

**A framework for reliability-based weather-responsive speed management
for rural highways located in extremely cold regions**

by

NAVODA YASANTHI RILLAGODAGE

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Civil Engineering

University of Manitoba

Winnipeg

Abstract

Weather-responsive variable speed limit (WRVSL) systems attempt to regulate drivers' speed choice in adverse road-weather conditions (RWCs) by executing a set of differentiated speed limits implemented under different RWCs. This thesis attempts to propose an effective approach to set WRVSLs in cold region rural highways based on the reliability of a WRVSL — the probability of a WRVSL being safe and complied by drivers.

This thesis addresses four research questions: *(i)* What is the impact of different RWCs on drivers' speed choice? *(ii)* What is the most appropriate approach to evaluate drivers' speed behaviour in adverse RWCs? *(iii)* Are there any specific driving conditions which intensify safety risks when indicated by speed and speed variability? *(iv)* What is an effective approach to regulate speed in adverse RWCs to mitigate safety risks induced by such RWCs? The first research question is addressed by proposing an approach to model speed distributions for different combinations of traffic and road-weather conditions using the central limit theorem (CLT). The second research question is addressed by comparing the CLT-based speed distribution modelling approach proposed in this thesis with a regression modelling based approach. To address the third research question, a holistic crash indicative measure is proposed based on both within- and across-lane crash risks (estimated based on speed variability) for different driving conditions such as prevailing RWCs, vehicle type, travel lane, and truck payload condition. The fourth research question is addressed by proposing an effective approach to set WRVSLs based on the reliability theory.

Overall, this thesis demonstrates that extreme RWCs such as snow, and icy pavements notably affect drivers' speed choice, resulting increased speed variability, thus intensified crash risks in such RWCs. The results of this research also revealed that tractor-trailer combinations are highly vulnerable when travelling in extreme RWCs. Practical applications of the proposed research methods include *(i)* accurately understanding driver's speed choice patterns in different RWCs, *(ii)* developing RWC-specific speed distributions that could be used as inputs in microsimulation studies, *(iii)* identifying RWCs with intensified crash potential based on speed and speed variability, and *(iv)* effectively regulating drivers' speed choice in extreme RWCs.

Acknowledgement

This thesis is the ultimate product of an unforgettable four-year-long journey, which would have not been possible without a special set of individuals. First and foremost, I'm deeply indebted to my academic advisor Dr. Babak Mehran for everything — for providing me the opportunity to pursue higher studies, for making me feel welcomed in Canada, for his excellent academic directives, and most importantly, his consistent patience that cannot be underestimated. I could not have imagined having a better advisor, a mentor and a role model to guide me on climbing this mountain of Ph.D.

I gratefully acknowledge the help from the members of my academic advisory committee, Dr. Jeannette Montufar and Dr. Ahmed Ashraf, for their invaluable constructive comments and enduring support throughout my Ph.D. Many thanks go to Dr. Wael K.M. Alhajyaseen for his insightful directions on research article preparation. I would also like to express my heartfelt gratitude to Dr. Jonathan Regehr, for supporting throughout my Ph.D. journey in many ways.

I cannot leave the University of Manitoba without acknowledging my fellow research mates with whom I have shared both laughter, and tears. All our memories from the good times we spent at the university will be forever cherished.

Many thanks go to Alberta Transportation for providing the data used in this research.

Lastly, to my loving family. Words can't express how grateful I am to my parents, sister, and brother-in-law, for everything, starting from motivating me to enroll in grad school to graduating with a Ph.D. To my beloved husband, Dhammika — thank you for your countless sacrifices, and being an incredibly supportive partner!

Dedication

To my husband — *Dhammika*

For your love, patience, and never-failing encouragement.

Table of Contents

Abstract.....	i
Acknowledgement.....	ii
Dedication	iii
Table of Contents	iv
List of Tables	viii
List of Figures.....	ix
List of Copyrighted Materials.....	xii
Contribution of Authors.....	xiii
List of Abbreviations	xv
Chapter 1. Introduction	1
1.1 Background	1
1.1.1 State of the practice WRSM strategies	1
1.1.2 Importance of developing and implementing effective WRSM strategies	1
1.1.3 Conceptual framework for developing WRVSLs.....	4
1.2 Problem statement, research needs and research questions	5
1.3 Goal, objectives, definitions, and scope	7
1.4 Thesis organization	8
Chapter 2. Modelling speed behaviour in rural highways: Safety analysis of driving under adverse road-weather conditions.....	10
2.1 Abstract	10
2.2 Introduction.....	11
2.3 Research objectives.....	14
2.4 Study data.....	15

2.4.1 Data collection	15
2.4.2. Data preparation.....	17
2.4.3 Preliminary analysis.....	21
2.5 Population, Sample and Sampling Distributions	24
2.6 Methodology.....	28
2.6.1 Mean of the of desired speeds’ distribution (μ) of individual vehicles’ population....	29
2.6.2 Identification of road-weather conditions with intensified safety risks.....	36
2.7 Modeling results.....	38
2.7.1 Impact of road-weather and traffic conditions on estimated desired speed distributions	40
2.7.2 Adverse road-weather conditions with intensified road safety risks	44
2.7.3 Comparative analysis	47
2.8 Conclusion and future directions	50
Chapter 3. Application of Different Data Analytics for Evaluation of Heavy Vehicle Vulnerability in Cold-Region Rural Highways.....	52
3.1 Abstract.....	52
3.2 Introduction.....	53
3.3 Background and literature review	56
3.4 Study data.....	59
3.4.1 Data collection	59
3.4.2 Data preparation.....	61
3.5 Methodology.....	64
3.5.1 Modelling speed distributions.....	64
3.5.2 Identifying Road-Weather-Specific Collision Risk Levels for Different Vehicle Types	68

3.6 Modelling results	70
3.6.1 Speed distribution models.....	70
3.6.2 Road-weather Conditions with Increased Collision Potential	75
3.7 Conclusion & future directions.....	79
Chapter 4. A Reliability-based Weather-Responsive Variable Speed Limit System to Improve the Safety of Rural Highways.....	82
4.1 Abstract.....	82
4.2 Introduction.....	83
4.3 Literature review	84
4.4 Study data.....	88
4.4.1 Data collection	88
4.4.2 Preliminary data analysis	89
4.5 Methodology	92
4.5.1 Assumptions.....	92
4.5.2 Modelling methodology.....	93
4.6 Implementation: Case study.....	107
4.7 Results and discussion	109
4.7.1 Drivers’ response to the existing fixed SL.....	109
4.7.2 Reliability-based Weather-Responsive Variable Speed Limit Estimations.....	110
4.8 Conclusion and future directions	112
Chapter 5. Conclusions.....	115
5.1 Summary of key findings.....	115
5.1.1 Impacts of adverse RWCs on drivers’ speed choice.....	115
5.1.2 Variability of modelling speed distributions based on different data analytics.....	116

5.1.3 Driving conditions with intensified crash risks	117
5.1.4 Importance of setting WRVSLs in cold region rural highways.....	117
5.2 Contributions.....	117
5.3 Limitations and future research	118
References	123
Appendix.....	133
A: Kolmogorov-Smirnoff (K-S) test results	133
B: Estimating the mean (μ) and the set of weight factors representing the contribution of each five-minute aggregate speed distribution in the desired speed distribution.....	137
B.1 Estimating the mean of the desired speed distribution (μ).....	137
B.2 Estimating the set of weight factors representing the contribution of each five-minute aggregate speed distribution in estimating the desired speed distribution.....	139

List of Tables

Table 2.1: Road-weather condition categories.....	18
Table 2.2: Influencing factors' statistical information.....	19
Table 2.3: Classification criteria for combinations of road-weather conditions to identify road-weather conditions with intensified safety risks	38
Table 2.4: Comparison of results with past literature	48
Table 3.1: Study Data Categories	62
Table 3.2: Details of the Three Cluster Centroids	76
Table 4.1: Summary of literature review	87
Table 4.2: Terminological mapping between reliability theory and WRSM.....	103
Table 5.1: Overall summary of the thesis	119

List of Figures

Figure 1.1: The three types of WRSM strategies.....	2
Figure 1.2: Relationship between crash involvement rate and travel speed variation.....	3
Figure 1.3: A stepwise conceptual framework for developing WRVSLs	4
Figure 2.1: Study site location and the positions of the RWIS & the WIM station.....	16
Figure 2.2: Speed-flow variation in the eastbound shoulder and median lanes at the study site..	21
Figure 2.3: Distribution of speeds in the study site	23
Figure 2.4: Hourly variation of the Heavy vehicles' percentage and the lane utilization factor at the study site	24
Figure 2.5: Categories of road-weather and traffic conditions to identify populations	25
Figure 2.6: Illustration of a population and samples.....	26
Figure 2.7: Sampling distribution of mean speeds for a population represented by eight sampling distributions (Population ID: D-NP-N-T:GIII-TF:GI-HV:GI)	27
Figure 2.8: Illustration of the chaining process of k sampling distributions of five-minute aggregate speeds.....	33
Figure 2.9: Algorithm for estimation of weight factors and μ	33
Figure 2.10: Weight factor VS the variance of the distributions of five-minute aggregate speeds	39
Figure 2.11: Number of populations recorded in terms of the traffic flow and HV categories....	40
Figure 2.12: Desired speed distribution characteristics; mean (μ) and standard deviation (σ) of road-weather combinations under the prevailing traffic conditions	41
Figure 2.13: Crash Severity and Crash Exposure Factors for each road-weather combination under the prevailing traffic conditions	45
Figure 2.14: Classification of each road-weather combination under the prevailing traffic conditions.....	46
Figure 3.1: Summary of related literature, and research gaps	59

Figure 3.2: Study site	60
Figure 3.3: Descriptive statistics of speed data.....	63
Figure 3.4: Algorithm for estimating parameters of within- and across-lane speed differentials	69
Figure 3.5: Algorithm for assigning collision risk levels	70
Figure 3.6: Model coefficients for Models 1, and 2	71
Figure 3.7: Speed distribution parameters	73
Figure 3.8: Vehicle-specific collision risks estimated for different road-weather conditions.....	77
Figure 4.1: Study site	88
Figure 4.2: Study site traffic conditions.....	91
Figure 4.3: Descriptive statistics of speeds at the study site.....	91
Figure 4.4: Proposed framework of the developed WRVSL system.....	94
Figure 4.5: A two-vehicle setting with a hypothetical disturbance introduced to the preceding vehicle.....	98
Figure 4.6: Modelling data generation process for a given combination of RWCs and C_r	101
Figure 4.7: Defining success and failure.....	104
Figure 4.8: Desired speed distribution characteristics (μ and σ) and the 85 th percentile speeds under different combinations of pavement and precipitation conditions.....	109
Figure 4.9: Reliability curves and the proposed reliability-based WRVSLs under different combinations of pavement and precipitation conditions	111
Figure A. 1: K-S test results for temperature category levels.....	133
Figure A. 2: K-S test results for time of the day categories.....	134
Figure A. 3: K-S test results for precipitation condition levels	134
Figure A. 4: K-S test results for pavement surface condition levels	134
Figure A. 5: K-S test results for traffic flow categories.....	135
Figure A. 6: K-S test results for heavy vehicle (HV) percentage categories.....	135

Figure A. 7: K-S test results for travel lane categories	136
Figure A. 8: K-S test results for vehicle type categories	136

List of Copyrighted Materials

Maps included in this thesis (Figures 2.1, 3.2, and 4.1) were produced using two types of spatial data:

1. Boundary shapefiles (polygon type shapefiles) used to create Alberta's provincial boundary (Statistics Canada, 2016a), and census divisions (Statistics Canada, 2016b) downloaded from the Statistics Canada website:

<https://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2016-eng.cfm>

2. Road network files extracted from the National Road Network GeoBase Series (Government of Canada, 2021) available for download at:

<https://open.canada.ca/data/en/dataset/3d282116-e556-400c-9306-ca1a3cada77f>.

The road network files used in this thesis contain information licensed under the Open Government Licence – Canada (<https://open.canada.ca/en/open-government-licence-canada>).

Contribution of Authors

This thesis is prepared as a grouped-manuscript thesis. The material presented in chapters two through four of this thesis are reproduced from published journal articles. Details of the articles presented in each chapter are as follows:

1. Chapter 2 reprints the material published in the article titled “Modelling speed behaviour in rural highways: Safety analysis of driving under adverse road-weather conditions” which was published in the journal *PLOS ONE* in 2021 (Yasanthi et al., 2021). The authors and their contributions to this paper are as follows:
 - Rillagoda G.N. Yasanthi (the thesis author): Formal analysis, Methodology, Visualization, Writing – original draft
 - Babak Mehran: Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing
 - Wael K. M. Alhajyaseen: Conceptualization, Writing – review & editing
2. Chapter 3 reprints the material published in the article titled “Application of Different Data Analytics for Evaluation of Heavy Vehicle Vulnerability in Cold-Region Rural Highways” which was published in the *Focus Issue on Freight Transportation* in the *Journal of the Transportation Research Record* in 2022 (Yasanthi & Mehran, 2022). The authors and their contributions to this paper are as follows:
 - Rillagoda G.N. Yasanthi (the thesis author): Research conception and design, Analysis and interpretation of results, Draft manuscript preparation
 - Babak Mehran: Research conception and design, Data collection, Analysis and interpretation of results, Manuscript reviewing and editing
3. Chapter 4 reprints the material published in the article titled “A Reliability-based Weather-Responsive Variable Speed Limit System to Improve the Safety of Rural Highways” which was published in the journal *Accident Analysis and Prevention* in 2022 (Yasanthi et al., 2022). The authors and their contributions to this paper are as follows:
 - Rillagoda G.N. Yasanthi: Formal analysis, Methodology, Visualization, Interpretation of results, Writing – original draft
 - Babak Mehran: Conceptualization, Funding acquisition, Data collection, Methodology, Supervision, Interpretation of results, Writing – review & editing

- Wael K.M. Alhajyaseen: Conceptualization, Writing – review & editing

In summary, publications resulting from the research conducted in each chapter are as follows:

- Chapter 2: Yasanthi, R.G., Mehran, B. and Alhajyaseen, W.K., 2021. Modelling speed behaviour in rural highways: Safety analysis of driving under adverse road-weather conditions. *PLoS one*, 16(8), p.e0256322.
- Chapter 3: Yasanthi, R.G. and Mehran, B., 2022. Application of Different Data Analytics for Evaluation of Heavy Vehicle Vulnerability in Cold-Region Rural Highways. *Transportation Research Record*, p.03611981221111353.
- Chapter 4: Yasanthi, R.G., Mehran, B. and Alhajyaseen, W.K., 2022. A reliability-based weather-responsive variable speed limit system to improve the safety of rural highways. *Accident Analysis & Prevention*, 177, p.106831.

List of Abbreviations

a.k.a.	Also Known As
AADT	Annual Average Daily Traffic
ASOS	Automated Surface Observing System
BRT	Brake Reaction Time
CBA	Central Limit Theorem-based Approach
CEF	Crash Exposure Factor
CLT	Central Limit Theorem
CMF	Crash Modification Factor
CRB	Cramér–Rao Bound
CSF	Crash Severity Factor
E-MTT	Empty Multi-Trailer Truck
E-STT	Empty Single-Trailer Truck
FHWA	Federal Highway Administration
GMM	Gaussian Mixture Model
GVW	Gross Vehicle Weight
HCM	Highway Capacity Manual
HV	Heavy Vehicle
K-S	Kolmogorov–Smirnov Test
L-MTT	Loaded Multi-Trailer Truck
LOS	Level of Service
L-STT	Loaded Single-Trailer Truck
MTT	Multi-Trailer Trucks
MUTCD	Manual on Uniform Traffic Control Devices for Streets and Highways
MVUE	Minimum Variance Unbiased Estimation
PC	Passenger Car
PV	Passenger Vehicle
RBA	Regression modelling-based approach
RS	Recommended Speed

RW	Road-Weather
RWC	Road-Weather Condition(s)
RWIS	Road-Weather Information System
SD	Standard Deviation
SL	Speed Limit
SSA	Safe System Approach
SSD	Stopping Sight Distance
STT	Single-Trailer Trucks
SUT	Single-Unit Truck
TRB	Transportation Research Board
TTC	Tractor Trailer Combination
VSL	Variable Speed Limit
WD	Weibull distribution
WIM	Weigh-In-Motion
WRSM	Weather Responsive Speed Management
WRVSL	Weather Responsive Variable Speed Limit

Chapter 1. Introduction

Weather-responsive speed management (WRSM) strategies attempt to improve road safety in adverse road-weather conditions (RWCs) mainly through the application of weather-responsive variable speed limits (WRVSLs) — differentiated speed limits implemented under different RWCs. Such WRSM techniques are particularly important for improving road safety in cold region rural highways, where (i) speeding is prevalent due to frequent uncongested traffic conditions (Yasanthi & Mehran, 2020), and (ii) driving is often challenging due to natural hazards such as recurrent extreme RWCs. This thesis attempts to improve road safety in cold region rural highways by proposing an effective approach for WRSM in such highways.

1.1 Background

Developing an effective WRSM strategy to improve road safety in cold region rural highways starts with (i) recognizing the state of the practice WRSM strategies, (ii) acknowledging the need for WRSM techniques, and (iii) identifying the steps involved in developing an effective WRSM strategy.

1.1.1 State of the practice WRSM strategies

WRSM is a branch of Weather-responsive traffic management (WRTM) — a road safety countermeasure focusing on the development of coherent strategies to manage transportation operations during different RWCs (Pisano & Goodwin, 2004). According to Pisano & Goodwin (2004), the state-of-the-art WRSM techniques are threefold (Figure 1.1): (i) advisory, (ii) treatment, and (iii) control. While both advisory, and treatment WRSM strategies attempt to mitigate the impacts of adverse RWCs on drivers' speed choice, control strategies are perceived to be more effective due to the enforcement nature of control WRSM techniques. In fact, among the widespread types of such control WRSM strategies, many countries in Europe, North America, and the Australasian region have successfully adopted WRVSL systems to regulate speed in extreme RWCs (Robinson, 2000). Therefore, this thesis focuses on developing WRVSL systems.

1.1.2 Importance of developing and implementing effective WRSM strategies

Speeding — driving above the speed limit (SL) or driving too fast for the prevailing driving conditions (Forbes et al., 2012) — is hazardous, particularly in adverse RWCs such as rain, icy

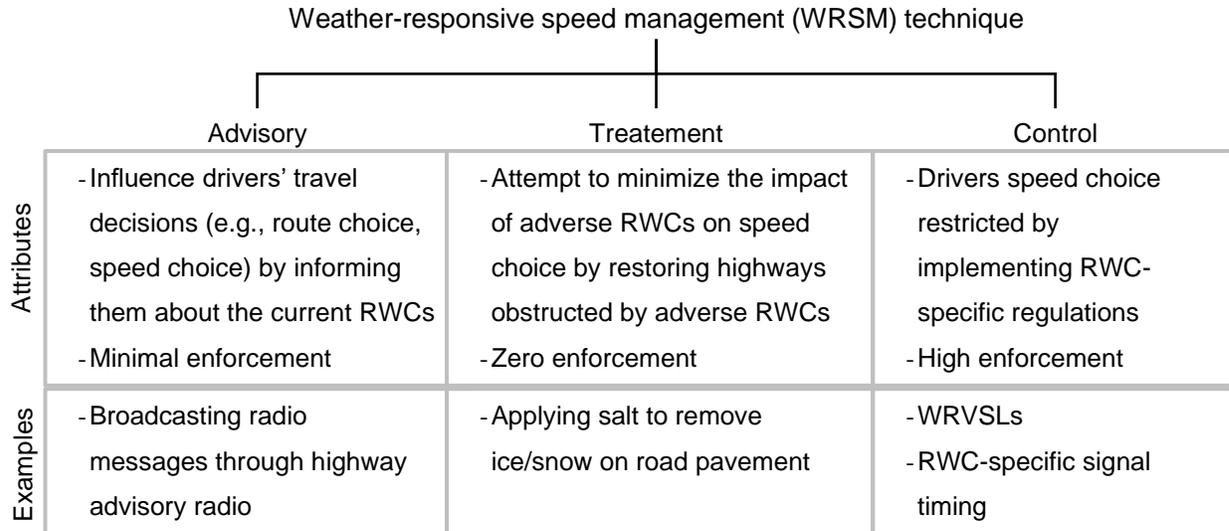


Figure 1.1: The three types of WRSM strategies (Pisano & Goodwin, 2004)

pavements, poor visibility, and strong wind. In fact, 22.3% of fatal crashes reported in 2019 in Alberta, Canada, were attributed to travelling too fast for the prevailing driving conditions (Alberta Transportation, 2022). On one hand, square of speed is directly proportional to the stopping sight distance (SSD); thus, high speeds are associated with longer SSDs (AASHTO, 2018). On the other hand, adverse RWCs deteriorate driving conditions. For instance, formation of ice and/or frost on road pavements substantially reduce pavement friction coefficients (Ivey et al., 1975). Pavement surface friction coefficient, however, is inversely proportional to SSD; thus, adverse RWCs also yield longer SSDs (Kordani et al., 2018). Therefore, speeding in adverse RWCs may intensify propensity for weather-related crashes — motor vehicle collisions attributed to at least one adverse RWC (Pisano et al., 2008).

Speeding, irrespective of the prevailing RWCs, reflects drivers' speed behaviour — choice of speed from several possible speeds (Letirand & Delhomme, 2005). Speed behaviour is often perceived as idiosyncratic, i.e., subjective to an individual. According to Hauer (1971), drivers can be categorized into two types: (i) slow drivers, and (ii) fast drivers, depending on their speed choice patterns. Slow drivers typically choose low speeds (Hauer, 1971) because they are reluctant to be exposed to risks associated with high speeds (e.g., increased crash potential, severe injuries). In contrast, fast drivers trade-off safety for lesser travel times, thus selecting higher travel speeds (Hauer, 1971). In fact, past research (Edwards, 1999) suggests that fast drivers may continue to select high speeds even in extreme RWCs although such drivers acknowledge the increased crash

potential in adverse RWCs. Slow drivers, however, may further reduce speed while travelling in adverse RWCs. Such divergent speed choice patterns present in inclement RWCs result in increased speed variability, and thus increased crash potential (Solomon, 1964). In fact, the Solomon curve (Figure 1.2) — a graphical representation of the relation between speed variation and crash involvement — developed by Solomon (1964) demonstrates how high speed variability contributes to increased crash risk, particularly during nighttime when visibility is poor. The Solomon curve is critiqued by past literature (Research Triangle Institute, 1970; Hauer, 2009) predominantly due to the heavy usage of crashes involving left-turning vehicles in the left half of the curve (the slow-moving vehicle crashes). Yet, speed variability may lead to increased crash potential. For instance, there is a possibility of collision when the succeeding vehicle of a two vehicle platoon drives faster than the preceding vehicle. Therefore, speed variability is often used as a surrogate measure of crash propensity (Cirillo, 2003).

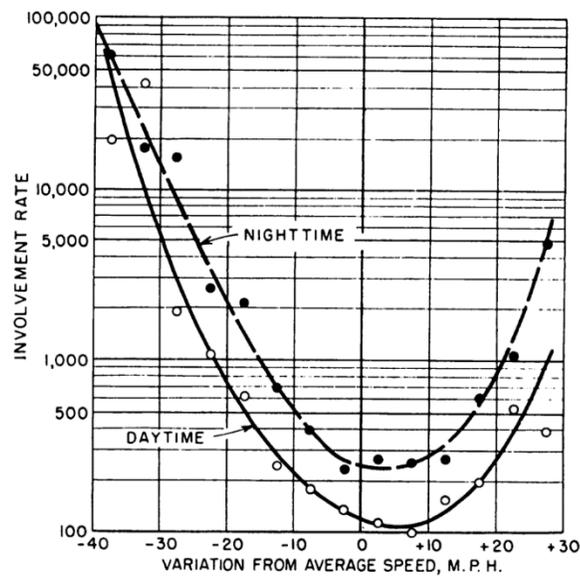


Figure 1.2: Relationship between crash involvement rate and travel speed variation (Solomon, 1964)

Note: Involvement rate represents the number of vehicles involved in crashes per 100 million vehicle-miles travelled.

In fact, while most drivers acknowledge the risks associated with driving in inclement RWCs (e.g., hydroplaning, poor visibility), some drivers may (i) not be able to apprehend the safest speed to travel in adverse RWCs, and/or (ii) disregard the risks associated with such RWCs, (Edwards, 1999). In such conditions, WRVSLs can be used by transport authorities as a medium to effectively communicate the safest RWC-specific travel speeds to drivers. On the other hand, adverse RWCs

may cause substantial speed variabilities, thus alarming crash risks (Figure 1.2). Yet, the intensities of such crash risks are dependent on the magnitude of speed variability present in a specific combination of RWCs (e.g., no precipitation in synchrony with wet pavements). Therefore, it is important to design WRSM strategies to effectively alleviate crash risks induced by different combinations of RWCs.

1.1.3 Conceptual framework for developing WRVSLs

Developing a WRVSL system is a three-step process (Figure 1.3).

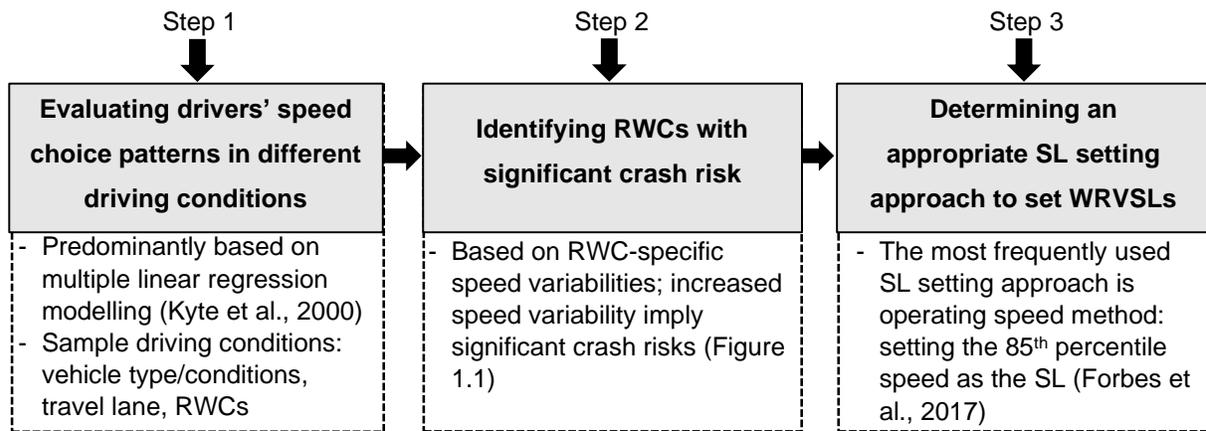


Figure 1.3: A stepwise conceptual framework for developing WRVSLs

The first step assumes that drivers do not have knowledge on the safest speed to drive in adverse RWCs and/or drivers may take risks when selecting a speed to travel. Accordingly, the first step involves evaluating drivers' speed behaviour in different driving conditions such as (i) different RWCs (Kyte et al., 2000; Yasanthi & Mehran, 2020), (ii) travel lane (Yasanthi & Mehran, 2020), and (iii) vehicle type (Catbagan & Nakamura, 2008; Yasanthi & Mehran, 2020), i.e., passenger vehicle or heavy vehicle — vehicles with more than four tires touching the ground (HCM, 2022). When evaluating speed behaviour of heavy vehicles (HVs), it is also important to consider other HV-related characteristics such as truck type and payload hauled (i.e., the weight of commodity hauled by a truck) which might (i) have a considerable impact on speed behaviour of HV drivers particularly when driving under adverse RWCs, and/or (ii) lead to severe consequences in a potential HV-involved crash. In fact, some HV-involved crashes (e.g., crashes involving hazardous material transport trucks) may result in severe consequences such as leakage of toxic gasses and/or release of flammable substance. The conventional methods used to understand context-specific speed behaviour (e.g., speed choice in adverse RWCs) in road safety research are largely based on

multiple linear regression modelling (Kyte et al., 2000; Yasanthi & Mehran, 2020). Yet, there is an emerging trend of developing speed distributions for different driving conditions (e.g., prevailing RWCs, vehicle type driven) to evaluate speed behaviour in different contexts (Hoogendoorn, 2005a). The second step of the WRVSL development process (Figure 1.3) involves identifying RWCs with intensified crash risks. Such hazardous RWCs could be identified by quantifying crash risk based on the RWC-specific speed variabilities estimated in step 1 of the WRVSL development process (Figure 1.3). The third and final step in developing a WRVSL system corresponds to selecting an effective approach to set WRVSLs. In general, SLs are set based on the operating speed approach, in which, the 85th percentile speed (i.e., the speed at or below which is selected by 85 percent of drivers) is adopted as the SL (FHWA, 2009).

1.2 Problem statement, research needs and research questions

Each step of the stepwise WRVSL development approach (Figure 1.3) endures a unique set of problems.

- Problems related to step 1: Evaluating drivers' speed choice patterns in different driving conditions

The conventional approach to evaluate drivers' speed behaviour is largely based on regression modelling. Such regression-based speed behaviour models attempt to estimate average travel speed chosen by a driver to drive in a specific combination of driving conditions (e.g., speed chosen by a passenger car (PC) driver travelling in snow precipitation and icy pavements). Fundamental problems associated with such regression-based speed models are twofold. First, extreme RWCs are rare events resulting reduced traffic throughput in such RWCs (Goodwin, 2002); thus, weather-related speed studies are often confined to limited sampling conditions. Therefore, generalizing speed behaviour of an entire driver population using average speed estimates is questionable. Second, selecting a travel speed represents a driver's cognitive ability in rational decision making. In other words, evaluating the impacts of a set of driving conditions on drivers' speed behaviour involves comprehending the typically versatile human psychology. In fact, modelling human psychology through multiple linear regression modelling might be challenging, particularly under limited sampling conditions.

Such problems associated with the conventional approach of modelling RWC-specific speed behaviour have resulted in inconsistent RWC-specific speed behaviour estimations as suggested by past literature. For instance, some studies (Ibrahim & Hall, 1994; Kyte et al., 2000) suggest that extreme RWCs considerably affect drivers' speed behaviour; in contrast, some studies (Edwards, 1999; Yasanthi & Mehran, 2020) demonstrate marginal impacts of RWCs on drivers' speed choice. Such controversies in past literature emphasize the need for a robust approach to model RWC-specific speed behaviour. In addition, drivers' speed choice is also hypothesized to be dependent on the vehicle type driven (Hoogendoorn, 2005b), i.e., PC or heavy vehicles — vehicles with more than four tires touching the ground (HCM, 2022), and/or the payload conditions of a heavy vehicle (HV). For instance, speed choice of a PC driver travelling in snow precipitation and an icy pavement might be differ from the speed choice of the driver of a loaded HV travelling in the same RWCs. Yet, past literature on the impacts of RWCs on the speed choice of drivers driving different vehicle types/HV loading conditions is limited.

- Step 2: Identifying RWCs with intensified crash risk

An intuitive approach to identify RWCs with intensified crash risk is to evaluate the impacts of different RWCs on crash frequency. In practice, however, road safety researchers may not have access to accurate and complete crash datasets. When historical crash data are not available, speed variability is conventionally used as a metric of crash propensity (Solomon, 1964). Yet, past literature does not provide guidance on developing a holistic crash indicative measure to evaluate RWC-specific crash potential in different driving conditions. On the other hand, crashes are rare events; motor vehicle collisions in extreme RWCs are particularly scarce. Identifying RWCs with intensified crash risks based on such limited crash events is questionable because (i) developing weather-related crash models with statistically significant model coefficients under limited sampling conditions is challenging, and (ii) differentiating the impacts of RWCs on crash risk from the other factors (e.g., driver fatigue, vehicle condition) contributing to observed crashes is arguable particularly under limited sampling conditions. Therefore, there is a need to develop a robust crash indicative measure which could help identify RWCs with intensified crash risks, particularly when historical crash data are not available.

- Step 3: Determining an appropriate SL-setting approach to set WRVSLs

Conventionally, fixed posted SLs, such as the SLs implemented on most highways in North America, are (i) designed for ideal RWCs (i.e., no precipitation, and dry pavement conditions), and (ii) often set based on the 85th percentile speed (Forbes et al., 2012). Yet, such fixed speed limits set for normal RWCs may not be appropriate for extreme RWCs, because (i) some drivers may not be able to comprehend the safest speed to travel in inclement RWCs, (ii) some RWCs may lead to increased speed variabilities which may result in increased crash potential in such RWCs, and (iii) the degraded driving conditions resulted by most adverse RWCs warrant reduced travel speeds. As a solution, WRVSLs attempt to communicate an estimation of the safest speed to travel in different RWCs, thus mitigate weather-related crash risks. Despite the potential safety benefits of WRVSL systems (e.g., reduced crash risks resulting from harmonization of speeds), the existing WRVSL systems are deemed ineffective irrespective of the SL-setting approach used to set the WRVSLs (Downey, 2015; Elvik et al., 2009). Therefore, it is important to develop a robust approach to set effective WRVSLs.

In summary, there are four research questions arising from the problems related to the WRVSL development process:

1. What is the impact of different RWCs on drivers' speed behaviour?
2. What is the most appropriate approach to evaluate drivers' speed behaviour in adverse RWCs?
3. Are there any specific driving conditions (e.g., RWCs, vehicle types) which notably intensify safety/crash risks when indicated by speed and speed variability?
4. What is an effective approach to regulate speed in adverse RWCs to mitigate safety risks induced by such RWCs?

1.3 Goal, objectives, definitions, and scope

The goal of this thesis is to improve road safety of cold region rural highways in extreme RWCs.

The goal of this research is achieved through four main objectives:

- **Objective 1 (O₁)** : To model desired speed distributions under the combined effect of specific road-weather and traffic conditions
- **Objective 2 (O₂)** : To investigate the variability of modelling speed distributions based on different data analytics

- **Objective 3 (O₃)** : To identify adverse RWCs which notably intensify safety risks based on RWC-specific speed variabilities
- **Objective 4 (O₄)** : To propose an effective approach to set WRVSLs based on the reliability theory

The scope of this thesis is defined by the study data which are collected on rural highways. Therefore, methods proposed in this study are only applicable to rural highways. In addition, the scope of this thesis includes heavy due to the scarcity of past literature focusing on the impacts of RWCs on heavy vehicle drivers' speed choice. Further, it should be noted that, in this thesis, the safety considerations are based on speed and speed variabilities.

The proposed framework to develop an effective WRVSL system would help road safety researchers and transport authorities located in cold regions to (i) understand safety risks associated with extreme RWCs, (ii) identify the need for WRVSL systems in cold region rural highways, and (iii) develop an effective WRVSL system thus mitigate weather-related crashes in cold region rural highways.

1.4 Thesis organization

This thesis is organized as a sandwich-style (a.k.a. a grouped-manuscript) thesis; thus, each chapter reproduces published peer-reviewed journal articles. It should be noted that the content presented in chapters two through four are slightly changed from their original papers to facilitate a smooth flow of the thesis. All papers reproduced in this thesis are reprinted with the authorization of the corresponding publishers and co-authors. Each chapter includes its own abstract, introduction, literature review, and conclusion sections. The acknowledgment section, and author contributions of the original papers are included in the thesis preface; references are provided at the end of the thesis.

Chapter 2 presents a research article focusing on developing an approach to model speed distributions in different RWC and traffic combinations (O₁) followed by identifying adverse RWCs intensifying safety risks (O₃). Chapter 3 pertains to a research paper on (i) comparing the speed distribution modelling approach developed in chapter 2 with a conventional speed modelling approach (i.e., regression modelling) (O₂), and (ii) developing a holistic crash indicative measure to identify RWCs with intensified safety risks using RWC-speed variabilities (O₃), in the context

of different vehicle types including freight transport trucks. Chapter 4 pertains to a research paper focusing on the development of an effective WRVSL system based on the reliability theory (O₄).

Chapter 5 concludes the thesis by (i) summarizing research findings, (ii) highlighting contributions of this thesis, (iii) outlining research limitations, and (iv) recommending future research, pertaining to each research article presented in chapters 2 through 4.

Chapter 2. Modelling speed behaviour in rural highways: Safety analysis of driving under adverse road-weather conditions

This chapter initiates the process of developing an effective WRVSL system for cold region rural highways by evaluating drivers' speed behaviour in different RWCs. More specifically, this chapter presents an innovative approach to (i) model speed distributions in different RWC and traffic condition combinations (O_1 ; step 1 in Figure 1.3), and (ii) identify RWCs with intensified safety risks (O_3 ; step 2 in Figure 1.3). The speed distribution modelling approach proposed in this chapter can be effectively used as an alternative to the conventional regression-modelling-based approach used to evaluate speed behaviour in different conditions.

2.1 Abstract

This research proposes a methodical approach to model desired speed distributions under different road-weather and traffic conditions followed by the identification of road-weather conditions with potentially higher safety risks in rural divided highways located in extremely cold regions. Distributions of drivers' desired speed (i.e., speed chosen at driver's discretion at specific traffic and roadway conditions) for unique combinations of adverse road-weather and traffic conditions are modelled as normal distributions characterized by their means and standard deviations formulated based on two principal statistical theorems and techniques i.e., Central Limit Theorem and Minimum Variance Unbiased Estimation. Combination of the precipitation conditions, road surface conditions, time of the day, temperature, traffic flow and the heavy vehicle percentage at the time of travel were considered in defining the combinations of road-weather and traffic conditions. The findings reveal that simultaneous occurrence of particular precipitation and pavement conditions notably affect the characteristics of the desired speed distribution and potentially expose drivers to elevated safety risks. Jurisdictions experiencing extreme road-weather conditions may adopt the proposed methodology to assess speed behaviour under different road-weather conditions to establishing and deploying weather-responsive traffic management strategies such as variable speed limit to regulate speeding and improve traffic safety in winter.

Keywords: Adverse road-weather conditions; Speed behaviour; Desired speed; Highway speed management; Traffic safety

2.2 Introduction

Highway safety, characterized by the ability of a person to travel freely without injury or death, has always been the primary objective of traffic engineering and is typically measured by the rate of crashes belonging to different severity levels (Carter et al., 2017). As defined by Hauer (2020), a cause of a crash can be defined as “a circumstance or an action that, were it different, the frequency of crashes and/or their severity would be different”. The contributing factors to crashes can be divided into four major categories: human factors (driver, pedestrian, etc. behavior), vehicle conditions, roadway conditions and environmental conditions (Garber & Hoel, 2004). A considerable proportion of crashes includes weather-related crashes, which are defined as crashes occurring in the presence of rain, sleet, snow, fog, wet pavement, snowy/slushy pavement, and/or icy pavement (Pisano et al., 2008). In fact, Transport Canada (2018) identified 307 fatalities resulting from weather-related crashes in Canada in 2016. It is well documented in the literature that weather-related crashes mostly stem from the impacts of adverse road-weather conditions on driver behaviour (Pisano et al., 2008; Andrey & Yagar, 1993).

Speed is considered as one of the major contributing factors to crashes in terms of frequency and severity, thus it has been widely used as a surrogate measure in assessing highway safety (Abdel-Aty et al., 2008). Speed limits posted by jurisdictions regulate and influence drivers' speed choice and attempts to mitigate crash risks (Elvik, 2012). It is hypothesized that particular road-weather conditions may alter drivers' speed choice and intensify crash risks. Therefore, fixed speed limits may not effectively mitigate crash risks under adverse road-weather conditions. Variable Speed Limit (VSL) which is implemented for real-time crash mitigation (Abdel-Aty et al., 2008) is a viable weather-responsive traffic management strategy to alleviate safety risks imposed by adverse road-weather conditions. Thus, understanding the speeding behaviour under adverse road-weather conditions is essential for successful implementation of a weather-responsive VSL program. Desired speed, which is the speed chosen at driver's discretion at specific traffic and roadway conditions directly reflects driver behaviour and is extensively used in studying driver behaviour (Yasanthi & Mehran, 2020; Catbagan & Nakamura, 2008). Desired speed is manifested through the free-flow speed of vehicles (TRB, 2016) and is referred to as “speed” hereafter in. In fact, the highway capacity manual (HCM) (TRB, 2016) recommends taking the speed observed under base conditions at traffic flow volumes less than 1,000 passenger cars per hour per lane, as a reliable

measure of free-flow speed (i.e., a measure of average vehicle speed on a specific facility) as traffic becomes insensitive to flow rate at this much low traffic volumes. Accordingly, the lack of negligible vehicle interactions in such free-flow conditions substantiate approximating observed vehicle speeds as drivers' desired speeds. Distribution of desired speeds at a specific location varies in a wide range mainly attributing to geometric characteristics of the road, existing road-weather conditions and the demographics of the drivers. In other words, desired speed explains the speed at which the drivers are comfortable under the prevalent driving conditions. Adverse road-weather conditions such as icy road surface conditions and poor visibility are proven to affect the desired speed notably (Yasanthi & Mehran, 2020; Kyte et al., 2000). Accordingly, speed behavior, which is defined as the choice between several possible speeds (Letirand & Delhomme, 2005), is widely chosen as a measure to study the impacts of adverse road-weather conditions on driver behaviour.

Past literature explored several approaches to investigate the impacts of adverse road-weather conditions on speed behaviour including statistical modelling and machine learning techniques (Yasanthi & Mehran, 2020; Ibrahim & Hall, 1994; Thakuriah & Tilahun, 2013). For instance, Yasanthi & Mehran (2020) developed a group of statistical models to estimate speed variations in different road-weather conditions by developing six multiple linear regression models where separate linear and non-linear models were developed for three data treatment approaches: (i) treating individual vehicle speeds as regression model's dependent variable, (ii) treating 20-minute aggregate speeds as regression model's dependent variable, and (iii) adopting a data elimination technique. According to the study results of Yasanthi & Mehran (2020), models with 20-minute aggregate speeds as the dependent variable (Group II models) was concluded as the best performing model. Hoogendoorn (2005a) developed a new unified approach to estimate the free speed distributions based on the composite headway distribution model and the method of censored observations.

While some studies confirmed crucial changes in speed behaviour under adverse road-weather conditions (Kyte et al., 2000; Ibrahim & Hall, 1994; Oh et al., 2002), other studies contradict these observations, mainly in terms of the intensity of speed dispersion (Yasanthi & Mehran, 2020; Edwards, 1999). Further, the performance of conventional approaches such as regression modelling to study speed behaviour has been criticized (Yasanthi & Mehran, 2020). The

underlying reasons for discrepancies in the literature are mainly threefold. The main prospective reason discussed is limited sampling to represent specific road-weather conditions (Yasanthi & Mehran, 2020; Hall & Barrow, 1988). On one hand, adverse road-weather conditions are intrinsically scarce, restricting the number of samples, which can be collected under specific inclement road-weather conditions. On the other hand, the travel patterns are considerably affected by adverse road-weather conditions resulting in reduced traffic volumes under adverse road-weather conditions (Goodwin, 2002). Limited sample rate under adverse road-weather conditions results in inconsistent speed reduction estimations and limits the applicability of desired speed distribution modelling approaches developed by previous studies. For instance, free speed distribution models developed in Catbagan & Nakamura (2008) and Hoogendoorn (2005a) require sufficiently large samples to minimize the standard error of the estimates, which may not be feasible under adverse road-weather conditions. Second, most of the studies relied on road-weather data collected from data collection devices located considerably distant from the traffic data collection stations, which potentially reduces the representativeness of prevailing road-weather conditions of the collected data. Third, treatment of speed data (e.g. aggregation interval) also appears to impact the outcomes (Yasanthi & Mehran, 2020; Kyte et al., 2000). While it is common to analyze speed behaviour through statistical modelling, it should be noted that speed behavior is primarily a human response. Thus, modelling human psychology through regression modelling, especially with a limited number of samples is challenging and may lead to questionable results (Dhami & Harries, 2001). Further, the form of the dependent variables in the statistical models, whether aggregated or disaggregated as well as the aggregation interval is proven to impact the study results (Yasanthi & Mehran, 2020). In fact, the optimum aggregation interval for loop detector data (i.e., speed data) depends on the purpose of the application and traffic conditions (Park et al., 2009). Moreover, many researchers queried about the functional form to be used in modelling the impacts of road-weather conditions on traffic stream characteristics of uninterrupted flow (Yasanthi & Mehran, 2020; Ibrahim & Hall, 1994).

