. 9, 1992.
128. Speed choice versus celeration behavior as traffic accident predictor. *Journal of Safety Research*, Vol. 37, No. 1, 2006, pp. 43–51.
129. Jerky driving—An indicator of accident proneness? *Accident Analysis & Prevention*, Vol. 43, No. 4, 2011, pp. 1359–1363.
130. Hankey, J. M., M. A. Perez, and J. A. McClafferty. *Description of the SHRP 2 Naturalistic Database and the Crash, Near-Crash, and Baseline Data Sets*. Virginia Tech Transportation Institute, 2016.
131. Performance of basic kinematic thresholds in the identification of crash and near-crash events within naturalistic driving data. *Accident Analysis & Prevention*, Vol. 103, 2017, pp. 10–19.
132. Naturalistic assessment of novice teenage crash experience. *Accident Analysis & Prevention*, Vol. 43, No. 4, 2011, pp. 1472–1479.
133. Malta, L., M. Ljung Aust, F. Faber, B. Metz, G. Saint Pierre, M. Benmimoun, and R. Schäfer. *EuroFOT final results: impacts on traffic safety*. Germany, 2012.
134. Fitch, G. M., S. A. Soccolich, F. Guo, J. McClafferty, Y. Fang, R. L. Olson, M. A. Perez,

R. J. Hanowski, J. M. Hankey, and T. A. Dingus. The Impact of Hand-Held and Hands-Free Cell Phone Use on Driving Performance and Safety-Critical Event Risk. 2013.

135. Poch, M., and F. Mannering. Negative Binomial Analysis of Intersection-Accident Frequencies. *Journal of Transportation Engineering*, Vol. 122, No. 2, 1996, pp. 105–113.

136. Love, P. E. D., and P. Teo. Statistical Analysis of Injury and Nonconformance Frequencies in Construction: Negative Binomial Regression Model. *Journal of Construction Engineering and Management*, Vol. 143, No. 8, 2017, p. 5017011.

137. Imprialou, M.-I. M., M. Quddus, D. E. Pitfield, and D. Lord. Re-visiting crash–speed relationships: A new perspective in crash modelling. *Accident Analysis & Prevention*, Vol. 86, 2016, pp. 173–185.

138. Wang, K., J. N. Ivan, N. Ravishanker, and E. Jackson. Multivariate poisson lognormal modeling of crashes by type and severity on rural two lane highways. *Accident Analysis & Prevention*, Vol. 99, 2017, pp. 6–19.

139. de Oña, J., G. López, R. Mujalli, and F. J. Calvo. Analysis of traffic accidents on rural highways using Latent Class Clustering and Bayesian Networks. *Accident Analysis & Prevention*, Vol. 51, 2013, pp. 1–10.

140. Kumar, S., D. Toshniwal, and M. Parida. A comparative analysis of heterogeneity in road accident data using data mining techniques. *Evolving Systems*, Vol. 8, No. 2, 2017, pp. 147–155.

141. Kumar, S., and D. Toshniwal. A data mining approach to characterize road accident locations. *Journal of Modern Transportation*, Vol. 24, No. 1, 2016, pp. 62–72.

142. Kumar, S., and D. Toshniwal. A novel framework to analyze road accident time series data. *Journal of Big Data*, Vol. 3, No. 1, 2016, p. 8.