While using identical study data as used by Yasanthi & Mehran (2020), this research intends to fill the gaps identified in past literature by proposing a novel approach for modeling distributions of desired speeds in uncongested highways under different combinations of road-weather and traffic conditions, rather than quantifying speed variations resulted by adverse road-weather conditions. The proposed methodology embodies two novelties. First, by modelling desired speed

distributions rather than quantifying speed variations resulted by adverse road-weather conditions through statistical modelling (e.g., regression analysis), the proposed methodology in this research attempts to avoid the well-documented issues in RWC-speed modelling such as poor goodness of fit and statistically insignificant model coefficients which are attributed to the low sample sizes and the divergent nature of human behaviour. This research attempts to resemble the speed choice of a specific driver population rather than attempting to confine drivers' speed behavior by a specific mathematical criterion such as ordinary least squares method. Accordingly, the proposed methodology can be practiced under limited sampling conditions as it preserves all observations. Second, this research attempts to evaluate the impacts of different road-weather conditions on a population-level in contrast to the conventional sample-level evaluations. One of the main advantages of the proposed novel approach is that modelling the desired speed distributions eliminates the need to define a specific function or an aggregation interval. Moreover, road-weather and traffic data collection devices used in this research are located alongside each other providing highly accurate data.

2.3 Research objectives

This research attempts to address two research questions related to rural divided highways: (i) How does the distribution of desired speeds change in the presence of adverse road-weather conditions? and (ii) Are there specific adverse road-weather conditions, which notably intensify safety risks?

The research has two objectives focusing on addressing the research questions:

- i. To model the desired speed distributions under the combined effect of specific road-weather and traffic conditions, and
- ii. To identify specific adverse road-weather conditions which notably intensify the safety risks based on speed and speed variability.

The distribution of traffic speeds is primarily dependent on drivers' speed choice ranging from very low to very high speeds. According to Hauer (1971), slow drivers habitually select lower speeds because they believe that slow driving is safe driving. On the other hand, fast drivers trade off safety for lesser travel times (Hauer, 1971). In this research, it is hypothesized that the population of drivers under specific road-weather conditions choose a safe speed to travel based on their comfort, convenience and confidence to travel under prevailing road-weather conditions.

Further, it is hypothesized that the divergent speed selection patterns under specific adverse road-weather conditions increase the variability of desired speeds, which is an indication of increased safety risks. The research attempts to test these hypotheses by estimating desired speed distribution parameters for different populations of drivers under different road-weather and traffic conditions. Results of this research will help authorities establish weather-responsive traffic management schemes i.e., VSL to improve traffic safety and operations under adverse road-weather conditions in cold regions.

2.4 Study data

2.4.1 Data collection

The study data were collected with the courtesy of Alberta Transportation and are of two types; road-weather and traffic data. Road-weather data were collected by using a Road Weather Information System (RWIS) and traffic data were acquired by a Weigh-In-Motion (WIM) station installed alongside of each other (148.7m apart) on a four-lane, two-way divided highway segment (Figure 2.1). The study site is located on Highway 16 i.e., a major inter-provincial highway in Western Canada a.k.a “Yellowhead highway”. It connects Jasper and Lloydminster via Edmonton, and the study site lies west to the city of Edmonton. The location configuration of RWIS and WIM sensors (Figure 2.1) enabled collecting real-time, highly accurate and representative road-weather information for each vehicle recorded by the WIM station, adding a distinct feature to this research. The study site was subjected to an Annual Average Daily Traffic (AADT) of 8,120 vehicles in 2015 (Alberta Transportation, 2016). Moreover, the study site is in a level (i.e., negligible grade), straight road section without any on/off ramps nearby. The study data were collected for 15 months ranging from October 2014 to December 2015.

The RWIS station records the pertaining road-weather conditions including air temperature, atmospheric precipitation situation, pavement surface condition, pavement temperature, and average wind speed in every 20 minutes. In terms of traffic data, the WIM station detects the date, time and vehicle-by-vehicle information including travel lane, travel speed, axle weight and interaxial spacing.

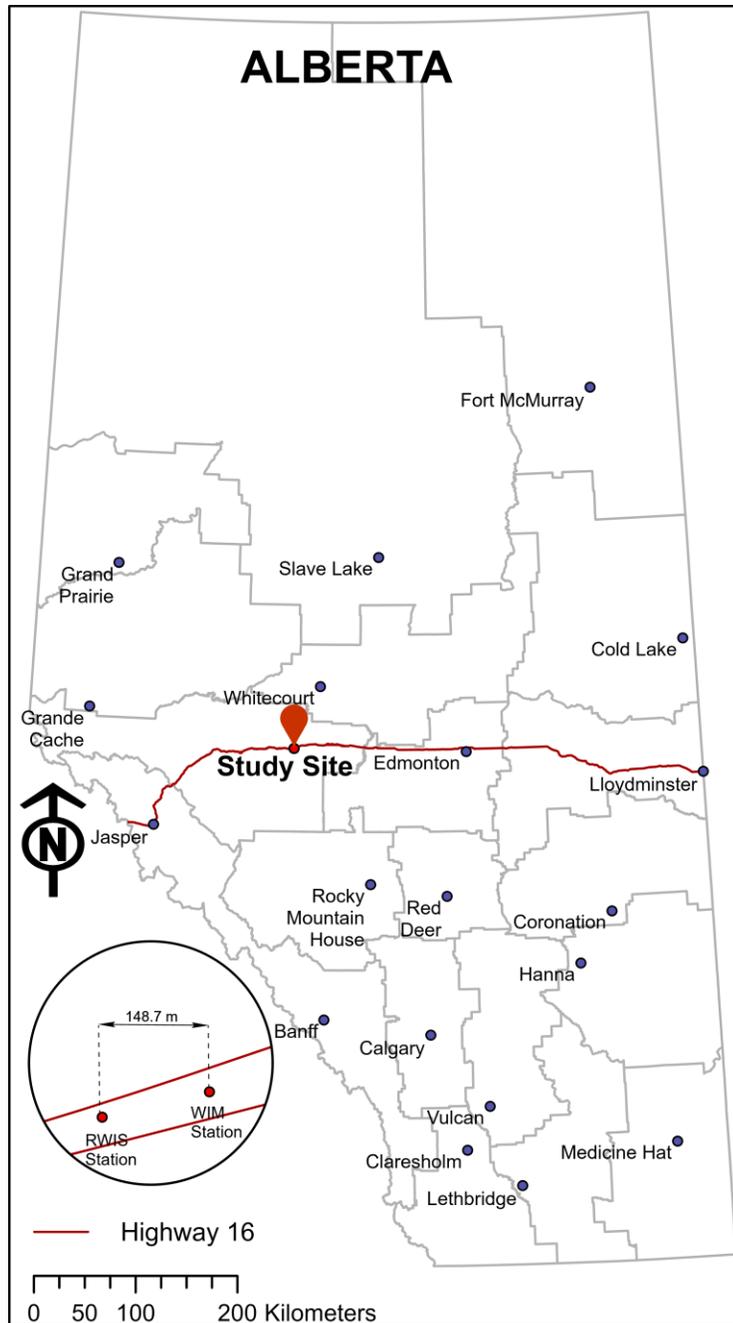


Figure 2.1: Study site location and the positions of the RWIS & the WIM station (Yasanthi et al., 2021)

Notes:

- Map generated using data retrieved from Statistics Canada (2016a), Statistics Canada (2016b), and Government of Canada (2021).
- Contains information licensed under the Open Government Licence-Canada.

2.4.2. Data preparation

Table 2.1 summarizes the details of the road-weather conditions as recorded by the RWIS station and the alterations made to the raw data recorded by the RWIS station, in terms of the categorization of continuous data and/or re-grouping of the categorical data records.

Initially, atmospheric temperature, precipitation situation, pavement surface condition, wind speed and time of the day were selected as independent variables representing road-weather conditions as discussed in Yasanthi & Mehran (2020). However, wind speed was not considered in the analysis since the maximum wind speed recorded at the study site is 45.36 km /h, which is lower than the reported minimum wind speed required (51 km/h) to impact the behavior of passenger car drivers (Schmidlin et al., 2003). The next step of data preparation includes evaluating whether speeds in each road-weather attribute level recorded by the RWIS station (Table 2.1) are statistically different from the speeds recorded under the remaining levels of the same road-weather attribute (e.g., whether speeds recorded under slight rain condition follow the same distribution as speeds observed under moderate rain). Therefore, a series of two-sample Kolmogorov–Smirnov (K-S) tests was conducted to (i) re-group the levels of road-weather condition categories originally recorded by the RWIS by statistically evaluating the difference of speeds in each level of the categorical road-weather attributes, and (ii) categorize the continuous variables (i.e., temperature). Reasons for choosing K-S test to evaluate the statistical independence of speeds observed in each road-weather attribute level are twofold. First, the K-S test is a non-parametric test (Engmann & Cousineau, 2011) which does not require any presumptions about the underlying distribution(s) of the tested samples. Second, the K-S test is particularly recommended when testing samples with small and/or possess unequal sample sizes (Engmann & Cousineau, 2011). Results of the K-S tests conducted in this thesis are presented in Appendix A. After analyzing the K-S test results (Appendix A), the following alterations were executed:

- Original precipitation conditions recorded by the RWIS station were combined in terms of the intensities due to the statistically insignificant difference of speeds recorded under most “Moderate” and “Heavy” precipitation levels (Figure A1.2). The resultant precipitation conditions only have two levels labeled “slight” and “Moderate & Heavy” each for rain, frozen precipitation and snow (Table 2.1).

- Air temperature was categorized into three groups (GI, GII and GIII) based on the numerical value of the air temperature as shown in Table 2.1.
- Precipitation condition, pavement surface condition, temperature and time of the day were modified to have seven, six, three and two levels, respectively (see Appendix A for a detailed explanation of the re-grouping process).

Table 2.1: Road-weather condition categories

RW condition	Description of the raw data as recorded by RWIS		Alterations to the raw data	
Air temperature	Wet bulb temperature falling between minus and plus 40 recorded as a continuous variable		GI: Temperature $\leq -10^{\circ}\text{C}$ GII: $-10^{\circ}\text{C} < \text{Temperature} \leq 0^{\circ}\text{C}$ GIII: Temperature $> 0^{\circ}\text{C}$	
Pavement surface condition	Dry	No moisture or unusual condition detected	No alterations	
	Ice warning	Detection of ice or black ice		
	Trace moisture	Detection of isolated moisture on pavement surface		
	Wet	Wet roadway with considerable moisture detection		
	Ice watch	The risk of the formation of ice or black ice on the roadway is elevated, but its occurrence, location, and/or timing is still uncertain		
Pavement surface condition	Frost	Detection of frost formation	No alterations	
Time of the day	Day	Time from the sunrise to sunset		No alterations
	Night	Time from the sunset to sunrise		
Precipitation condition	No precipitation	0 mm/h	No precipitation	No alterations
	Rain	<2 mm/h	Slight rain	No alterations
		≥ 2 and <8 mm/h	Moderate rain	Moderate & heavy rain
		≥ 8 mm/h	Heavy rain	
	Snow	<2 mm/h	Slight snow	No alterations
Precipitation condition	Snow	≥ 2 and <8 mm/h	Moderate snow	Moderate & heavy snow

RW condition	Description of the raw data as recorded by RWIS			Alterations to the raw data		
Precipitation condition	snow	≥ 8 mm/h	Heavy snow			Moderate & heavy snow
	Frozen precipitation	< 2 mm/h	Slight precipitation	frozen		No alterations
		≥ 2 and < 8 mm/h	Moderate precipitation	frozen		Moderate & heavy frozen precipitation
		≥ 8 mm/h	Heavy precipitation	frozen		Moderate & Heavy Frozen Precipitation

As for traffic data, erroneous records including “Error”, “Other” and “Not Applicable” entries as well as vehicles with speeds higher than 200km/h were removed from the analysis. Further, the study data were aggregated into five-minute intervals, which is suggested as the optimum interval to investigate traffic operations (Oh et al., 2005). Consequently, traffic flow and the percentage of heavy vehicles values for each five-minute interval count were estimated. Traffic flow was grouped into eight levels with a bin size of 100 veh/h. Likewise, heavy vehicle percentage was grouped into 10 levels with a bin size of 10%. Traffic flow and heavy vehicles percentage categories were also subjected to a series of two-sample K-S tests to verify the statistical significance of the difference of speeds in each category (see Appendix A). Accordingly, each five-minute interval during the analysis period was tagged with the prevailing precipitation, pavement condition, temperature, time of the day, heavy vehicle percentage and traffic flow categories at the time of travel.

Table 2.2 tabulates statistical information of the influencing factors with respect to several statistical attributes; number of vehicles recorded under each level of the influence factor along with the mean and standard deviation where applicable.

Table 2.2: Influencing factors’ statistical information

Influencing factor			Vehicles observed	Mean	Standard deviation
Road- weather conditions	Air temperature	GI ($\leq -10^{\circ}\text{C}$)	230,489	-16.58 °C	4.77 °C
		GII ($-10^{\circ}\text{C} < \text{ \& } \leq 0^{\circ}\text{C}$)	441,030	-3.98 °C	2.79 °C
		GIII ($> 0^{\circ}\text{C}$)	880,823	8.74 °C	6.35 °C
	Pavement surface condition	Dry	958,017	Not Applicable	
		Ice warning	269,603	Not Applicable	
		Trace moisture	41,498	Not Applicable	

Influencing factor			Vehicles observed	Mean	Standard deviation	
Road-weather conditions	Pavement surface condition	Wet	134,556	Not Applicable		
		Ice watch	132,540	Not Applicable		
		Frost	16,128	Not Applicable		
Time of the day	Time of the day	Day	1,159,360	Not Applicable		
		Night	392,982	Not Applicable		
Precipitation condition	Precipitation condition	No precipitation	1,448,239	Not Applicable		
		Slight rain	21,429	0.48 mm/h	0.54 mm/h	
		Moderate & heavy rain	46,397	6.6 mm/h	4.29 mm/h	
		Slight snow	23,332	0.55 mm/h	0.47 mm/h	
		Moderate & heavy snow	1,685	2.96 mm/h	2.73 mm/h	
		Slight frozen precipitation	9,578	0.63 mm/h	0.75 mm/h	
		Moderate & heavy frozen precipitation	1,682	5.28 mm/h	5.27 mm/h	
Traffic conditions	Traffic flow category	TF: GI (≤ 100 veh/h)	38,923	48 veh/h	27 veh/h	
		TF: GII (101 veh/h-200 veh/h)	26,544	152 veh/h	28 veh/h	
		TF: GIII (201 veh/h-300 veh/h)	27,953	248 veh/h	31 veh/h	
		TF: GIV (301 veh/h-400 veh/h)	10,962	346 veh/h	27 veh/h	
		TF: GV (401 veh/h-500 veh/h)	3,107	440 veh/h	27 veh/h	
		TF: GVI (501 veh/h-600 veh/h)	767	538 veh/h	29 veh/h	
		TF: GVII (601 veh/h-700 veh/h)	170	646 veh/h	29 veh/h	
		TF: GVIII (701 veh/h-800 veh/h)	94	752 veh/h	30 veh/h	
	Heavy vehicles percentage category	Heavy vehicles percentage category	HV: GI ($\leq 10\%$)	19,492	3.1%	3.8%
			HV: GII (11%-20%)	20,410	15.8%	3.0%
			HV: GIII (21%-30%)	19,032	25.4%	2.7%
			HV: GIV (31%-40%)	16,847	35.2%	23.0%
			HV: GV (41%-50%)	13,623	47.5%	3.1%
			HV: GVI (51%-60%)	3,919	57.2%	2.5%
Traffic conditions	Heavy vehicles percentage category	HV: GVII (61%-70%)	4,275	66.1%	1.7%	
		HV: GVIII (71%-80%)	2,656	75.8%	2.7%	
		HV: GIX (81%-90%)	576	84.6%	1.9%	
		HV: GX ($>90\%$)	7,690	99.9%	0.3%	

2.4.3 Preliminary analysis

A preliminary data analysis was conducted on the eastbound data aiming to: (i) identify the level of vehicle interactions at the study site, (ii) understand the overall speed behaviour in different road-weather conditions using descriptive statistics of speeds in different road-weather conditions, and (iii) analyze traffic composition and lane utilization patterns at the study site. A past study conducted using identical study data (Yasanthi & Mehran, 2020) revealed that speed behaviour in the eastbound and westbound directions at the study site were marginally different. Accordingly, in this research, study data collected from the eastbound data were used (a similar analysis could be conducted on westbound data).

Figure 2.2 depicts the speed-flow relationship in the eastbound shoulder and median lanes separately, where flow and speed are expressed in the hourly equivalent of the five-minute passenger car volumes per lane and the corresponding five-minute aggregate speed (i.e., five-minute mean speed) respectively. The flow values presented in Figure 2.2 were estimated by considering the standard passenger car equivalent (PCE) units — the number of passenger cars needed to replace the capacity of the considered vehicle (HCM, 2022) — for freeways with a negligible terrain. In this thesis, motorcycles were assigned with a PCE unit of 0.5 (Parvathy et al.,

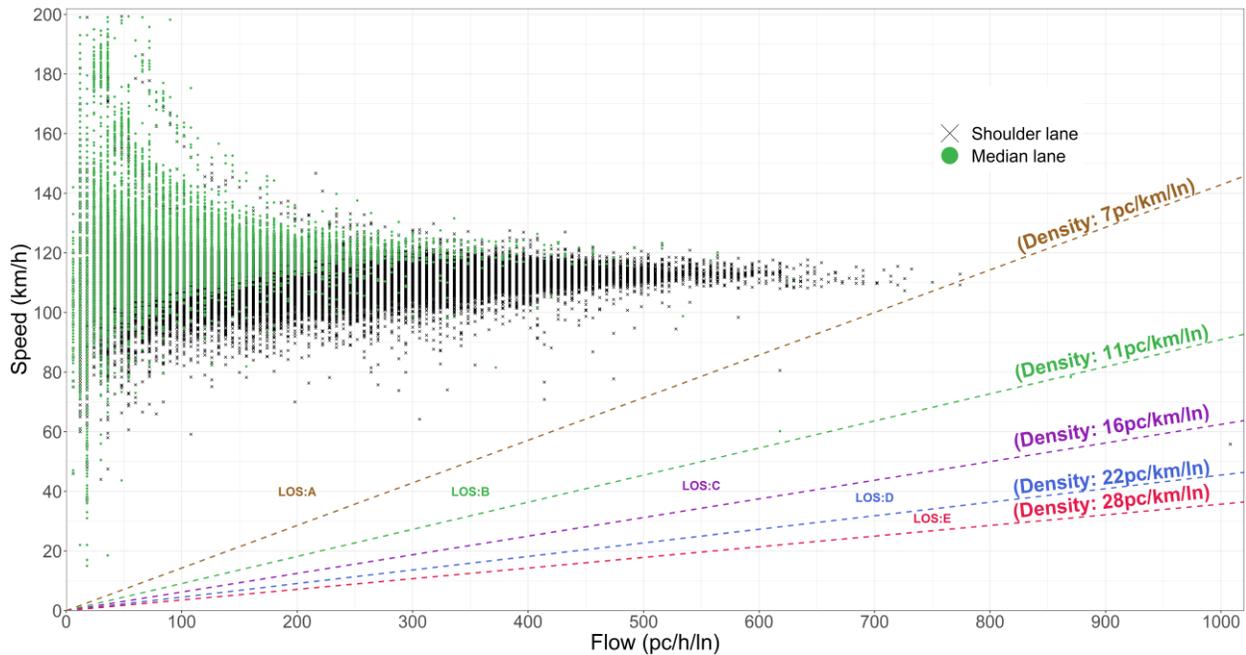


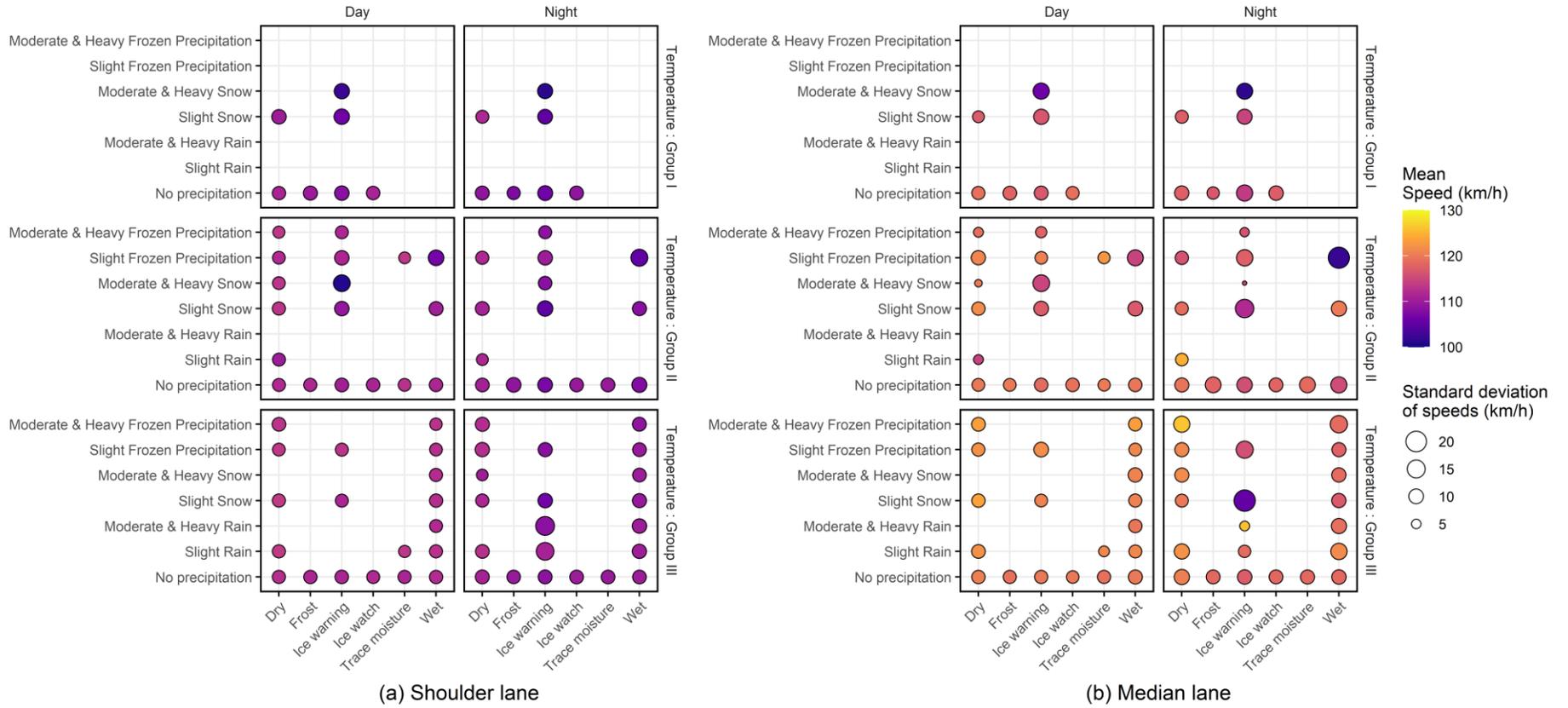
Figure 2.2: Speed-flow variation in the eastbound shoulder and median lanes at the study site (Yasanthi et al., 2021)

2013), passenger cars with a PCE unit of 1, and trucks with a PCE unit of 1.5 (HCM, 2010). Figure 2.2, the study site mostly operates under a density of 7pc/h/ln and experiences Level of Service (LOS) “A”, which exhibits free-flow conditions and minimal vehicle interactions between the vehicles as characterized by Oh et al. (2005).

The median lane, however, exhibits a lower range of traffic flow values and higher speeds compared to the shoulder lane. While presenting two types of descriptive statistics; mean and standard deviation of the speeds observed in the eastbound shoulder and median lanes, Figure 2.3 confirms the observation of higher speeds in the median lane while highlighting the variations in the mean and standard deviation of speeds under adverse road-weather conditions in both lanes.

Further, vehicles travelling in daytime, especially in the shoulder lane, travel at slightly higher speeds. Adverse pavement surface conditions, when coupled with adverse precipitation conditions tend to reduce speeds considerably. For instance, presence of an ice warning pavement condition under a slight snow recorded notably lower mean speeds and considerably high standard deviation of speeds irrespective of the temperature conditions, time of day and the lane type, suggesting that the drivers are more sensitive to perceptible hazards such as adverse pavement and precipitation conditions. Figure 2.3 suggests a few counterintuitive implications potentially due to the small sample sizes. For example, the mean and standard deviation of speeds of vehicles travelling in the median lane at nighttime in temperature values more than 0 °C under a dry pavement and a moderate and heavy frozen precipitation recorded a comparatively higher mean speed which can potentially be attributed to the low sample size (44 vehicle speeds only).

Figure 2.4 presents the hourly variation of the heavy vehicles’ percentage and the lane utilization factor in the eastbound shoulder and median lanes, where the lane utilization factor is expressed in terms of the proportion of hourly vehicular flow of each lane to the respective total hourly vehicular flow of the eastbound direction. According to Figure 2.4, a substantial portion of the hourly vehicular flow in the shoulder lane constitutes of heavy vehicles, especially around the midnight. Moreover, the high hourly lane utilization factors of the shoulder lane further imply the preference of the majority of the drivers to use the shoulder lane. Therefore, only the vehicles observed in the shoulder lane are considered in this research to typify the pragmatic heterogeneous traffic composition in study site.



**Figure 2.3: Distribution of speeds in the study site
(Yasanthi et al., 2021)**

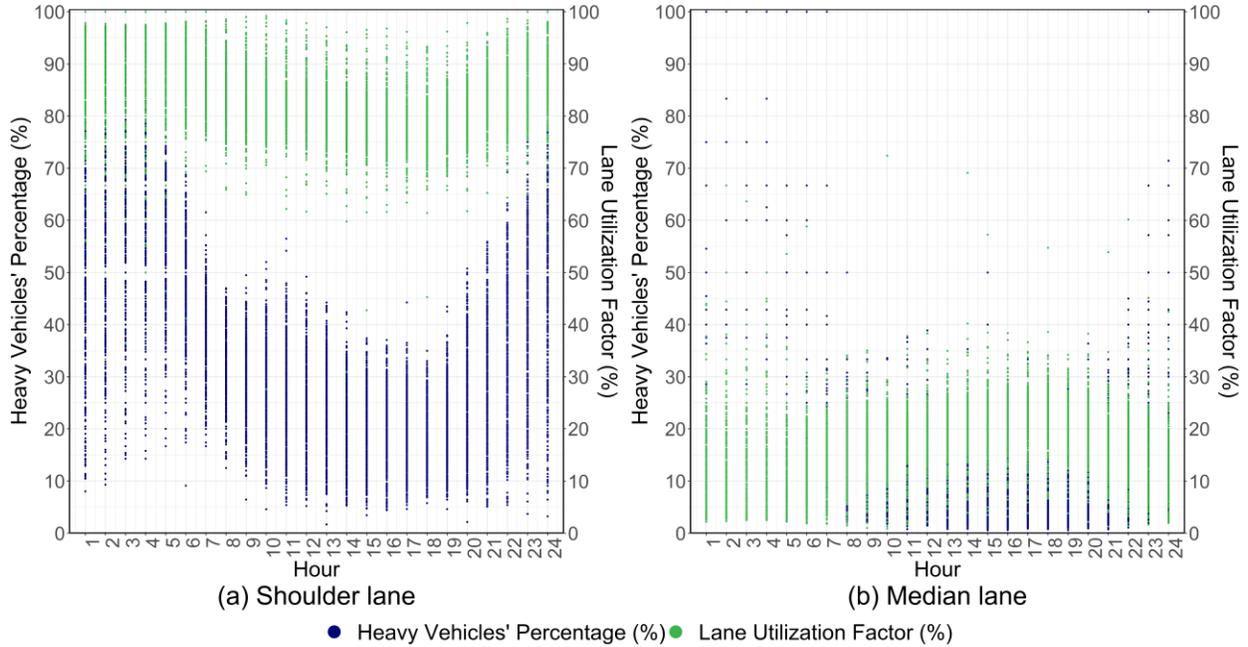


Figure 2.4: Hourly variation of the Heavy vehicles' percentage and the lane utilization factor at the study site (Yasanthi et al., 2021)

2.5 Population, Sample and Sampling Distributions

It is hypothesized that drivers possess different speed behaviors under various inclement road-weather and traffic conditions, characterized by a “population” to which they belong to. Three important aspects in modelling the distribution of desired speeds are defined namely, population, sample and sampling distribution.

Population, in the context of this research, encompasses the desired speeds of individual vehicles observed at the study site under a specific combination of road-weather and traffic conditions. Population groups are identified based on the combination of six criteria explaining the road-weather and traffic conditions at the time of travel (Figure 2.5). The distribution of desired speeds of the individual vehicles observed under a unique combination of road-weather and traffic conditions is referred to as “population distribution” hereinafter.

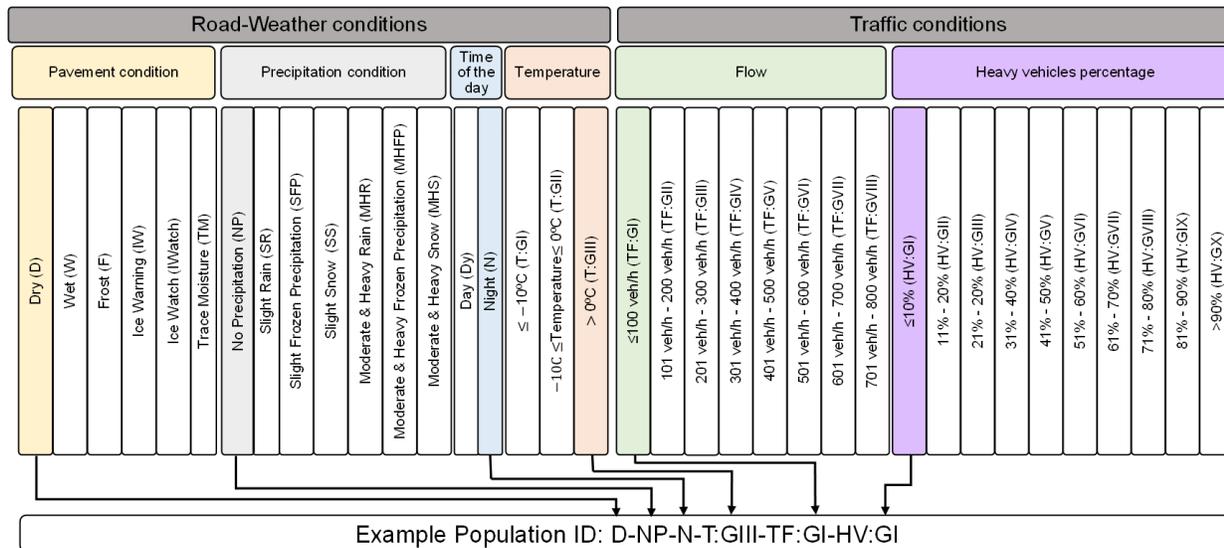


Figure 2.5: Categories of road-weather and traffic conditions to identify populations (Yasanthi et al., 2021)

The study period is divided into consecutive five-minute intervals analogous to samples drawn from the speed population in equal intervals. Each five-minute interval belongs to a specific population labeled with a proper population ID based on the prevailing road-weather and traffic conditions. The individual speeds in each time interval are considered as “observations” while the respective five-minute intervals are considered as “samples” drawn from a population. Accordingly, the number of vehicles observed in a particular five-minute interval is referred to as the “sample size” denoted by n ($n \in \mathbb{Z}^+$) i.e., the five-minute traffic volume. The number of unique sample sizes that belong to a particular population is denoted by k ($k \in \mathbb{Z}^+$). The average speed of all vehicles observed during a five-minute interval represents the sample mean and is referred to as five-minute aggregate speed. For a particular population, the sampling distribution of mean speed refers to the distribution of the sample mean speeds with a specific sample size n . Thus, each speed population is represented by k sampling distributions.

For demonstration purpose, Figure 2.6 presents a graphical illustration of a speed population of 202 individual vehicle speeds recorded under a unique combination of road-weather and traffic conditions, enclosed in 38 five-minute intervals belonging to three unique sample sizes ($k=3$). Each five-minute aggregate interval in the population is assigned with a unique sample ID denoted by $S_{n,m}$, where n denotes the sample size and m ($1 \leq m \leq M$) denotes the sample number pertaining to the sample size n . The number of five-minute intervals recorded with identical five-minute

volumes is denoted by M ($M \in \mathbb{Z}^+$). The five-minute aggregate speed in each five-minute interval represents the sample mean speed and is denoted by $\bar{v}_{n,m}$. The five-minute aggregate intervals observed with identical number of vehicles are clustered together to produce sampling distributions of mean speed. Accordingly, V_n represents the sampling distribution of five-minute aggregate speeds pertaining to intervals with traffic volume n . Five-minute intervals with $M=1$ were removed (to prevent population being represented by a sample distribution) and the rest of the samples are sorted in an ascending order based on n .

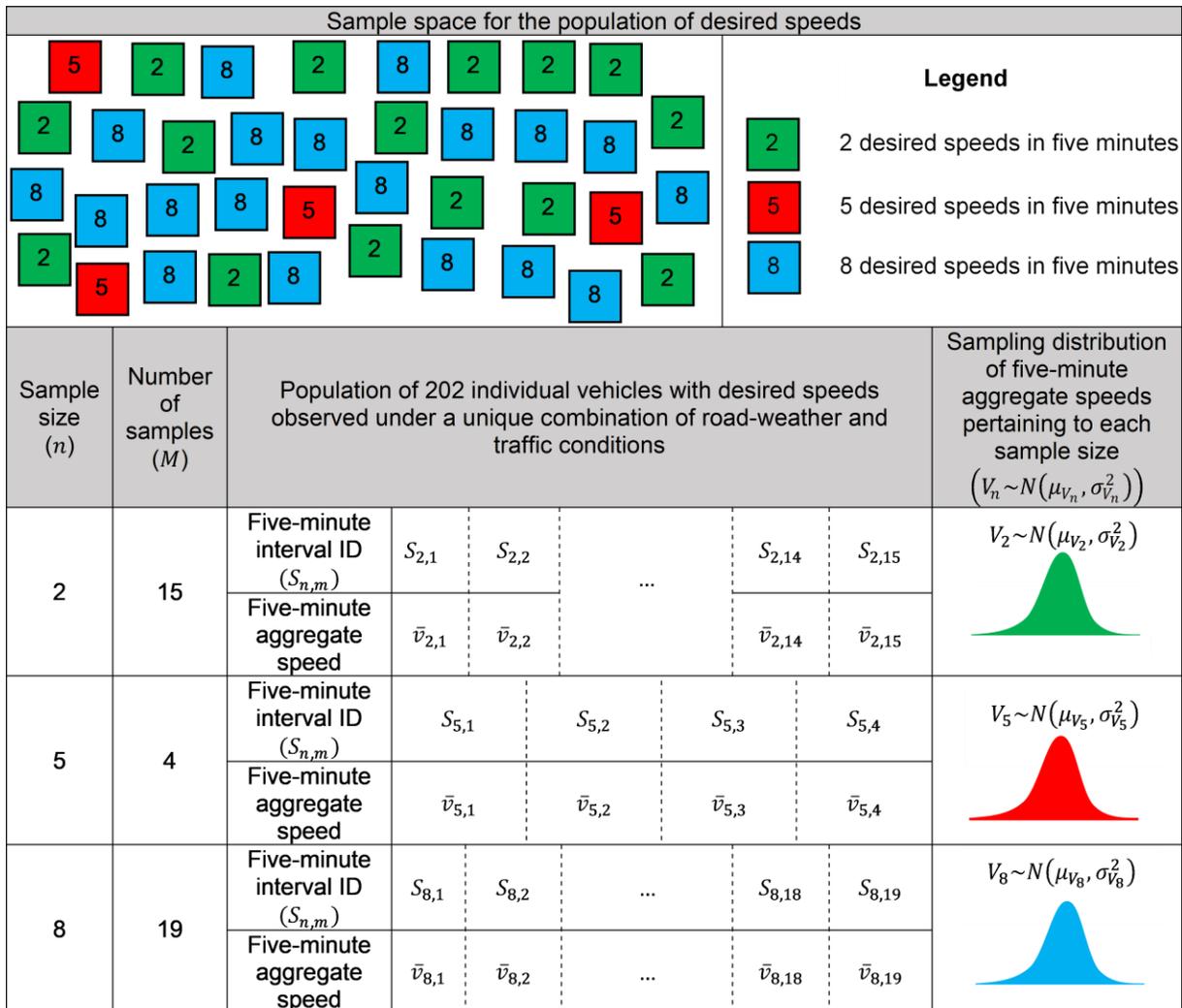


Figure 2.6: Illustration of a population and samples (Yasanthi et al., 2021)

Figure 2.7 shows an example of a speed population (ID: D-NP-N-T:GIII-TF:GI-HV:GI) i.e. dry pavement, no precipitation, temperature more than 0°C, $q < 100$ veh/h and a HV < 10% in nighttime) represented by 8 different sampling distributions of speeds ($k=8$) along with the respective number of five-minute intervals (M). For instance, the fourth sampling distribution ($n=4, M=267$) in Figure 2.7 corresponds to the distribution of the five-minute aggregate speeds encompassing 267 five-minute intervals with four vehicles observed in each five-minutes. The frequency of samples for each sample size n tends to decrease as the sample size increases implying comparatively low traffic volumes observed at the study site, which is in an uncongested rural highway.

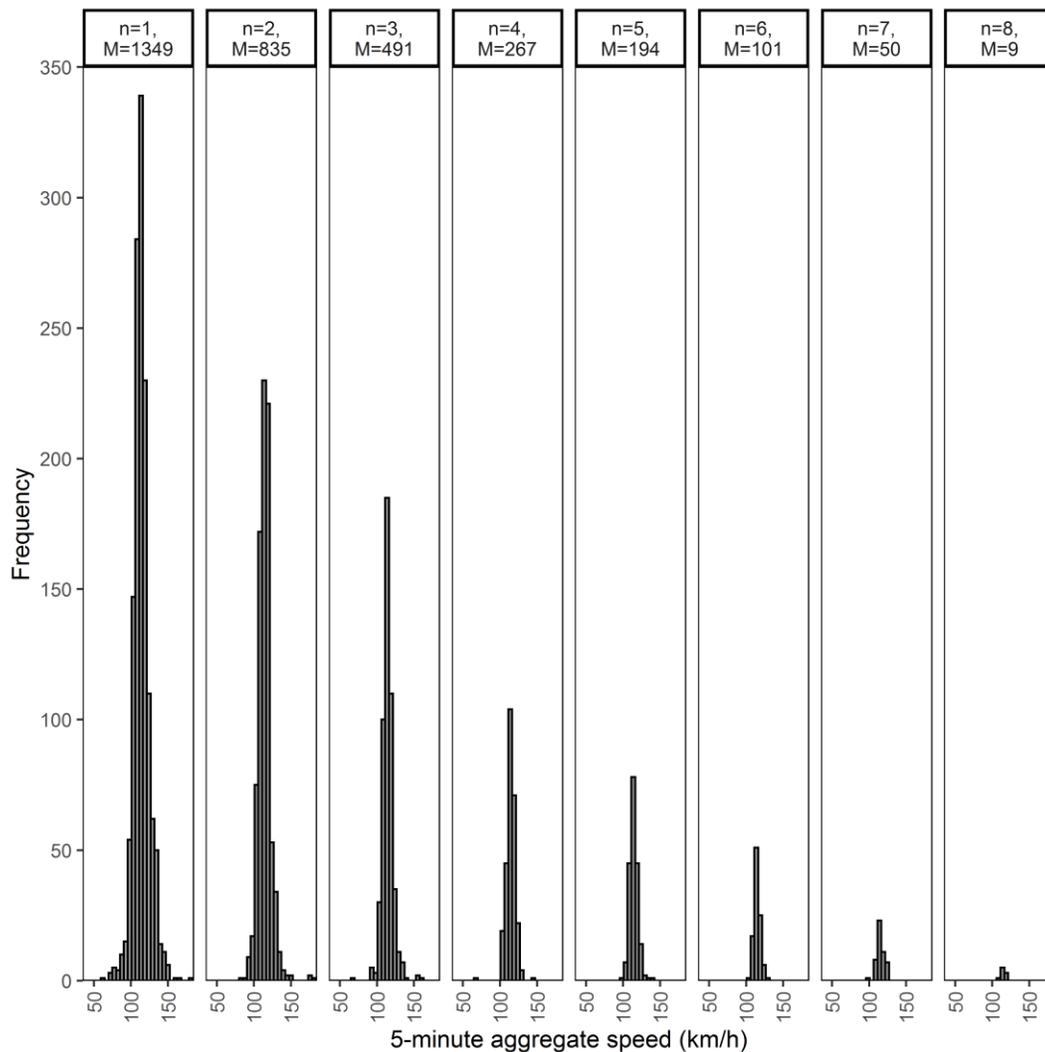


Figure 2.7: Sampling distribution of mean speeds for a population represented by eight sampling distributions (Population ID: D-NP-N-T:GIII-TF:GI-HV:GI) (Yasanthi et al., 2021)

2.6 Methodology

The conventional approaches of evaluating the impacts of RWCs on drivers' speed choice (e.g., regression modelling) attempt to evaluate the impacts of different RWCs on speed behaviour by inferring population parameters using sample-level statistics (e.g., using descriptive statistics to express the impacts of RWCs on speed behaviour). Using such sample-level-based statistics to infer population parameters is questionable particularly when some RWCs may be under-represented in collected study data indicated by considerably low sample sizes in some RWCs. In such cases, estimating population-level parameters by approximating a sample-level distribution to its population-level distribution is questionable. In contrast, according to the Central Limit Theorem (CLT), for normally distributed populations, a sampling distribution of means is highly representative of its population distribution irrespective of its sample size (Hayter, 2012) unlike a sample distribution which may or may not be representative of its population. Accordingly, this research attempts to estimate a population-level speed distribution's parameters based on the mean and standard deviation of a sampling distribution (i.e., a five-minute aggregate speed distribution). In other words, the speed distribution modelling approach proposed in this research takes inputs — samples (i.e., speeds observed in a five-minute interval) — to produce sampling distributions of means (i.e., five-minute aggregate speed distributions) which are thus used to produce the ultimate output of the speed distribution modelling process — population-level speed distribution for a unique combination of road-weather and traffic conditions.

The proposed methodology to estimate the mean and standard deviation of the distribution of desired speeds of individual vehicles is founded upon two main assumptions; (i) observed speeds on the study segment represent drivers' desired speed, and (ii) the population of observed individual vehicle speeds belonging to a unique combination of traffic and road-weather conditions is normally distributed and is represented by a mean of μ and a standard deviation of σ . Nevertheless, the study data were carefully checked to scrutinize the contrarities between the theoretical postulations and the empirical indications. To implement the proposed methodology, the observed speeds should be first affirmed as desired speeds, characterized as the speeds of lead vehicles i.e., non-followers or observed speed of vehicles with minimal interactions (Catbagan & Nakamura, 2008). For congested highways with frequent vehicle interactions, a “follower/non-follower” identification algorithm such as the composite headway distribution model

(Hoogendoorn, 2005) can be applied to identify “follower” and “non-follower” vehicles. Consequently, the analysis can be conducted using the desired speed observations from non-follower vehicles. Alternatively, uncongested traffic streams operating at Level of Service (LOS) “A” or “B” with traffic flow values less than 1,000 veh/h/lane can be reasonably considered to represent “free-flow” conditions (TRB, 2016), where drivers are free to adopt their desired speeds. The shoulder lane of the highway segment in this research mainly operates at LOS “A” and “B” with traffic flow values less than 1,000 veh/h/lane, implying minimal interactions between vehicles (Figure 2.2). Thus, observed speeds are assumed to represent desired speeds of drivers due to prevailing free-flow conditions at the study site.

On the other hand, normal distribution is often used to represent desired speeds (Taylor & Bonsall, 2017) except for deviations under certain situations such as special site characteristics resulting from road geometry (García-Jiménez et al., 2016) and highly heterogeneous traffic such as traffic compositions with very high HV percentages (Hashim, 2011). According to Central Limit Theorem (CLT), V_n (sampling distribution of mean speeds) will be nearly normal regardless of the sample sizes given that the speed populations are normally distributed (Mendenhall & Sincich, 2016). Accordingly, the characteristics of the study data substantiate the main assumptions of the research.

2.6.1 Mean of the of desired speeds’ distribution (μ) of individual vehicles’ population

In this research, it is hypothesized that the distribution of desired speeds under a specific combination of road-weather and traffic conditions is represented by a random variable Y following a normal distribution with a mean of $\mu_{\bar{y}}$ and a standard deviation of $\sigma_{\bar{y}}$. As indicated by CLT, for a normally distributed population, the population mean is equal to the mean of the sampling distribution even for small sample sizes. However, as each speed population is represented by k sampling distributions, a methodology is proposed to estimate the mean of the speed population by combining k sampling distributions (k denotes the number of different distributions of five-minute aggregate speeds with specific traffic volumes (sample sizes), comprising a population). In other words, $Y \sim N(\mu_{\bar{y}}, \sigma_{\bar{y}}^2)$ is analogous to the ultimate result of the linear combination of k independent normal random variables, i.e. k number of sampling distributions of five-minute aggregate speeds, where the i^{th} independent normal random variable

and the associated weight factor are denoted by $V_i \sim N(\mu_i, \sigma_i^2)$ and a_i respectively, where $i \in \mathbb{Z}^+ \mid i \leq k$ (Hayter, 2012).

Theorem: Linear combination of k distributions of five-minute aggregate speeds

If $V_i \sim N(\mu_i, \sigma_i^2)$, $1 \leq i \leq k$, are independent random variables and if a_i , $1 \leq i \leq k$ are constants, then

$$Y = a_1V_1 + \dots + a_iV_i + \dots + a_kV_k \sim N(\mu_{\bar{y}}, \sigma_{\bar{y}}^2) \quad \text{Eq. 2.1}$$

where,

$$\mu_{\bar{y}} = a_1\mu_1 + \dots + a_i\mu_i + \dots + a_k\mu_k \quad \text{Eq. 2.2}$$

and

$$\sigma_{\bar{y}}^2 = a_1^2\sigma_1^2 + \dots + a_i^2\sigma_i^2 + \dots + a_k^2\sigma_k^2 \quad \text{Eq. 2.3}$$

The mean of the distribution of individual vehicle speeds observed under a unique combination of road-weather and traffic conditions (analogous to a normally distributed population $\sim N(\mu, \sigma^2)$), is inferred by $\mu_{\bar{y}}$ i.e., the weighted combination of the mean of the distribution of k different five-minute aggregate speeds (sampling distributions) representing the population. Nevertheless, estimating $\mu_{\bar{y}}$ evokes two challenges.

First, the aforementioned Theorem only holds true for mutually independent random variables. In the context of this research, the existence of vehicle interactions violates the condition of the observations being mutually independent random variables. Therefore, the mutual independence of the observations must be confirmed prior to applying the linear combination of the distribution of five-minute aggregate speeds. As explained earlier, this research investigates the speed behavior observed in a rural highway, which mostly encounters traffic flow volumes less than 1,000 passenger cars per hour per lane (Figure 2.2). Hence, it is reasonable to assume negligible vehicle interactions in this analysis. In mathematical terms, recorded speed observations in five-minute intervals can be considered as independent random variables. Thus, k distributions of five-minute aggregate speeds are mutually independent.

Second, estimation of $\mu_{\bar{y}}$ also requires the estimation of the weight factors $a_{i=1 \text{ to } k}$ (Eq. 2.2). This research proposes to adopt the “Minimum Variance Unbiased Estimation” (MVUE) technique to estimate the mean of the weighted combination of the distributions of five-minute aggregate speeds ($\mu_{\bar{y}}$) and the respective weight factors ($a_{i=1 \text{ to } k}$) pertaining to Y bearing the minimum variance ($\sigma_{\bar{y}}^2$). MVUE represents the unbiased point estimate with a minimum variance of all the possible unbiased point estimates for a particular parameter (Hayter, 2012). Therefore, the MVUE of $\mu_{\bar{y}}$ is represented by the mean of a unique distribution i.e., derived by linearly combining k distributions of five-minute aggregate speeds, possessing the minimum variance among all other linear combinations. Accordingly, the precise population mean is epitomized by the MVUE of the population mean. Yet, the widely used methods of estimating the MVUEs such as the Cramér–Rao bound (CRB) directly estimates the MVUE without estimating the weight factors (Kay, 1993). CRB estimates the theoretical lower bound for the variance of the unbiased estimator and a particular estimate is recognized efficient if the CRB is met by the estimate (Kay, 1993). Nevertheless, MVUE, which matches the lower bound proposed by the CRB may not exist in certain cases (Kay, 1993). Therefore, a more efficient and practical alternative approach is proposed to estimate the MVUE and the corresponding weight factors for population mean. The precise value of the MVUE of $\mu_{\bar{y}}$ can be obtained by minimizing the variance of the weighted combination of distributions of five-minute aggregate speeds $\sigma_{\bar{y}}^2$ as explained below.

2.6.1.1 Estimating the Minimum Variance Unbiased Estimate (MVUE) of $\mu_{\bar{y}}$

In the proposed methodology, a population of individual vehicles with the desired speeds observed under a unique combination of road-weather and traffic conditions is represented by k sampling distributions of five-minute aggregate speeds, and Y is defined in terms of its mean $\mu_{\bar{y}}$ and variance $\sigma_{\bar{y}}^2$ as in Eqs. 2.1 through 2.3, respectively. The weight factors $a_{i=1 \text{ to } k}$, represent the contribution of the individual sampling distribution of five-minute aggregate speeds V_i in the resulting population distribution for Y . Therefore, the weight factors sum up to 1.

$$\sum_{i=1}^k a_i = 1 \quad \text{Eq. 2.4}$$

A statistical approach to estimate $\mu_{\bar{y}}$ involves solving Eqs. 2.2, 2.3 and 2.4 while minimizing $\sigma_{\bar{y}}^2$ which eventually yields the MVUE of $\mu_{\bar{y}}$. Nevertheless, the aforementioned system of equations consists of $k+2$ degree of freedom including $\mu_{\bar{y}}$, $\sigma_{\bar{y}}^2$ and $a_{i;i=1 \text{ to } k}$, with only four constraints (i.e., Eqs. 2.2 through 2.4 and variance minimization equation) expressing the relationship between the aforementioned variables. Accordingly, the system of equations becomes indeterminate for $k > 2$. However, the system of equations can be still resolved for $k = 2$ by treating the cluster of sampling distributions of five-minute aggregate speeds as a chain of the products of two distributions at a time. Figure 2.8 graphically illustrates the sequential combination procedure of k sampling distributions of five-minute aggregate speeds, where all k distributions represent an identical population. First, k distributions are sorted ascendingly in terms of the five-minute traffic volume (sample size). Subsequently, the first two out of k distributions; V_1 and V_2 as depicted in Figure 2.8, are combined linearly yielding an intermediate hypothetical distribution of five-minute aggregate speeds i.e., $Y'_{1,2}$. Afterwards, $Y'_{1,2}$ is chained with the third distribution of five-minute aggregate speeds V_3 , yielding the subsequent intermediate hypothetical distribution of five-minute aggregate speeds $Y'_{1,2,3}$. The chaining process is continued until all k individual sampling distributions of five-minute aggregate speeds representing the population of individual vehicle speeds are encompassed in the intermediate hypothetical distributions, eventually producing $Y'_{1,2,3,4,\dots,i,\dots,k}$ bearing the MVUE of $\mu_{\bar{y}}$. Accordingly, an algorithm is proposed to estimate the weight factors $a_{i=1 \text{ to } k}$ and the MVUE of $\mu_{\bar{y}}$ which is equivalent to μ (Figure 2.9).

For a given combination or road-weather and traffic conditions, the algorithm presented in Figure 2.9 initially combines the first two individual sampling distributions of five-minute aggregate speeds; V_1 and V_2 according to Eq. 1, 2 and 3, such that the linear combination of $V_1 \sim (\mu_1, \sigma_1^2)$ and $V_2 \sim (\mu_2, \sigma_2^2)$ yields a normally distributed hypothetical distribution of five-minute aggregate speeds; $Y'_{1,2} \sim (\mu'_{1,2}, (\sigma'_{1,2})^2)$ possessing the MVUE of $\mu'_{1,2}$, which corresponds to the most representative unbiased estimate of the mean of $Y'_{1,2}$. In other words, $Y'_{1,2}$ holds the minimum variance among all distributions representing the linear combination of V_1 and V_2 . The proposed methodology for estimating the exact value of the MVUE of $\mu'_{1,2}$ is presented below.

The linear combination of V_1 and V_2 corresponds to the bivariate linear combination of two (i.e., $k=2$) sampling distributions of five-minute aggregate speeds. The linear combination ultimately

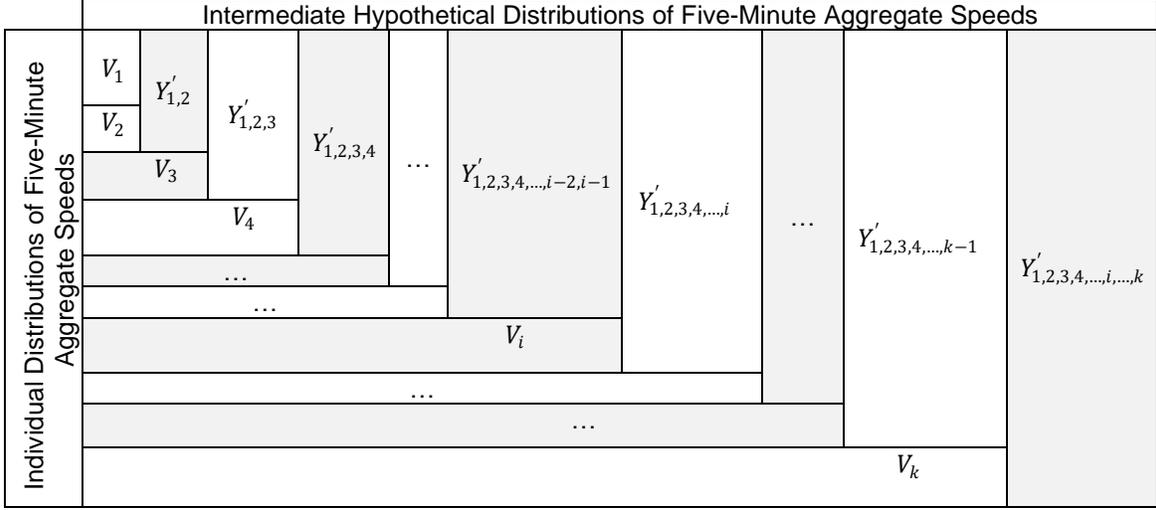


Figure 2.8: Illustration of the chaining process of k sampling distributions of five-minute aggregate speeds (Yasanthi et al., 2021)

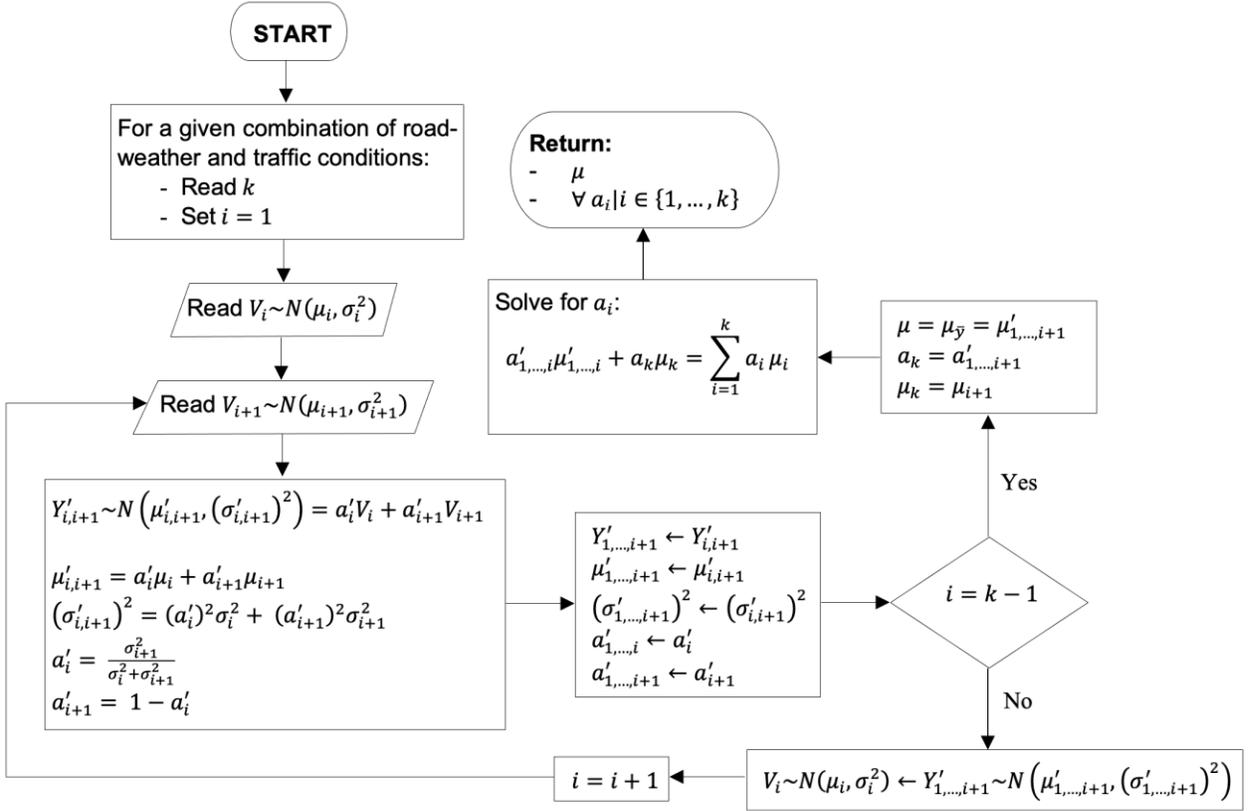


Figure 2.9: Algorithm for estimation of weight factors and μ (Yasanthi et al., 2021)

produces $Y'_{1,2}$ (Eq. 2.1 with $k=2$) which is a normally distributed intermediate random variable characterized by its mean $\mu'_{1,2}$ and variance $(\sigma'_{1,2})^2$ estimated according to Eq. 2.2 and 2.3 respectively with $k=2$. The values of $\mu'_{1,2}$, a'_1 , a'_2 and $(\sigma'_{1,2})^2$ are estimated as explained below, while the coefficients in Eq. 2.2 and 2.3 i.e., μ_1, σ_1^2, μ_2 and σ_2^2 (mean and variance of sampling distributions) are known for a specific population.

The weight factors a'_1 and a'_2 represent the respective contribution of V_1 and V_2 in $Y'_{1,2}$. Therefore, the two weight factors a'_1 and a'_2 sum up to 1 (Eq. 2.4). Accordingly, the variance of $Y'_{1,2}$ is expressed by:

$$(\sigma'_{1,2})^2 = (a'_1)^2 \sigma_1^2 + (1 - a'_1)^2 \sigma_2^2 \quad \text{Eq. 2.5}$$

The weight factors a'_1 and a'_2 are estimated by setting the first derivative of $(\sigma'_{1,2})^2$ with respect to a'_1 (i. e. $d(\sigma'_{1,2})^2/da'_1$) to zero to obtain the minimum variance of $Y'_{1,2}$ yielding,

$$a'_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \quad \text{Eq. 2.6}$$

Substituting a'_1 in Eq. 2.4 ($k=2$) yields:

$$a'_2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \quad \text{Eq. 2.7}$$

Substituting a'_1 and a'_2 in Eq. 2.1, Eq. 2.2 and Eq. 2.3 ($k=2$) yields:

$$Y'_{1,2} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} V_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} V_2 \quad \text{Eq. 2.8}$$

$$\mu'_{1,2} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \mu_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \mu_2 \quad \text{Eq. 2.9}$$

$$(\sigma'_{1,2})^2 = \left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \right)^2 \sigma_1^2 + \left(\frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \right)^2 \sigma_2^2 \quad \text{Eq. 2.10}$$

$Y'_{1,2}$ is then chained with the next sampling distribution of five-minute aggregate speeds; V_3 and a new intermediate hypothetical normally distributed distribution of five-minute aggregate speeds $Y'_{1,2,3}$ is produced. This process is continued until all k original individual distributions of five-minute aggregate speeds are integrated in the chain of the resulting hypothetical intermediate distributions. Consequently, the ultimate hypothetical distribution of five-minute aggregate speeds; $Y'_{1,2,3,4,\dots,i,\dots,k}$ conveys the characteristics of the population of five-minute aggregate speeds representing the distribution of individual desired speeds observed under a unique combination of road-weather and traffic conditions. Therefore, $Y'_{1,2,3,4,\dots,i,\dots,k}$ is equivalent to Y as expressed in Eq. 2.1 and referred to as Y hereinafter. Moreover, Y possesses the minimum variance ($\sigma_{\bar{y}}^2$) among all combinations of k sampling distributions of five-minute aggregate speeds leading Y to be the most representative representation of the distribution of individual speeds under a particular combination of road-weather and traffic conditions (a population). Consequently, the mean of Y i.e., $\mu_{\bar{y}}$ is concluded as the MVUE for the population mean, μ .

Estimating the MVUE of $\mu_{\bar{y}}$ according to the proposed algorithm for a population represented by three sampling distributions ($k=3$) is elaborated in the supporting information (Appendix B).

It should be noted that the process of determining weight factors proposed in this research ensures that each weight factor is directly proportional to the variance of a five-minute aggregate speed distribution. In other words, the weight factor estimation process proposed in this research attempts to optimize the representativeness of the combined five-minute aggregate speed distribution (Y) to its population distribution. Accordingly, the MVUE of μ and the weight factors $a_{i=1 \text{ to } k}$ can be considered as performance measures of the speed distribution modelling approached proposed in this thesis.

2.6.1.2. Standard deviation of the desired speeds' distribution (σ) of individual vehicles' population

As explained before, the linear combination of k sampling distributions of five-minute aggregate speeds yields to Y (Eq. 2.1), which is characterized by a variance of $\sigma_{\bar{y}}^2$ (Eq. 2.3). On the other hand, σ_i , denotes the standard deviation of the i^{th} sampling distribution of five-minute aggregate speeds (V_i) pertaining to a specific sample size (five-minute traffic volume n_i). V_i is one out of k sampling distributions representing the population of desired speeds of individual vehicles

observed under a unique combination of road-weather and traffic conditions with a standard deviation σ . Eq. 2.11 defines the relationship between the standard deviation of the population (σ) and that of i^{th} sampling distribution (σ_i):

$$\sigma_i = \sigma / \sqrt{n_i} \quad \text{Eq. 2.11}$$

Substituting σ_i from Eq. 2.11 to Eq. 2.3 yields,

$$\sigma_{\bar{y}}^2 = a_1^2 \left(\frac{\sigma^2}{n_1} \right) + \dots + a_i^2 \left(\frac{\sigma^2}{n_i} \right) + \dots + a_k^2 \left(\frac{\sigma^2}{n_k} \right) \quad \text{Eq. 2.12}$$

After rearranging Eq. 2.12, the standard deviation of the distribution of desired speeds of individual vehicles observed under a unique combination of road-weather and traffic conditions, i.e., the population standard deviation can be expressed as:

$$\sigma = \sqrt{\frac{\sigma_{\bar{y}}^2}{\sum_{i=1}^k a_i^2 / n_i}} \quad \text{Eq. 2.13}$$

2.6.2 Identification of road-weather conditions with intensified safety risks

To investigate the impacts of adverse road-weather conditions on the desired speed distribution characteristics, the combinations of road-weather conditions under prevailing traffic conditions, i.e., traffic flow values from 100 veh/h to 200 veh/h and heavy vehicle percentages from 20% and 30% are explored.

Safety risks in this research are evaluated in terms of two important aspects of a potential crash i.e., severity and propensity. Speed and the variability of speeds are often acknowledged as appropriate measures in evaluating crash severity and crash involvement respectively, where high speeds and high standard deviation of speeds are associated with severe injuries occurred at a potential crash and high crash involvement rates respectively (Solomon, 1964). A weather-related crash (hereinafter crash), however, involves the presence of adverse road-weather conditions. Moreover, combination of the most frequent road-weather conditions (dry pavement surface condition, no precipitation, daytime and temperature values between -10°C and 0°C in the context of this research) can be considered as normal road-weather conditions due to drivers' frequent exposure to such road-weather conditions. Therefore, the severity and the propensity of a weather-related crash can be accurately evaluated by assessing the difference in the desired speed

distribution characteristics under normal and adverse road-weather conditions considering prevailing traffic conditions.

The distributions of desired speeds under prevailing traffic conditions encompass two types of road-weather conditions: normal and adverse. The term “adverse” refers to any combination of road-weather conditions under prevailing traffic conditions. In this research, the mean and standard deviation of a desired speed distribution belonging to prevailing traffic and normal road-weather conditions are denoted by μ_{Normal} and σ_{Normal} respectively, where the mean and standard deviation of a desired speed distribution belonging to prevailing traffic and adverse road-weather conditions are denoted respectively by; $\mu_{Adverse}$ and $\sigma_{Adverse}$. In this research, weather-related collision risks are estimated by defining (i) Crash Severity Factor (CSF): proportion between $\mu_{Adverse}$ and μ_{Normal} (Eq. 2.14), and (ii) Crash Exposure Factor (CEF): proportion between $\sigma_{Adverse}$ and σ_{Normal} (Eq. 2.15).

$$\mu_{Adverse} = \alpha_{Adverse} \times \mu_{Normal} \quad \text{Eq. 2.14}$$

$$\sigma_{Adverse} = \beta_{Adverse} \times \sigma_{Normal} \quad \text{Eq. 2.15}$$

where,

$\alpha_{Adverse}$: Crash Severity Factor (CSF)

$\beta_{Adverse}$: Crash Exposure Factor (CEF)

Mean speed is well-acknowledged as an indication of crash severity (Solomon, 1964). Therefore, CSF is considered as a surrogate measure for crash severity under a particular combination of road-weather conditions. Considering mean desired speed under normal road-weather conditions as a reference, road-weather conditions with $CSF \leq 1$ imply lower (or equal) mean desired speed ($\mu_{Adverse} \leq \mu_{Normal}$). Thus, potential crashes under such road-weather conditions are classified as “low severity” as compared to normal road-weather conditions. In contrast, potential crashes under road-weather conditions with $CSF > 1$ are classified as “high severity”, due to higher mean desired speed ($\mu_{Adverse} > \mu_{Normal}$).

Similarly, CEF i.e., defined based on standard deviation of desired speed under a particular combination of road-weather conditions, is considered as a surrogate measure for crash propensity

(Peng et al., 2017). Road weather conditions with $CEF \leq 1$ are classified as “low exposure” due to lower variability of desired speeds ($\sigma_{Adverse} \leq \sigma_{Normal}$). On the contrary, road-weather conditions with $CEF > 1$ are identified as “high exposure” due to high standard deviations of desired speeds as compared to normal road-weather conditions ($\sigma_{Adverse} > \sigma_{Normal}$).

Subsequently, each combination of road-weather conditions with the most frequent traffic conditions are classified into four categories and labelled in terms of crash severity and exposure depending on the values of CSF and CEF as shown in Table 2.3.

Table 2.3: Classification criteria for combinations of road-weather conditions to identify road-weather conditions with intensified safety risks

	$CEF \leq 1$	$CEF > 1$
$CSF \leq 1$	Low severity, Low exposure	Low severity, High exposure
$CSF > 1$	High severity, Low exposure	High severity, High exposure

Conditions labelled as “High severity, High exposure” are identified as the road-weather conditions imposing the highest safety risks followed by conditions labelled “High severity, Low exposure”, “Low severity, High exposure” and “Low severity, Low exposure” in a descending order with respect to safety risks.

2.7 Modeling results

Data collected from the study site encompassed 933 unique combinations of road-weather (precipitation condition, pavement surface condition, time of the day, air temperature) and traffic conditions (traffic flow, heavy vehicles percentage) as shown in Figure 2.5. Accordingly, the desired speed distributions of each of the 933 combinations were modelled as normal distributions characterized by mean (μ) and standard deviation (σ).

To recall, a weight factor (a_i) depending on the variance (σ_i^2) was assigned to each sampling distribution of five-minute aggregate speeds representing a specific desired speed population pertaining to a particular combination of road-weather and traffic conditions. With a few exceptions, the sampling distributions with a larger variance were assigned with smaller weight factors, i.e., contributing less toward estimating the desired speed distribution characteristics (Figure 2.10) that is the direct consequence of implementing MVUE. In fact, the observation of small weight factors estimated for five-minute aggregate speed distributions with a larger variance

(i.e., the proposed performance measure for CLT-based speed distribution modelling) confirms the effectiveness of the speed distribution modelling approach proposed in this research. Thus, the proposed methodology assures that more emphasis is placed on sampling distributions with more stable speed observations (e.g., lower variance due to larger sample size) while preserving and using all observations.

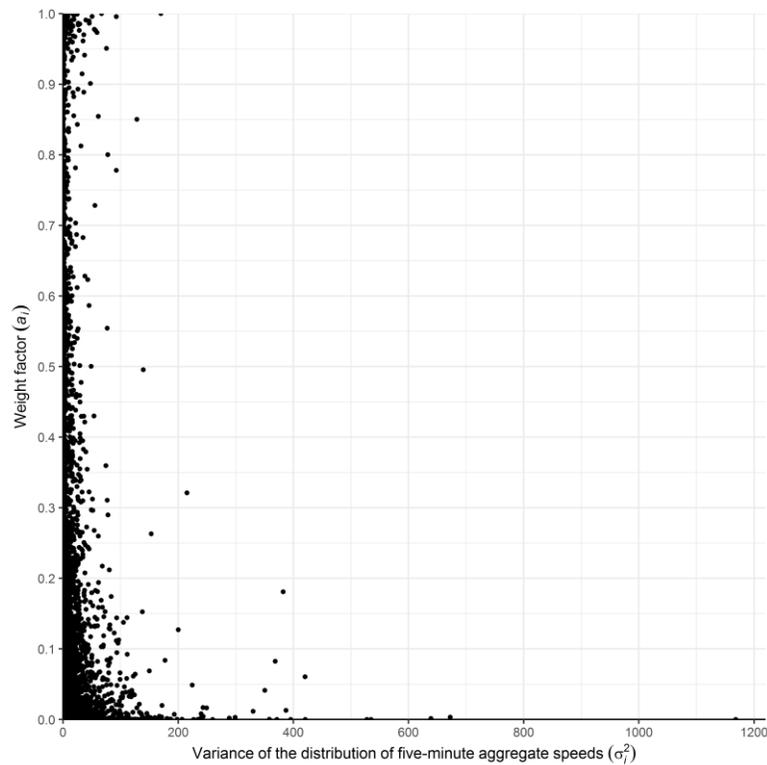


Figure 2.10: Weight factor VS the variance of the distributions of five-minute aggregate speeds (Yasanthi et al., 2021)

Figure 2.11 presents the number of populations recorded at the study site in terms of traffic flow and Heavy Vehicle (HV) percentage conditions. For instance, majority of the populations observed at the study site correspond to normal (i.e., prevailing) traffic conditions: the traffic flow category of “100 veh/h – 200 veh/h” and the heavy vehicle percentage category of “20% - 30%”. In fact, 44 out of the 933 populations (i.e., combinations of RWC and traffic conditions) recorded at the study site corresponds to the prevailing traffic conditions — traffic flow category of “100 veh/h – 200 veh/h” and the heavy vehicle percentage category of “20% - 30%”. The next section presents and discusses the characteristics of desired speed distributions of those 44 populations which correspond to the prevailing traffic conditions.

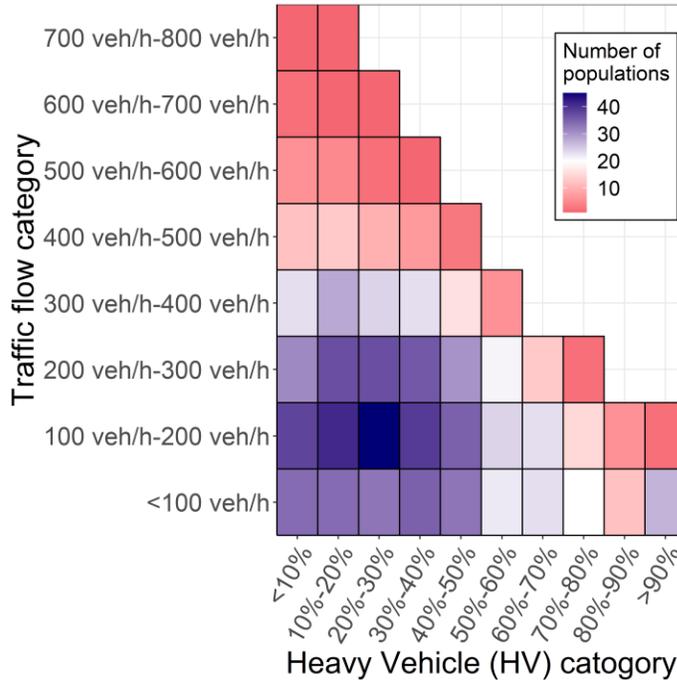


Figure 2.11: Number of populations recorded in terms of the traffic flow and HV categories (Yasanthi et al., 2021)

2.7.1 Impact of road-weather and traffic conditions on estimated desired speed distributions

Figure 2.12 depicts the mean (μ) and standard deviation (σ) of the desired speed distributions under 44 unique combinations of precipitation conditions, pavement surface conditions, time of the day and air temperature categories observed under prevailing traffic conditions. Distribution of desired speeds under the combination of the most frequent road-weather conditions i.e., dry pavement, no precipitation, daytime and temperature values between -10°C and 0°C (Group II) is characterized by a normal distribution with a mean of 112 km/h and a standard deviation of 6.5 km/h (Figure 2.12). Figure 2.12 reveals interesting speed behaviour idiosyncrasies. For instance, the average speed chosen by drivers travelling under ice warning pavement conditions is considerably lower than the average speed chosen by drivers travelling under other road surface conditions (dry, wet, frost, trace moisture and ice watch), irrespective of weather conditions prevailing at the time of travel. A closer inspection of Figure 2.12 further reveals that the desired speed distributions of vehicles travelling under ice warning road surface conditions possess comparatively higher values of standard deviation, signifying the diversified speed choices under ice warning pavement conditions. The lower numerical values for means coupled with higher standard deviation values of the desired speed distributions under ice warning road surface

conditions is intuitive since majority of the drivers are very much attentive towards the perceptible roadway hazards during driving maneuvers. Nevertheless, each individual driver perceives the risk on driving on such perceptible hazards at different levels which is manifested by the high standard deviation of speeds eventually intensifying the crash risks. In general, drivers are well aware of the reduced road surface friction under icy pavement conditions and typically incline towards

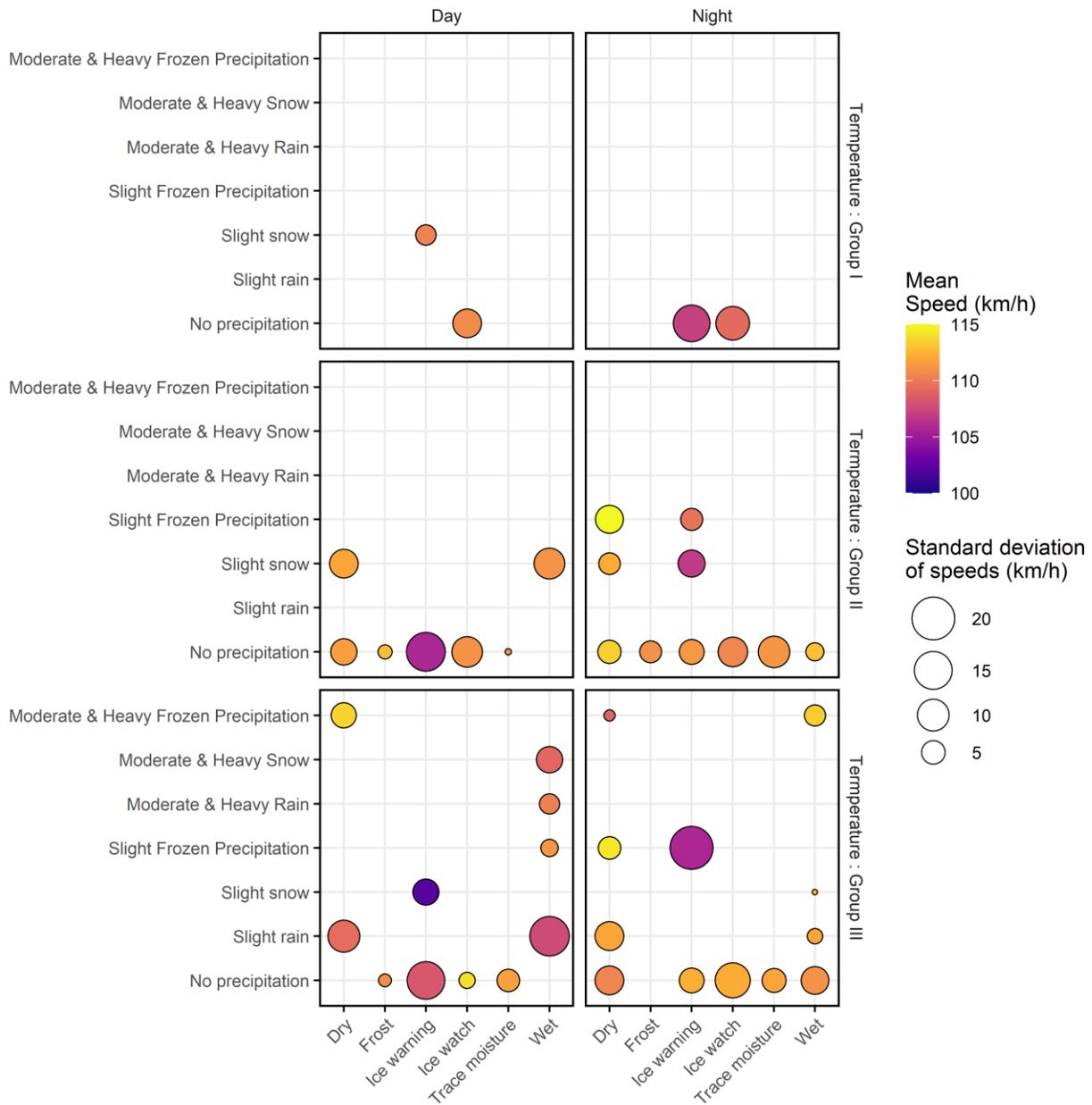


Figure 2.12: Desired speed distribution characteristics; mean (μ) and standard deviation (σ) of road-weather combinations under the prevailing traffic conditions (Yasanthi et al., 2021)

driving at a lower speed compared to the fixed posted speed limit of 110 km/h. The choice of such lower speed is determined by drivers predominantly based on their driving experience, confidence, and comfort levels. Accordingly, some experienced drivers may drive at higher speeds even though they are aware of the deteriorated driving conditions, leading to higher means of the desired speed distributions (e.g., under frost and or ice watch conditions) and high crash severities.

The aforesaid argument is further validated in the presence of acute precipitation. For instance, the desired speed distributions in daytime recorded with a road surface condition of ice warning, temperature values above 0°C (Group III) and slight snow evinced a phenomenally low mean with a numerical value of 102 km/h, which is the lowest mean among the 44 desired speed distributions, and a considerably low standard deviation (6.39 km/h) compared to the remaining 43 combinations of road-weather conditions (Figure 2.12). This observation supports the hypothesis that the majority of drivers are not comfortable driving at or near the speed limit of 110 km/h, which is considerably higher than 102 km/h. Moreover, the low standard deviation of the aforementioned desired speed distribution suggests a comparatively low variability in the desired speeds implying the alike collective judgment of a safe speed among the majority of the drivers. In contrast, the highest variability of speeds prevails under ice warning pavement, slight frozen precipitation, nighttime and temperature values above 0°C (Group III) which is manifested through a standard deviation of 20.07 km/h. Nevertheless, mean of the aforementioned desired speed distribution was estimated as 106 km/h, which is comparatively low with respect to the rest of the desired speed distribution means presented in Figure 2.12. This inconsistency is potentially due to the unexpected slight frozen precipitation occurring at peculiar temperature values, leading the drivers to prefer speeds in a wide range depending on their confidence and comfort of driving at a particular speed, yielding a considerably higher standard deviation of the distribution of desired speeds.

Surprisingly, however, the presence of precipitation alone does not seem to jeopardize drivers' uncertainty towards adopting a safe speed to travel under most precipitation conditions. This is evident from the comparatively higher means of the desired speed distributions under dry pavement conditions even in the presence of frozen precipitation and snowy conditions. In fact, the highest mean desired speed (115 km/h) among the combinations of road-weather conditions considered in Figure 2.12 was under dry pavement, temperature values between 0°C and -10°C (Group II), nighttime and slight frozen precipitation conditions. Nevertheless, the aforementioned

desired speed distribution possesses a comparatively low standard deviation (7.49 km/h) signaling a moderate homogeneity of desired speeds under particular combination of road-weather conditions. On the other hand, the combination of road-weather conditions; wet pavement, slight snow, nighttime and temperature values above 0°C (Group III) recorded the minimum standard deviation of desired speed distributions in Figure 2.12 with a value of 0.39 km/h. Meanwhile, the mean of the same desired speed distribution is surprisingly high with a value of 113 km/h. These rather contradictory results can be attributed to driver experience as a result of frequent exposure to driving under inclement road-weather conditions. As discussed before, the study site is located in an extremely cold region which is subjected to frequent adverse road-weather conditions. Hence, it could conceivably be hypothesized that frequent exposure to hazardous driving conditions could be a major factor in the inferred weak link between the adverse road-weather conditions and the selection of a lower speed as a safe speed to travel, which is manifested by the high mean and the extremely low standard deviation of the aforesaid desired speed distribution.

With successive increases in the intensity of the atmospheric temperature, the mean speed of the desired speed distributions gradually increases in the case of ice warning, no precipitation and nighttime conditions. Interestingly, however, both the minimum and the maximum standard deviation of desired speed distributions under prevailing traffic conditions emerged under identical time of the day and temperature group, which are nighttime and temperature values above 0°C (Group III) respectively. This is rather an important outcome as it reveals that the variability of speeds is particularly affected by the combination of precipitation and road surface condition. Nevertheless, there is no convincing evidence to conclude a notable relationship between the desired speed distribution characteristics and the two road-weather conditions, time of the day and temperature.

In summary, the characteristics of the desired speed distributions suggest that there is a strong association principally with the two road-weather conditions i.e., road surface and precipitation conditions. In particular, the speed considered safe by each driver under identical road-weather conditions seems to be dependent on their personal comfort levels and confidence to travel, specially under perilous pavement conditions which is manifested through divergent values of desired speed distribution means and standard deviations.

2.7.2 Adverse road-weather conditions with intensified road safety risks

To identify the specific road-weather conditions with intensified safety risks, mean and standard deviation of desired speed distribution pertaining to a normal road-weather and normal traffic conditions; μ_{Normal} and σ_{Normal} were first estimated as 112 km/h and 6.5 km/h, respectively. Consequently, CSF and CEF values estimated for each combination of road-weather conditions under prevailing traffic conditions were estimated as presented in Figures 2.13 (a) and 2.13 (b), respectively. The maximum CSF corresponds to the dry pavement, temperature values between 0°C and -10°C (Group II), nighttime and slight frozen precipitation conditions, while the minimum CSF corresponds to ice warning road surface conditions, temperature values above 0°C (Group III), daytime and slight snow conditions (Figure 2.13 (a)). In terms of the CEF, the maximum crash exposure is estimated for combination of ice warning road surface conditions, temperature values above 0°C (Group III), slight frozen precipitation and nighttime. The minimum CEF is estimated for the combination of wet road surface conditions, temperature values above 0°C (Group III), slight snow and nighttime (Figure 2.13 (b)).

Figure 2.14 classifies each combination of road-weather condition observed under prevailing traffic conditions in terms of potential safety risks (estimated based on speed and speed variability) as defined in Table 2.3. According to Figure 2.14, only two combinations of road-weather conditions are classified in the category of extremely high potential safety risks, i.e., the combination of (i) ice watch road surface, temperature values above 0°C (Group III), nighttime and no precipitation conditions, and (ii) dry pavement, temperature values between 0°C and -10°C (Group II), nighttime and slight frozen precipitation conditions. Moreover, 14 out of the 44 combinations of road-weather conditions considered in the analysis were classified in the category with no potential safety risks compared to reference conditions while 10 combinations were classified in the category with high potential risk of severe crashes. Finally, 18 combinations were classified in the category with high potential risk of crash occurrence.

Overall, 68% of road-weather combinations under prevailing traffic conditions are classified to be hazardous in terms of either potential crash severity or potential exposure to crashes, which conveys the vulnerability of traffic safety under adverse road-weather conditions. More specifically, the CSF for four combinations of RWCs were estimated as 1.01 whereas the CSF for five RWC combinations were estimated as 1.02.

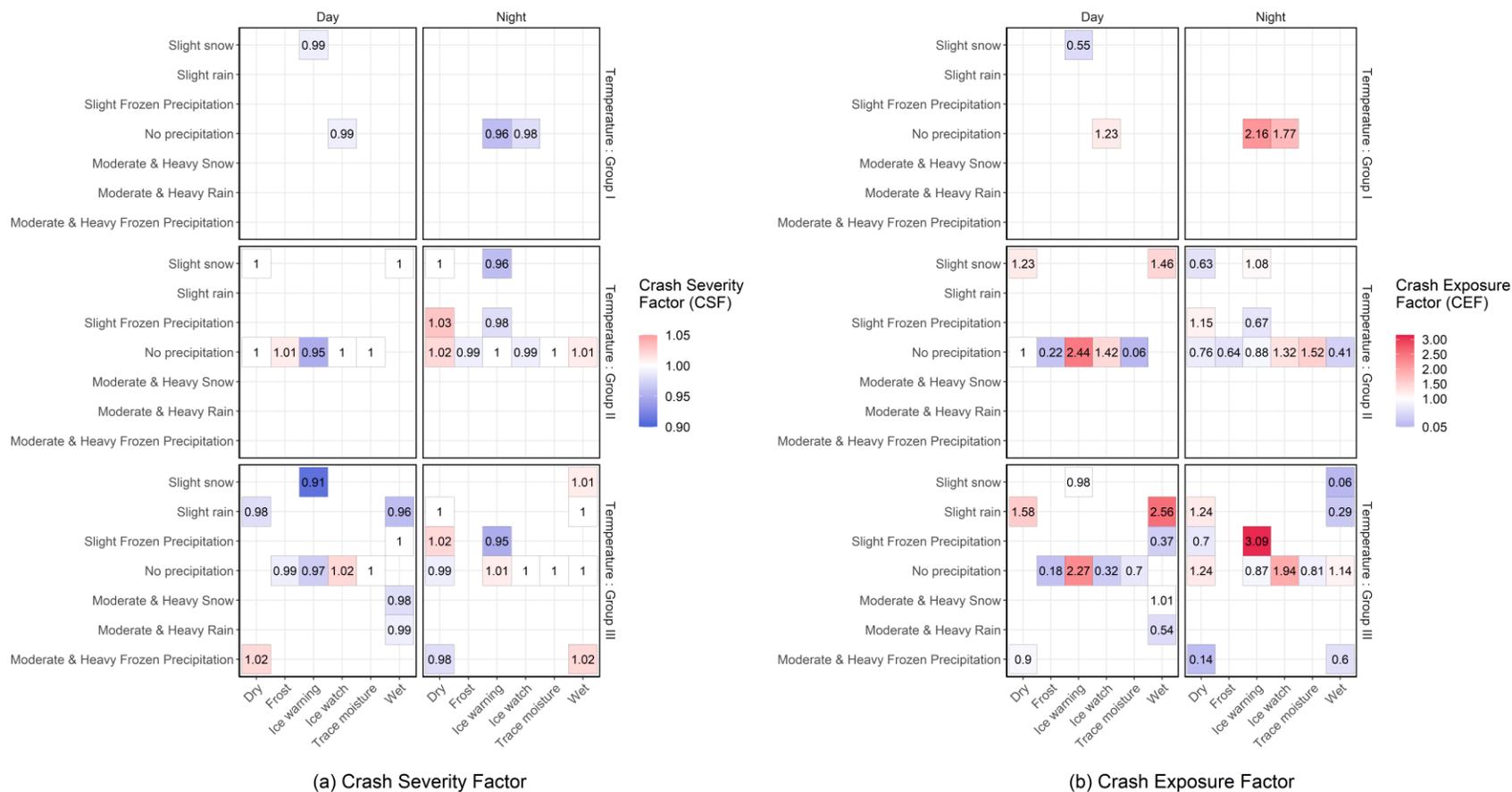


Figure 2.13: Crash Severity and Crash Exposure Factors for each road-weather combination under the prevailing traffic conditions (Yasanthi et al., 2021)

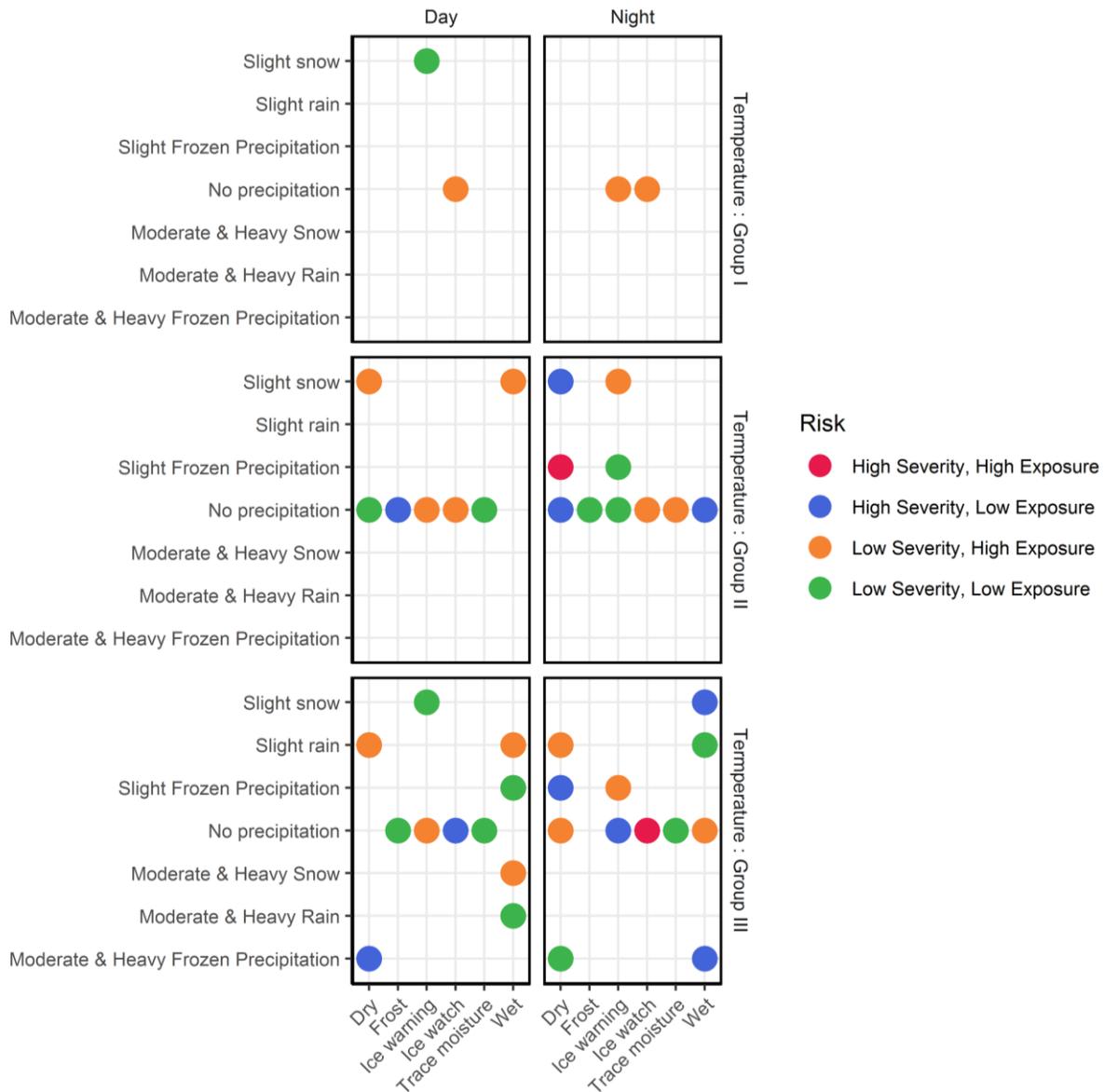


Figure 2.14: Classification of each road-weather combination under the prevailing traffic conditions (Yasanthi et al., 2021)

One combination of RWCs was estimated to have a CSF of 1.03. The number of RWC combinations with specific CEFs are as follows: (i) four RWC combinations for $1 < CEF \leq 1.5$, (ii) five RWC combinations for $1.5 < CEF \leq 2.0$, (iii) two RWC combinations for $2.0 < CEF \leq 2.5$, (iv) one RWC combinations for $2.5 < CEF \leq 3.0$, and (v) one RWC combinations for $3.0 < CEF \leq 3.5$.

2.7.3 Comparative analysis

According to past literature, no previous research has modelled the desired speed distributions under the combination of road-weather and traffic conditions for rural highways in cold regions even though several past studies have estimated the speed variations resulted by adverse road-weather conditions (Yasanthi & Mehran, 2020; Kyte et al., 2000). Nevertheless, the results of this research relating to the most frequent combinations of traffic conditions (i.e., traffic flow values between 100-200 veh/h and heavy vehicle percentages between 20%-30%) are compared with Yasanthi & Mehran (2020) and the HCM (TRB, 2016). Yasanthi & Mehran (2020) used regression modelling to investigate the same study site and data as in the current research which provides consistent basis for comparison.

Therefore, speed reduction factors for vehicles travelling in the shoulder lane under specific combinations of road-weather and traffic conditions estimated by Group II models (linear regression models with 20-minute aggregate speeds as the dependent variables) as suggested by Yasanthi & Mehran (2020) were considered for comparative evaluation. Further, speed reduction factors suggested by the HCM (TRB, 2016) were also considered for evaluation and comparison. It should be noted that numerical values for the temperature, heavy vehicle percentage and traffic flow are used to allow better comparison as well as to represent a practical application of the populations. For instance, a numerical value of 5°C is used for the comparison as compared to using the temperature group of “GIII” (Table 2.1) which enables direct estimation of the speed reduction suggested by Yasanthi & Mehran (2020) and the HCM (TRB, 2016).

Table 2.4 presents speed reduction factors (comparing to reference normal road-weather conditions) under four combinations of road-weather conditions. To allow better comparison, only slight and heavy snow are considered. The results show that desired speed distributions under different road-weather conditions can be characterized by considerably distinctive means and standard deviations which implies that the speed behaviour of drivers travelling in rural divided highways under adverse road-weather conditions is affected by pertaining road-weather conditions, which is consistent with Yasanthi & Mehran (2020) and the HCM (TRB, 2016) observations. Both the present research and Yasanthi & Mehran (2020) revealed that the choice of desired speed is predominantly influenced by precipitation and perceptible road surface conditions at the time of travel.

Table 2.4: Comparison of results with past literature

		Combination of Road-weather conditions				
		Combination I	Combination II	Combination III	Combination IV	
Road-weather condition	Pavement condition	Ice warning	Wet	Wet	Ice warning	
	Precipitation Intensity (mm/h)	1	1	9	1	
	Temperature (°C) ¹	5	5	5	-15	
	Time of the day	Daytime	Nighttime	Daytime	Daytime	
Traffic condition¹	Traffic flow (veh/h)	150 veh/h	150 veh/h	150 veh/h	150 veh/h	
	Heavy vehicles percentage (%)	20%	20%	20%	20%	
Desire speed distribution characteristics	Mean (Km/h)	102	113	109	110	
	Standard deviation (Km/h)	6.39	0.39	6.53	3.58	
Speed reduction factors²	Present research ³	8.93%	-0.89%	2.68%	1.79%	
	Yasanthi & Mehran (2020)	Light vehicles ⁴	2.15%	2.16%	3.79%	3.73%
		Heavy vehicles ⁵	4.64%	3.47%	1.67%	6.03%
	HCM (2016) ⁶		13%	13%	16%	13%

Notes:

¹ Numerical values used to allow better comparison of the existing research's results with [9] and [11].

² Speed reduction factor = $\frac{(\text{Speed under normal road-weather}) - (\text{Speed under adverse road-weather})}{\text{Speed under normal road-weather}} \times 100$

³ Speed under normal road-weather conditions: 112 km/h

⁴ Speed under normal road-weather conditions: 117 km/h

⁵ Speed under normal road-weather conditions: 116 km/h

⁶ Speed reduction factors based on HCM (2016) only based on the precipitation condition and a base free-flow

First, it should be noted that the speed reduction factors (Table 2.4) estimated by Yasanthi & Mehran (2020), HCM (TRB, 2016) and the current research are founded upon fundamentally different approaches, which in turn lead to different speed reduction factors as anticipated. For instance, Yasanthi & Mehran (2020) inferred free-flow speed reductions under different road-weather conditions based on linear regression models while HCM (TRB, 2016) proposed free-flow speed reductions under different precipitation conditions based on stepwise regression models developed by Rakha et al. (2008). Yet, as explained earlier, representing speed behaviour under adverse road-weather conditions through regression models is challenging especially considering restricted sampling conditions occurring under adverse road-weather conditions. For example,

each individual 20-minute aggregate speed used in the “Group II” models in Yasanthi & Mehran (2020) equally contribute to estimating the regression coefficients irrespective of the number of vehicles (sample size) observed in the respective 20-minute intervals. Therefore, the suitability of a sample-level analysis such as regression modelling to study the impacts of adverse road-weather conditions on desired speed is questionable. In contrast, the current research results are based on a population-level analysis that differentiates the contribution of each five-minute aggregate speed observed in a particular combination of road-weather conditions upon the variance of the distributions of five-minute aggregate speeds representing a specific population. Yasanthi & Mehran (2020) concluded one of the linear regression models produced in their research as the best performing model to represent the relationship between free-flow speed and road-weather conditions, the dependent variable of the regression models (free-flow speed) only provides an average estimation of the free-flow speed under specific adverse road-weather conditions unlike the current research, which estimates the distribution of all possible desired speeds under specific road-weather conditions. Further, the regression models developed by Yasanthi & Mehran (2020) include some statistically insignificant regression coefficients pertaining to certain road-weather conditions implying the absence of a linear relationship between such road-weather conditions and desired speed. In contrast, the present research only uses four road-weather conditions with statistically different levels (in terms of the difference of speeds) for each road-weather condition (Table 2.1). Moreover, it should be noted that the speed reduction factors estimated according to Yasanthi & Mehran (2020), HCM (TRB, 2016) and the present research (Table 2.4) respectively consider six, one and four road-weather conditions leading the speed reductions estimated by the three studies to spread among a different number of factors.

The speed reduction factors suggested in the HCM (TRB, 2016) are based on road-weather data collected from an automated surface observing system (ASOS) located in nearby airports as compared to the alongside data collection devices used in both the present research and Yasanthi & Mehran (2020). In fact, Rakha et al. (2008), the underlying source for the speed reduction factors proposed in the HCM (TRB, 2016), highlighted the issue of obtaining highly representative microclimate data to represent the prevailing road-weather conditions at the traffic counters. Besides, the HCM (TRB, 2016) estimates correspond to a different geographical region possessing unique characteristics of the driver population, resulting considerably divergent speed reduction factors compared to the present research.

2.8 Conclusion and future directions

This research proposed a novel approach based on Central Limit Theorem to model desired speed distributions (mean (μ) and standard deviation (σ)) of vehicles travelling in rural divided highways under different combinations of road-weather and traffic conditions, followed by identification of the combinations of road-weather conditions imposing considerable safety risks under prevailing traffic conditions. Often, the impacts of adverse road-weather conditions on the speed behaviour are evaluated largely by enumerating the absolute speed reductions under adverse-road weather conditions estimated largely through regression analysis at a sample-level. Speed behaviour, however, reflects the intrinsically divergent driver psychology. This research, therefore, proposed a robust methodology to model the desired speed distributions at a population-level, which can in turn be used to identify road-weather and traffic conditions with potential safety risks. The theoretical models are calibrated with road-weather and traffic data collected from a study site in Alberta, Canada where a fixed speed limit of 110 km/h is implemented irrespective of the prevailing road-weather. Yet, the proposed methodology is adoptable for uncongested divided highways in other geographical locations with similar road-weather conditions.

The outcomes of the research highlight the importance of paying special attention to traffic safety under the combination of precipitation and atypical road surface conditions while the impacts of temperature and time of day were deemed negligible. This research identified two specific combinations of road-weather conditions potentially imposing higher crash severity and involvement risks: (i) ice watch pavements, temperature values of above 0°C, nighttime and no precipitation conditions and (ii) dry pavements, temperature values between 0°C and -10°C, nighttime and slight frozen precipitation conditions.

An extensive literature review conducted on the research topic revealed that this research is the first comprehensive attempt in proposing a coherent methodology to model the desired speed distributions under different road-weather and traffic conditions for rural divided highways located in cold regions. The research's contributions are twofold. First, it evaluates the impacts of different road-weather and traffic conditions on the desired speed distributions in uncongested rural highways while identifying road-weather conditions with potentially higher safety risks. Second, a methodological contribution is made by proposing an innovative approach to model desired speed distributions which can be used under different sampling rates and conditions. Thus, the

research findings contribute toward understanding the divergent speed behaviour which is often critiqued as a task not trivial due to (i) versatile nature of driver psychology and (ii) limited sample sizes observed under adverse road-weather conditions. Transportation authorities which experience extreme road-weather conditions may adopt the methodology to understand the speed behaviour to identify prominent road-weather and traffic conditions needing urgent safety precautions such as implementing a reliable weather-responsive variable speed limit under potentially high-risk road-weather and traffic conditions. The desired speed distributions can be further used as input driver behavior parameters in defining the speed distributions in microsimulation applications to realistically simulate traffic operations under different traffic and road-weather conditions.

The generalization of these results, however, is subject to certain limitations. First, the proposed methodology is only applicable under uncongested traffic conditions. Second, the methodology is not applicable for study sites with special features such as atypical road geometry and extremely heterogeneous traffic, which may violate the assumption of a normally distributed desired speed distribution. Notwithstanding these limitations, the results of this research support the idea that drivers select different speeds under different road-weather conditions depending on their attitude about a safe speed to travel irrespective of the posted speed limit. This is particularly manifested through the considerably high standard deviations of the desired speed distributions in the study area estimated for particular road-weather conditions. In other words, higher variability of speeds caused by lack of proper communication about the safe speed to travel, imposes considerable safety risks by paving the way toward elevated crash involvement. Consequently, the research findings provide the following insights for future research: (i) How reliable is the existing speed limit in terms of consistent communication of a safe speed in different road-weather conditions? (ii) How to propose a reliable speed limit to be implemented under the combinations of road-weather and traffic conditions which are characterized by potentially higher safety risks? and (iii) How to predict the performance of a specific speed limit prior to applying it? Accordingly, further research focusing on proposing a robust methodology to develop a reliable weather-responsive variable speed limit system which can be effectively used in uncongested rural highways in extremely cold regions are recommended.

Chapter 3. Application of Different Data Analytics for Evaluation of Heavy Vehicle Vulnerability in Cold-Region Rural Highways

This chapter compares the performance of two speed distribution modelling approaches (O_2): (i) the speed distribution modelling approach presented in chapter 1, and (ii) the conventional regression-based approach used to model speed distributions. Further, this chapter also involves developing a holistic crash indicative measure based on the context-specific speed variabilities estimated using speed distribution parameters modelled for different driving conditions such as RWCs, vehicle type, and travel lane (O_3). Thus, the research work presented in this chapter can be used to (i) understand vehicle-type-specific speed behaviour, and (ii) identify RWCs posing considerable safety risks for different vehicle types.

It should be noted that the crash indicative measure proposed in this chapter is an extension of the methodology proposed to identify RWCs with intensified safety risks in chapter 2. More specifically, chapter 2 proposes identifying RWCs with intensified safety risks based on within-lane crash risks estimated using speed and speed variability; chapter 3 presents an approach to identify such RWCs based on both within- and across-lane crash risks estimated using speed variability. The scope of this chapter includes tractor-trailer combinations, i.e., vehicles belonging to FHWA vehicle classes eight through thirteen.

The contents of this chapter correspond to the first and second steps of the stepwise conceptual framework for developing WRVSLs (Figure 1.3).

3.1 Abstract

Understanding the impacts of different driving conditions on truck speed is critical to the development, and maintenance of resilient highway freight transportation systems. This research attempts to evaluate the combined impact of road-weather, travel lane, vehicle type, and truck payload conditions on drivers' speed choice by modelling speed distributions as normal distributions. Two data analytics: a regression-based approach, and a central limit theorem-based approach, are adopted to model context-specific speed distributions. The regression-based approach models population-level speed distributions by considering samples of individual speed data, whereas the central limit theorem-based approach uses sampling distributions produced according to the central limit theorem. A holistic approach is proposed to identify overall vehicle-

specific collision risks imposed by different road-weather conditions, based on the speed distribution parameters estimated. Implications of the results of this research pertaining to trucking industry are threefold. First, adopting different data analytics lead to different results; yet, the central limit theorem-based approach is recommended to model speed distributions. Second, truck speeds are considerably affected by the presence of adverse road-weather conditions; yet marginally varied under different loading conditions. Third, overall, tractor-trailer combinations entail high collision risks, particularly when transporting a freight load under adverse road-weather conditions. The results of this research would be useful to policy makers particularly for effective speed management in extremely cold regions. Trucking companies may use the study results to identify the least risk posing road-weather conditions to deploy safe freight transport operations. The resulting speed distribution models are also useful as input to calibrate traffic micro-simulation models.

Keywords: Speed distributions, Adverse road-weather, Payload, Collision risk

3.2 Introduction

Adverse road-weather is often perceived as a risk to highway freight transportation systems (Chu, 2016) due to several reasons. For example, maneuvering freight transport trucks (hereinafter trucks) in winter storms can be challenging due to their vehicle characteristics (Naik et al., 2016) such as size, weight, structure, and vehicle performance attributes (e.g., braking, acceleration, off-tracking). Although truck drivers are compelled to drive slow under inclement road-weather conditions, it is hypothesized that drivers' speed choice is dictated by (i) prevailing road-weather (Yasanthi & Mehran, 2020; Yasanthi et al., 2021, Kyte et al., 2001), (ii) vehicle type (Catbagan & Nakamura, 2008; Jägerbrand & Sjöbergh, 2019), (iii) travel lane (Yasanthi & Mehran, 2020, Hoogendoorn, 2005b), and (iv) personal traits such as mood, driving experience, etc. (Zimasa, 2018). Payload hauled: the mass of commodity hauled by a truck (Regehr et al., 2020), is perceived as a safety concern by truck drivers (Blower & Woodrooffe, 2012), thus may also affect truck drivers' speed choice. For example, a truck driver driving an empty single-unit truck may choose to drive at 120 km/h in a dry-pavement shoulder lane, while the same driver may reduce speed to 80 km/h when driving a loaded multi-trailer truck on an icy-pavement median lane. In fact, speed choice is idiosyncratic, and each driver representing the population of truck drivers may therefore select a unique speed from a pool of context-specific speeds. Such divergent speed choices may

increase within- and across-lane speed variation which are well-acknowledged crash precursors associated with increased crash potential (Lee et al., 2002). Thus, understanding the impact of different factors on truck drivers' speed choice is critical in establishing resilient freight transportation systems, particularly in extremely cold regions.

While the choice of speed is intuitive under certain settings, the combined effect of several speed-decisive factors (e.g., road-weather conditions, truck type, truck payload, and travel lane) on truck drivers' speed choice can be comprehensive. In fact, speed data can be highly variable as they represent interactions of many factors (Hoogendoorn, 2005b). Advanced automated traffic data collection techniques practiced today (e.g., loop detectors) enable reliable collection of vehicle-by-vehicle speed data (Zhu et al., 2018) leading to large speed data sets. Accordingly, modelling population-level speed distributions is recommended in exploring the impacts of different driving conditions on drivers' speed choice to effectively identify speed patterns present in such versatile and large speed data sets (Yasanthi et al., 2021). Although several data analytics exist to model population-level speed distributions (speed distributions hereinafter), the modelling results may be influenced by the *(i)* analysis approach, and *(ii)* data treatment methods (e.g., definition of sample), adopted. Thus, there is an apparent need to evaluate the suitability of different statistical inference techniques to explore the impacts of different driving externalities such as prevailing road-weather conditions and vehicle type, on drivers' speed.

This research intends to study drivers' speed choice under four driving conditions: road-weather, vehicle type, truck payload, and travel lane. The aforementioned driving conditions are referred to as speed-decisive factors hereinafter. First, it is hypothesized that the speed choice of the population of drivers is dependent on the speed-decisive factors. Second, it is assumed that increased within- and across-lane speed differential leads to higher collision potential within- and across-lanes. This paper focuses on four research questions: *(i)* What is the combined impact of road-weather, vehicle type, truck payload, and travel lane on drivers' speed choice? *(ii)* Does data analytic approach adopted influence desired speed distribution parameters? *(iii)* To what level are trucks vulnerable under adverse road-weather conditions? and *(iv)* Which speed-decisive factors increase collision potential for trucks? The four research questions are addressed by three research objectives: *(i)* to model desired speed distributions under different combinations of road-weather, vehicle type, truck payload, and travel lane conditions, *(ii)* to investigate the variability of

modelling speed distributions based on different data analytics, and *(iii)* to propose a holistic approach to identify vehicle-specific collision risks imposed under different road-weather conditions based on the distributions of within- and across-lane speed differentials.

Traffic and road-weather data collected on an uncongested segment of highway 16 located west of Edmonton, Canada are used to model desired speed (hereinafter speed) distributions. The speed distributions are modelled as normal distributions characterized by: *(i)* mean (μ), and *(ii)* standard deviation (σ). Two data analytic approaches: regression modelling-based approach (RBA), and Central Limit Theorem i.e., CLT-based approach (CBA), are used to model speed distribution parameters μ and σ .

- Regression modelling-based approach (RBA): In the RBA, a set of multiple linear regression models are developed to model speed distribution parameters μ and σ based on collected speed samples. Sample, in the context of the RBA, is defined as individual speeds observed under a specific combination of speed-decisive factors. Accordingly, regression models are calibrated with descriptive statistics of speed samples belonging to different combinations of speed-decisive factors to infer μ and σ .
- CLT-based approach (CBA): The CBA — as proposed by Yasanthi et al., (2021) — attempts to model μ and σ based on statistical inference techniques including CLT and Minimum Variance Unbiased Estimation (MVUE). Vehicle speeds collected within five-minutes under a unique combination of speed-decisive factors is defined as a sample in the context of CBA. In brief, CBA combines multiple five-minute aggregate speed distributions (i.e., sampling distributions) to infer μ and σ by assigning relatively larger weights to five-minute aggregate speed distributions characterized by large samples.

A unified approach is proposed to identify vehicle-specific collision risks under different road-weather conditions based on the speed distribution parameters μ and σ . More specifically, two-dimensional naïve k-means clustering is used to define and identify such context specific collision risks at three levels: high, moderate, and low. The freight industry and policy makers may use the study results *(i)* to understand the impact of truck payload, truck type, travel lane, and road-weather conditions on truck drivers' speed choice, *(ii)* to understand the variability of speed distribution

estimates caused by adopting different data analytic techniques to model speed distributions, (iii) to apprehend the vulnerability of different vehicle types under adverse road-weather conditions, and (iv) as direct input variables or surrogate safety measures in traffic simulation studies.

3.3 Background and literature review

Several past studies have modelled drivers' speed choice as a function of different speed-decisive factors (Hoogendoorn, 2005b; Yasanthi et al., 2021, Kyte et al., 2001). For instance, by developing free-speed distributions, Hoogendoorn (2005b) revealed that trucks travel faster in median lane; yet, they travel much slower than passenger cars irrespective of the travel lane. Aggregating passenger cars and trucks resulted in asymmetry in the speed distribution for median lane developed by Hoogendoorn (2005b), emphasizing the impact of vehicle class on speed distributions. Speed distributions produced by Catbagan & Nakamura (2008) revealed that road-weather conditions impose notable impacts on heavy vehicles' speed, specially while following similar heavy vehicles. Adverse pavements i.e., slippery road surface conditions due to weather events (Yasanthi & Mehran, 2020; Kyte et al., 2001), and snow (Yasanthi & Mehran, 2020; Catbagan & Nakamura, 2008) were found to have notable impacts on truck speed. While Yasanthi et al. (2021) stresses on the impact of the combined effect of adverse pavement, and precipitation conditions on speed, they concluded that time of the day or atmospheric temperature do not crucially influence speed distribution parameters. While estimating the impacts of different road-weather conditions on speed, only limited number of studies (Yasanthi & Mehran, 2020; Yasanthi et al., 2021; Jägerbrand & Sjöbergh, 2019) used Road-Weather Information Systems (RWIS) to collect road-weather data in close proximity to traffic data collection devices.

Past literature (Hamdar et al., 2016; Chen et al., 2017; Lefevre, 1953) also identified road grade: the ratio of rise (or drop) in the vertical distance to the horizontal distance, as an important factor among different roadway geometric characteristics affecting speed. For instance, Hamdar et al., (2016) showed that, for passenger cars, the impacts of the presence of frequent changes in road grade may be more severe than the impacts of adverse road-weather conditions. By evaluating the impact of road grade on free-flow speeds, Chen et al., (2017) concluded that road grade is negatively associated with drivers' speed choice. In fact, the impacts of upslope segments (i.e., positive slopes) on drivers' speed choice were deemed more influential than that of downslope

segments (Chen et al., 2017). Further, it is documented that drivers tend to reduce speed ahead of vertical curves with short sight distances (Lefevre, 1953).

The intensity of speed reductions resulted by different driving conditions seem to vary among past studies. For example, when compared with normal road-weather, Yasanthi et al. (2021) estimated a speed reduction factor of 2.68% for vehicles travelling on wet pavements in daytime under heavy snow, and a temperature of 5°C; while Highway Capacity Manual (HCM) (TRB, 2016) recommends a speed reduction factor of 16% for comparable road-weather conditions. While such differences are predominantly attributed to adopting different data analytics (Yasanthi et al., 2021), some authors argue that the number of speed-decisive factors considered in each research may also cause such controversies (Yasanthi & Mehran, 2020). It is apparent that adopting different data treatment methods to study drivers' speed choice may also limit each research to a confined analysis-level. For example, using descriptive statistics to explain drivers' speed choice under different road-weather conditions (TRB, 2016) convey sample-specific observations. Although popular, regression modelling with individual vehicle speeds as dependent variable is deemed controversial in modelling the impacts of different driving conditions on speed due to several reasons (Yasanthi et al., 2021). For example, it is argued that some drivers may not perceive adverse road-weather as a substantial hazard-risk (TRB, 2016); thus, the average speed reductions suggested by regression models' estimates might not reasonably reflect the impact of road-weather conditions on speed choice of such drivers (Yasanthi et al., 2021; Yasanthi & Mehran, 2020). To address such issues, past studies have suggested comparing speed distributions under different driving conditions (Catbagan & Nakamura, 2008; Hoogendoorn, 2005b) rather than apprehending the impacts of different driving conditions on speed based on average speed reductions suggested by regression models' estimates. Literature suggests several approaches to estimate speed distribution parameters (Yasanthi et al., 2021; Jägerbrand & Sjöbergh, 2019; Catbagan & Nakamura, 2008). For example, Hoogendoorn (2005b) and Catbagan & Nakamura (2008) used Kaplan-Meier estimation to model lane- and vehicle-specific desired speed distributions; yet, application of this methodology is restrained under limited sampling conditions due to high standard error associated with small samples (Catbagan & Nakamura, 2008). Moreover, speed distributions developed by Hoogendoorn (2005b) and Catbagan & Nakamura (2008) do not infer population parameters. In contrast, Yasanthi et al. (2021) proposed a methodology to estimate population-level speed distribution parameters based on the CLT and MVUE irrespective of the

sampling condition. Yet, it is important to note that findings of Yasanthi et al. (2021) are controlled for the impact of traffic conditions, and vehicle type.

Despite controversial conclusions about road-weather-specific speed reduction factors, most past studies (Yasanthi et al., 2021; Kyte et al., 2001; Hoogendoorn, 2005b) support the hypothesis that all speed-decisive factors prompt speed variation to some extent. It is well-documented that travelling at speeds considerably different from the average traffic speed increase crash involvement (Lee et al., 2002; Edwards, 1999; Solomon, 1964). This is rather intuitive because high speed differentials trigger passing maneuvers, particularly for fast drivers who are willing to tradeoff safety for lesser travel times. By comparing speed distribution parameters under adverse, and normal road-weather conditions, Yasanthi et al. (2021) identified combinations of road-weather conditions with high crash severity, and high crash exposure for shoulder lane traffic. Yet, across-lane speed differential was not considered as an indication of intensified safety risks. It is important to note that understanding differences between within- and across-lane collision risks considering different vehicle types, and road-weather conditions is crucial (Lee et al., 2002; Shi, & Liu, 2019), particularly in weather-responsive speed management.

Figure 3.1 presents a summary of the reviewed literature, and the research gaps identified. First, literature on the extent of the impact of different road-weather conditions on speed is less consistent. Although evaluating road-weather-specific speed distributions is recommended in apprehending road-weather-specific speed choice (Yasanthi et al., 2021; Hoogendoorn, 2005b), past literature do not provide a conclusive reference to the most appropriate data analytic approach (e.g., RBA versus CBA) and data treatment method (e.g., definition of a sample), leading to the second research gap. Third, an extensive literature review conducted on the research topic revealed that the combined impact of road-weather, travel lane, truck type (single- versus multi-trailer), and payload conditions (empty versus loaded) on truck speed has not been studied. Fourth, past literature partially addresses the problem of identifying crash indicative road-weather conditions when collision statistics are not available. To fill the first three research gaps, this research intends to evaluate the combined impact of road-weather, travel lane, vehicle type (focusing on different truck types) and truck payload condition, on speed distribution parameters using (i) RBA, and (ii) CBA, with a unique definition of sample used in each approach. Modelling speed distributions under different driving conditions indeed provides a deeper insight to drivers' speed choice, rather

than estimating average speed reductions under different driving conditions. The fourth research gap is addressed by proposing a holistic approach to identify road-weather conditions with escalated collision risks for different vehicle types based on both within- and across-lane speed differentials.

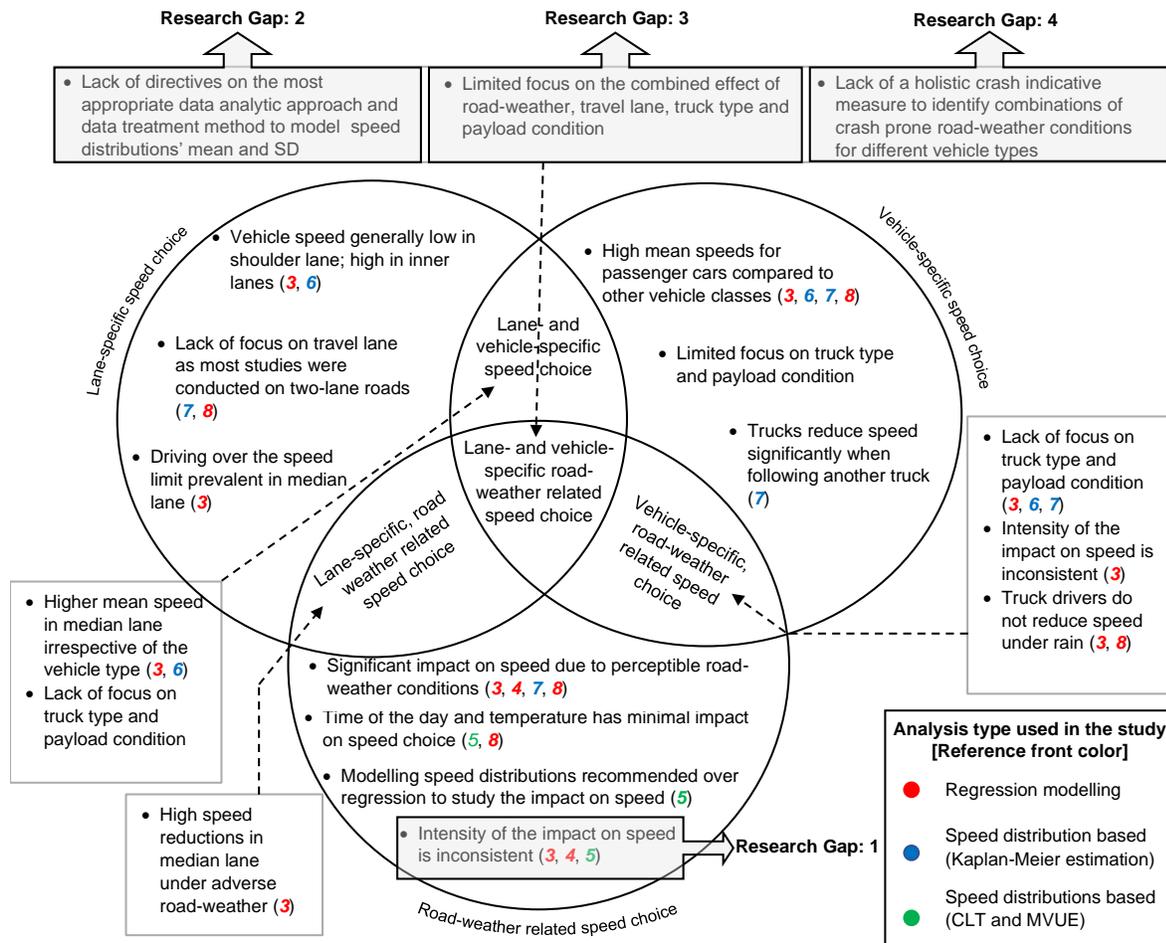


Figure 3.1: Summary of related literature, and research gaps (Yasanthi & Mehran, 2022)

3.4 Study data

3.4.1 Data collection

Study data used in this research consists of road-weather and traffic data, provided by Alberta Transportation. Road-weather, and traffic data were collected by a RWIS station and a Weigh-In-Motion (WIM) station respectively, for 15 months from October 2014 to December 2015. The

study site (Figure 3.2) is located west of Edmonton, Alberta on a four-lane, divided segment of highway 16-a provincial highway, which is part of yellowhead highway between Jasper and Lloydminster. The study site operates at fixed speed limit of 110 km/h. According to Alberta Transportation, the Annual Average Daily Traffic (AADT) at the study site in 2015 was 8,120 vehicles (Alberta Transportation, 2016). The study site is located on a road section with negligible longitudinal grade and with no appreciable horizontal curves. The site (Figure 3.2) was selected for this research due to two reasons. First, it typically operates under levels of service (LOS) “A” and “B” (Figure 2.2) i.e., uncongested; thus, vehicle speeds collected reasonably represent drivers’ desired speed (TRB, 2016). Second, the RWIS station precisely captures the microclimate data for each vehicle recorded by the WIM station as it is only 148.7 m away from the WIM (Figure 3.2).

The WIM station recorded vehicle-by-vehicle speed, axel weight, inter-axle spacing, and timestamp for each vehicle passing the WIM station. The RWIS station recorded several road-weather attributes including precipitation condition (type and intensity), pavement surface condition, and temperature in every 20-minutes. A recent research by Yasanthi et al. (2021) revealed that pavement surface condition and precipitation condition at the time of travel considerably affect drivers’ speed choice although the time of day, and temperature do not

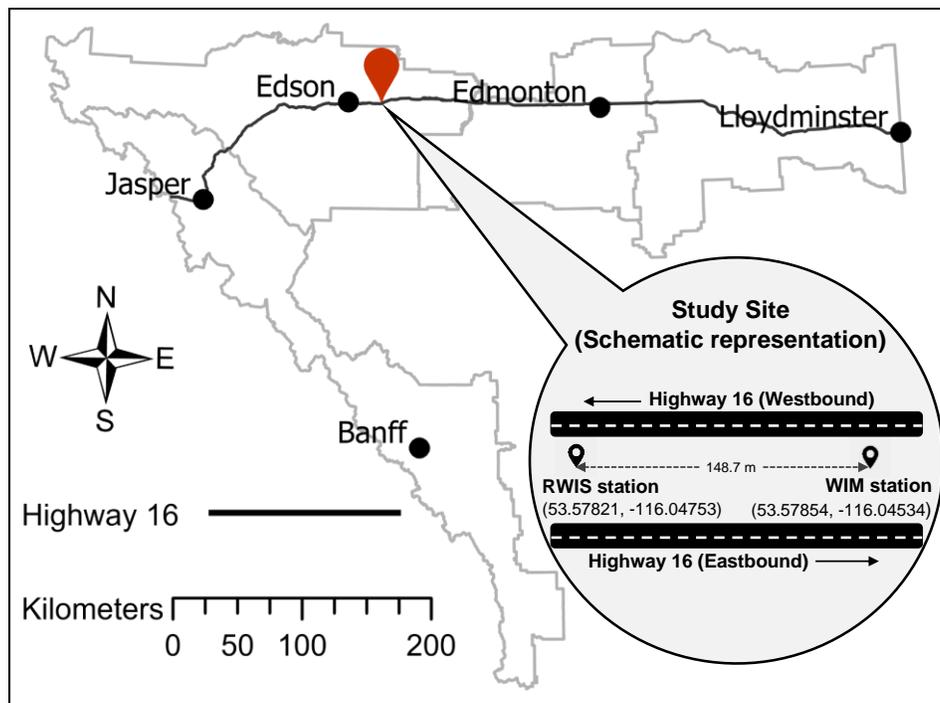


Figure 3.2: Study site
(Yasanthi & Mehran, 2022)

crucially affect speed distribution parameters. Accordingly, this research only focusses on two road-weather attributes: pavement surface, and precipitation condition.

3.4.2 Data preparation

First, the collected data were cleaned by (i) removing “N/A” and “Error” entries in the road-weather data set, and (ii) removing vehicle records with “null” speeds and speeds greater than 200 km/h i.e., considered as error. Thereafter, each vehicle was categorized into one of the 13 vehicle classes proposed in the Federal Highway Administration (FHWA) scheme “F” vehicle classification system (FHWA, 2001). Tractor-trailer combinations (TTC), i.e., trucks belonging to FHWA vehicle classes eight through thirteen, were further divided into two categories based on their payload condition. According to Regehr et al. (2020), mean tare weight of a specific truck configuration can be estimated by approximating its Gross Vehicle Weight (GVW) distribution by a Gaussian Mixture Model (GMM). Accordingly, the mean tare weight for each tractor-trailer configuration observed at the study site was estimated by approximating the GVW distribution of each tractor-trailer configuration as a GMM. Similar to Regehr et al. (2020), the 99th percentile of the first GMM component was concluded as the tare weight of each TTC. The TTC with a GVW more than its tare weight were labelled as “Loaded”, or “Empty” otherwise. As a result, the 13 FHWA classes were aggregated into six categories, each representing different vehicle characteristics. A series of Kolmogorov–Smirnov (K-S) tests conducted between each level of vehicle type, precipitation type, pavement surface condition, and travel lane categories concluded that the speeds in all levels in each study data category are statistically independent from the rest of levels at a significance level of 5% except for a few levels where the original levels were kept for practicality and consistency (see detailed results in Appendix A1.2). Underlying reasons for choosing K-S test to evaluate the statistical difference of speeds in different attribute levels are presented in section 2.4.2. Details of different study data categories used in this research are tabulated in Table 3.1.

Descriptive statistics of the speed data are presented in Figure 3.3 in terms of mean, Standard Deviation (SD) of speeds and the number of vehicle observations for each combination of study data categories. According to Figure 3.3, speed is apparently affected by travel lane, vehicle type, and road-weather (the numbers in brackets show the sample size).

Table 3.1: Study Data Categories

Category	Description
Vehicle Type	
Passenger vehicle (PV)	Vehicles belonging to FHWA classes 1-4
Single-Unit Truck (SUT)	Vehicles belonging to FHWA classes 5-7
Empty Single-Trailer Truck (E-STT)	Vehicles belonging to FHWA classes 8-10 with a GVW less than its tare weight
Loaded Single-Trailer Truck (L-STT)	Trucks belonging to FHWA classes 8-10 with a GVW more than its tare weight
Empty Multi-Trailer Truck (E-MTT)	Trucks belonging to FHWA classes 11-13 with a GVW less than its tare weight
Loaded Multi-Trailer Truck (L-MTT)	Trucks belonging to FHWA classes 11-13 with a GVW more than its tare weight
Precipitation Condition	
No precipitation	Zero precipitation
Slight rain	
Slight snow	Precipitation intensity < 2 mm/h
Slight frozen precipitation	
Moderate & heavy rain	
Moderate & heavy snow	Precipitation intensity ≥ 2 mm/h
Moderate & heavy frozen precipitation	
Pavement Surface Condition	
Dry	No moisture or unusual condition detected
Ice warning	Detection of ice or black ice
Trace moisture	Detection of isolated moisture on pavement surface
Wet	Wet roadway with considerable moisture detection
Ice watch	The risk of the formation of ice or black ice on the roadway is elevated, but its occurrence, location, and/or timing is still uncertain
Frost	Detection of frost formation
Travel Lane	
Shoulder	The rightmost lane in each travel direction
Median	The leftmost lane in each travel direction

As expected, most vehicles seem to travel faster in median lane irrespective of the vehicle type, and prevailing road-weather conditions. Similarly, the higher SD of speeds recorded under all road-weather, and vehicle type combinations in median lane imply comparatively higher speed

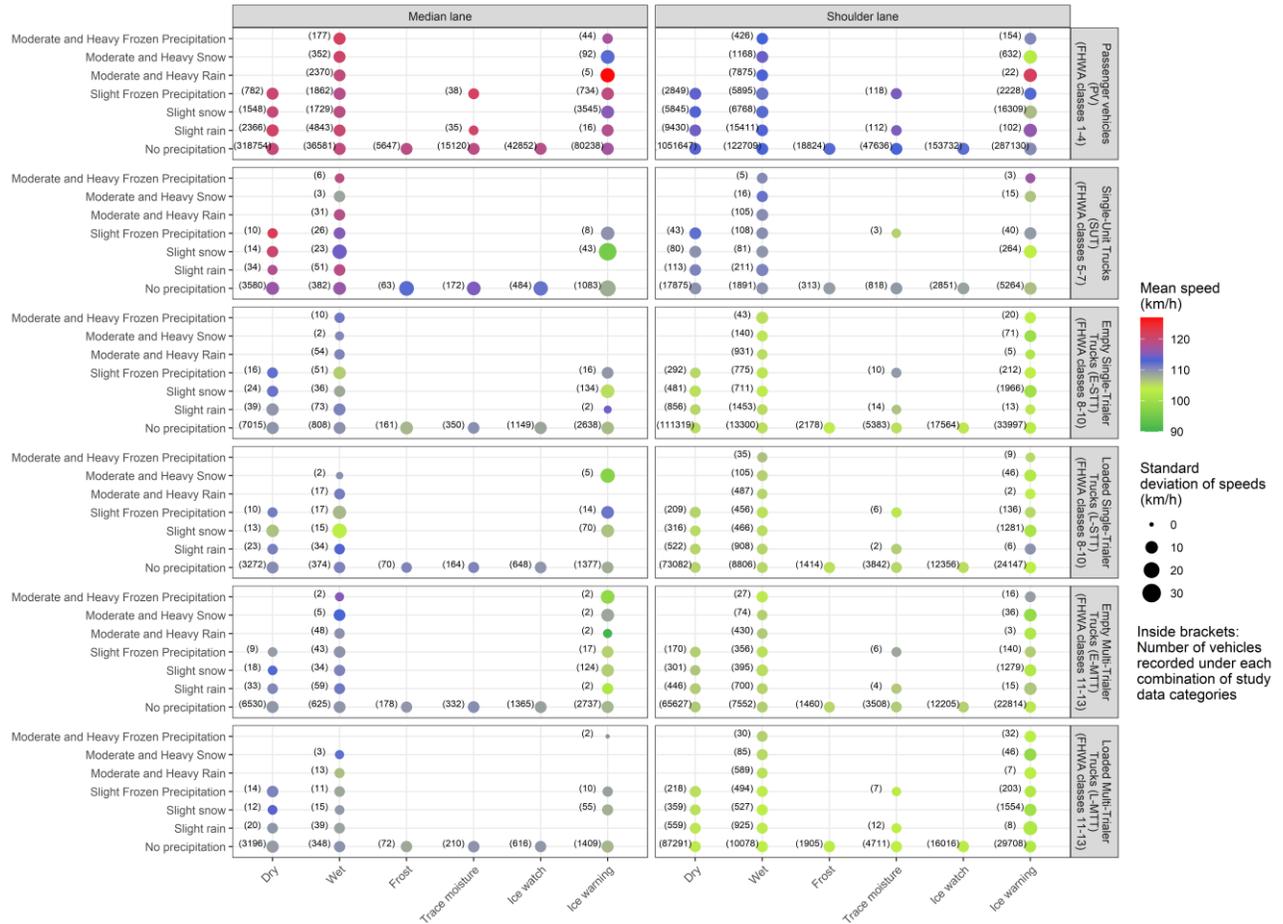


Figure 3.3: Descriptive statistics of speed data (Yasanthi & Mehran, 2022)

variations in median lane as compared to the shoulder lane. The notably higher mean speeds for Passenger Vehicles (PV), followed by Single-Unit Trucks (SUT) signal the importance of differentiating speed distributions by vehicle type. Nevertheless, Figure 3.3 reports approximately similar mean, and SD of speeds for Single-Trailer Trucks (STT), and Multi-Trailer Trucks (MTT) suggesting that TTC speeds might not be affected by the number of trailers attached to a tractor. Overall, the synchrony of snow precipitation (irrespective of intensity) and ice warning pavements crucially affect mean and SD of speeds particularly for PV, and SUT travelling in both lanes. Such speed variations imply drivers' attentiveness while driving under adverse road-weather conditions.

Yet, it is important to note that Figure 3.3 merely exhibit descriptive statistics based on small samples, potentially leading to counterintuitive observations.

3.5 Methodology

The methodology proposed to evaluate the combined impact of four speed-decisive factors (road-weather, travel lane, vehicle type, and truck payload) on speed consists of two steps. The first step focuses on modelling speed distributions under different combinations of speed-decisive factors. The second step attempts to identify combinations of road-weather conditions posing high collision risks for different vehicle types, based on the speed distributions modelled in the first step.

3.5.1 Modelling speed distributions

In this research, speed distributions are modelled based on two assumptions: (i) observed speeds reasonably represent drivers' desired speeds, and (ii) the population of individual desired speeds under a specific combination of speed-decisive factors can be modelled as a normal distribution characterized by a mean of μ , and a SD of σ . Study data used in this research support both assumptions (Yasanthi et al., 2021). In fact, it is well-documented that desired speed distributions can be modelled as normal distributions under most conditions (Yasanthi et al., 2021; Taylor & Bonsall, 2017). Yet, in this research, it is hypothesized that drivers' speed choice is dependent on the prevailing driving conditions; thus, the speed distribution parameters are expected to vary under different combinations of speed-decisive factors. However, this research argues that the speed distribution parameters might also be affected by the data analytic approach considered. Therefore, this research intends to compare speed distributions developed using two data analytic approaches: RBA, and CBA. Population is defined as the distribution of individual vehicle speeds under a specific combination of speed-decisive factors (e.g., L-MTT travelling in shoulder lane under slight rain, and wet pavements) in both approaches. Yet, there are two principal differences between the two approaches. First, definition of sample adopted in each approach is unique. Second, the theoretical foundation of the two analysis approaches is fundamentally different i.e., RBA infers the population-level speed distribution parameters μ and σ based on samples, while CBA infers μ and σ based on sampling distributions.

3.5.1.1 Modelling μ and σ based on the RBA

Sample, in the context of RBA, is defined as individual vehicle speeds observed under a specific combination of speed-decisive factors during the entire study period. For instance, sample considered for the speed distribution representing PV traveling in shoulder lane under no precipitation, and dry pavements consists of all individual PV speeds observed under such driving conditions during the study period. To model speed distributions according to RBA, this research assumes that a speed distribution's mean μ and SD σ are linearly dependent on the set of speed-decisive factors considered in this research. Accordingly, two multiple linear regression models (Models 1, and 2) are developed to estimate the relationships between each speed distribution parameter, and the set of speed-decisive factors. Multiple linear regression modelling is used to develop the two speed models due to being highly explainable (Eberly, 2007), and its widespread use in modelling the impacts of adverse road-weather on speed (Yasanthi et al., 2021; Edwards, 1999; Kyte et al., 2001). Model 1 estimates the relationship between the four speed-decisive factors and the mean of a speed distribution μ ; whereas Model 2 estimates the relationship between the four speed-decisive factors and the SD of a speed distribution σ . It should be noted that levels of each independent variable (Table 3.1) are coded as dummy variables where the i^{th} independent variable is considered to have r number of levels. The individual desired speeds of PV travelling in shoulder lane under no precipitation, and dry pavements (hereinafter normal road-weather) are considered as the reference condition (i.e., intercept) for both models. Modal formats used in the two models are expressed in Eq. 3.1:

$$y = \beta_0 + \sum_{i=1}^4 \beta_i^{j=1 \text{ to } r} x_i^{j=1 \text{ to } r} \quad \text{Eq. 3.1}$$

where,

$$y = \begin{cases} \text{Speed distribution mean } (\mu) \text{ for Model 1} \\ \text{Speed distribution SD } (\sigma) \text{ for Model 2} \end{cases}$$

$$\beta_0 = \begin{cases} \text{The mean value of speed distribution mean } (\mu) \text{ for reference conditions for Model 1} \\ \text{The mean value of speed distribution SD } (\sigma) \text{ for reference conditions for Model 2} \end{cases}$$

$$i = \begin{cases} 1 \text{ for Precipitation condition} \\ 2 \text{ for Pavement surface condition} \\ 3 \text{ for Vehicle type/loading condition} \\ 4 \text{ for Travel lane} \end{cases}$$

$$r = \begin{cases} 6 \text{ for } i = 1 \\ 5 \text{ for } i = 2 \\ 5 \text{ for } i = 3 \\ 1 \text{ for } i = 4 \end{cases}$$

$x_i^{j=1 \text{ to } r}$ = Independent variable level (see Table 3.1)

$\beta_i^{j=1 \text{ to } r}$ = Regression coefficient for independent variable level $x_i^{j=1 \text{ to } r}$

Modelling speed distributions using RBA is straightforward. Yet, the explanatory power and statistical significance of model coefficients may vary depending on the samples collected, particularly under small sample sizes.

3.5.1.2 Modeling μ and σ based on the CBA

The CBA applies the approach proposed by Yasanthi et al., (2021) to estimate μ and σ . The CBA is based on two fundamental statistical theorems: CLT, and MVUE. While the following section briefly explains the fundamental steps, comprehensive details of the methodology adopted in the CBA were discussed extensively in Yasanthi et al., (2021).

- Step 1: Defining speed populations and samples

In the context of the CBA, a speed population is defined as all collected individual speeds under a specific set of speed-decisive factors. Sample is defined as vehicle speeds collected during a five-minute interval. Two sample statistics: (i) n : sample size (i.e., the number of vehicles observed in five minutes), and (ii) v : sample mean (i.e., five-minute aggregate speed), are used in the CBA.

- Step 2: Producing five-minute aggregate speed distributions (sampling distribution of speed)

According to the CLT, a population of individual desired speeds is represented by a normally distributed five-minute aggregate speed distribution i.e., the sampling distribution of speed

(Yasanthi et al., 2021). Such five-minute aggregate speed distributions are characterized by unique sample sizes. Yet, due to the definition of sample adopted in the CBA, a specific population of speeds may encompass several sample sizes. Accordingly, some populations are represented by multiple sampling distributions (see Figure 2.5). For example, Figure 2.7 shows a population represented by eight different sampling distributions of observed speeds (the number of sampling distributions (k) = 8), where M represents the number of five-minute intervals (i.e., samples) contributing to each sampling distribution.

- Step 3: Linear combination of sampling distributions of speed

All k sampling distributions representing a specific speed population can be linearly combined to produce \bar{Y} that is a random variable possessing the minimum variance of such combination. MVUE technique is applied to estimate a_i : estimated weight factor assigned to the i^{th} sampling distribution to ensure minimum variance property of \bar{Y} . By applying estimated weight factors, each sampling distribution will be represented proportional to their sample size while producing \bar{Y} .

- Step 4: Estimating μ and σ

According to CLT, the mean (μ) and SD (σ) of a particular speed population are related to $\mu_{\bar{Y}}$ and $\sigma_{\bar{Y}}$ i.e., the mean and SD of \bar{Y} as expressed in Eqs 3.2 and 3.3, where n_i is the sample size used in i^{th} sampling distribution combined to produce \bar{Y} .

$$\mu = \mu_{\bar{Y}} \tag{Eq. 3.1}$$

$$\sigma = \sqrt{\frac{\sigma_{\bar{Y}}^2}{\sum_{i=1}^k a_i^2/n_i}} \tag{Eq. 3.2}$$

By estimating speed population parameters using all representative sampling distributions, the CBA is particularly advantageous for investigating the impacts of extreme road-weather conditions in rural highways, where the sample sizes are inherently small due to limited traffic throughput under different adverse road-weather conditions.

3.5.2 Identifying Road-Weather-Specific Collision Risk Levels for Different Vehicle Types

In this research, “speed differential” is defined as the difference between two population-level speed distributions. If the two speed distributions correspond to identical travel lanes, the distribution of speed differentials is categorized as within-lane speed differential, or across-lane speed differential otherwise. This research develops within- and across-lane speed differentials for different combinations of road-weather conditions, vehicle types, and travel lanes (in case of within-lane speed differential). Speed differential is often used as a surrogate safety measure indicative of crash potential (Gettman & Head, 2003). Thus, for a specific combination of driving conditions, this research assumes that (i) within-lane collision risk for a particular travel lane is implied by corresponding within-lane speed differential, and (ii) across-lane collision risk is reflected by corresponding across-lane speed differential between two adjacent lanes. According to the CLT, the subtraction of normally distributed random variables (i.e., speed distributions) yields a new normally distributed random variable, irrespective of the sample size (Hayter, 2012). Accordingly, this research assumes that the population of the distribution of speed differentials for a specific combination of road-weather, and vehicle type conditions is represented by a normally distributed random variable characterized by its mean and SD.

Theorem: Difference of two population-level speed distributions

If $V_a \sim N(\mu_a, \sigma_a^2)$, $a \in \mathbb{Z}^+$, and $V_b \sim N(\mu_b, \sigma_b^2)$, $b \in \mathbb{Z}^+$, represent two independent population-level speed distributions, then

$$V_{a-b} = V_a - V_b \sim N(\mu_{a-b}, \sigma_{a-b}^2) \quad \text{Eq. 3.4}$$

where,

$$\mu_{a-b} = \mu_a - \mu_b \quad \text{Eq. 3.5}$$

$$\sigma_{a-b} = \sqrt{(\sigma_a)^2 + (\sigma_b)^2} \quad \text{Eq. 3.6}$$

For a specific combination of L : travel lane, α : vehicle type, and β : adverse road-weather condition, Algorithm 3.1 (Figure 3.4) explains the process of modelling (i) $D_{\alpha,\beta,L}^W$: within-lane speed differential distribution, and (ii) $D_{\alpha,\beta}^A$: across-lane speed differential distribution, for the combination of α , β , and L . More specifically, Algorithm 3.1 outputs four matrices providing: (i) $\mu_{\alpha,\beta,L}^W$: mean of $D_{\alpha,\beta,L}^W$, (ii) $\sigma_{\alpha,\beta,L}^W$: SD of $D_{\alpha,\beta,L}^W$, (iii) $\mu_{\alpha,\beta}^A$: mean of $D_{\alpha,\beta}^A$, and (iv) $\sigma_{\alpha,\beta}^A$: SD of $D_{\alpha,\beta}^A$.

Algorithm 3.1 Pseudocode for estimating speed differential distribution parameters

Input: $N_{vt}, N_{rw}, \alpha, \beta, L, \mu_{\alpha,\beta,L}, \sigma_{\alpha,\beta,L}$

1. Estimate parameters of the within-lane speed differential distributions
for $L = 1$ to 2
for $\alpha = 1$ to N_{vt}
for $\beta = 1$ to N_{rw}
$$\mu_{\alpha,\beta,L}^W = |\mu_{\alpha,\beta,L} - \mu_{\alpha,0,L}|$$
$$\sigma_{\alpha,\beta,L}^W = \sqrt{(\sigma_{\alpha,\beta,L})^2 + (\sigma_{\alpha,0,L})^2}$$

end
end
end
 2. Estimate parameters of the across-lane speed differential distributions
for $\alpha = 1$ to N_{vt}
for $\beta = 1$ to N_{rw}
$$\mu_{\alpha,\beta}^A = |\mu_{\alpha,\beta,1} - \mu_{\alpha,\beta,2}|$$
$$\sigma_{\alpha,\beta}^A = \sqrt{(\sigma_{\alpha,\beta,1})^2 + (\sigma_{\alpha,\beta,2})^2}$$

end
end
- Output:** The matrices of $\mu_{\alpha,\beta,L}^W, \sigma_{\alpha,\beta,L}^W, \mu_{\alpha,\beta}^A, \sigma_{\alpha,\beta}^A$
-

Figure 3.4: Algorithm for estimating parameters of within- and across-lane speed differentials

(Yasanthi & Mehran, 2022)

Algorithm 3.1 takes seven inputs: (i) N_{vt} : number of vehicle types considered (6 types in this research), (ii) N_{rw} : number of adverse road-weather combinations considered (41 combinations in this research), (iii) vehicle type indicator α (coded from 1 through N_{vt} ; e.g., 1 for PV), (iv) road-weather condition indicator β (coded from 1 through N_{rw} with 0 representing normal road-weather conditions i.e., dry pavements and no precipitation), (v) lane indicator L (coded as 1 for shoulder lane and 2 for median lane), (vi) $\mu_{\alpha,\beta,L}$: speed distribution mean for combination of α , β , and L , and (vii) $\sigma_{\alpha,\beta,L}$: speed distribution SD for combination of α , β , and L . For a specific combination of α , β , and L , Algorithm 3.1 first estimates $\mu_{\alpha,\beta,L}^W$ according to Eq. 3.5. Speed differential for a particular combination of α , β , and L , irrespective of its sign, indicates increased crash potential; thus, the algorithm only retains the absolute value of $\mu_{\alpha,\beta,L}^W$. Thereafter, $\sigma_{\alpha,\beta,L}^W$ is estimated according to Eq. 3.6. Similarly, the algorithm estimates $\mu_{\alpha,\beta}^A$ and $\sigma_{\alpha,\beta}^A$ for all combinations of α and β observed at the study site. The distribution parameters of within- and across-lane speed differentials output by Algorithm 3.1 are explored separately to identify clusters of road-weather

conditions, and vehicle types which are hypothesized to impose different levels of collision risk: high, moderate, and low. Two-dimensional naïve k-means clustering (Hackeling, 2017) is adopted to partition each set of speed differential distributions into the three collision risk levels based on the SD of the speed differential distributions. Accordingly, a holistic collision potential indicative measure is proposed by considering both within- and across-lane collision potential for each combination of road-weather conditions, and vehicle type categories.

Algorithm 3.2 explains the criteria for assigning collision risk level based on the within- and across-lane collision potential (Figure 3.5). Algorithm 3.2 takes four inputs: (i) N_C : the number of combinations of road-weather conditions and vehicle type categories, (ii) R_W^S : within-lane collision risk for shoulder lane (high, moderate, or low), (iii) R_W^M : within-lane collision risk for median lane (high, moderate, or low), and (iv) R_A : Across-lane collision risk (high, moderate, or low). To preserve indication of high collision risk for either within- or across-lanes, the algorithm assigns the collision risk for each combination of road-weather conditions, and vehicle type with the highest collision risk level among R_W^S , R_W^M , or R_A . Consequently, algorithm 3.2 outputs a unified collision risk level for each combination of road-weather, and vehicle type category observed.

Algorithm 3.2 Pseudocode for assigning collision risk levels

Input: N_C, R_W^S, R_W^M, R_A

1. Initialize matrix R with N_C random variables
2. Assign collision risk level for each road-weather, and vehicle type combination
 - for $c = 1$ to N_C
 - if $(R_W^S \text{ OR } R_W^M \text{ OR } R_A) == \text{“High”}$
 - $R_C == \text{“High”}$
 - else if $(R_W^S \text{ OR } R_W^M \text{ OR } R_A) == \text{“Moderate”}$
 - $R_C == \text{“Moderate”}$
 - else
 - $R_C == \text{“Low”}$
 - end

Output: Collision risk level matrix R

Figure 3.5: Algorithm for assigning collision risk levels (Yasanthi & Mehran, 2022)

3.6 Modelling results

3.6.1 Speed distribution models

In total, speed data belonging to 248 combinations of speed-decisive factors were used while modelling the speed distributions. To recall, Models 1 and 2 (Eq. 3.1) attempt to estimate the linear

relationship between a set of speed-decisive factors, and speed distribution parameters μ and σ . Figure 3.6 graphically presents the regression model coefficients for Models 1 and 2, where the intercept of both models corresponds to PV travelling in shoulder lane under no precipitation and dry pavement conditions.

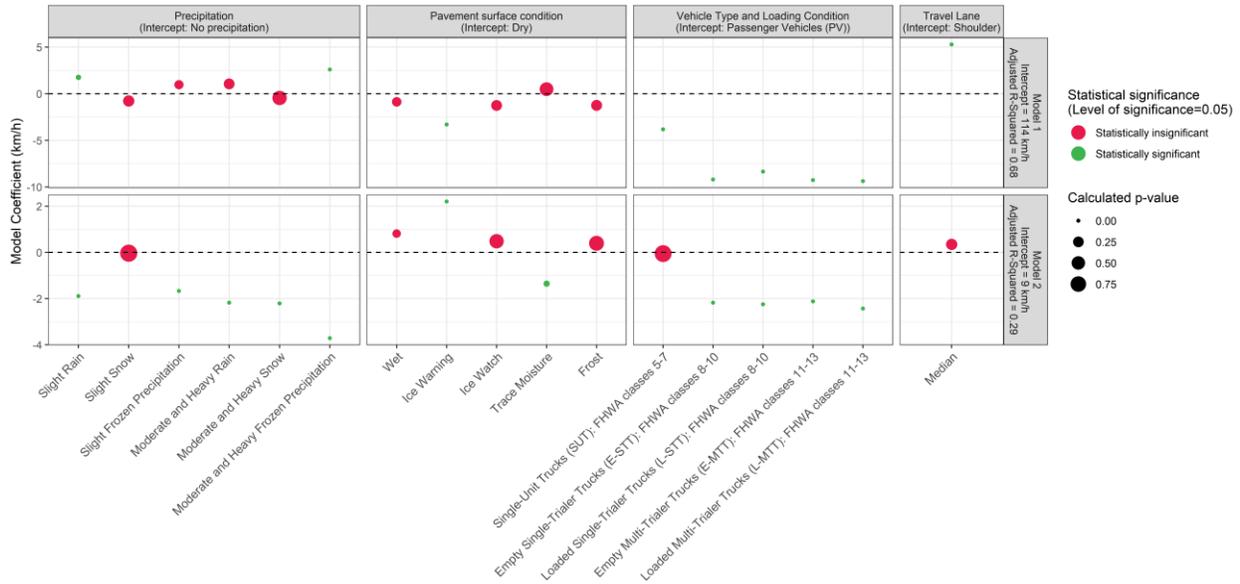


Figure 3.6: Model coefficients for Models 1, and 2 (Yasanthi & Mehran, 2022)

The adjusted R^2 values of the two models suggest that 68% of variations in μ and 29% of variations in σ are explained by the independent variables of the models (Figure 3.6). The poor prediction power of Models 1 and 2 indicated by such comparatively low R^2 values imply that the two regression models may not be suitable to make predictions. However, according to Figure 3.6, most model coefficients (e.g., coefficient for slight snow, wet pavements) in both models are reported as statistically insignificant at a significance level of 5%. Such statistically insignificant model coefficients might result due to small sample sizes for some combinations as presented in Figure 3.3. For instance, only three SUTs were observed on wet pavements under moderate, and heavy snow in median lane. Yet, samples with such small sample sizes may not be representative of the population thus violating one of the assumptions in regression modeling approaches. On the other hand, another potential reason for the statistically insignificant modelling coefficients is that the relationship between the dependent variable in question (i.e., either μ or σ) and the independent variables may not be linear. However, it is well-documented that very small sample sizes are

associated with type II error (Lochner et al., 2001) which is often referred to as false positive — accepting a false null hypothesis. Accordingly, deriving conclusions about the impacts of different driving conditions on μ and σ based on indications of the acceptance of the null hypothesis in linear regression modeling (i.e., no statistically significant linear relationship between the dependent and independent variable) is questionable in limited sampling conditions. Limited sampling conditions might have also led to some counterintuitive suggestions implied by statistically insignificant model coefficients. For example, Model 1 suggests that on average, PV travelling on dry pavements in shoulder lane drive approximately 2.5 km/h faster under moderate and heavy frozen precipitation, as compared to travelling under no precipitation. In contrast, the statistically significant model coefficients convey intuitive suggestions. As expected, the average speed of trucks travelling in shoulder lane under normal road-weather conditions are considerably lower than that of PV. For instance, the model coefficient for E-STT (-9.2km/h) suggests that, on average, E-STT travel approximately 9km/h slower than PV when travelling under normal road-weather conditions. Similarly, travelling in median lane under normal road-weather conditions is estimated to increase μ by 5.3km/h for PV. However, the impact of payload condition on both μ and σ seem marginal for all TTC. In fact, the difference of μ between E-STT and L-STT was estimated as 0.85km/h, whereas the difference of σ was estimated as 0.07km/h. Such marginal effects of payload condition on truck speed may be potentially attributed to firm schedules practiced in trucking industry prompting truck drivers to prioritize travel time irrespective of truck loading condition, leading to similar speed choices under all loading conditions.

Figure 3.7 presents and compares the speed distribution parameters estimated according to the RBA and CBA. Speed distribution parameters estimated based on RBA are presented irrespective of their statistical significance. Yet, discussions on RBA estimates are limited to cases where both μ and σ are estimated based on statistically significant coefficients (RBA estimates marked with “*,**”) to allow for factual comparison. Thus, discussion on the impacts of travel lane on speed choice is confined to the CBA.

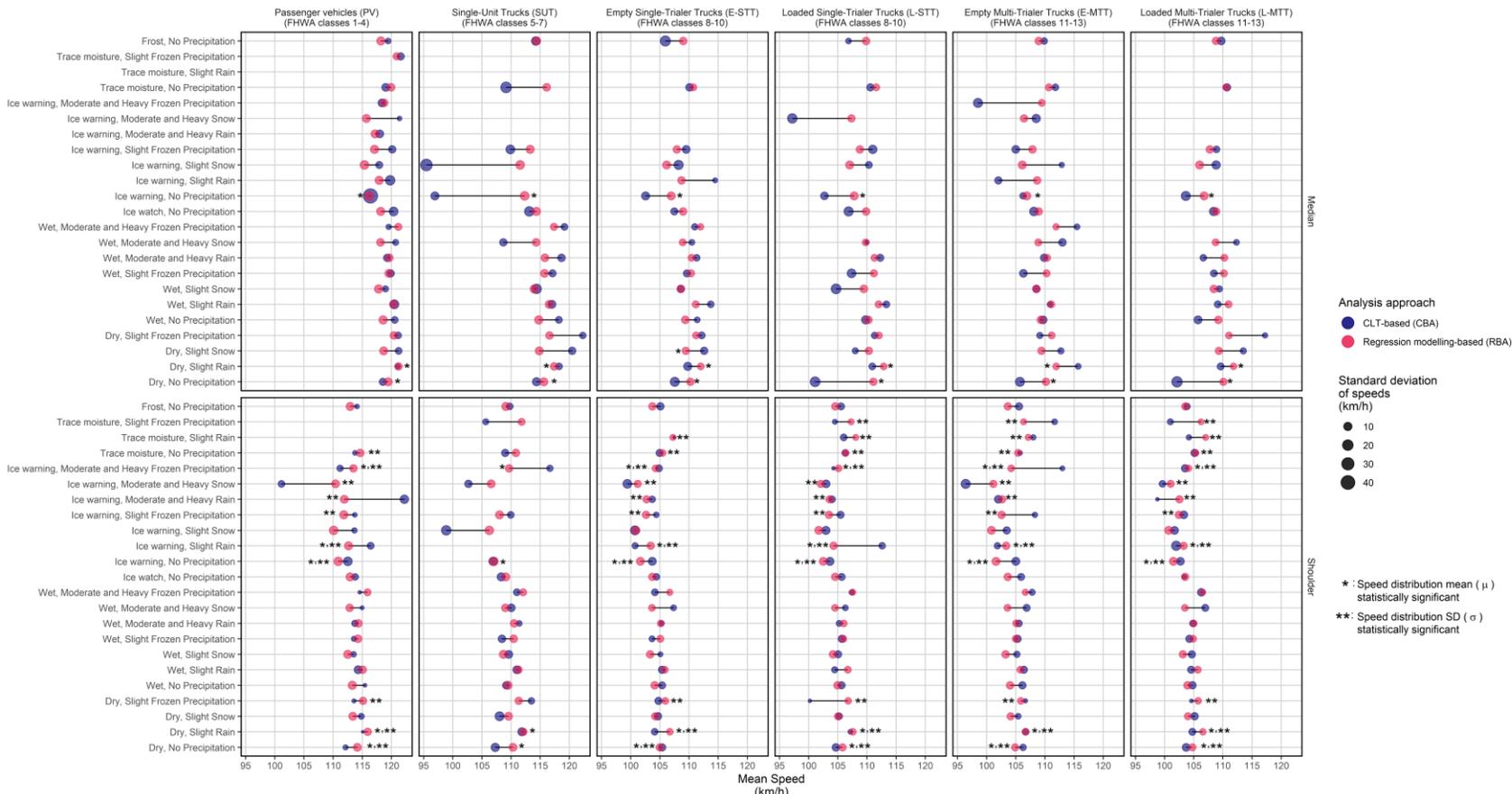


Figure 3.7: Speed distribution parameters
(Yasanthi & Mehran, 2022)

According to Figure 3.7 as per CBA estimations, TTC travelling in median lane are described by considerably low μ as compared to TTC travelling in shoulder lane under extreme road-weather conditions (e.g., μ of L-STT travelling in median lane on ice warning pavements and moderate and heavy snow is approximately 6km/h less than the μ of L-STT travelling in shoulder lane under similar road-weather conditions). Although counter-intuitive, such low values of μ imply that truck drivers are vigilant while driving under inclement road-weather conditions irrespective of their travel lane. Yet, TTC travelling in median lane seem to choose their speed in a wide range, reflected by relatively high SD values characterizing most speed distributions irrespective of the speed-decisive factor combination considered. As anticipated, PV travelling in median lane are estimated to drive faster than their counterparts travelling in shoulder lane under most road-weather conditions (e.g., PV travelling on ice warning pavements and moderate & heavy snow).

Vehicle type, in particular, notably impact drivers' speed choice. While the impact of vehicle type on σ , as estimated by both RBA and CBA, is marginal under normal road-weather conditions, TTC combinations are characterized by considerably low values of μ in comparison with PV travelling under such road-weather conditions. For instance, both RBA and CBA estimations suggest that L-STT travelling in shoulder lane on wet pavements and moderate and heavy snow will travel approximately 7km/h less than PV travelling under the same driving conditions. Similarly, both RBA and CBA estimated comparatively low values of μ for all vehicle types (except for SUT) travelling in shoulder lane on ice warning pavements with no precipitation. The increase in σ is marginal for the aforementioned combination of speed-decisive factors. For instance, according to CBA estimates, σ for PV travelling in shoulder lane on ice warning pavements with no precipitation was 3.7km/h greater than that for L-MTT travelling in shoulder lane under identical road-weather conditions. However, even a small increase in σ imply that drivers perceive the risk of adverse pavement conditions differently. Although marginal in most cases, loaded trucks appear to have higher μ under most road-weather conditions. For instance, μ estimated according to CBA for E-STT travelling in shoulder lane under slight rain and ice warning pavements ($\mu = 100.8\text{km/h}$) is considerably lower than that of L-STT ($\mu = 112.6\text{km/h}$). It is possible that truck drivers tend to drive faster irrespective of prevailing road-weather conditions, when delivering a freight load. According to the CBA, L-MTT in median lane under dry pavements and slight frozen precipitation, are estimated to drive above the speed limit on average. This is

rather an important observation because such speed choices may lead to additional severe consequences (e.g., load spill) in the event of a collision.

According to Figure 3.7, the data analytic approach adopted notably affects speed distribution parameters, particularly in the simultaneous presence of adverse road-weather conditions. For example, while the RBA estimated a considerably low μ (104 km/h) for E-MTT travelling in shoulder lane under ice warning pavements and moderate & heavy frozen precipitation, the CBA estimated a considerably high μ (113 km/h). Yet, it should be noted that the sample size for the aforementioned combination of speed-decisive factors is very limited (i.e., $n=16$), which potentially explains the substantial difference in estimation. Further, it is important to recall that the two approaches-RBA and CBA, had adopted fundamentally different analysis approaches while modelling speed distributions. While both approaches use identical source of data, the differentiation between the definition of a sample used in each approach potentially contribute towards the differences observed in speed distribution estimates. The CBA, in a way, subdivides the sample used in the RBA into smaller groups to develop sampling distributions which are highly representative of a population. Thus, while RBA infers μ and σ based on samples (which might not be representative of their respective populations), the CBA benefits from sampling distributions while inferring μ and σ . On the other hand, the statistically insignificant model coefficients reflected in study results question the application of RBA, particularly under limited sample size conditions due to the possibility of type II error associated with small sample sizes. In contrast, the CBA provides a robust approach to model speed distributions by optimizing the representativeness of a sampling distribution of means to its population distribution and thus eliminating problems (e.g., statistically insignificant model coefficients) associated with applying RBA in limited sampling conditions. Accordingly, the CBA is deemed as a more appropriate approach to infer μ and σ under different combinations of speed-decisive factors.

3.6.2 Road-weather Conditions with Increased Collision Potential

Speed distributions modelled according to the CBA were used to model the distributions of within- and across-lane speed differentials for 135 combinations of road-weather, and vehicle type categories. K-means clustering labelled each combination of road-weather, and vehicle type category based on the parameters of the distributions of speed differentials (Table 3.2). It should be noted that the collision risk level was assigned to each cluster based on the SD of speed

differential distributions due to the marginal difference in means of the speed differential distributions among the three clusters (Table 3.2).

Table 3.2: Details of the Three Cluster Centroids

Collision Risk	Within-lane speed differentials distribution parameters				Across-lane speed differentials distribution parameters	
	Median		Shoulder		Mean (km/h)	Standard deviation (km/h)
	Mean (km/h)	Standard deviation (km/h)	Mean (km/h)	Standard deviation (km/h)		
High	2.12	44.24	3.37	15.06	3.89	41.97
Moderate	5.78	21.06	1.89	11.07	3.31	14.17
Low	3.96	13.86	1.86	6.96	5.86	3.07
BSS ¹ /TSS ²	80.4%		84.3%		92.1%	

Notes:

¹BSS=Between Sum of Squares

²TSS=Total Sum of Squares

Figure 3.8 demonstrates the collision risk levels estimated for different combinations of road-weather, and vehicle type categories in terms of within-lane, across-lanes, and overall collision risks (see Figure 3.5 for estimation of overall collision risks). The silhouette coefficient: a measure evaluating the compactness or separation of the clusters, typically ranges from zero to one; a silhouette coefficient of zero indicates poorly separated clusters whereas one indicates properly assigned clusters (Hackeling, 2017). The high silhouette coefficients (shown in brackets) pertaining to each classification implies that collision risks are well clustered. Figure 3.8 highlights important implications. First, the within-lane collision risks in shoulder lane seem divergent as compared to the within-lane collision risks present in the median lane, particularly for SUT. Presence of such comparatively low within-lane collision risks in the median lane is rather intuitive because, unlike shoulder lane, median lane in a rural highway is predominantly used for passing with more homogeneous passing speeds. In contrast, shoulder lane is used by different vehicles (Figure 3.3) potentially travelling at their desired speed; thus, shoulder lane entails higher within-lane collision risks, particularly for SUT travelling under adverse road-weather conditions. Number of trailers attached to trucks seem to reduce the within-lane collision risk for trucks

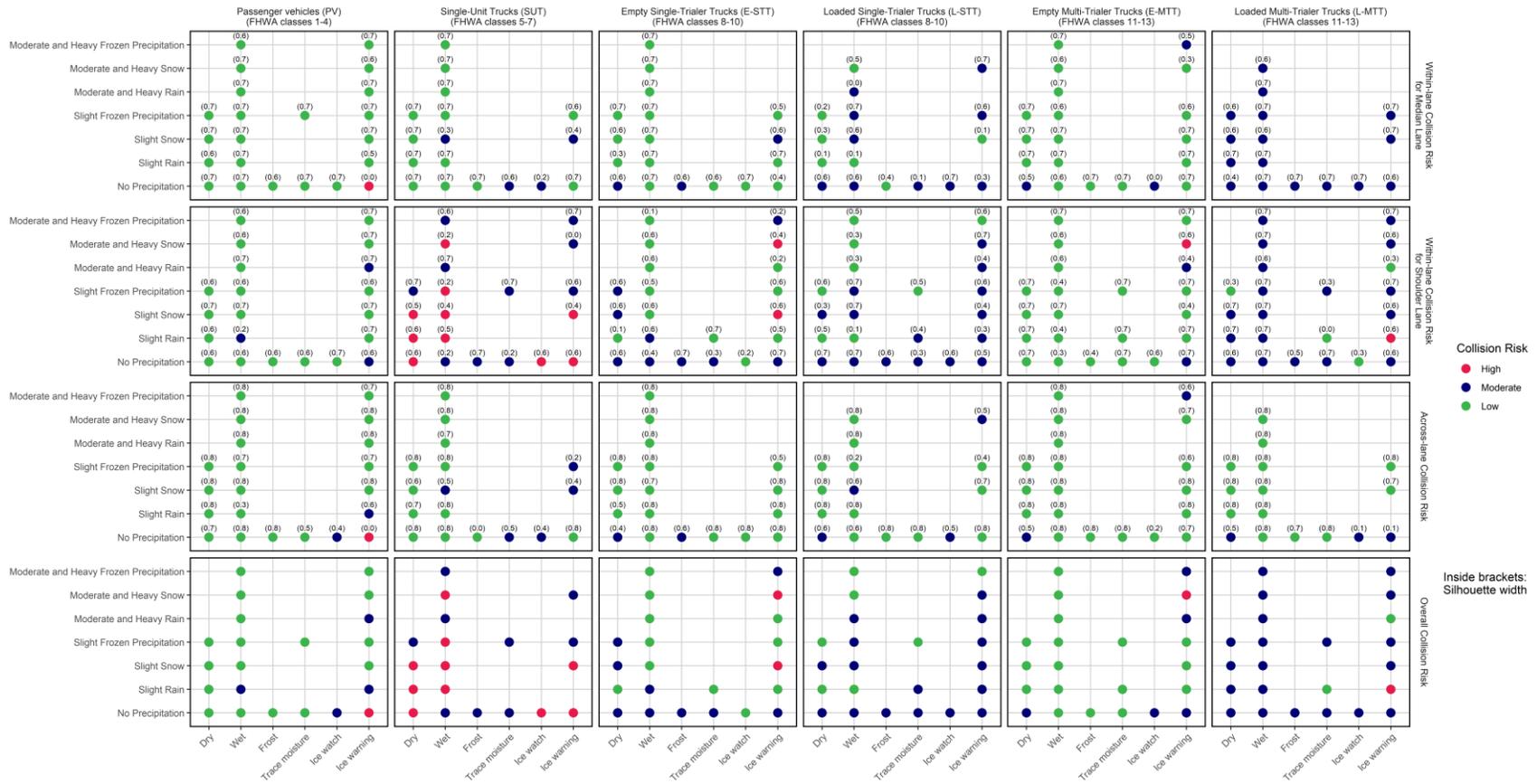


Figure 3.8: Vehicle-specific collision risks estimated for different road-weather conditions (Yasanthi & Mehran, 2022)

irrespective of the travel lane, and prevailing road-weather conditions. For instance, the within-lane collision risks for SUT (zero trailers), E-STT, and E-MTT travelling in shoulder lane under normal road-weather conditions are estimated as high, moderate, and low respectively. In fact, past literature (Regher et al., 2009) suggests that longer combination vehicles (i.e., multi-trailer trucks), when evaluated based on collision rate, are safer than other articulated trucks (i.e., single-trailer trucks). Overall, truck payload condition increases within-lane collision risk, particularly for multi-trailer trucks. Unlike for PV, the across-lane collision risk for all types of trucks appears to be low under most road-weather conditions. Trucks predominantly occupy the shoulder lane, which potentially explains the low across-lane collision risk for trucks. The only high across-lane collision risk is present for PV travelling under no precipitation, and ice warning pavements. The comparatively low across-lane collision risks may be attributed to the predominant usage of shoulder lane in rural highways.

According to Figure 3.8, the within-lane collision risk consists of a diverse mix of low, moderate, and high collision risks. In contrast, the across-lane collision risks are described by either low or moderate collisions risks except for PV travelling under no precipitation and ice warning pavements. Thus, overall risk mainly reflects the within-lane collision risks due to the low variation of across-lane collision risks irrespective of the road-weather condition and vehicle type. According to Figure 3.8, trucks (particularly SUT and L-MTT) entail considerably higher collision risks as compared to PV under most road-weather conditions. Such high overall collision risks confirm the vulnerability of freight transport vehicles in comparison to PV. In fact, icy pavement conditions seem to intensify the overall collision risk particularly for TCC. For example, the overall collision risk is estimated to be low for empty TTC travelling on wet pavements under moderate & heavy snow; yet, the collision risk is high when travelling on ice pavements. Importantly, travelling loaded appears to increase collision potential for both TTCs under most road-weather conditions.

In summary, the results of this research unveil important findings. First, the CBA, in comparison with RBA, is deemed as a more appropriate approach to model speed distributions under different driving conditions. In fact, this thesis does not recommend the RBA to model speed distributions particularly under limited sampling conditions. Second, results of this research support the hypothesis that drivers' speed is crucially affected by the prevailing road-weather, travel lane, and

vehicle type. Yet, the impact of truck payload condition on truck speeds seems to be marginal under most road-weather conditions. Although the across-lane collision risk seems marginal for most road-weather conditions, the notable within-lane collision risks signal the need for advanced speed management techniques. One potential approach to minimize within-lane collision risks is to differentiate fast-moving vehicles from slow-moving vehicles in each lane under adverse road-weather conditions. This could be achieved by implementing Lane-based Weather-responsive Variable Speed Limit (L-WR-VSL) systems together with lane advisories for slow moving vehicles. Such L-WR-VSL systems attempt to minimize speed variation induced by different road-weather, vehicle type, and travel lane conditions by imposing unique, differentiated per-lane speed limits under different road-weather conditions (Shi & Liu, 2019). For example, imposing a comparatively low speed limit for shoulder lane under snowy conditions constrains drivers who are not comfortable to drive in high speeds to shoulder lane; it further enables fast drivers driving in inner lanes to maintain high speeds without being constrained by slow-moving vehicles (Shi & Liu, 2019).

3.7 Conclusion & future directions

Inclement road-weather undoubtedly impose notable collision risks to highway freight transportation. Yet, how to identify road-weather conditions posing crucial truck collision risks remains a critical question. On the other hand, while past studies recommend modelling context-specific speed distributions to evaluate the impacts of different driving conditions on speed, the most appropriate data analytic approach to model such speed distributions needs to be evaluated. To address this question, this research (i) evaluates the combined impact of four speed-decisive factors: road-weather, travel lane, vehicle type, and truck loading condition (empty or loaded) on drivers' desired speed, (ii) explores the variability caused by adopting a regression-based approach (RBA) and a central limit theorem-based approach (CBA) to model speed distributions, and (iii) proposes a unified approach to identify road-weather conditions imposing higher collision risks based on both within- and across-lane collision potential.

The study results recommend modelling speed distributions using the CBA as compared to evaluating context-specific speed choice according to the RBA. Further, findings of this research support the hypothesis that truck speeds are mainly affected by road-weather conditions, travel lane, and vehicle type. Particularly, the impact of road-weather conditions was crucial when

adverse precipitation, and pavement conditions are present simultaneously. Travelling loaded appears to marginally increase truck speed under most road-weather conditions. Further, the study results suggest low to moderate across-lane collision risks for all truck types irrespective of the prevailing road-weather conditions. In contrast, the within-lane, and overall truck collision risks are alarmingly inflated, particularly under adverse road-weather conditions. Such collision risks emphasize the importance of implementing proactive countermeasures to mitigate potential truck collisions under adverse road-weather conditions. Thus, this research recommends evaluating the suitability of advanced speed management strategies such as lane-based weather-responsive variable speed limits for cold-region rural highways.

Contributions of this research are mainly threefold. First, jurisdictions located in extremely cold regions may adopt the methodology to evaluate the combined impact of different speed-decisive factors on drivers' speed. Understanding truck drivers' speed choice is also important to the trucking industry, as drivers' speed choice reveal their speeding patterns to some extent. Second, the collision risks estimated by this research manifest the vulnerability of trucks under different combinations of road-weather conditions, particularly when loaded. Thus, transport authorities may apply the proposed methodology to form the basis for important policy implications such as weather-specific truck weight restrictions to mitigate truck collisions potentially leading to severe consequences. In addition, trucking companies may also use the study results to strategize their freight operations based on road-weather conditions, using the road-weather specific truck collision risks unveiled by this research. For instance, transporting freight loads only in road-weather conditions with low collision risks may help trucking companies to minimize weather-related risks imposed on their combination trucks. Further, transportation authorities may use the study results to explore the need for advanced weather-responsive speed management systems. Third, the population-level speed distributions can be used for calibration of traffic simulation models considering different vehicle types and road-weather conditions, while overall collision risks estimated in this research can be used as surrogate safety measures when collision data are not available.

Despite the important implications associated with this research, adoptability of the proposed methodology is limited to uncongested, four-lane divided highways. Adopting the methodology in congested highways violates the two basic assumptions: (i) observed speeds

represent drivers' desired speeds, and (ii) speed distributions are represented by the normal distribution, formulated while developing speed distributions. First, congested traffic entails frequent vehicle interactions. In such traffic conditions, drivers' speed choice is notably affected by the speed of surrounding vehicles. Thus, speeds observed in congested traffic conditions may not reasonably represent drivers' desired speeds. Second, speed distribution under congested traffic conditions might not be represented by normal distribution (Jun, 2010). The overall collision risks estimated in this research are based only on two lanes; thus, the methodology may not be adopted for multilane highways with more than two lanes in one travel direction. Accordingly, future research may focus on extending the proposed methodology to congested traffic, and multi-lane conditions. The study results highlight the limitation of linear regression to model speed distributions under limited sampling conditions; thus, future research may focus on evaluating the minimum sample size required to model speed distributions using regression modelling. Future research may also focus on validating the overall collision risks estimated in this research using collision statistics associated with similar study site characteristics. In addition, future research may focus on empirically validating CBA using extensive datasets collected in various highway locations with similar features. While this could be challenging due to limited availability of RWIS stations in the vicinity of traffic detectors, implementation of connected vehicle technologies to collect required speed and road-weather data at regional scale should be explored further.

Chapter 4. A Reliability-based Weather-Responsive Variable Speed Limit System to Improve the Safety of Rural Highways

This chapter develops an effective approach to set WRVSLs to improve the safety of rural highways using the reliability theory. The content presented in this chapter corresponds to the last step of the WRVSL development approach (Figure 1.3). Transport authorities located in extremely cold jurisdictions may adopt the proposed methodology to set WRVSLs in their rural highways. Although this chapter develops WRVSLs for all RWCs observed at the study site, in practice, it is recommended to prioritize setting WRVSLs for RWCs with notable safety risks (step 2 in Figure 1.3) particularly when resources (e.g., human resources, funds) are limited.

4.1 Abstract

Weather-responsive Variable Speed Limit (WRVSL) systems treat speed limits as weather-dependent random variables, as opposed to the conventional static speed limits. This research (i) evaluates drivers' response to a fixed speed limit in different road-weather conditions, and (ii) proposes an effective approach to set WRVSLs, for rural divided highways located in extremely cold regions. Study data: road-weather, and speed data, collected from a rural highway (fixed speed limit=110km/h), are used to (i) estimate the 85th percentile speeds of population-level speed distributions, and (ii) develop WRVSLs based on the reliability theory. More specifically, the WRVSLs are set based on reliability: the probability of a speed being (i) likely complied by drivers, and (ii) adequate to avoid a rear-end collision. The study results reveal that merely 73% of the drivers at the study site comply with the existing posted speed limit under normal road-weather conditions i.e., no precipitation and dry pavements. The reliability of the current speed limit is revealed to be approximately one under normal road-weather conditions; thus, the current speed limit is perceived credible under such road-weather conditions. Yet, reliability of the current speed limit is substantially reduced in the presence of slight snow, and ice warning pavement conditions. A set of reliability-based WRVSLs ranging from 80 to 110 km/h is proposed. Jurisdictions experiencing extreme road-weather conditions may adopt the proposed methodology to effectively manage speed, particularly in rural highways under adverse road-weather conditions to enhance the probability of speed limits being safe and complied by drivers and as a result reduce crash propensity.

Keywords: Variable speed limit; adverse road-weather; rear-end collision; rural highway

4.2 Introduction

Weather-responsive variable speed limit (WRVSL) systems consist of differentiated speed limits, implemented under different road-weather conditions (RWC) (Downey, 2015). More specifically, WRVSL systems treat desired speed (hereinafter speed), thus speed limit (SL) as a weather-dependent random variable rather than setting SLs as fixed deterministic values established irrespective of the prevailing RWC (Downey, 2015). WRVSL systems attempt to regulate speeding i.e., driving above the speed limit (SL) or driving too fast for the prevailing driving conditions (Forbes et al., 2012), particularly under inclement RWC. Speeding under adverse RWC is perceived hazardous due to two reasons. First, in theory, braking distance is directly proportional to the square of speed (Garber & Hoel, 2014); thus, speeding is associated with longer stopping sight distances (SSD). On the other hand, adverse RWC such as heavy precipitation and icy pavements lead to degraded driving conditions described by limited visibility (OECD, 1976) and reduced pavement friction (Ivey et al., 1975); thus, SSDs are further elongated under adverse RWC. Accordingly, speeding in the presence of adverse RWC undoubtedly jeopardize road safety particularly in rural highways, where speeding is prevalent due to low traffic throughput (Yasanthi & Mehran, 2020). Second, drivers' speed choice is essentially idiosyncratic (Yasanthi et al., 2021; Anastasopoulos & Mannering, 2016), and is argued to be dependent on drivers' level of confidence, and comfort to travel under the prevailing driving conditions (Yasanthi et al., 2021; Hauer, 1971). In fact, desired speed distributions under extreme RWC are generally represented by (i) comparatively lower mean speeds, and (ii) substantial speed variations, with respect to normal RWC i.e., no precipitation and dry pavements (Yasanthi et al., 2021). Such versatile RWC-specific speed patterns signal (i) stimulated crash propensity under adverse RWC (Xing et al., 2019; Solomon, 1964), and (ii) drivers' ambiguity in appraising driving externalities while selecting a safe travel speed (Forbes et al., 2012). Accordingly, WRVSL systems attempt to communicate commuters with a safe speed to travel under different RWC, thus promoting road safety particularly under adverse RWC (Downey, 2015).

This research aims to improve traffic safety in rural highways located in cold regions by proposing a methodological approach to set effective WRVSLs using the reliability theory. Two research questions are addressed in this research: (i) Is drivers' response to a fixed SL dependent on RWCs?

and (ii) What is an effective way to set a credible WRVSL on rural highways under extreme RWCs? The research questions are addressed through two research objectives: (i) To evaluate drivers' response to a fixed SL under different combinations of RWCs, and (ii) To propose a reliability based WRVSL system to regulate speed under adverse RWCs. The research methodology is divided into two phases. First, drivers' response to a fixed SL is evaluated using the 85th percentile speed of population-level speed distributions under different RWCs. Second, a set of WRVSLs is developed based on the reliability theory. More specifically, this research identifies WRVSLs, based on the probability of each WRVSL being (i) adequate to avoid a rear-end collision, and (ii) likely compliant by drivers, under a given combination of RWCs. The research methodology is demonstrated using road-weather and traffic data collected from a rural highway in Alberta, Canada.

The key contribution of this research is that it provides an effective approach to implement WRVSL systems in rural highways located in cold regions. The proposed WRVSL-setting approach focuses on collision mitigation, and SL compliance. Thus, practical application of the proposed methodology may (i) require lesser SL enforcement efforts, and (ii) mitigate rear-end collisions, in extreme RWCs.

4.3 Literature review

Speed limit: the maximum (or minimum) threshold vehicular speed allowed by law in a particular highway segment (AASHTO, 2009), attempts to improve road safety by regulating drivers' speed choice. Highways are often legislated by maximum SLs set according to one of the four universally adopted SL setting approaches (Forbes et al., 2012): (i) engineering approach, (ii) expert system approach, (iii) optimization, and (iv) safe system approach (SSA). Depending on the consistency of operation, SLs can be divided into two types: fixed, and variable. As opposed to the fixed SL systems, variable SL systems implement dynamic SLs set based on the prevailing driving conditions (Forbes et al., 2012).

Engineering approach — the most common methodology used to set SLs worldwide—establishes SLs based on (i) the 85th percentile free-flow speed (Gayah et al., 2018), or (ii) road characteristics such as adjacent land use and/or anticipated traffic conditions (Forbes et al., 2012). The 85th percentile free-flow speed (V_{85}): speed at or below which is adopted by 85 percent of drivers (FHWA, 2009), is widely used in setting SLs particularly in rural highways (NRC-US, 1998)

particularly in North America (Forbes et al., 2012). The V_{85} is frequently used to set SLs due to two reasons: (i) travelling at or near V_{85} yields minimum collision risks for drivers (Solomon, 1964), and (ii) V_{85} represents the collective judgment of drivers (Forbes et al., 2012). According to the Manual on Uniform Traffic Control Devices for Streets and Highways (MUTCD), a SL set based on the V_{85} is acquiescent if V_{85} is within 8km/h (~ 5 mi/h) of the posted SL (FHWA, 2009). Yet, adverse RWCs are typically described by substantial speed variabilities indicated by comparatively high standard deviation values of speed distributions (Yasanthi et al., 2021). Standard deviation is directly proportional to V_{85} . Accordingly, V_{85} in some adverse RWCs may be greater than the SL in effect. For instance, Yasanthi et al. (2021) modelled the speed distribution for dry pavement, temperature values between 0°C and -10°C (Group II), nighttime and slight frozen precipitation conditions which was described by a mean of 115 km/h and a standard deviation of 7.49 km/h (Figure 2.12); the V_{85} for the above combination of RWCs is 122 km/h which is considerably higher than the speed limit at the study site (i.e., 110 km/h). Thus, implementing a fixed V_{85} -based SL may violate the MUTCD criterion for an acceptable SL under some RWCs. Accordingly, SLs for such RWCs should either be lowered, or increased to complement the V_{85} of corresponding speed distributions. Yet, setting very high SLs, specially under inclement RWCs, undeniably jeopardize traffic safety. On the other hand, empirical evidence suggests limited compliance to SLs on rural highways (Johnson & Murray, 2010). Thus, however widespread, setting SLs based on the V_{85} is questioned in the literature mainly because (i) some drivers may not efficiently identify the safest speed to drive particularly under degraded driving conditions (NRC-US, 1998), and (ii) using V_{85} as the SL may escalate vehicular operating speeds (i.e., the speeds selected by drivers to operate their vehicles under free-flow conditions (AASHTO, 2018)) in the long term (Hauer, 2009).

To address this issue, the expert system approach uses algorithms (e.g., USLIMITS2) which takes multiple inputs (e.g., V_{85} , crash history, etc.) to mimic the experts' thought process in setting SLs (Forbes et al., 2012). Yet, such algorithms are still in their infancy with exclusive focus on fixed SLs (Fitzpatrick et al., 2019). Although well-established, optimization: setting SLs based on the optimal speed possessing minimum total societal cost (e.g., crash costs, travel time costs), is seldom used due to the difficulty in accurate estimation of societal costs (Forbes et al., 2012). Speed limits set according to the SSA are designed to provide adequate SSD resulting minimal crash and injury potential (Arason, 2019). Crash injury severity, however, is dependent on the

kinetic energy released in a crash event which is directly proportional to a vehicle's operating speed (Donnell et al., 2018; Alhajyaseen, 2015). Thus, injury minimization approaches in setting SL such as SSA often output SLs which are typically lower than the conventional SLs (Forbes et al., 2012); thus, not perceived suitable for rural roads (NRC-US, 1998).

Adverse RWCs are often described as vulnerable driving conditions (Shaheed et al., 2016) due to the increased speed variability present in such RWCs (Yasanthi et al., 2021). Yet, the intensities of such speed variabilities are dependent on the prevailing RWCs (Yasanthi et al., 2021). On the other hand, substantial speed variabilities result increased crash potential (Solomon, 1964). Thus, frequent occurrence of adverse RWCs warrant variable SLs differentiated by the prevailing RWCs i.e., WRVSLs (Downey, 2015). Yet, in practice, the existing WRVSL are reported to be ineffective (Downey, 2015; Elvik et al., 2009) despite the methodology adopted in setting such SLs (Sisiopiku, 2011). Potential reasons for documented inefficacies are twofold. First, most existing WRVSL are developed based on sample-level statistical measures (e.g., V_{85} of RWC-specific speed samples). Yet, past literature argues the credibility of using such sample-level appraisals in weather-responsive speed management (WRSM) due to limited sampling conditions prevalent under adverse RWC (Yasanthi et al., 2021). Second, although the effectiveness of a SL is heavily dependent on SL compliance (Hellinga & Mandelzys, 2011), most SL setting approaches attribute a notably low weight to SL compliance while setting the SLs. To address these issues, this research proposes to incorporate (i) population-level speed distributions rather than establishing SLs based on an individual sample-level statistical measure, (ii) crash mitigation measures, and (iii) SL compliance, when setting WRVSL. More specifically, a methodology is proposed to set distinctive SLs for different combinations of RWCs by incorporating the concept of reliability with the (i) maximum speed a vehicle can travel avoiding a rear-end collision, and (ii) compliance of drivers to the SL in effect at the time of their travel (compliance hereinafter). Thus, the safety- and compliance-driven SL setting approach proposed in this research could effectively improve road safety, particularly in cold region rural highways. An additional advantage of the proposed methodology is that it allows evaluating the performance of a particular proposed WRVSL prior to practicing, by considering the probability of such a SL scheme being safe and complied by drivers simultaneously under given RWCs. Table 4.1 presents a summary of literature review, identified research gaps, and research gap fills proposed in this research.

Table 4.1: Summary of literature review

SL setting approach		Strength(s)	Shortcoming (s)/ research gap (s)	Proposed research gap fill (s)
Engineering	Operating speed method (SL set based on V_{85})	Reflects actual travel speeds; thus, minimal enforcement required (Forbes et al., 2012)	Identifying a safe speed based on drivers' judgment is questionable (Forbes et al., 2012)	Safe speed defined in terms of the minimum SSD required to prevent a rear-end collision
	Road risk method (SL set based on road characteristics)	Considers functional classification of road; thus, agrees with the road design (Forbes et al., 2012)	Typically set below V_{85} ; thus, heavy enforcement required to ensure compliance (Forbes et al., 2012)	Population-level speed distributions used to set SLs
Expert system		Multiple factors considered while setting SLs (Forbes et al., 2012)	Not applicable to setting variable SLs (Forbes et al., 2012; Fitzpatrick et al., 2019)	WRVSLs determined based on the reliability theory
Optimization		Incorporates several aspects of speed on society including societal costs (Forbes et al., 2012)	Difficult to estimate societal costs accurately (Forbes et al., 2012)	Proposed SL setting approach attempts collision prevention, which eliminates the need to estimate societal costs while indirectly reducing crash cost
Safe system		Prioritizes on serious crash/injury prevention (Forbes et al., 2012)	Typically yield low SLs; thus, not appropriate for rural highways (Forbes et al., 2012)	<ul style="list-style-type: none"> • SL setting approach demonstrated based on a rural highway • A range of predefined speeds considered when setting WRVSLs, enabling practitioners to select a rational speed as the SL

4.4 Study data

4.4.1 Data collection

Study data: traffic and road-weather data used in this research were collected over 15 months (October 2014 to December 2015) from a weigh-in-motion (WIM) sensor and a road-weather information system (RWIS). The WIM and RWIS devices are stationed alongside each other at the study site

(Figure 4.1) located on highway 16: an east-west four-lane rural highway located in Alberta, Canada. The study site operates under a fixed SL of 110 km/h. The data collection site's topography can be described as flat with negligible slope and no sharp horizontal curves within a radius of 5 km (Figure 4.1).

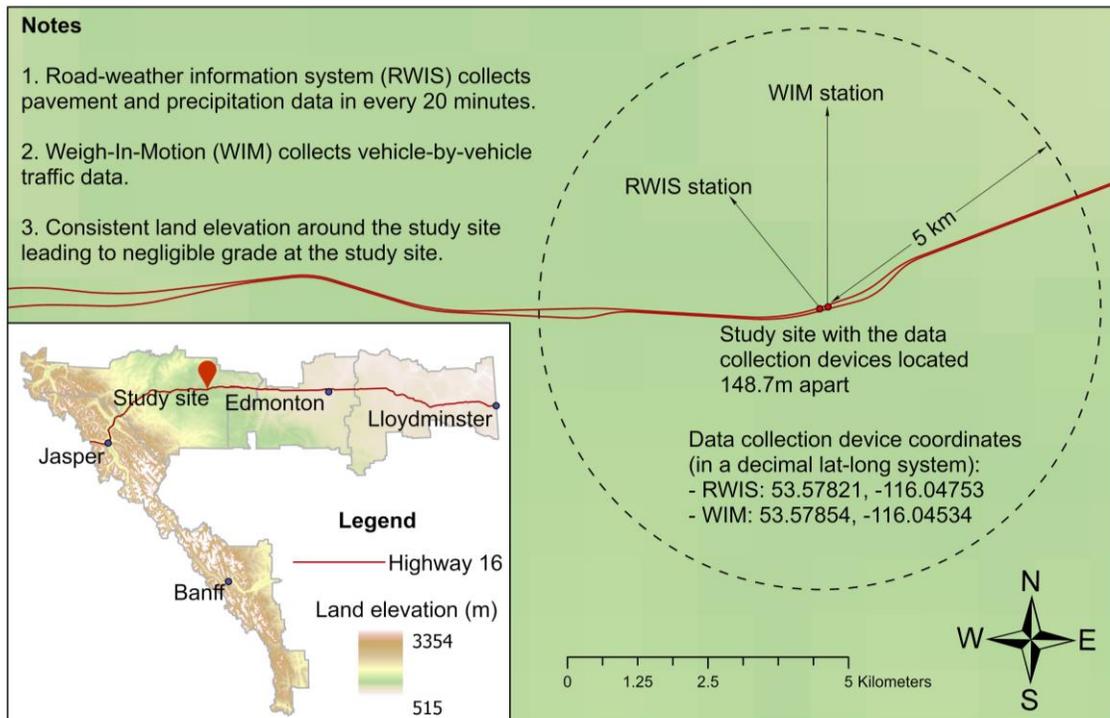


Figure 4.1: Study site (Yasanthi et al., 2022)

In total, six pavement surface conditions: dry (no ice or moisture detected on pavement), wet (considerable amount of moisture on the entire pavement), trace moisture (moisture detected on parts of the pavement), ice watch (potential ice formation detected), ice warning (ice present on pavement), and frost (frost present on the pavement), were observed at the study site. Further, three

precipitation types: snow, rain, and frozen precipitation, were observed in two precipitation intensity levels: (i) slight (intensity $<2\text{mm/h}$), and (ii) moderate and heavy (intensity $\geq 2\text{mm/h}$). Traffic records collected by the WIM station includes (i) spot speed, (ii) date and time, (iii) axle weight, and (iv) inter-axle spacings, for each vehicle passing the traffic counter.

4.4.2 Preliminary data analysis

The impacts of travel direction (i.e., eastbound, or westbound for this research) on speed is documented as negligible under most RWCs at the study site (Yasanthi & Mehran, 2020). In fact, geometric conditions at the study site are consistent in both travel directions (Figure 4.1). Therefore, there is no obvious difference in the speed distributions in the east or west direction. Thus, only eastbound traffic data were used in this research. Further, this research only considers shoulder-lane-traffic due to its (i) divergent traffic composition implied by substantial hourly heavy vehicles' percentages, and (ii) high hourly lane utilization factors which manifest drivers' preference to travel in shoulder lane, at the study site (Figure 2.4). A Kolmogorov–Smirnov (K-S) test was conducted to statistically evaluate the independence of desired speeds in each RWC category. Underlying reasons for choosing K-S test to evaluate the statistical difference of speeds in different attribute levels are presented in section 2.4.2. According to the K-S test results (see Figures A1.1 and A1.2 presented in Appendix A1), desired speed in all road-weather attribute categories are statistically independent except for some cases including the precipitation intensity levels “moderate” and “heavy”. Therefore, the precipitation intensity levels “moderate” and “heavy” were combined into one category: “moderate & heavy”.

Frequent, and fast transitioning of SLs, particularly in a WRVSL system, lead to recurrent speed recoveries (Katz et al., 2017). In fact, implementing excessive number of WRVSLs may decrease drivers' compliance. Thus, the number of unique WRVSLs (i.e., the number of RWCs considered) implemented in WRSM needs to be prudently selected (Katz et al., 2017). Accordingly, the study data were first filtered to include RWCs with notable safety risks. While the presence of solitary RWCs (e.g., icy pavements with no precipitation) impose impacts on drivers' speed choice (Yasanthi et al., 2020; Kyte et al., 2000), it is reported that the synchronicity of combinations of different RWCs (e.g., icy pavements and snow precipitation) crucially intensifies safety risks when estimated based on speed and speed variabilities (Yasanthi et al., 2021; Fountas et al., 2020). Particularly, combinations of adverse precipitation and pavement surface conditions are often

recognized as vulnerable (Yasanthi & Mehran, 2020) due to the substantial speed variations under such RWC combinations (Yasanthi et al., 2021). Therefore, only two RWC indicators: pavement surface condition, and precipitation condition, were considered in this research. The impact of time of the day (day/night) on drivers' speed choice at the study site was found to be marginal (Yasanthi et al., 2021). Further, road alignment of the study site does not include any special features such as steep grades or sharp curves (Figure 4.1). Therefore, this research does not consider the impact of road alignment as a factor influencing drivers' speed choice.

To explore the combinations of traffic and road-weather conditions at the study site, each five-minute interval during the study period was labelled with the (i) hourly equivalent of five-minute traffic volumes categorized into eight groups (bin size=100 veh/h), (ii) hourly equivalent of five-minute heavy vehicles percentages categorized into 10 groups (bin size=10%), (iii) precipitation condition, and (iv) pavement surface condition, reported within the five minutes. Figure 4.2 presents the number of RWC combinations observed under different traffic flow and heavy vehicle percentage categories at the study site. Implications of Figure 4.2 are twofold. First, the study site operates uncongested under all RWC combinations indicated by traffic flow values less than 800 veh/h. In fact, low to moderate traffic throughput i.e., traffic flow less than 1,000 veh/h, facilitates driving maneuvers with negligible vehicle interactions (TRB, 2016). Thus, each vehicle recorded by the WIM station corresponds to a driving maneuver with negligible vehicle interactions. In such free-flow settings, the observed speeds reflect drivers' speed choice under the prevailing traffic and road-weather conditions. Second, most RWC combinations (44 combinations) were observed operating with (i) traffic flow values between 100 veh/h and 200 veh/h, and (ii) heavy vehicles' percentages between 20% and 30%, which indicates the most prevalent traffic conditions at the study site. Effective WRVSL systems attempt to prioritize on safety risks induced by inclement RWCs. Therefore, it is important to distinguish the impacts of different RWCs on drivers' speed choice which could be obtained by analyzing desired speed under steady traffic conditions (Yasanthi et al., 2021). Accordingly, traffic records reported during the most prevalent traffic conditions were used in this research.

Figure 4.3 presents the descriptive statistics of speeds: mean and standard deviation (SD) of speeds, observed at the study site under different RWCs (sample size indicated within brackets). The RWC-specific means and SDs of speeds (Figure 4.3), confirm the impacts of prevailing RWCs

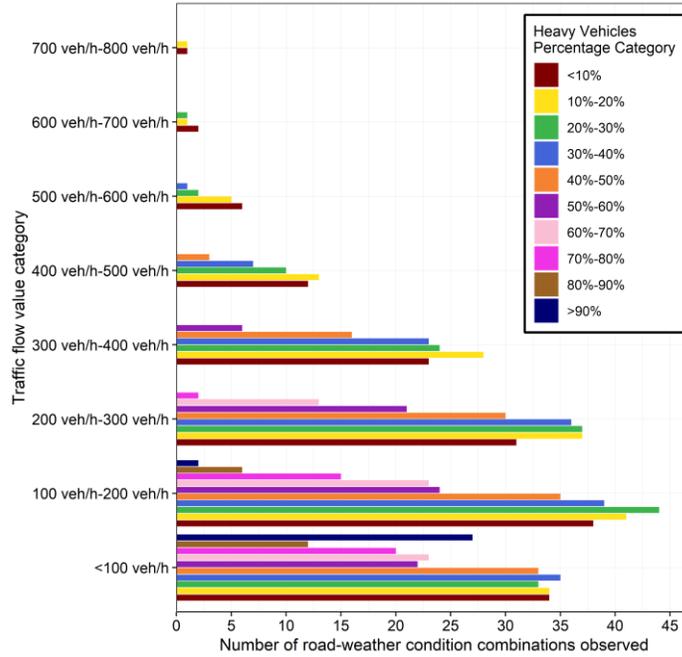


Figure 4.3: Study site traffic conditions (Yasanthi et al., 2022)

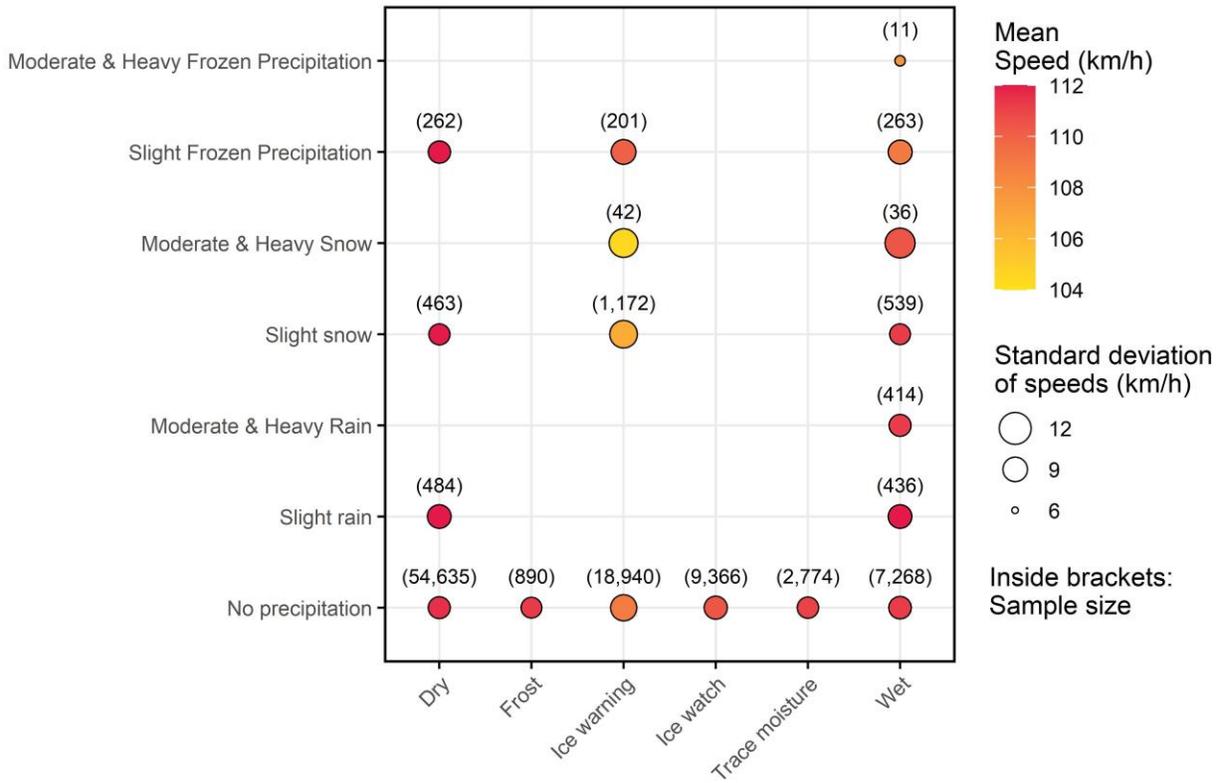


Figure 4.2: Descriptive statistics of speeds at the study site (Yasanthi et al., 2022)

on speed. For example, the comparatively low mean speeds and high SDs of speeds under ice warning pavements imply drivers' attentiveness under ice warning pavements. The lowest mean speed (104 km/h) was recorded under the combination of ice warning pavements, and moderate and heavy snow precipitation. Adverse precipitation conditions, particularly moderate and heavy snow, seemingly diversify drivers' speed choice, which is indicated by notably higher SDs of speeds observed under such precipitation conditions. The highest SD of speeds (11.2 km/h) was observed when moderate and heavy snow occurs in synchrony with wet pavements. Figure 4.3 also presents a few counter-intuitive observations potentially due to the limited sampling conditions. For instance, the minimum SD of speeds i.e., 6.1 km/h (Figure 4.3), was estimated based on 11 speed observations recorded under the concurrent presence of wet pavements, and moderate and heavy frozen precipitation conditions. While Figure 4.3 presents drivers' speed choice under different RWCs, it is important to note that sample-level speed statistics such as the descriptive statistics presented in Figure 4.3, may not effectively communicate the impacts of different RWCs on speed, particularly under limited sampling conditions (Yasanthi et al., 2021).

4.5 Methodology

4.5.1 Assumptions

Drivers' speed behaviour: the choice of speed among potential speeds (Letirand & Delhomme, 2005), implies (i) drivers' response to the existing SL (Yasanthi & Mehran, 2020), and (ii) the stochasticity of speed choice particularly under adverse RWCs (Yasanthi et al, 2021). Thus, understanding RWC-specific speed behaviour is a critical precedent in weather-responsive speed management. Drivers' speed behaviour, particularly under inclement RWCs, can be reasonably evaluated by exploring population-level speed distributions under such RWCs. This research proposes a methodology to set effective WRVSLs based on three main assumptions:

- i. A population-level speed distribution representing a unique combination of RWCs can be modelled as a normal distribution represented by a mean of μ , and a SD of (σ).

Population-level speed distributions are often modelled as normal distributions for rural highway segments which are not characterized by special properties such as distinctive geometric features (e.g., sharp curves) or diversified traffic (Yasanthi et al., 2021; García-Jiménez et al., 2016; Hashim, 2011).

- ii. The V_{85} of a population-level speed distribution pertaining to specific RWCs represents drivers' collective judgment of a safe speed under the considered road-weather attributes.

The V_{85} is often used as a measure in evaluating operating speed. The study site in this research operates under free-flow conditions with negligible vehicle interactions (Figure 4.2). Thus, assessing V_{85} under different combinations of RWCs at the study site reflects the preferred safe speed adopted by most drivers to travel under the prevailing RWCs.

- iii. A WRVSL can be modelled as a weather-dependent discrete random variable.

Desired speed, particularly under inclement RWCs, is often modelled as a discrete random variable (Yasanthi et al., 2021) due to the (i) idiosyncrasies of speed behaviour, and (ii) randomness of the presence of adverse RWCs. The process of setting an effective WRVSL attempts to identify the most appropriate speed among a set of such discrete random variables representing desired speed under different RWCs. Therefore, the candidate WRVSLs assume the properties of a discrete random variable; thus, can be modelled as weather-dependent discrete random variables.

4.5.2 Modelling methodology

The general framework of the proposed methodology (Figure 4.4) is divided into two phases. In phase I, drivers' response to the current fixed SL at the study site is evaluated under different combinations of precipitation, and pavement surface conditions. Phase II of the methodology attempts to propose an effective WRVSL system based on the reliability theory.

Fundamental steps of the proposed methodology to set a WRVSL for a given combination of RWCs is explained in the following section.

4.5.2.1 Step 0: Data collection and preparation

Setting WRVSLs require two types of data: (i) road-weather data, and (ii) traffic data. Traffic data required consists of vehicle-by-vehicle data. Road-weather data corresponds to the microclimatic environment for each vehicle observed. In terms of data preparation, first, the collected data are cleaned by removing erroneous data entries (e.g., "N/A", "Error", "Other" entries in road-weather data, and unrealistic vehicle speeds in traffic data i.e., speeds greater than 200 km/h in this

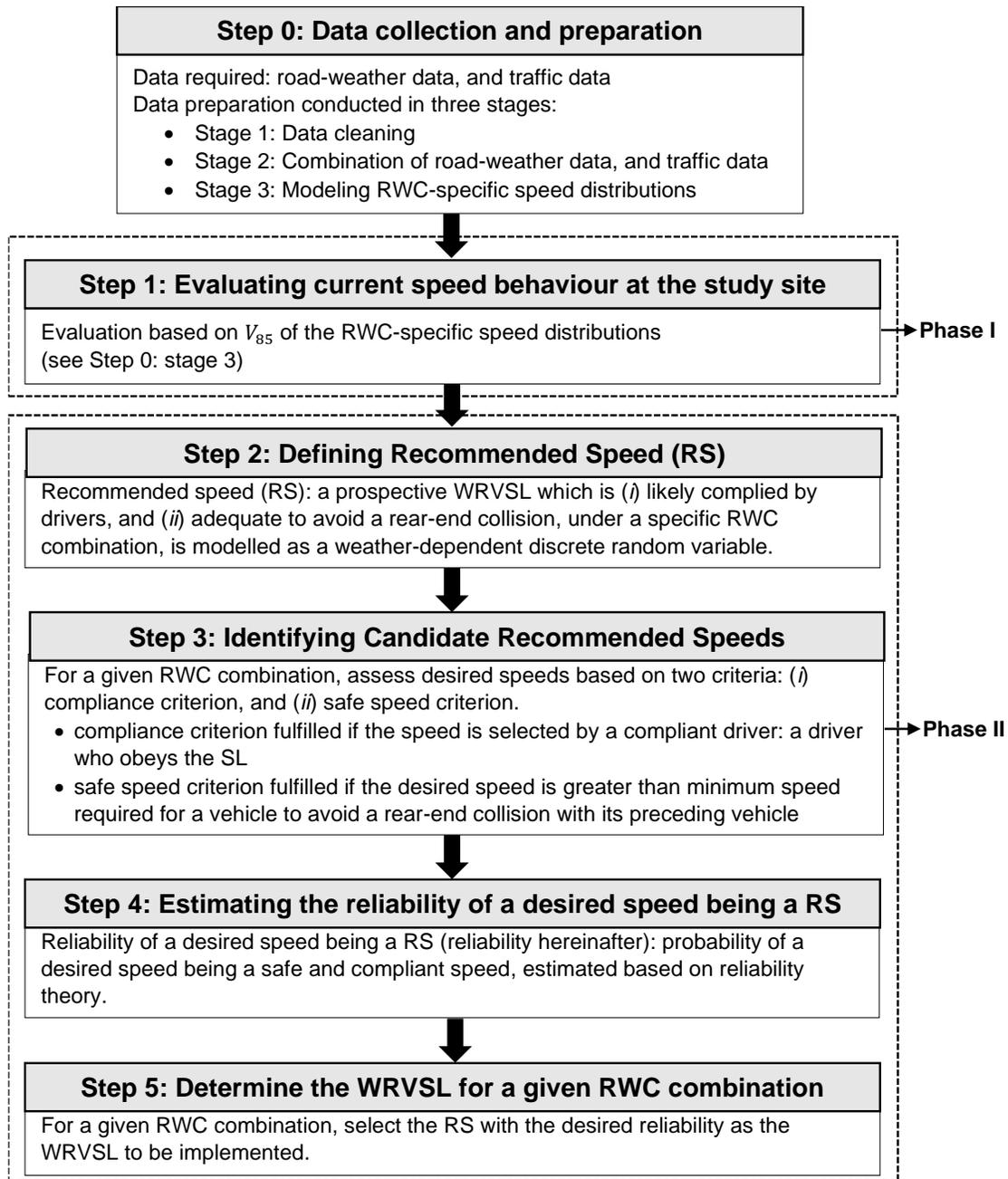


Figure 4.4: Proposed framework of the developed WRVSL system (Yasanthi et al., 2022)

research). Second, each traffic entry in the cleaned traffic dataset is assigned with the prevailing RWCs at the time of the traffic entry. Third, a methodology previously developed by authors (Yasanthi et al., 2021) is adopted in this research to estimate mean (μ) and standard deviation (σ) of a population-level speed distribution for a unique RWC combination. In brief, speed distributions are developed for each speed population: individual vehicle speeds under specific

traffic and road-weather conditions, by considering speed samples: individual vehicle speeds collected within five-minute intervals. Sample size represents the number of vehicles present within such five-minute intervals. Thus, an individual population may encompass samples with different sample sizes. According to the Central Limit Theorem (CLT), for normally distributed speed populations, the sampling distributions of mean speeds, i.e., five-minute aggregate speed distributions, are normally distributed irrespective of the sample size. Moreover, a speed population's mean (μ) is represented by the mean of its five-minute aggregate speed distribution (Hayter, 2012). However, the presence of multiple sample sizes leads some populations to be represented by multiple sampling distributions. In such cases, the normally distributed five-minute aggregate speed distributions are linearly combined to produce \bar{Y} : the resultant five-minute aggregate speed distribution of the linear combination of all five-minute aggregate speed distributions representing a population. Mean of \bar{Y} ($\mu_{\bar{Y}}$) is estimated using the Minimum Variance Unbiased Estimation (MVUE) technique where each five-minute aggregate speed distribution is assigned with a proper weight factor proportionate to its variance. A speed population's standard deviation (σ) is thereafter estimated using (i) variance of \bar{Y} , (ii) a_i : weight factor assigned to the i^{th} five-minute aggregate speed distribution during the process of estimating $\mu_{\bar{Y}}$, and (iii) n_i : sample size of the i^{th} five-minute aggregate speed distribution.

Step 0 outputs population-level speed distribution (speed distribution hereinafter) parameters μ and σ for different combinations of RWCs.

4.5.2.2 Step 1: Evaluating current speed behaviour at the study site

Drivers' response to the existing fixed SL at the study site is evaluated based on V_{85} (Assumption 2). Thus, V_{85} can be estimated by considering the Z score corresponding to the 85th percentile in normal distribution, i.e., 1.036. Outputs of step 0 i.e., μ and σ of a particular RWC-specific speed distribution are used to estimate V_{85} for a given combination of RWC (properties of normal distribution):

$$V_{85} = \mu + 1.036 \times \sigma \quad \text{Eq. 4.1}$$

4.5.2.3 Step 2: Defining Recommended Speed (RS)

Determining the anticipated properties of a SL is crucial in developing effective SLs. While ensuring drivers' safety is the prime objective of a SL (Forbes et al., 2012), drivers' compliance is

a key determinant of the effectiveness of a SL (Hellinga & Mandelzys, 2011). Thus, in this research, a credible SL is characterized by the synchrony of two features: safety, and compliance. The safety objective of a SL is typically achieved by setting SLs based on (i) V_{85} (e.g., engineering approach application), or (ii) crash mitigation (e.g., SSA). For example, travelling at or near V_{85} yields the lowest crash risk for drivers (Solomon, 1964). Yet, adverse RWCs increase speed variability (Yasanthi et al., 2021) leading to notably high V_{85} under such RWCs. In contrast, crash mitigation approach attempts to prevent collisions by warranting adequate SSDs under prevailing RWCs (Arason, 2019). Accordingly, a potential WRVSL can be represented by a weather-dependent discrete random variable namely recommended speed (RS): a desired speed which is (i) likely complied by drivers when implemented as a SL, and (ii) adequate to avoid a rear-end collision, under a specific RWC combination. Thus, any u : desired speed of the succeeding vehicle in a platoon of two vehicles travelling under a unique combination of RWCs, fulfilling the compliance and safety criteria may represent a RS. Thus, RS for a specific RWC combination may be represented by a distribution.

4.5.2.4 Step 3: Identifying candidate recommended speeds

A RS is defined based on the synchrony of two criteria; *i.*) compliance criterion and *ii.*) safe speed criterion. Thus, identification of candidate RSs consists of two tasks:

i. Evaluating u for compliance

Drivers' speed choice is typically individualistic; yet, largely depends on driving externalities (e.g., RWCs, speed regulations, etc.) and driver's mood at the time of travel. To allow for such contingencies intrinsic to speed behaviour, in this research, C : a dichotomous random variable representing compliance, is modelled with two distinct levels — compliant, or non-compliant. A driver who obeys SL, is specified as a compliant driver; otherwise labelled as a non-compliant driver. For instance, two drivers travelling at 100 km/h and 120 km/h at the study site (SL=110 km/h) are identified as compliant, and non-compliant respectively.

ii. Assessing the safety property of u

Crash avoidance ability of a SL has long been used to indicate the credibility of SLs (Arason, 2019). In this research, the safety property of u i.e., the ability of a vehicle

traveling at u to avoid a rear-end collision (collision hereinafter) with its preceding vehicle, is assessed by considering the maximum speed allowed for such avoidance. Safe speed is defined as any u adequate to avoid a collision under the prevailing RWCs. Stopping maneuver of a vehicle consists of two phases: (i) break reaction time (BRT), and (ii) braking process. The BRT of the i^{th} vehicle ($i = 1$ for preceding vehicle, $i = 2$ for succeeding vehicle) $t_{i=1,2}$: time interval from a driver identifying an obstacle to successfully applying brake, is dependent on driver's psychological competence (AASHTO, 2018). In contrast, braking process is governed by driving externalities (i.e., vehicle condition, prevailing RWCs, headway) (AASHTO, 2018). For a vehicle i driving at a speed u_i in a level road with a negligible grade, SSD_i : the SSD for vehicle i , can be expressed according to AASHTO (2018) as:

$$SSD_i = 0.278 u_i t_i + 0.039 \frac{u_i^2}{a} \quad \text{Eq. 4.2}$$

Where,

- u_i : Initial speed of vehicle i ; $i = \{1,2\}$ in km/h
- t_i : Brake reaction time of the driver of vehicle i ; $i = \{1,2\}$ in seconds
- a : Deceleration rate in ms^{-2}

However, a vehicle's deceleration rate (and thus SSD) is dependent on its vehicle characteristics such as weight and braking capacity (AASHTO, 2018). In fact, HVs are anticipated to have longer SSDs particularly due to their large weight and brake system characteristics (AASHTO, 2018). For instance, HVs are often manufactured with air brakes which prolong braking maneuvers (leading to lower HV deceleration rates) as compared to hydraulic braking systems which are often used in passenger cars. Yet, the eye level of a HV driver is substantially higher than the eye level of a passenger car driver due to the comparatively higher position of the driving seat of a HV. Accordingly, a HV driver would be able to identify obstructions downstream of a highway more effectively and promptly as compared to a passenger car driver, which negates the need to consider vehicle-specific braking distances and thus vehicle-specific SSDs. To mitigate the impact of vehicle type on SSD, in this research, a in Eq. 4.2 was represented by the product of f : the friction

coefficient of a pavement and g : the acceleration of gravity (Garber & Hoel, 2014). Accordingly, by considering a value of 9.81 ms^{-2} for g , Equation 4.2 can be modified as:

$$SSD_i = 0.278 u_i t_i + \frac{u_i^2}{254 (f)} \quad \text{Eq. 4.3}$$

Figure 4.5 demonstrates safe stopping maneuvers in a two-vehicle setting, where an abrupt stop is triggered due to an obstacle downstream the leading vehicle when travelling at a space headway d : the distance between the fronts of vehicle 1 and 2 at time $T = 0$ (Garber & Hoel, 2014).

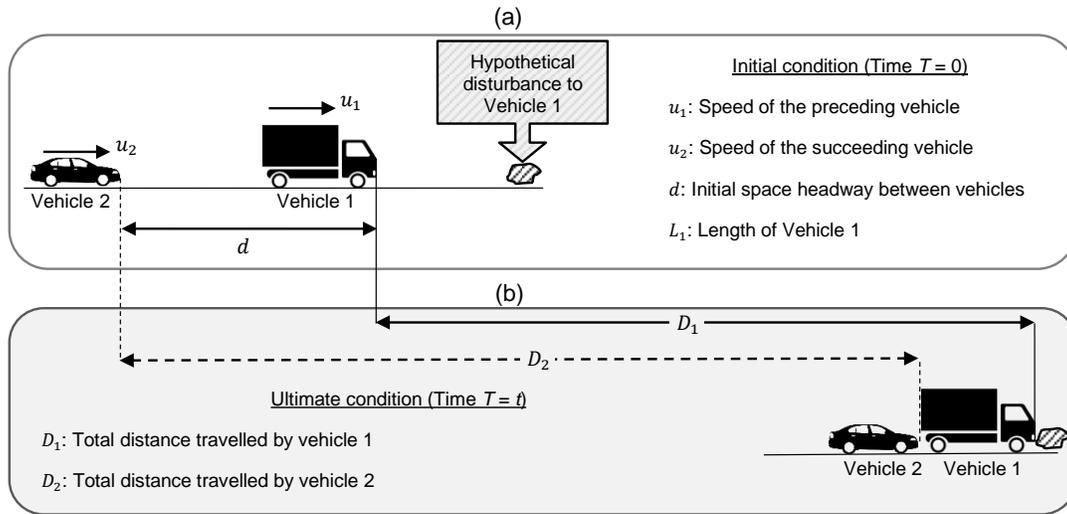


Figure 4.5: A two-vehicle setting with a hypothetical disturbance introduced to the preceding vehicle (Yasanthi et al., 2022)

At the initial condition (Time $T = 0$), vehicle 1: the preceding vehicle, identifies an interruption downstream (Figure 4.5 (a)), and stops after time t by applying breaks (Figure 4.5 (b)). Accordingly, vehicle 2: the succeeding vehicle, is prompted to stop within a time interval t . In such cases, the possibility of a collision can be determined by evaluating $D_{i=1,2}$: the total distance travelled by vehicle i during time t . However, determinants of D_i depends on the position of the vehicle i.e., preceding or succeeding. For the preceding vehicle (i.e., $i = 1$), D_1 is equivalent to SSD_1 (see Eq. 4.3). While the stopping maneuver for vehicle 1 starts at time $T = 0$, stopping action for vehicle 2 is triggered at time $T = t_1$; i.e., immediately after vehicle 1's driver applies brake.

Thus, D_2 consists of (i) SSD_2 (Eq. 4.3), and (ii) S_{t_1} : distance traversed by vehicle 2 within time interval t_1 . Accordingly, S_{t_1} can be estimated as:

$$S_{t_1} = u_1 \times t_1 \quad \text{Eq. 4.4}$$

The total distances traversed by vehicles 1 and 2 during time period t are expressed according to Eq. 4.5 and Eq. 4.6 respectively.

$$D_1 = 0.278 u_1 t_1 + \frac{u_1^2}{254 (f)} \quad \text{Eq. 4.5}$$

$$D_2 = 0.278 u_2 t_2 + \frac{u_2^2}{254 (f)} + u_1 t_1 \quad \text{Eq. 4.6}$$

For vehicle 2 to safely stop while avoiding a collision with vehicle 1, a distance of L_1 : the length of vehicle 1, must be left between the fronts of the two vehicles after a time t . It is important to note that vehicle 2 also travels a distance equivalent to the initial space headway d . For both vehicles to stop safely while avoiding a collision, relationship between (i) $D_{i=1,2}$, (ii) L_1 , and (iii) $d = u_2 \times h$, can be formulated as

$$D_1 + d \geq D_2 + L_1 \quad \text{Eq. 4.7}$$

Thus, any u_2 satisfying Eq. 4.7 is designated as a safe speed (u_{safe}) in the context of this research. In other words, u_{safe} corresponds to the maximum speed the succeeding vehicle of a two-vehicle platoon can travel under the prevailing pavement conditions to avoid a rear-end collision (i.e., $u_2 \leq u_{safe}$).

A simulation modelling approach was adopted to generate modelling data used to identify candidate RSs (step 3 of Figure 4.4). For a given combination of RWCs, modelling data used in this research includes a set of 10,000 two-vehicle scenarios produced using the Monte Carlo simulation method. Each simulation modelling (simulation hereinafter), by considering a two-vehicle scenario (Figure 4.5), attempts to evaluate the succeeding vehicle's speed for safety and compliance, thus RS. The inputs required to simulate a two vehicle scenario: (i) vehicle speeds ($u_{i=1,2}$), (ii) brake reaction times ($t_{i=1,2}$), (iii) pavement friction coefficient (f), (iv) preceding vehicle's length (L_1), (v) initial space headway between vehicles (d), and (v) compliance (C), are

generated by considering the study data collected and some additional assumptions. Algorithm 4.1 (Figure 4.6) presents the modelling data generation process for a given combination of RWCs, and C_r : average SL compliance ratio at the study site. For each simulation, Algorithm 4.1 takes nine inputs: (i) K : number of simulation trials, (ii) $PAVE$: pavement surface condition considered, (iii) μ , (iv) σ , (v) $PREC$: precipitation condition considered, (vi) C_r , (vii) λ : mean vehicle arrival rate at the study site, (viii) τ : time headway shift, and (ix) Pr_{HV} : average proportion of heavy vehicles at the study site.

For each simulation, Algorithm 4.1 executes the following tasks:

1. f^k : f for the k^{th} simulation, is generated based on the input $PAVE$. In this research, f values proposed by Kordani et al., (2018) are used to generate three f values: (i) 0.6 for dry pavements, (ii) 0.5 for wet, frost, or trace moisture pavements, and (iii) 0.18 for ice warning or ice watch pavements.
2. $u_{i=1,2}^k$: vehicle speeds for the k^{th} simulation, are generated as random speeds drawn from corresponding modelled speed distribution (i.e., $N(\mu, \sigma)$) for the considered combinations of RWCs.
3. Driver specific $t_{i=1,2}^k$: BRTs for the k^{th} simulation, are drawn from a lognormal distribution based on empirical evidence suggesting that BRTs are represented by a lognormal distribution (Koppa, 2017). The average precipitation-specific BRTs proposed by Hammit et al. (2018) were considered as the means of BRT distributions for each precipitation type considered in this research. The standard deviations of the lognormal BRT distributions were estimated in three steps. First, the standard deviation of the BRT distribution under no precipitation condition was estimated by considering the 90th percentile of a BRT distribution which is 2.5 seconds according to AASHTO (2018), and the mean BRT as suggested by Hammit et al. (2018). Second, for each adverse precipitation condition considered, p_{BRT_m} : the percentage change in the mean of a BRT distribution (as compared to no precipitation), was estimated based on the difference of BRT_{np} : mean BRT in no precipitation, and BRT_{ap} : mean BRT in an adverse precipitation condition, as

Algorithm 4.1 Pseudocode for generating modelling data for a given RWC combination and a compliance rate

Input: $K, PAVE, \mu, \sigma, PREC, C_r, \lambda, \tau, Pr_{HV}$

```
1. Generate  $f^k, u_{i=1,2}^k, t_{i=1,2}^k, C^k$ , and  $d^k$ 
  for  $k = 1$  to  $K$ 
    if  $PAVE = \text{"Dry"}$ 
       $f^k = 0.6$ 
    else if  $PAVE = \text{"Ice Warning" OR "Ice Watch"}$ 
       $f^k = 0.18$ 
    else
       $f^k = 0.5$ 
    end
    for  $i = 1$  to  $2$ 
       $u_i^k :=$  generate random number from the normal distribution  $N(\mu, \sigma)$ 
      if  $PREC = \text{"No Precipitation"}$ 
         $t_i^k = 0.8$  seconds
      else if  $PREC = \text{"Slight Rain"}$ 
         $t_i^k = 1$  second
      else if  $PREC = \text{"Moderate & Heavy Rain"}$ 
         $t_i^k = 1.7$  seconds
      else
         $t_i^k = 1.2$  seconds
      end
    end
     $h^k :=$  generate random number from the exponential distribution  $exp(\lambda)$  with a shift of  $\tau$ 
     $d^k = u_2^k \times h^k$ 
     $C^k :=$  generate binomial random individual according to the Bernoulli distribution  $Bern(C_r)$ 
    if  $C^k = 1$ 
       $C = \text{"Compliant"}$ 
    else
       $C = \text{"Non-Compliant"}$ 
    end
  end
2. Generate  $L_1^k$ 
  for  $k = 1$  to  $K \times Pr_{HV}$ 
    preceding vehicle considered as a heavy vehicle
     $L_1^k :=$  generate random number between 12.5 to 27.5 meters
  end
  for  $k = K \times Pr_{HV}$  to  $K$ 
    preceding vehicle considered as a light vehicle
     $L_1^k :=$  generate random number between 4 to 12.5 meters
  end
3. Evaluate  $u_2^k$  for RS
  for  $k = 1$  to  $K$ 
     $u^k = u_2^k$ 
     $D_1^k = 0.278 u_1^k t_1^k + \frac{(u_1^k)^2}{254 (f^k)}$ 
     $D_2^k = 0.278 u_2^k t_2^k + \frac{(u_2^k)^2}{254 (f^k)} + u_1^k t_1^k$ 
    if  $\{D_1^k + d^k \geq D_2^k + L_1^k\}$  AND  $\{C^k = \text{"Compliant"}\}$ 
       $E^k = \text{"RS"}$ 
    else
       $E^k = \text{"Non - RS"}$ 
    end
  end
end
Output: The matrices of  $u^k, E^k$ 
```

Figure 4.6: Modelling data generation process for a given combination of RWCs and C_r (Yasanthi et al., 2022)

suggested by Hammit et al. (2018). Estimation of p_{BRT_m} can be mathematically described

as: $\frac{BRT_{np} - BRT_{ap}}{BRT_{ap}} * 100$. Third, for each adverse precipitation condition considered, p_{BRT_sd} : the percentage change in the standard deviation of a BRT distribution (as compared to no precipitation), was estimated by assuming that p_{BRT_sd} is equivalent to p_{BRT_m} , i.e., $p_{BRT_sd} = p_{BRT_m}$. However, the minimum BRT value in this research was limited to 1.64 seconds to allow for the minimum BRT proposed by AASHTO (2018).

4. C^k : compliance for the k^{th} simulation, is generated using Bernoulli distribution represented by a discrete probability value of C_r . Bernoulli distribution is used to simulate compliance due to the dichotomy between the compliant and non-compliant behaviour of drivers. In this research, three C_r values are considered: 50%, 70%, and 95%. For instance, a driver population with a C_r value of 50% implies that 50% of drivers at the study site comply to the SL in effect; the rest 50% drivers tend to defy the SL.
5. h^k : h for the k^{th} simulation, is generated according to a shifted exponential distribution represented by λ with a shift of τ . In this research, λ was estimated based on the most frequent traffic flow value at the study site i.e., 150 veh/h. On the other hand, drivers prefer to maintain comparatively short headways, particularly in rural highways with frequent passing options (Al-Kaisy & Karjala, 2010). Thus, τ of three seconds was adopted to eliminate the presence of very short headways in the set of h^k values (HCM, 2016). Accordingly, d^k : d for the k^{th} simulation, is generated as the product of u_2^k , and h^k .
6. L_1^k : L_1 for the k^{th} simulation, is generated based on the type of preceding vehicle considered. The preceding vehicle's type is determined based on the traffic composition of the study site. In this research, the proportion of heavy vehicles in each modelling dataset is proportionate to the most frequent heavy vehicles' percentage at the study site ($Pr_{HV} = 30\%$). Accordingly, 30% of the preceding vehicles in the modelling data correspond to heavy vehicles (i.e., $12.5 m \leq L_1^k \leq 27.5 m$); the rest of vehicles correspond to light vehicles (i.e., $4 m \leq L_1^k \leq 12.5 m$). The extent of the vehicle lengths is determined based on the legal dimension regulations enacted in Alberta (Government of Alberta, 2021).

7. Finally, Algorithm 4.1 identifies candidate RSs in all K simulations by evaluating u_2^k for compliance and safety (see Step 3 in Figure 4.4). Thus, Algorithm 4.1 outputs u^k : k^{th} random speed, and E^k : evaluation of u^k for RS, i.e., RS, or non-RS.

4.5.2.5 Step 4: Estimating the reliability of a desired speed being a RS

Each u_{safe} fulfilling the compliance criterion is identified as a candidate RS (see step 3) i.e., a potential WRVSL; a non-RS corresponds to any u which fails to satisfy the compliance and/or safety properties simultaneously. Accordingly, RWC combinations are often represented by multiple RSs. Yet, for effective speed management, each RWC combination should be regulated by only one WRVSL representing the epitome of all RSs representing an identical combination of RWCs. Thus, in this research, the reliability theory is used to evaluate the reliability of a particular desired speed u : the probability of u being a RS under a certain combination of RWCs. The reliability theory provides a mathematical framework to model the ability of a technical system to function according to a set of specifications (Mendenhall & Sincich, 2016). Estimating the reliability of such technical systems consists of three stages: (i) terminological mapping from reliability theory to WRSM, (ii) formulating reliability, and (iii) determining parameters of the reliability function.

- i. Terminological mapping from the reliability theory to WRSM

Table 4.2 analogizes the conventional reliability theory terminology to WRSM nomenclature used in this research. Technical systems are often represented by formal study objects — i.e., a credible WRVSL in this research. A credible WRVSL system is expected to function according to the fundamental criteria used in developing its SLs (compliance, and safety properties in this research); thus, can be treated as a technical system.

Table 4.2: Terminological mapping between reliability theory and WRSM

Reliability theory	Weather-responsive speed management
Technical system	WRVSL system
System unit	u
Success	Recommended Speed
Reliability	Probability of u being a RS under a certain combination of RWCs

In reliability theory, success represents the event of fulfilling a set of specifications (Mendenhall & Sincich, 2016). On the other hand, RSs are identified based on two specifications (i.e., compliance, and safety); thus, each RS corresponds to a success. Accordingly, failure represents any speed which is not a RS i.e., a non-RS. Figure 4.7 demonstrates the phenomenon of “Success” and “Failure” in terms of the RS in the context of this research.

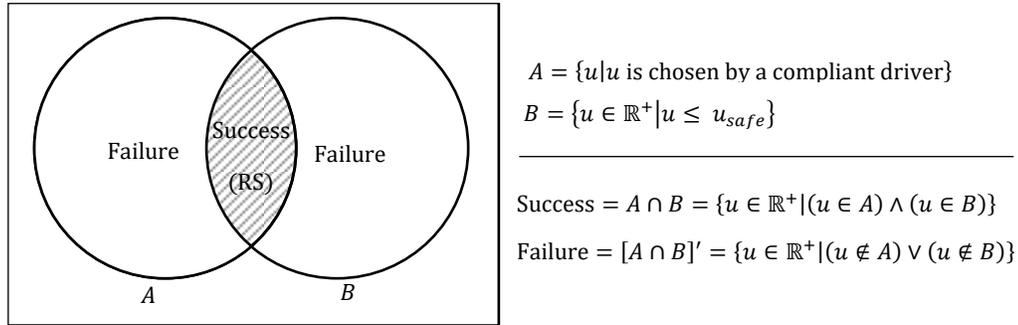


Figure 4.7: Defining success and failure
(Yasanthi et al., 2022)

ii. Formulating reliability

For technical systems, reliability (a.k.a. survival function) represents the probability of such systems functioning according to their specifications (Mendenhall & Sincich, 2016). In this research, reliability $R(u)$ represents the probability of u being a RS under a certain combination of RWCs. Reliability modelling approaches are twofold: (i) parametric, and (ii) non-parametric. Parametric reliability modelling approaches assume that reliability can be modelled according to a specific distribution; non-parametric approaches do not consider such assumptions. Parametric approaches offer a wide range of benefits (Zhang, 2016; Crowther & Lambert, 2014). First, pre-defining a particular distribution (e.g., exponential, Weibull, or Gamma distribution) to model reliability facilitates estimating reliability for all system units (Andersson et al., 2013); thus, reliability, in this research, can be estimated for any given speed. Second, parametric reliability modelling is widespread due to its availability in standard software (Crowther & Lambert, 2014). Weibull distribution (WD) is frequently used in parametric reliability modelling (Zhang, 2016), mainly due to its flexibility in modelling (Mendenhall & Sincich, 2016). More

specifically, the WD parameters: (i) α : shape parameter, and (ii) β : scale parameter, can be often manipulated to represent a unique distribution. For instance, WD is identical to the exponential distribution when represented by a shape parameter of 1 i.e., $\alpha = 1$. In fact, WD can be used to accurately represent the reliability of a wide range of systems and products (Mendenhall & Sincich, 2016).

This research models $R(u)$ according to Weibull distribution represented by α , and β . Mathematical formulation of $R(u)$ can be expressed as:

$$R(u) = e^{-u^\alpha/\beta} \quad \text{Eq. 4.8}$$

iii. Determining parameters of the reliability function

Parameters of the reliability function: α , and β , can be estimated by (i) maximum-likelihood estimation, or (ii) method of least squares. While maximum-likelihood method is often used in parameter estimation, its application is limited to sufficiently large samples (Mendenhall & Sincich, 2016). Inclement RWCs, however, are often attributed as seldom events leading to limited sampling conditions characterized by small sample sizes (Yasanthi & Mehran, 2020). In such limited sampling conditions, maximum-likelihood estimates can be heavily biased (Mendenhall & Sincich, 2016). Thus, in this research, α , and β are estimated in a three-step process according to the method of least squares (Mendenhall & Sincich, 2016).

First, $R(u)$ is rearranged by transforming Eq. 4.8:

$$-\ln R(u) = \frac{u^\alpha}{\beta} \quad \text{Eq. 4.9}$$

Second, Eq. 4.9 is obtained by taking the natural logarithm of both sides of Eq. 4.9.

$$\ln[-\ln R(u)] = -\ln(\beta) + \alpha \ln(u) \quad \text{Eq. 4.10}$$

Third, α and β are estimated by approximating Eq. 4.10 as a simple linear regression model $M: y = c + mx$, where (i) dependent variable (y): $\ln[-\ln R(u)]$, (ii) independent variable (x): $\ln(u)$, (iii) gradient (m): α , and (iv) intercept (c): $-\ln(\beta)$. The characteristics of modeling data required to calibrate M are twofold. First, the modelling dataset includes

individual random speeds collected under a given combination of RWCs. Second, each speed in the modelling dataset needs to be evaluated for the properties of RS. Therefore, this research uses the Monte Carlo simulation method to evaluate each speed in S : a set of n random speeds belonging to a population of speeds derived for a specific RWC combination, for its safety, and compliance properties. Accordingly, each speed in S is represented by a RS or a non-RS (see step 3 in Figure 4.4). The modelling data generation process used in this research is discussed in detail in the next section. Consequently, point reliability $\hat{R}(i)$: the proportion of RSs at the i^{th} observed speed in S (Mendenhall & Sincich, 2016), is estimated for each speed in S (Eq. 4.11).

$$\hat{R}(i) = \frac{n_i}{n} \quad \text{Eq. 4.11}$$

Where,

n_i : the number of RSs at the i^{th} speed

n : sample size of S

Accordingly, α and β are estimated according to Eq. 4.12 and Eq. 4.13 respectively, by considering m , and c :

$$\alpha = m \quad \text{Eq. 4.12}$$

$$\beta = e^{-c} \quad \text{Eq. 4.13}$$

Speeds ranging from 50km/h to 150km/h are considered in this research for the reliability analysis.

4.5.2.6 Step 5: Determine the WRVSL for a given RWC combination

A credible WRVSL is defined based on its reliability i.e., probability of the WRVSL being a RS under the considered RWCs. Thus, the WRVSL for a given combination of RWCs is estimated (see Eq. 4.8) according to a user-defined reliability (preferably closer to one). It is important to note that RWCs may be represented by multiple potential WRVSLs as per desired reliability levels. Thus, it is possible that an estimated WRVSL is greater than the existing SL. In such cases, this thesis recommends limiting WRVSL to the existing SL to discourage over speeding. It is

recommended to select WRVSLs in the form of multiples of five for jurisdictions using imperial units (Noyce, 2019), and ten for jurisdictions using metric units (Forbes et al., 2012).

The proposed methodology to set effective WRVSLs can be validated by simulating drivers' speed behaviour in a computer-based environment before and after implementing the WRVSLs suggested in this thesis. For instance, traffic flow simulation software (e.g., PTV VISSIM) can be used to calibrate TM_1: a base microsimulation model simulating drivers' current speed behaviour in a unique combination of RWCs under the current speed limit. In fact, inputs (i.e., base data) to TM_1 include: (i) the RWC-specific population-level speed distribution (step 0 in Figure 4.4) added as desired speed distribution, (ii) speed limit (i.e., speed limit fixed at 110 km/h for this research), and (iii) the prevailing traffic composition at the study site (e.g., a HV percentage of 30% for this research). Similarly, a separate model TM_2: a microsimulation model simulating drivers' expected speed behaviour in a unique combination of RWCs under the proposed WRVSL can be developed. Inputs to TM_2 include: (i) the RWC-specific population-level speed distribution (step 0 in Figure 4.4) added as desired speed distribution, (ii) the proposed WRVSL for the RWC combination considered, and (iii) the prevailing traffic composition at the study site. The outputs of TM_1 and TM_2 (i.e., vehicle speeds under a unique combination of RWCs and two different SLs) can be used to estimate V_{85} (Eq 4.1) and thus evaluate the impact of SL on drivers' speed. In fact, vehicle speeds output by TM_2 are anticipated to be comparatively lower than the vehicle speeds output by TM_1 because the proposed WRVSLs are developed by incorporating drivers' compliance with the SL in effect. In practice, the proposed WRVSL setting approach can be validated by conducting a before- and after-study focusing on (i) V_{85} , and (ii) crash statistics at the study site, before and after implementing the proposed WRVSLs.

4.6 Implementation: Case study

This section presents a sequential stepwise approach for setting a WRVSL for a given combination of RWCs, and a C_r value:

1. Task 1: Data collection, cleaning, and preparation (see Step 0 in Figure 4.4)

This step takes two inputs: (i) road-weather data, and (ii) traffic data, and outputs μ , and σ for the combination of RWCs considered. Road-weather data used in this research were collected from a RWIS station located on a straight segment of highway 16 (SL= 110km/h),

a rural multilane divided highway located in Alberta, Canada (Figure 4.1). Traffic data were collected from a WIM station located alongside the RWIS station. The study data were collected for a period of 15 months from October 2014 to December 2015.

2. Task 2: Evaluate the current speed behaviour at study site (see Step 1 in Figure 4.4)

Drivers' response to the SL in effect may be influenced by the type of SL implemented, i.e., fixed or variable. The study site considered in this thesis operates under a fixed SL. Therefore, in this research, drivers' response to the current SL under a unique combination of RWCs is evaluated based on V_{85} (Eq. 4.1) estimated using speeds recorded under the RWC combination in question. In other words, current speed behaviour evaluations for fixed speed limit study sites can be expressed using an individual V_{85} because the SL is not varied. In contrast, variable speed limit (VSL) systems may implement different SLs even under the same RWCs. Therefore, for study sites operating under VSL systems, driver's response to each VSL should be evaluated based on the observed speeds collected during each VSL implementation. Drivers' response to each VSL in a study site operated with a VSL system can be evaluated according to Eq. 4.1.

3. Task 3: Defining RS (see Step 2 in Figure 4.4)

A RS in this research was defined based on the synchrony of two criteria: (i) safety, i.e., the ability of a vehicle to avoid a rear-end collision in this research, and (ii) compliance to the SL in effect. However, practitioners may consider additional site-specific attributes in defining the RS in practice. For instance, for study sites encountering excessive head-on collisions, the proposed methodology can be successfully adopted by amending the definition of RS to include a third criteria: the ability of a vehicle to avoid a head-on collision.

4. Task 4: Generate modelling data (see Algorithm 4.1)

Modelling data are generated according to Algorithm 4.1 (Figure 4.4).

5. Task 5: Identifying candidate RS (see Step 3 in Figure 4.4)

Candidate RSs in the set of modelling data generated in step 3 are identified based on the synchrony of compliance and safety properties of the succeeding vehicle's speed.

6. Task 6: Estimating the reliability of a desired speed being a RS (see Step 4 in Figure 4.4)
Reliability can be modelled according to different distribution functions (e.g., WD, gamma distribution, normal distribution, etc.). In this research, reliability is represented by WD.
7. Task 7: Set WRVSL for the considered combination of RWCs (see Step 5 in Figure 4.4).

4.7 Results and discussion

4.7.1 Drivers' response to the existing fixed SL

In total, 18 different combinations of precipitation, and pavement surface conditions were observed at the study site. Figure 4.8 presents (i) μ , (ii) σ , and (iii) V_{85} , for each combination of RWCs.

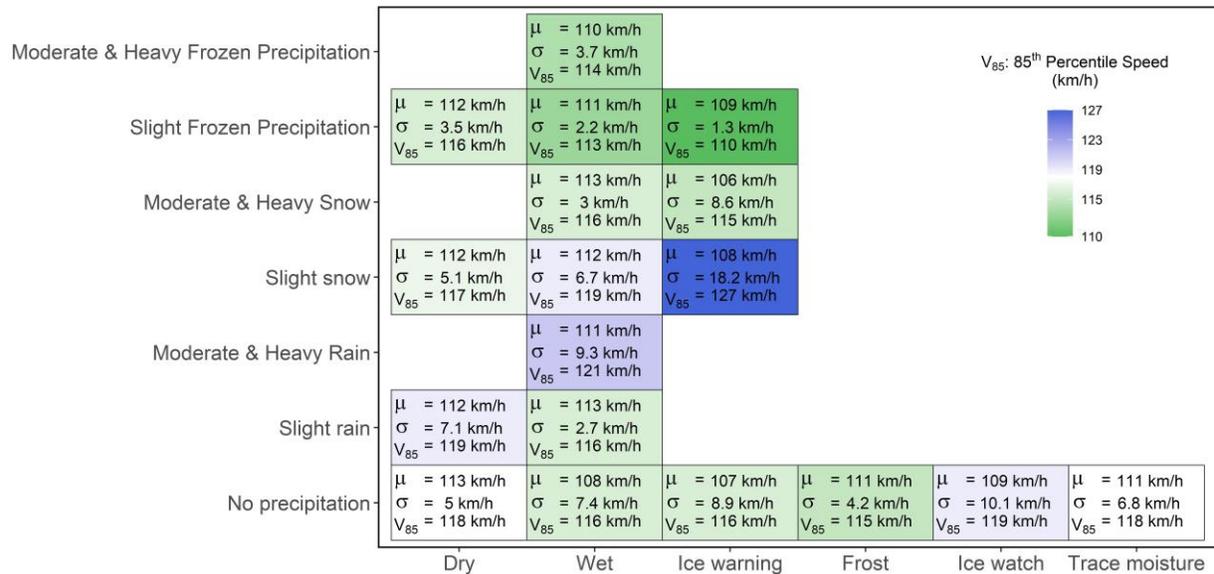


Figure 4.8: Desired speed distribution characteristics (μ and σ) and the 85th percentile speeds under different combinations of pavement and precipitation conditions (Yasanthi et al., 2022)

As anticipated, the speed distribution parameters μ and σ , thus V_{85} , are governed by the prevailing RWCs. Importantly, V_{85} , under all RWCs except for the combination of slight frozen precipitation and ice warning pavements, are greater than the current SL implemented at the site (110 km/h). Such high V_{85} values indicate the lack of drivers' collective judgment about a safe speed to drive under most RWCs. In fact, only 73% of drivers comply with the existing SL at the study site under normal road-weather conditions i.e., no precipitation and dry pavements. The percentage of compliant drivers is further reduced to 54% in the presence of slight snow, and ice warning

pavements. On the other hand, the notably high V_{85} (118 km/h) observed under normal RWCs questions the credibility of the current SL implemented at the study site. Further, substantially lower V_{85} values under some combinations of RWCs (e.g., 110 km/h under slight frozen precipitation and ice warning pavements) implies most drivers' lack of confidence to drive at high speeds in the presence of some adverse RWCs. Yet, the presence of aggressive drivers: motorists who are willing to trade-off safety for lesser travel times (Hauer; 1971), may increase σ of a speed distribution. For instance, the speed distribution under slight snow, and ice warning pavements is represented by a notably low μ (108 km/h); yet a considerably high σ (18.2 km/h). In fact, some combinations of RWCs (e.g., moderate and heavy snow, and ice warning pavements) are represented by speed distributions with notably higher values of σ , although the V_{85} under such RWCs are less than that under normal RWCs. Such high values of σ demonstrate the versatile speed behaviour, indicating higher crash propensity at the study site under specific adverse RWCs. Thus, observations of this research imply that formulating WRVSLs based on V_{85} is questionable.

According to MUTCD, a V_{85} -based SL is acceptable if the operating speed at the study site i.e., V_{85} , is within 8 km/h of the posted SL. Accordingly, the current SL at the study site is admissible under normal RWCs. Yet, implementing a SL of 110 km/h under some RWCs violates the MUTCD criterion for acceptable SLs (Figure 4.8). For instance, implementing a fixed SL of 110 km/h under the combination of slight snow, and ice warning pavements ($V_{85}=127$ km/h) is particularly questionable. Theoretically, in such cases ($V_{85} \gg SL$), the operating speed method in setting SLs prompts increasing SL at the study site (FHWA, 2009). Yet, it is apparent that raising SL may provoke drivers into speeding, thus elevating traffic safety risks (Garber & Hoel, 2014). Therefore, implementing a WRVSL system would be vastly beneficial.

4.7.2 Reliability-based Weather-Responsive Variable Speed Limit Estimations

In total, reliability was evaluated for 18 combinations of RWCs corresponding to three different mean compliance rates (C_r); 50%, 70% and 95%, and the results are presented in Figure 4.9. According to Figure 4.9, the reliability of the current SL (110 km/h) under normal RWCs (dry, no precipitation) is approximately equal to one irrespective of C_r , implying the suitability of implementing a fixed SL of 110km/h under normal RWCs at the study site. Similarly, a SL of 110 km/h proves to be reliable under most combinations of RWCs, when C_r is greater than 70%. Yet, $R(110)$: reliability of the current SL, is notably reduced for driver populations where C_r is 50%

particularly under extreme RWCs. For instance, $R(110)$ is alarmingly low for a C_r of 50% ($R(110) = 0.78$) under ice warning pavements, and slight snow precipitation. In fact, reliability is largely dependent on C_r for most RWCs. For example, under slight snow and ice warning pavements, C_r values of 50%, 70% and 95% correspond to $R(110)$ values of 78%, 90% and 98% respectively. Accordingly, it is practical to implement a WRVSL of 90 km/h under slight snow, and ice warning pavements to ensure a high reliability in SL considering different rates of C_r . Similarly, using the reliability curves in Figure 4.9, for each combination of RWCs, a new practical WRVSL can be determined to satisfy the desired reliability level and assumed mean C_r . The reliability curves can be reproduced for additional mean C_r values. However, realistic determination of C_r is challenging. In summary, the study results reveal the following observations:

- i. Adverse RWCs crucially affect population-level speed distribution parameters i.e., μ and σ . Thus, V_{85} is dependent on the prevailing RWCs. In particular, adverse RWCs increase speed variability. Such variation in speed distribution characteristics indicate the versatile speed behaviour under different RWCs, particularly in rural highways.
- ii. Drivers' compliance, when evaluated in terms of V_{85} , is notably affected by the prevailing RWCs at the time of travel.
- iii. Conventional approaches in setting SLs (e.g., the operating SL method), prompts implementing considerably high SLs under combinations of RWCs represented by high V_{85} . Thus, setting WRVSLs according to such conventional approaches, particularly under adverse RWCs, might be controversial.
- iv. Setting WRVSLs for cold-region rural highways based on the proposed reliability-based approach is promising and convenient, as it offers flexibility to consider the trade-off between desired reliability level and compliance rate.

4.8 Conclusion and future directions

Weather-responsive Variable Speed Limit systems, a rather neoteric realm in speed management, attempts to mitigate the well-acknowledged impacts of inclement road-weather conditions on drivers' speed choice by regulating speed behaviour under such road-weather conditions. Yet, it is well-documented that the existing WRVSLs are mostly ineffective despite the methodology

adopted in setting such SLs. This research (i) evaluates drivers' response to a fixed SL under different RWCs, and (ii) presents an effective approach to set WRVSLs for rural highways located in extremely cold regions based on the concept of reliability. In this research, an approach to set WRVSLs is proposed based on the reliability theory as an alternative to the conventional SL-setting approaches such as the engineering approach. The proposed methodology to set reliability-based WRVSLs is demonstrated using traffic, and road-weather data collected from a rural highway (fixed SL=110 km/h) located in Edmonton, Canada.

The research results reveal that drivers' speed choice is notably affected by the prevailing RWCs. In particular, V_{85} under most adverse RWCs were revealed to be considerably higher compared to normal RWCs. Such high V_{85} values imply that drivers are suspicious about the existing fixed SL at the study site under adverse RWCs. Such suspicions potentially lead motorists to select speeds according to their natural instincts particularly under adverse RWCs. As a result, drivers' compliance with the existing SL at the study site is seemingly dependent on RWCs. Thus, the current fixed SL system implemented at the study site is deemed ineffective. In contrast, a credible WRVSL system consisting of a set of 18 WRVSLs ranging from 80 km/h (e.g., for ice warning, and slight snow conditions) to 110 km/h (e.g., for normal RWCs) is proposed to effectively regulate RWC-specific speed behaviour at the study site. While the WRVSLs are set based on their reliability, it is important to note that the existing SL was deemed credible under some RWCs (e.g., wet pavements, and slight rain). On the contrary, the SL is proposed to be substantially reduced under some RWCs (SL=90 km/h for ice warning pavements and no precipitation) to achieve a satisfactory reliability of the SL.

Contributions of this research are twofold. First, transport authorities may adopt the research methodology to evaluate drivers' compliance to current fixed SLs implemented on cold region rural highways. In fact, such evaluations may help policymakers to identify vulnerable RWCs demanding proactive safety countermeasures such as WRVSLs. Second, this research proposes a robust approach to set WRVSLs in cold region rural highways. Thus, practical application of the reliability-based WRVSLs proposed in this research, helps policymakers to effectively regulate speed behaviour under different RWCs. Despite the broad practical applications, the proposed methodology is subject to the availability of population-level desired speed distributions representing different RWCs. Thus, application of the methodology is limited to (i) uncongested

highways, or *(ii)* congested highways with population-level speed distribution information under different RWCs. Further, this research focusses on developing WRVSLs for uncongested highways; thus, mobility was not considered as a property of WRVSLs. Thus, future research may focus on *(i)* expanding the proposed methodology to congested highways, *(ii)* incorporating mobility — a critical objective of congested highway SLs — as a property of WRVSLs, and *(iii)* evaluating the suitability of different collision types (e.g., single vehicle head-on collisions, multi-vehicle collisions, etc.) to model reliability of WRVSLs. In addition, this research does not differentiate braking distance between HVs and passenger vehicles which may limit the practical application of the proposed methodology to a certain extent; future research may focus on evaluating the safety criterion of RS (Eq. 4.7) based on vehicle-specific deceleration rates and/or braking distances (Eq. 4.2). Further, future research may adopt the proposed methodology to develop lane-based WRVSLs which could be used to differentiate fast-moving vehicles from slow-moving vehicles and thus improve road safety in extreme RWCs.

Chapter 5. Conclusions

Adverse road-weather conditions often result in degraded driving conditions such as poor visibility, and reduced pavement surface friction. Such deteriorated driving conditions prevalent in adverse RWCs may crucially impact the speed behaviour of some drivers. For instance, slow drivers and/or inexperienced drivers may choose to substantially reduce speed in adverse RWCs. Yet, some drivers may choose to trade-off safety for lesser travel times, thus select higher travel speeds although such drivers acknowledge risks associated with speeding in adverse RWCs. Such divergent speed choices prevalent in adverse RWCs may result in considerably high RWC-specific speed variabilities, thus increased RWC-specific crash potential. WRSM strategies such as WRVSL systems attempt to mitigate the RWC-specific crash propensity, conventionally by harmonizing drivers' speed behaviour through the implementation of weather-dependent SLs. In practice, however, WRVSL systems are deemed ineffective. Potential reasons for the inefficacies of the existing WRVSL systems are twofold. First, SLs, including WRVSLs, are typically set based on the 85th percentile speed which represents the speed at which or below 85 percent of drivers prefer to drive. However, such a sample-level speed statistic may not accurately represent speed behaviour of some driver populations. Second, most SL-setting approaches do not consider drivers' compliance to the prevailing SL into consideration.

This thesis attempts to (i) evaluate drivers' speed behaviour using population-level speed distribution parameters (chapters 2 and 3; step 1 in Figure 1.3; O₁ and O₂), (ii) identify adverse RWCs with intensified crash risks indicated by substantial speed variabilities (chapters 2 and 3; step 2 in Figure 1.3; O₃), and (iii) propose an innovative approach to set WRVSLs (chapter 4; step 3 in Figure 1.3; O₄) in the context of cold region rural highways. This chapter presents a summary of key findings related to each research objective, followed by contributions of this research, research limitations, and future research recommendations.

5.1 Summary of key findings

5.1.1 Impacts of adverse RWCs on drivers' speed choice

Setting an effective WRVSL system starts with understanding drivers' speed choice in different RWCs. In this research, a methodology was proposed to estimate the properties (i.e., mean and SD) of normally distributed speed distributions in different driving conditions. According to the

results of this thesis, as presented in chapters two, three, and four, drivers' speed choice is notably affected by the presence of adverse precipitation conditions (i.e., rain, snow, and frozen precipitation) in synchrony with atypical pavement conditions (i.e., wet, ice warning, ice watch, frost, and trace moisture). Time of the day (day/night) and the atmospheric temperature is deemed to have marginal effect on drivers' speed behaviour. In addition, drivers' speed choice is also revealed to be dependent on the vehicle type driven, particularly in terms of trucks (e.g., speed behaviour of single-trailer truck drivers in adverse RWCs is revealed to be different from the speed behaviour of multi-trailer truck drivers travelling in the same driving conditions). Yet, according to the research results (chapter 3), the loading condition of trucks (loaded versus empty) does not considerably affect truck drivers' speed behaviour in adverse RWCs.

5.1.2 Variability of modelling speed distributions based on different data analytics

The conventional approach to evaluate drivers' speed choice is predominantly based on regression modelling. Yet, a regression modelling based approach to model drivers' speed behaviour is critiqued in past literature, particularly in the context of evaluating the impacts of extreme RWCs characterized by limited sampling conditions. Therefore, this research compared the performance of two speed distribution modelling approaches: (i) a speed distribution modelling approach proposed based on the CLT (CBA), and (ii) the conventional regression-modelling-based approach (RBA).

According to the study results presented in chapter 3, adopting different data analytic approaches in modelling speed distributions in different driving conditions notably affects the resulting speed distribution parameter estimates. For instance, according to the RBA, the mean speed of loaded single-trailer trucks travelling in median lane in no precipitation and dry pavement conditions drive is approximately 111 km/h; yet, the CBA provides a mean speed estimation of approximately 101 km/h for the same driving conditions.

Among the two approaches compared in this thesis, the CBA is deemed a more appropriate method to evaluate drivers' speed choice in extreme RWCs as compared to RBA. More specifically, according to the study results, the RBA leads to statistically insignificant estimations of speed distribution properties (mean, and SD) for most driving conditions, thus deemed not appropriate to model speed distributions, particularly in limited sampling conditions.

5.1.3 Driving conditions with intensified crash risks

This research attempted to identify driving conditions with intensified safety risks based on a holistic crash indicative method developed based on speed distribution parameters (chapter 3). The crash indicative metric proposed in chapter 3 suggests that adverse RWCs may impose crucial threats to all road users irrespective of the vehicle type driven. Yet, tractor-trailer combinations are particularly deemed vulnerable, specially when travelling loaded. In terms of RWCs with intensified safety risks, the presence of icy pavements (i.e., ice watch, and ice warning pavements), and adverse precipitation conditions (particularly snow, and frozen precipitation), lead to substantial crash risks (chapters 2 and 3).

5.1.4 Importance of setting WRVSLs in cold region rural highways

The overall goal of this thesis was to propose a robust approach to set effective WRVSLs for cold region rural highways. The proposed approach to set an effective WRVSL system was demonstrated using a study site located in Alberta, Canada, operating at a fixed SL of 110 km/h. The study results presented in chapter 4 reveal that the current fixed SL implemented at the study site is not appropriate for extreme RWCs, although it is credible and safe to be implemented in normal RWCs (i.e., dry pavements and no precipitation conditions). In fact, the 85th percentile speed (V_{85}) at the study site is considerably high in some RWCs (e.g., $V_{85} = 127$ km/h in slight snow and ice warning pavements). Such a high value of V_{85} , particularly in extreme RWCs, imply that most drivers are not able to comprehend the safest speed to drive in such RWCs. Accordingly, this research recommends implementing a WRVSL system with notably low SLs in some RWCs (e.g., proposed WRVSL for ice warning and slight snow conditions=80 km/h) to increase the probability of the WRVSLs being safe (i.e., no potential read-end crashes) and complied by drivers.

5.2 Contributions

The main contributions of this thesis are threefold.

First, the research methodologically presented in this thesis contributes to the road safety literature by proposing a CLT-based approach to evaluate the impacts of different driving conditions (e.g., RWCs, vehicle type driven) on speed behaviour. More specifically, as opposed to the conventional RBA used to evaluate RWC-speed behaviour, the CLT-based speed distribution modelling

approach proposed in this thesis is applicable to limited sampling conditions. In fact, to the author's best knowledge, the CLT-based speed distribution methodology presented in chapter 2 (Yasanthi et al., 2021) is *sui generis*, particularly in the context of its methodological contributions. In terms of the practical applications, road safety researchers and/or practitioners may use the CLT-based speed distribution modelling approach to (i) understand the impacts of different RWCs on speed behaviour, and (ii) to develop input driver behavior parameters in defining speed distributions in microsimulation applications to realistically simulate traffic operations under different traffic and road-weather conditions.

Second, a holistic crash indicative metric is proposed to identify driving conditions possessing intensified safety risks based on context-specific speed distributions (chapter 3; Yasanthi et al., (2022)). In practice, such a crash indicative measure would be highly useful for cold region jurisdictions to identify driving conditions resulting in increased crash potential, particularly when historical crash data are not available. The truck-specific crash risks revealed in chapter 3 also highlights the importance of paying special attention to improving truck safety in adverse RWCs.

Third, this thesis demonstrates a robust approach to set effective WRVSLs. In fact, the study results presented in chapter 4 highlight the importance of adopting effective WRVSLs in extreme RWCs. Transport authorities may refer to the WRVSL-setting approach proposed in this thesis (i) to understand the importance of implementing WRVSLs as opposed to enforcing conventional statistic SLs, and (ii) as a guide to set effective WRVSLs, particularly in cold region rural highways.

5.3 Limitations and future research

Limitations, as well as the future research recommendations for each research method proposed in this thesis are presented in the "Conclusions and future directions" sub-sections in each chapter (see chapters two through five). Table 5.1 presents an overall summary of this thesis while highlighting the major research limitations and future directions associated with each research question (see section 1.2 for research questions), and the problems associated with the conventional approach to solve each research question.

Table 5.1: Overall summary of the thesis

Problems with the conventional approach used to address the research question	Alternative/solution recommended by this thesis	Research limitations	Future directions
Research question 1: What is the impact of different RWCs on drivers' speed behaviour? Addressed in chapter 1 (research objective: O₁)			
<ul style="list-style-type: none"> - Conventionally, regression modelling is used to evaluate RWC-speed behaviour. - Past literature on the impacts of different RWCs on speed behaviour is inconsistent. 	<ul style="list-style-type: none"> - Proposed an innovative approach to model population-level speed distributions using the CLT. 	<ul style="list-style-type: none"> - Only applicable for rural/uncongested roads without specific geometric conditions (e.g., sharp curves) 	<ul style="list-style-type: none"> - Expand the methodology to urban/congested highways and multilane highways - Evaluate the suitability of using connected vehicle data to model speed distributions using the proposed approach
Research question 2: What is the most appropriate approach to evaluate drivers' speed behaviour in adverse RWCs? Addressed in chapter 3 (research objective: O₂)			
<ul style="list-style-type: none"> - Extreme RWCs are rare, thus result in limited sampling conditions. Regression modelling with small samples is questionable. - Past literature mostly uses considerably distant road-weather and traffic data collection devices. - Speed choice represents human behaviour; attempting to estimate idiosyncratic human nature using linear regression models is questionable 	<ul style="list-style-type: none"> - Model population-level speed distributions using the CLT instead of estimating average RWC-specific speed reductions estimated by regression models. - Use data collection devices located alongside each other. 	<ul style="list-style-type: none"> - Only applicable for rural/uncongested roads without specific geometric conditions (e.g., sharp curves) 	<ul style="list-style-type: none"> - Expand the methodology to urban/congested highways and multilane highways - Explore the minimum sample size required to effectively evaluate speed behaviour using the conventional regression modelling approaches

Problems with the conventional approach used to address the research question	Alternative/solution recommended by this thesis	Research limitations	Future directions
Research question 3: Are there any specific driving conditions (e.g., RWCs, vehicle types) which notably intensify safety/crash risks?			
Addressed in chapter 3 (research objective: O₃)			
<ul style="list-style-type: none"> - Drivers' speed is highly versatile in adverse RWCs; such speed variability prevalent in inclement RWCs lead to increased crash potential in such RWCs. - Past literature does not provide guidance on quantifying crash risks under extreme RWCs, particularly when historical crash data are not accessible. 	<ul style="list-style-type: none"> - Proposed a holistic crash indicative measure to identify driving conditions with intensified crash risks based on within- and across-lane speed variability. 	<ul style="list-style-type: none"> - Only applicable for rural/uncongested roads without specific geometric conditions (e.g., sharp curves) - Only applicable for four lane divided highways 	<ul style="list-style-type: none"> - Expand the methodology to urban/congested highways - Expand the methodology to multilane highways
Research question 4: What is an effective approach to regulate speed in adverse RWCs to mitigate safety risks induced by such RWCs?			
Addressed in chapter 4 (research objective: O₄)			
<ul style="list-style-type: none"> - Conventional fixed SLs are usually designed for ideal RWCs (no precipitation, and dry pavements), thus, are not suitable for adverse RWCs. - A WRVSL is a unique SL designed for a specific combination of RWCs; yet most existing WRVSLs are deemed not effective. 	<ul style="list-style-type: none"> - Proposed an effective approach to WRVSLs based on reliability — the probability of a speed being safe and compliant by drivers 	<ul style="list-style-type: none"> - Only applicable for rural/uncongested roads without specific geometric conditions (e.g., sharp curves) - Only applicable for four lane divided highways - Reliability assumed to follow Weibull distribution 	<ul style="list-style-type: none"> - Expand the methodology to urban/congested highways - Expand the methodology to multilane highways - Evaluate the suitability of other statistical distribution types (e.g., Gamma distribution) to model reliability of speeds

In addition to the research limitations presented in Table 5.1, it should be noted that the study data used in this thesis were collected from a study site where the two data collection devices (RWIS and WIM) were located alongside each other (Figures 2.1, 3.2, and 4.1). Benefits associated with using such a location configuration of data collection devices are mainly twofold. First, because the RWIS is located very close to the WIM sensor, the road-weather data used in this thesis are highly representative of the microclimatic conditions for each vehicle recorded by the WIM detector. Second, such high representativeness of data can be approximated as connected vehicle data which can be used for real-time traffic activities such as real-time WRVSL systems in the near future. Despite such benefits, using study data collected from one study site may limit the application of the research methodology to a certain extent. In fact, the study site used in this research is located in a straight highway segment with (i) negligible grade and no sharp horizontal curves in the immediate vicinity of the study site, and (ii) comparatively low HV percentages. However, the presence of some highway geometric features (e.g., sharp horizontal curves, steep slopes) may compel some drivers to drive slow particularly in adverse RWCs. On the other hand, the presence of very high HV percentages in the traffic stream may violate the normally distributed populations assumption used in this thesis. Yet, methodology presented in this thesis can be expanded to eliminate such limitations. For instance, to incorporate highway geometric conditions into the speed distribution modelling process, future research may adopt a five-step speed distribution modelling approach:

- Step 1: Collecting road-weather and traffic data from study sites with different geometric conditions
- Step 2: Re-defining population by adding site geometric conditions as an additional parameter in defining populations (e.g., a sample population definition — the desired speeds of individual vehicles observed at a study site with a grade of 7% under a specific combination of road-weather and traffic conditions)
- Step 3: Categorizing samples (i.e., speeds observed in five-minute intervals) such that all samples belonging to identical road-weather, traffic, and site geometric conditions are grouped together
- Step 4: Producing five-minute aggregate speed distributions for each population using samples belonging to the road-weather, traffic, and site geometric conditions considered in the population definition

- Step 5: Linearly combining five-minute aggregate speed distributions using MVUE to model desired speed distributions for each population

It should be noted that population-level speed distributions modelled in this thesis represent all desired speeds that can potentially be collected in a unique combination of driving conditions. Accordingly, parameters of all speed distributions corresponding to identical driving conditions (e.g., similar RWC combination, identical site geometric characteristics) are anticipated to be approximately similar even though each speed distribution is modelled using study data collected from a unique study site.

Future research may also focus on developing non-parametric speed distribution modelling approaches. For instance, speed distributions may be modelling using Kaplan-Meier estimation which is based on the product limit theorem. Such non-parametric speed distribution approaches are beneficial when the population distribution is not represented by a normal distribution (e.g., speeds collected from a traffic stream with high HV percentages). Yet, such non-parametric approaches are typically highly sensitive to small sample sizes (Hoogendoorn, 2005a). Thus, future research may attempt to identify an approach to minimize the error associated with small sample sizes when modelling speed distributions using such non-parametric speed distribution modelling approaches.

References

- Abdel-Aty M, Cunningham RJ, Gayah VV, Hsia L. Dynamic variable speed limit strategies for real-time crash risk reduction on freeways. *Transportation Research Record*. 2008 Jan;2078(1):108-16.
- Alberta Transportation. Alberta Highways 1 to 986: Traffic Volume History 2007-2016; 2016
- Alberta Transportation. Traffic Safety: Alberta traffic collision statistics for 2019; 2022
- Alhajyaseen, W. K. (2015). The integration of conflict probability and severity for the safety assessment of intersections. *Arabian Journal for Science and Engineering*, 40(2), 421-430.
- Al-Kaisy, A. and Karjala, S., 2010. Car-following interaction and the definition of free-moving vehicles on two-lane rural highways. *Journal of Transportation Engineering*, 136(10), pp.925-931.
- American Association of State Highway and Transportation Officials (AASHTO). (2018). Policy on Geometric Design of Highways and Streets (7th Edition), including 2019 Errata. (pp. 3-4). American Association of State Highway and Transportation Officials (AASHTO). Retrieved from <https://app.knovel.com/hotlink/toc/id:kpPGDHSE12/policy-geometric-design/policy-geometric-design>
- Anastasopoulos, P.C. and Mannering, F.L., 2016. The effect of speed limits on drivers' choice of speed: a random parameters seemingly unrelated equations approach. *Analytic methods in accident research*, 10, pp.1-11.
- Andersson, T.M.L., Dickman, P.W., Eloranta, S., Lambe, M. and Lambert, P.C., 2013. Estimating the loss in expectation of life due to cancer using flexible parametric survival models. *Statistics in medicine*, 32(30), pp.5286-5300.
- Andrey J, Yagar S. A temporal analysis of rain-related crash risk. *Accident Analysis & Prevention*. 1993 Aug 1;25(4):465-72.
- Arason, N., 2019. Canada and the Safe Systems Approach to Road Safety. *Institute of Transportation Engineers. ITE Journal*, 89(4), pp.26-30.

- Blower, D. and Woodrooffe, J., 2012. Survey of the status of truck safety: Brazil, China, Australia, and the United States.
- Carter D, Gelinne D, Kirley B, Sundstrom C, Srinivasan R, Palcher-Silliman J. Road Safety Fundamentals: Concepts, Strategies, and Practices that Reduce Fatalities and Injuries on the Road. United States. Federal Highway Administration. Office of Safety; 2017 Nov 1.
- Catbagan, J.L. and Nakamura, H., 2008. Desired speed distributions on two-lane highways under various conditions. *Transportation research record*, 2088(1), pp.218-226.
- Chen X, Li Z, Wang Y, Cui Z, Shi C, Wu H. Evaluating the impacts of grades on vehicular speeds on interstate highways. *PLoS one*. 2017 Sep 1;12(9):e0184142.
- Chu, H.C., 2016. Effects of Extreme Weather and Economic Factors on Freight Transportation. *Advances in Management and Applied Economics*, 6(1), p.113.
- Cirillo, J.A., Former Assistant Administrator and Chief Safety Officer for the Federal Motor Carrier Safety Administration. Testimony Before the Senate Highways and Transportation Committee, Senate Bill 94, June 10, 2003
- Crowther, M.J. and Lambert, P.C., 2014. A general framework for parametric survival analysis. *Statistics in Medicine*, 33(30), pp.5280-5297.
- Dhami MK, Harries C. Fast and frugal versus regression models of human judgement. *Thinking & Reasoning*. 2001 Feb 1;7(1):5-27.
- Donnell, E.T., Kersavage, K. and Tierney, L.F., 2018. *Self-Enforcing Roadways: A Guidance Report* (No. FHWA-HRT-17-098). United States. Federal Highway Administration.
- Downey, M.B., 2015. Evaluating the Effects of a Congestion and Weather Responsive Advisory Variable Speed Limit System in Portland, Oregon.
- Eberly, L.E., 2007. Multiple linear regression. *Topics in Biostatistics*, pp.165-187.
- Edwards, J.B., 1999. Speed adjustment of motorway commuter traffic to inclement weather. *Transportation research part F: traffic psychology and behaviour*, 2(1), pp.1-14.
- Elvik R. Speed limits, enforcement, and health consequences. *Annual review of public health*. 2012 Apr 21;33:225-38.

- Elvik, R., Vaa, T., Høy, A. and Sørensen, M. eds., 2009. *The handbook of road safety measures*. Emerald Group Publishing.
- Engmann, S. and Cousineau, D., 2011. Comparing distributions: the two-sample Anderson-Darling test as an alternative to the Kolmogorov-Smirnov test. *Journal of applied quantitative methods*, 6(3).
- Federal Highway Administration (FHWA), Manual on Uniform Traffic Control Devices for Streets and Highways, 2009 Edition- 2B.13 Speed Limit Sign (R2-1). (pp. 57)
- FHWA. (2001). Traffic Monitoring Guide. U.S. Department of Transportation, Federal Highway Administration Office.
- Fitzpatrick, K., McCourt, R. and Das, S., 2019. Current attitudes among transportation professionals with respect to the setting of posted speed limits. *Transportation research record*, 2673(4), pp.778-788.
- Forbes, G., Gardner, T., McGee, H.W. and Srinivasan, R., 2012. Methods and practices for setting speed limits: An informational report (No. FHWA-SA-12-004). United States. Federal Highway Administration. Office of Safety.
- Fountas, G., Fonzone, A., Gharavi, N. and Rye, T., 2020. The joint effect of weather and lighting conditions on injury severities of single-vehicle accidents. *Analytic methods in accident research*, 27, p.100124.
- Garber NJ, Hoel LA. Traffic and highway engineering. Cengage Learning; 2014 Jan 30.
- García-Jiménez ME, Pérez-Zuriaga AM, Llopis-Castelló D, Camacho-Torregrosa FJ, García A. Examination of the Free-Flow Speed Distribution on Two-Lane Rural Roads. *Transportation Research Record*. 2016;2556(1):86-97.
- Gayah, V.V., Donnell, E.T., Yu, Z. and Li, L., 2018. Safety and operational impacts of setting speed limits below engineering recommendations. *Accident Analysis & Prevention*, 121, pp.43-52.
- Gettman, D. and Head, L., 2003. Surrogate safety measures from traffic simulation models. *Transportation Research Record*, 1840(1), pp.104-115.

- Goodwin LC. Weather impacts on arterial traffic flow. Mitretek systems inc. 2002 Dec 24:4-8.
- Government of Alberta, *Module 4: WEIGHTS AND DIMENSIONS*, Weight and dimension regulations, 2021
- Government of Canada. National Road Network - NRN - GeoBase Series - NRN Alberta SHAPE [Internet] 2021 [cited 2021 July 24]. Available at National Road Network - NRN - GeoBase Series - NRN Alberta SHAPE - Open Government Portal (canada.ca)
- Hackeling, G., 2017. *Mastering Machine Learning with scikit-learn*. Packt Publishing Ltd.
- Hall FL, Barrow D. Effect of weather on the relationship between flow and occupancy on freeways. *Transportation Research Record*. 1988;1194:55-63.
- Hamdar SH, Qin L, Talebpour A. Weather and road geometry impact on longitudinal driving behavior: Exploratory analysis using an empirically supported acceleration modeling framework. *Transportation research part C: emerging technologies*. 2016 Jun 1;67:193-213.
- Hammit, B.E., Ghasemzadeh, A., Ahmed, M.M. and Young, R.K., 2018. Evaluation of Weather-Related Freeway Car-Following Behavior Using the SHRP 2 Naturalistic Driving Study (No. 18-03287).
- Hashim IH. Analysis of speed characteristics for rural two-lane roads: A field study from Minoufiya Governorate, Egypt. *Ain Shams Engineering Journal*. 2011 Mar 1;2(1):43-52.
- Hauer E. Accidents, overtaking and speed control. *Accident Analysis & Prevention*. 1971 Jul 1;3(1):1-3.
- Hauer E. Crash causation and prevention. *Accident Analysis & Prevention*. 2020 Aug 1;143:105528.
- Hauer, E., 2009. Speed and safety. *Transportation Research Record*, 2103(1), pp.10-17.
- Hayter, A.J., 2012. *Probability and statistics for engineers and scientists*. Cengage Learning.
- Hellinga, B. and Mandelzys, M., 2011. Impact of driver compliance on the safety and operational impacts of freeway variable speed limit systems. *Journal of transportation engineering*, 137(4), pp.260-268.

- Hoogendoorn SP. 2005a. Unified approach to estimating free speed distributions. *Transportation Research Part B: Methodological*. 2005 Sep 1;39(8):709-27.
- Hoogendoorn, S.P., 2005b. Vehicle-Type and Lane-Specific Free Speed Distributions on Motorways: A Novel Estimation Approach Using Censored Observations. *Transportation research record*, 1934(1), pp.148-156.
- Ibrahim AT, Hall FL. Effect of adverse weather conditions on speed-flow-occupancy relationships. 1994.
- Ivey, D. L., Lehtipuu, E. K. & Button, J. (1975). Rainfall & visibility: the view from behind the wheel. *J. Safety Research*, 7: 156–169.
- Jägerbrand, A.K. and Sjöbergh, J., 2019. Speed responses of trucks to light and weather conditions. *Cogent Engineering*, 6(1), p.1685365.
- Johnson, S. and Murray, D., 2010. Empirical analysis of truck and automobile speeds on rural interstates: Impact of posted speed limits. In *Transportation Research Board 89th Annual Meeting* (No. 10-0833).
- Jun J. Understanding the variability of speed distributions under mixed traffic conditions caused by holiday traffic. *Transportation Research Part C: Emerging Technologies*. 2010 Aug 1;18(4):599-610.
- Katz, B., Ma, J., Rigdon, H., Sykes, K., Huang, Z., Raboy, K. and Chu, J., 2017. Synthesis of variable speed limit signs.
- Kay SM. Fundamentals of statistical signal processing. Prentice Hall PTR; 1993.
- Koppa, 2017, *Chapter 3: Human Factors*, Revised Monograph on Traffic Flow Theory, Federal Highway Administration Research and Technology
- Kordani, A.A., Rahmani, O., Nasiri, A.S.A. and Boroomandrad, S.M., 2018. Effect of adverse weather conditions on vehicle braking distance of highways. *Civil Engineering Journal*, 4(1), p.46.

- Kyte, M., Khatib, Z., Shannon, P. and Kitchener, F., 2000. Effect of environmental factors on free-flow speed. *Transportation Research Circular*, Transportation Research Board, Washington, DC.
- Kyte, M., Khatib, Z., Shannon, P. and Kitchener, F., 2001. Effect of weather on free-flow speed. *Transportation Research Record*, 1776(1), pp.60-68.
- Lee, C., Saccomanno, F. and Hellinga, B., 2002. Analysis of crash precursors on instrumented freeways. *Transportation Research Record*, 1784(1), pp.1-8.
- Lefevre, B.A., 1953. Speed characteristics on vertical curves. In *Highway Research Board Proceedings (Vol. 32)*.
- Letirand, F. and Delhomme, P., 2005. Speed behaviour as a choice between observing and exceeding the speed limit. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(6), pp.481-492.
- Lochner, H.V., Bhandari, M. and Tornetta III, P., 2001. Type-II error rates (beta errors) of randomized trials in orthopaedic trauma. *JBJS*, 83(11), pp.1650-1655.
- Mendenhall WM, Sincich TL. *Statistics for Engineering and the Sciences*. CRC Press; 2016 Apr 5.
- Naik, B., Tung, L.W., Zhao, S. and Khattak, A.J., 2016. Weather impacts on single-vehicle truck crash injury severity. *Journal of safety research*, 58, pp.57-65.
- National Research Council-United States (NRC-US). Transportation Research Board. Committee for Guidance on Setting and Enforcing Speed Limits, 1998. *Managing Speed: Review of Current Practice for Setting and Enforcing Speed Limits*.
- Noyce, David A.. (2019). International Conference on Transportation and Development 2019 - Smarter and Safer Mobility and Cities - 20.5.1 Actual Operating Speed. American Society of Civil Engineers (ASCE). Retrieved from <https://app.knovel.com/hotlink/pdf/id:kt0126ONUB/international-conference/actual-operating-speed>

- OECD (1976). Road Research: Adverse Weather, Reduced Visibility & Road Safety: Driving in Reduced Visibility Conditions due to Adverse Weather. OECD Publications, Paris.
- Oh C, Ritchie SG, Oh JS. Exploring the relationship between data aggregation and predictability to provide better predictive traffic information. *Transportation Research Record*. 2005;1935(1):28-36.
- Oh JS, Shim YU, Cho YH. Effect of weather conditions to traffic flow on freeway. *KSCE Journal of Civil Engineering*. 2002 Dec 1;6(4):413-20.
- Park D, Rilett LR, Gajewski BJ, Spiegelman CH, Choi C. Identifying optimal data aggregation interval sizes for link and corridor travel time estimation and forecasting. *Transportation*. 2009 Jan;36(1):77-95.
- Parvathy, R., Sreelatha, T. and Reebu, Z.K., 2013. Development of new PCU values and effect of length of passenger cars on PCU. *Int. J. Innovative Res. Sci. Eng. Technol*, 2(1), pp.344-351.
- Peng Y, Abdel-Aty M, Shi Q, Yu R. Assessing the impact of reduced visibility on traffic crash risk using microscopic data and surrogate safety measures. *Transportation research part C: emerging technologies*. 2017 Jan 1;74:295-305.
- Pisano PA, Goodwin LC, Rossetti MA. US highway crashes in adverse road weather conditions. In 24th conference on international interactive information and processing systems for meteorology, oceanography and hydrology, New Orleans, LA 2008 Jan 24.
- Pisano, P.A. and Goodwin, L.C., 2004. Research needs for weather-responsive traffic management. *Transportation research record*, 1867(1), pp.127-131.
- Rakha H, Farzaneh M, Arafeh M, Sterzin E. Inclement weather impacts on freeway traffic stream behavior. *Transportation Research Record*. 2008;2071(1):8-18.
- Regehr, J.D., Maranchuk, K., Vanderwees, J. and Hernandez, S., 2020. Gaussian mixture model to characterize payload distributions for predominant truck configurations and body types. *Journal of Transportation Engineering, Part B: Pavements*, 146(2), p.04020017.

- Regehr, J.D., Montufar, J. and Rempel, G., 2009. Safety performance of longer combination vehicles relative to other articulated trucks. *Canadian Journal of Civil Engineering*, 36(1), pp.40-49.
- Research Triangle Institute. *Speed and Accidents*, Vol. II. RTI Project SU-409. National Highway Safety Bureau, June 1970
- Robinson, M., 2000. Examples of variable speed limit applications: Speed management workshop.
- Schmidlin TW, Hammer BO, King PS, Miller LS. 2003. 2.2 WIND SPEEDS REQUIRED TO UPSET VEHICLES.
- Shaheed, M.S., Gkritza, K., Carriquiry, A.L. and Hallmark, S.L., 2016. Analysis of occupant injury severity in winter weather crashes: A fully Bayesian multivariate approach. *Analytic methods in accident research*, 11, pp.33-47.
- Shi, J. and Liu, M., 2019. Impacts of differentiated per-lane speed limit on lane changing behaviour: A driving simulator-based study. *Transportation research part F: traffic psychology and behaviour*, 60, pp.93-104.
- Sisiopiku, V.P., 2001. Variable speed control: technologies and practice. In *Proceedings of the 11th Annual Meeting of ITS America* (pp. 1-11).
- Solomon, D.H., 1964. *Accidents on main rural highways related to speed, driver, and vehicle*. US Department of Transportation, Federal Highway Administration.
- Statistics Canada. Census-Boundary files (Census divisions) [Internet] 2016b [cited 2021 July 24]. Available at: https://www12.statcan.gc.ca/census-recensement/alternative_alternatif.cfm?l=eng&dispxt=zip&teng=lcd_000b16a_e.zip&k=%20%20%20%2032295&loc=http://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/files-fichiers/2016/lcd_000b16a_e.zip
- Statistics Canada. Census-Boundary files (Provinces/territories) [Internet] 2016a [cited 2021 July 24]. Available at: https://www12.statcan.gc.ca/census-recensement/alternative_alternatif.cfm?l=eng&dispxt=zip&teng=lpr_000b16a_e.zip&k=%20%20%20%2027960&loc=http://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/files-fichiers/2016/lpr_000b16a_e.zip

- Taylor, M.A. and Bonsall, P.W., 2017. *Understanding traffic systems: data analysis and presentation*. Routledge.
- Thakuriah P, Tilahun N. Incorporating weather information into real-time speed estimates: comparison of alternative models. *Journal of transportation engineering*. 2013 Apr 1;139(4):379-89.
- Transport Canada, Collision statistics for 2018 [cited 2020 Nov 15]. Database: National Collision Database (NCDB) Online. Available at <https://wwwapps2.tc.gc.ca/Saf-Sec-Sur/7/NCDB-BNDC/p.aspx?l=en>
- Transportation Research Board, National Research Council,. *Highway Capacity Manual*; 2010.
- Transportation Research Board, National Research Council,. *Highway Capacity Manual*; 2016.
- Transportation Research Board. (2022). *Highway Capacity Manual (7th Edition) - A Guide for Multimodal Mobility Analysis*. Transportation Research Board (TRB). Retrieved from <https://app.knovel.com/hotlink/toc/id:kpHCMEAGM2/highway-capacity-manual/highway-capacity-manual>
- Xing, F., Huang, H., Zhan, Z., Zhai, X., Ou, C., Sze, N.N. and Hon, K.K., 2019. Hourly associations between weather factors and traffic crashes: non-linear and lag effects. *Analytic methods in accident research*, 24, p.100109.
- Yasanthi, R.G. and Mehran, B., 2020. Modeling free-flow speed variations under adverse road-weather conditions: Case of cold region highways. *Case studies on transport policy*, 8(1), pp.22-30.
- Yasanthi, R.G. and Mehran, B., 2022. Application of Different Data Analytics for Evaluation of Heavy Vehicle Vulnerability in Cold-Region Rural Highways. *Transportation Research Record*, p.03611981221111353.
- Yasanthi, R.G., Mehran, B. and Alhajyaseen, W.K., 2021. Modelling speed behaviour in rural highways: Safety analysis of driving under adverse road-weather conditions. *PLoS one*, 16(8), p.e0256322.

- Yasanthi, R.G., Mehran, B. and Alhajyaseen, W.K., 2022. A reliability-based weather-responsive variable speed limit system to improve the safety of rural highways. *Accident Analysis & Prevention*, 177, p.106831.
- Zhang, Z., 2016. Parametric regression model for survival data: Weibull regression model as an example. *Annals of translational medicine*, 4(24).
- Zhu, L., Yu, F.R., Wang, Y., Ning, B. and Tang, T., 2018. Big data analytics in intelligent transportation systems: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 20(1), pp.383-398.
- Zimasa, T., 2018. *The influence of mood and cognitive load on driver performance: using multiple measures to assess safety* (Doctoral dissertation, University of Leeds).

Appendix

A: Kolmogorov-Smirnoff (K-S) test results

In this thesis, the two-sample K-S test was used to (i) evaluate if speeds observed under each level of road-weather condition that is originally recorded by the RWIS is different from the remaining levels of that road-weather condition, and/or (ii) group continuous variables (e.g., temperature, traffic flow) such that the difference of speeds in each level of a particular variable is statistically different from the rest of the levels of that variable. In this thesis, H_0 : the null hypothesis, and H_1 : the alternative hypothesis, are formulated as:

H_0 : Speeds observed in both levels tested are drawn from the same statistical distribution

H_1 : Speeds observed in the two levels tested are drawn from different statistical distributions

It should be noted that only the speeds recorded in different levels of a specific road-weather/traffic attribute are subjected to the K-S tests in this thesis. For instance, slight rain conditions are only tested against moderate rain and heavy rain categories. The following section presents the results of the K-S tests in terms of (i) D : the K-S statistic, (ii) p : the probability of null hypothesis being true, and (iii) the final outcome (i.e., accept/reject null hypothesis) and its interpretation (i.e., if speeds in the two levels are drawn from the same statistical distribution or not). It should be noted that, categorical variable levels where speeds are drawn from the same statistical distribution are combined together to form a new group except in some cases where the original categorical variable levels are sustained for the purpose of practicality and/or consistency.

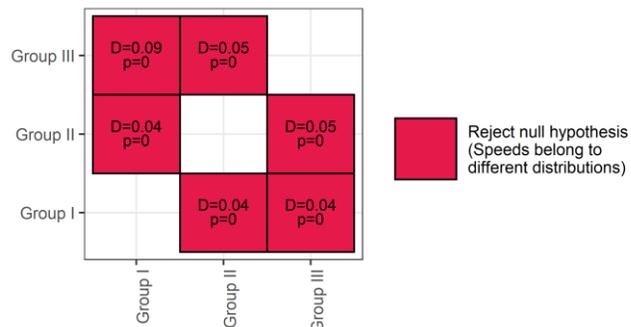


Figure A. 1: K-S test results for temperature category levels

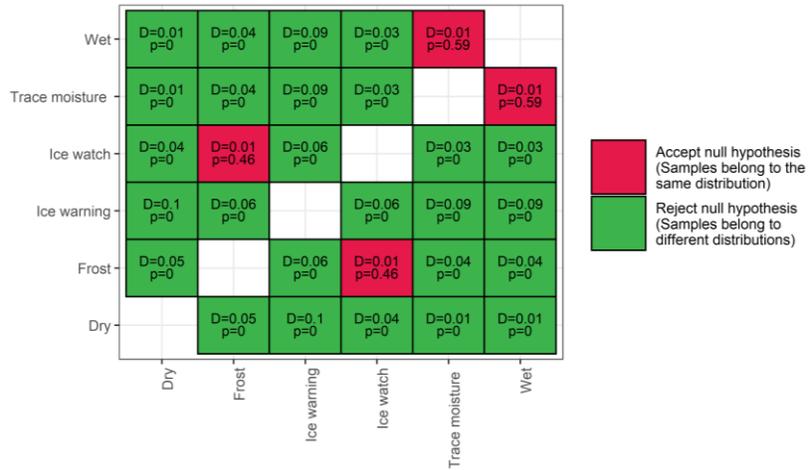


Figure A. 2: K-S test results for pavement surface condition levels

Note:

- Original pavement surface levels retained for practicality.

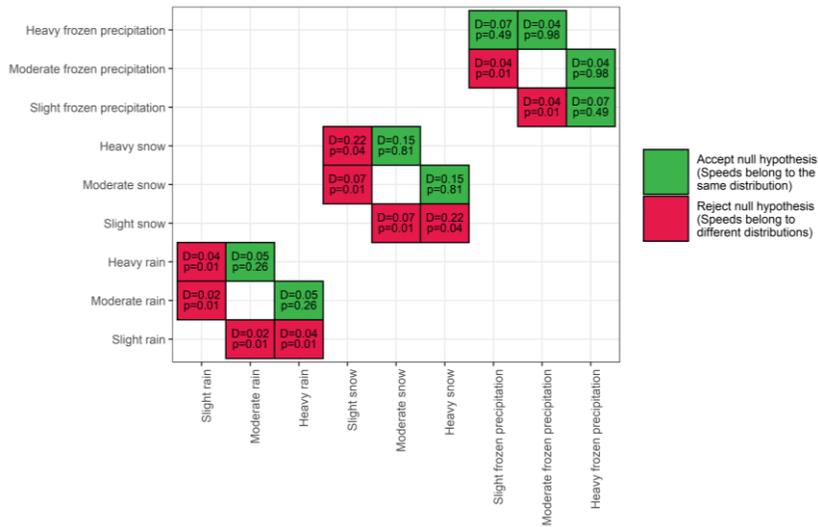


Figure A. 4: K-S test results for precipitation condition levels

Note:

- Moderate and heavy precipitation levels combined for all precipitation types for consistency.

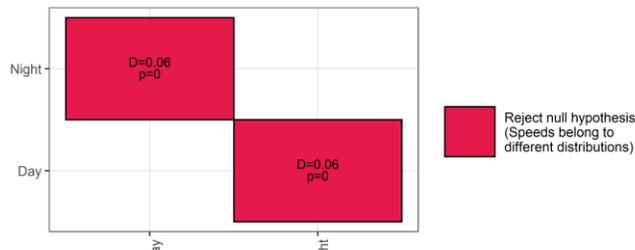


Figure A. 3: K-S test results for time of the day categories

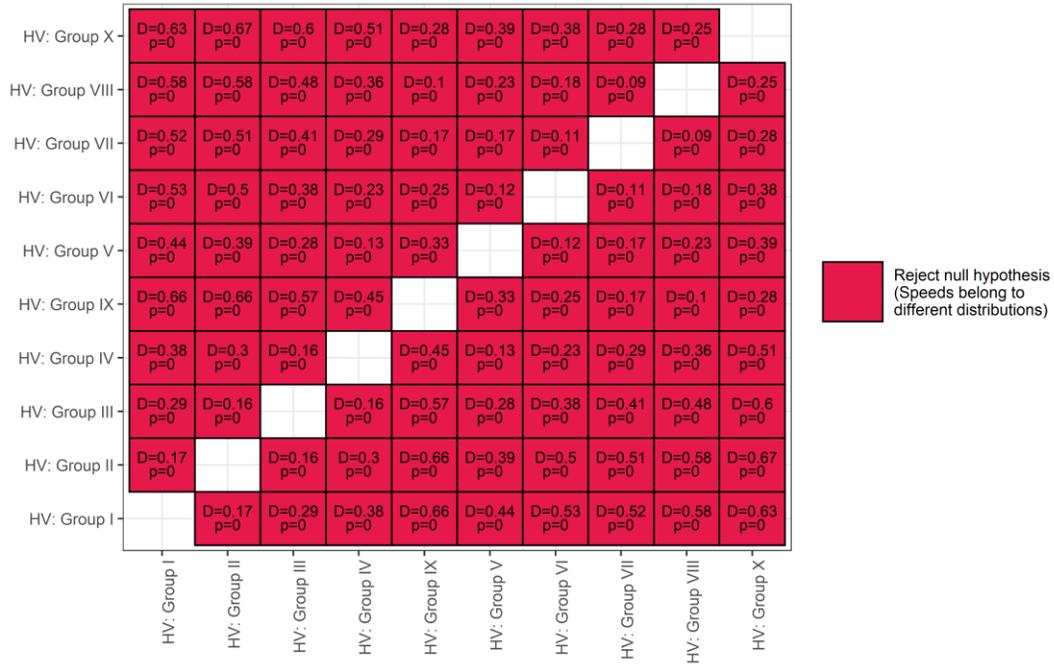


Figure A. 6: K-S test results for heavy vehicle (HV) percentage categories

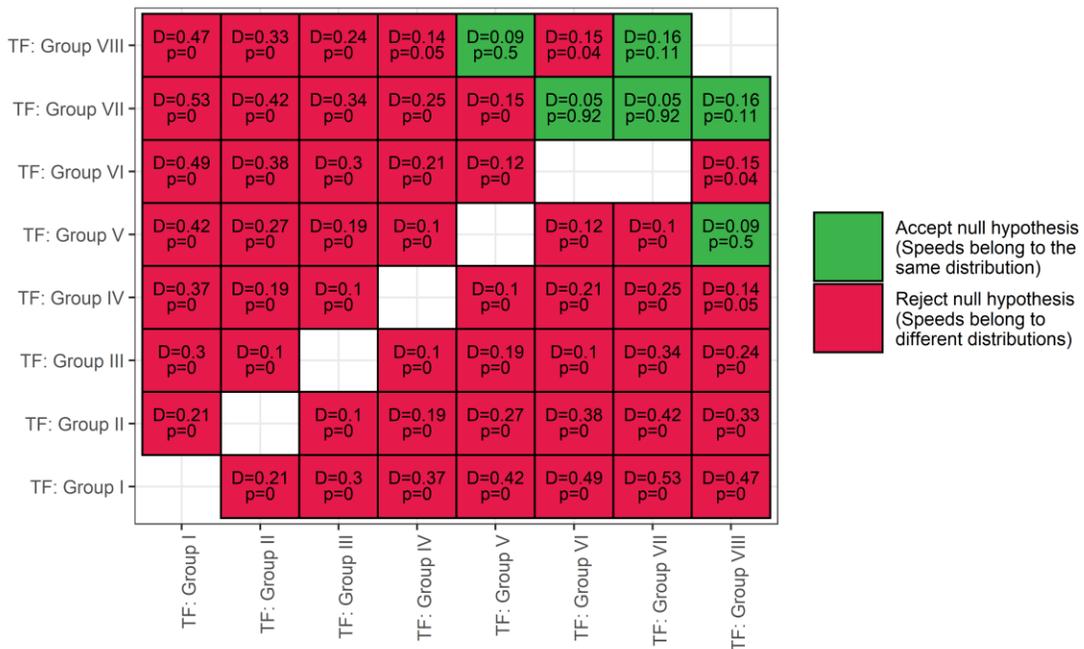


Figure A. 5: K-S test results for traffic flow categories

Note:

Traffic flow categories kept at 100 veh/h bins to retain consistency.



Figure A. 8: K-S test results for vehicle type categories

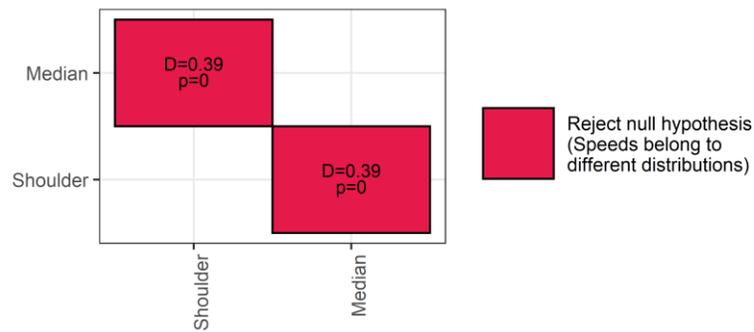


Figure A. 7: K-S test results for travel lane categories

B: Estimating the mean (μ) and the set of weight factors representing the contribution of each five-minute aggregate speed distribution in the desired speed distribution

A detailed demonstration of the algorithm proposed in Figure 2.9 where a particular desired speed distribution is represented by three five-minute aggregate speed distributions, is presented below. The algorithm presented in Figure 2.9, takes the parameters of the five-minute aggregate speed distributions $V_{i;i=1 \text{ to } k}$ as inputs for a given road-weather and traffic conditions. The algorithm outputs the MVUE of the mean of the desired speed distribution μ , resulting from the linear combination of the five-minute aggregate speed distributions $V_{i;i=1 \text{ to } k}$ along with weight factors $a_{i;i=1 \text{ to } k}$ representing the contribution of each five-minute aggregate speed distribution in the population-level desired speed distribution.

B.1 Estimating the mean of the desired speed distribution (μ)

Let a specific desired speed distribution under a particular combination of road-weather and traffic conditions (i.e., a population distribution) characterized by a mean of μ and a standard deviation of σ be represented by three ($k=3$ in Eq. 2.1) distinct distributions of five-minute aggregate speeds (i.e., sampling distributions). The value of k is read first ($k=3$ in this case), while an integer variable i is defined and initialized to 1 afterwards.

Thereafter, the input data; the five-minute aggregate speed distributions $V_i \sim N(\mu_i, \sigma_i^2)$ representing the desired speed distribution in question, are read and subjected to a chaining process where each iteration of the chaining process linearly combines at least one of the five-minute aggregate speed distributions with either another five-minute aggregate speed distribution or an intermediate hypothetical five-minute aggregate speed distribution. As stated earlier, the chaining process is initiated with $i= "1"$ yielding,

$$V_i = V_1 \sim N(\mu_1, \sigma_1^2) \quad \text{Eq. A-1}$$

and

$$V_{i+1} = V_2 \sim N(\mu_2, \sigma_2^2) \quad \text{Eq. A-2}$$

Linear combination of V_1 and V_2 yields the first intermediate hypothetical normal distribution $Y'_{1,2}$, representing the weighted combination of V_1 and V_2 , expressed by:

$$Y'_{i,i+1} = Y'_{1,2} = a'_1 V_1 + a'_2 V_2 \sim N(\mu'_{1,2}, (\sigma'_{1,2})^2) \quad \text{Eq. A-3}$$

$$\mu'_{i,i+1} = \mu'_{1,2} = a'_1 \mu_1 + a'_2 \mu_2 \quad \text{Eq. A-4}$$

$$(\sigma'_{i,i+1})^2 = (\sigma'_{1,2})^2 = (a'_1)^2 \sigma_1^2 + (a'_2)^2 \sigma_2^2 \quad \text{Eq. A-5}$$

The weight factors a'_1 and a'_2 respectively, represent the contribution of the five-minute aggregate speed distributions V_1 and V_2 in the resulting hypothetical five-minute aggregate speed distribution $Y'_{1,2}$. Thus, the weight factors a'_1 and a'_2 can be estimated by minimizing the variance of $Y'_{1,2}$ yielding:

$$a'_1 = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \quad \text{Eq. A-6}$$

and

$$a'_2 = 1 - a'_1 \quad \text{Eq. A-7}$$

Subsequently, the parameters of the first intermediate hypothetical five-minute aggregate speed distribution i.e., $Y'_{1,2}$ and the weight factors a'_1 and a'_2 which are respectively associated with the linear combination of V_1 and V_2 are stored before proceeding to the next step of the algorithm. The intermediated hypothetical five-minute aggregate speed distribution $Y'_{1,2}$ is chained with the next five-minute aggregate speed distribution until all k five-minute aggregate speed distributions are included in the chaining process. Accordingly, the need for the continuation of the chaining process is checked by evaluating the equality of the latest value of i . At the end of the first iteration, the value of i is equal to 1 and the value of “ $k-1$ ” is equal to 2, leading to $i \neq k-1$, which prompts the execution of the second iteration of the chaining process. In the second iteration, the first intermediate hypothetical five-minute aggregate speed distribution $Y'_{1,2}$ is linearly combined with the third five-minute aggregate speed distribution. Accordingly, variables i and V_i are assigned with new values as presented below.

$$V_i \sim N(\mu_i, \sigma_i^2) \leftarrow Y'_{1,2} \sim N(\mu'_{1,2}, (\sigma'_{1,2})^2) \quad \text{Eq. A-8}$$

$$i = 2 \quad \text{Eq. A-9}$$

Subsequently, V_3 is read as the first step of the second iteration. The second iteration produces the second intermediate hypothetical five-minute aggregate speed distribution $Y'_{1,2,3}$, representing the weighted combination of $Y'_{1,2}$ and V_3 , which is expressed by:

$$Y'_{1,\dots,i+1} = Y'_{1,2,3} = a'_{1,2}Y'_{1,2} + a'_3V_3 \sim N\left(\mu'_{1,2,3}, (\sigma'_{1,2,3})^2\right) \quad \text{Eq. A-10}$$

$$\mu'_{1,\dots,i+1} = \mu'_{1,2,3} = a'_{1,2}\mu'_{1,2} + a'_3\mu_3 \quad \text{Eq. A-11}$$

$$(\sigma'_{1,\dots,i+1})^2 = (\sigma'_{1,2,3})^2 = (a'_{1,2})^2(\sigma'_{1,2})^2 + (a'_3)^2\sigma_3^2 \quad \text{Eq. A-12}$$

The weight factors $a'_{1,2}$ and a'_3 respectively represent the contribution of the five-minute aggregate speed distributions $Y'_{1,2}$ and V_3 in the resulting hypothetical five-minute aggregate speed distribution $Y'_{1,2,3}$ bearing the minimum variance among the potential combinations of $Y'_{1,2}$ and V_3 . Thus, the weight factors $a'_{1,2}$ and a'_3 can be estimated by minimizing the variance of $Y'_{1,2,3}$ yielding:

$$a'_{1,2} = \frac{\sigma_3^2}{(\sigma'_{1,2})^2 + \sigma_3^2} \quad \text{Eq. A-13}$$

and

$$a'_3 = \frac{(\sigma'_{1,2})^2}{(\sigma'_{1,2})^2 + \sigma_3^2} \quad \text{Eq. A-14}$$

Similar to the first iteration, the values of the parameters $Y'_{1,2,3}$, $a'_{1,2}$ and a'_3 are stored prior to proceeding to the next step in the algorithm. For a desired speed distribution represented by three distributions of five-minutes aggregate speeds, i.e., $k=3$, the chaining process is completed at the end of the second iteration as $i=k-1=2$ upon the completion of the second iteration of the chaining process. Therefore, the chaining process is terminated, and the intermediate hypothetical five-minute aggregate speed distribution produced at the end of the second iteration is concluded as the mean of the random variable Y ; $\mu_{\bar{y}}$ (Eq. 2.2) representing the linear combination of the three five-minute aggregate distributions $V_{i,i=1 \text{ to } 3}$. Accordingly, the MVUE of $\mu_{\bar{y}}$ can be expressed in terms of the final intermediate hypothetical five-minute aggregate speed distribution (Eq. A-15), which is $\mu'_{1,2,3}$. As explained earlier, the MVUE of $\mu_{\bar{y}}$ is also recognized as the mean of the population mean μ . Therefore,

$$\mu = \mu_{\bar{y}} = \mu'_{1,2,3} \quad \text{Eq. A-15}$$

B.2 Estimating the set of weight factors representing the contribution of each five-minute aggregate speed distribution in estimating the desired speed distribution

Mean of the random variable Y ; $\mu_{\bar{y}}$ for a case of $k = 3$ in Eq. 2.2 can be expressed as:

$$\mu_{\bar{y}} = \sum_{j=1}^3 a_j \mu_j \quad \text{Eq. A-16}$$

On the other hand, $\mu_{\bar{y}}$ can be also expressed in terms of $\mu'_{1,2,3}$ (Eq. A-15) where $\mu'_{1,2,3}$ is primarily equivalent to the linear combination of $\mu'_{1,2}$ and μ_3 as expressed in Eq. A-11. Similarly, $\mu'_{1,2}$ (Eq. A-4) is equivalent to the linear combination of μ_1 and μ_2 . It should be noted that the weight factors $a_{i;i=1 \text{ to } k}$ are yet to be estimated even though the terms $a'_{1,2}$, $\mu'_{1,2}$ and a'_3 are known at this stage of the algorithm, as illustrated in Figure 2.9. The weight factor a_i representing the contribution of the mean of the five-minute aggregate speed distribution V_i in estimating $\mu_{\bar{y}}$ can be estimated by considering the coefficient of μ_i in Eq. A-16. Further, Eq. A-17 is derived by combining Eq. A-11, A-15 and A-16:

$$a'_{1,2}\mu'_{1,2} + a'_3\mu_3 = \sum_{j=1}^3 a_j \mu_j \quad \text{Eq. A-17}$$

While Eq. A-17 contains the terms $\mu'_{1,2}$ and μ_3 , it does not include the terms μ_1 and μ_2 restricting the estimation of the weight factors a_1 and a_2 . Therefore, Eq. A-17 is transformed to include the terms μ_1 , μ_2 and μ_3 . By substituting $\mu'_{1,2}$ from Eq. A-4 in Eq. A-17:

$$a'_{1,2}(a'_1\mu_1 + a'_2\mu_2) + a'_3\mu_3 = \sum_{j=1}^3 a_j \mu_j \quad \text{Eq. A-18}$$

Expanding the Right-Hand Side (R.H.S.) and rearranging Eq. A-18 yields:

$$(a'_1 a'_{1,2} - a_1)\mu_1 + (a'_2 a'_{1,2} - a_2)\mu_2 + (a'_3 - a_3)\mu_3 = 0 \quad \text{Eq. A-19}$$

Eventually, the weight factors a_1 , a_2 and a_3 can be estimated by evaluating the coefficients of the terms μ_1 , μ_2 and μ_3 in Eq. A-19 as presented below.

- By considering the coefficient of μ_1 :

$$a'_1 a'_{1,2} - a_1 = 0 \quad \text{Eq. A-20}$$

$$a_1 = a'_{1,2} a'_1 \quad \text{Eq. A-21}$$

Substituting for a'_1 and $a'_{1,2}$ respectively from Eq. A-6 and Eq. A-13 in Eq. A-21 yields:

$$a_1 = \frac{\sigma_2^2 \times \sigma_3^2}{(\sigma_1^2 + \sigma_2^2) \times ((\sigma'_{1,2})^2 + \sigma_3^2)} \quad \text{Eq. A-22}$$

- By considering the coefficient of μ_2 :

$$a'_2 a'_{1,2} - a_2 = 0 \quad \text{Eq. A-23}$$

$$a_2 = a'_2 a'_{1,2} \quad \text{Eq. A-24}$$

Substituting for a'_2 and $a'_{1,2}$ respectively from Eq. A-7 and Eq. A-13 in Eq. A-24 yields:

$$a_2 = \frac{\sigma_1^2 \times \sigma_3^2}{(\sigma_1^2 + \sigma_2^2) \times ((\sigma'_{1,2})^2 + \sigma_3^2)} \quad \text{Eq. A-25}$$

- By considering the coefficient of μ_3 :

$$a'_3 - a_3 = 0 \quad \text{Eq. A-26}$$

$$a_3 = a'_3 \quad \text{Eq. A-27}$$

Substituting for a'_3 from Eq. A-14 in Eq. A-27 yields:

$$a_3 = \frac{(\sigma'_{1,2})^2}{(\sigma'_{1,2})^2 + \sigma_2^2} \quad \text{Eq. A-26}$$

Ultimately, the algorithm returns the values of $\mu'_{1,2,3}$ as the mean of the desired speed distribution (i.e., the population) represented by the five-minute aggregate speed distributions $V_{i;i=1 \text{ to } 3}$ as well as the weight factors $a_{i;i=1 \text{ to } 3}$ representing the contribution of each five-minute aggregate speed distribution in the population's desired speed distribution.