©Copyright 2021
Huizhong Guo

Assessing Driver Behavior in the Context of Driving Environment

Huizhong Guo

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:
Linda Ng Boyle, Chair
Don MacKenzie
Youngjun Choe

Program Authorized to Offer Degree:
Civil & Environmental Engineering

University of Washington

Abstract

Assessing Driver Behavior in the Context of Driving Environment

Huizhong Guo

Chair of the Supervisory Committee:
Professor Linda Ng Boyle
Industrial & Systems Engineering, Civil & Environmental Engineering

Driver-related factors have long been an important component in traffic safety. Studies to assess driver behavior and the related safety concerns have primarily used data that does not capture the dynamic nature of driving tasks. The widespread use of naturalistic driving data in recent years allows researchers the capability to capture real-time driver behavior and be able to infer an individual's driving style. However, current studies focus largely on at-risk safety behavior that is often incomplete (e.g., does not consider all types of at-risk safety behavior) and broadly defined regardless of the driving environment. The goal of this dissertation is to assess driver behavior in the context of the driving environment. This is accomplished using data from the second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study, which includes more than 3,000 drivers on the road from 2010 to 2013. The concept of "abnormal" driving style is proposed as a complement to "normal" driving style. More specifically, the "abnormality" measures how much a driver deviates from the average driving behavior given the driving context. In this study, the average driving behavior is defined as the average of different vehicle kinematics for drivers that participated in SHRP2 and for a specific environmental context. The study thus aims to examine the association between driving "abnormality" and driver safety. Environmental factors that contribute to the formation of "normal" driving styles were identified in a systematic way through multivariate functional data clustering method and decision trees. The "abnormal-

ity" were described by a composite score as well as a set of statistical features that capture the different aspects of a driving style. Path analysis and Structural Equation Modeling method were used to reveal associations between driver safety and driving "abnormality". Results from the study provide insights into driver behavior and implications on driver safety in different environmental contexts. For example, the study showed that drivers who were more likely to crash were also more likely to have unstable lateral control on Urban Interstates. These findings can be integrated in autonomous vehicle algorithms where individual driving styles are considered. It can also provide insights on the development of new technologies to identify risky drivers and to quantify their risky levels.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

ADAS	Advanced Driver Assistance System
ADS	Automated Driving System
CAN	Controller Area Network
CB-SEM	Covariance-Based Structural Equation Modeling
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
DAS	Data Acquisition System
DUI	Driving Under the Influence
EFA	Exploratory Factor Analysis
FDA	Functional Data Analysis
FDR-BH	Benjamini and Hochberg's False Discovery Rate
GMM	Gaussian Mixture Model
LOS	Level of Services
MFPCA	Multivariate Functional Principal Components Analysis
ND	Naturalistic Driving
NDS	Naturalistic Driving Study
NIR	Non Information Rate
OBD	On-Board Diagnostics
PC	Principal Component
PLS-SEM	Partial Least Squares Structural Equation Modeling
RID	Roadway Information Database
RMSEA	Root Mean Square Error of Approximation
SEM	Structural Equation Modeling
SHRP2	Strategic Highway Research Program 2
SRMR	Standardized Root Mean Square Residual
UI	Environmental group #1 on Urban Interstate in Table 5.1
UMA	Environmental group #10 on Urban Minor Arterial in Table 5.1
UOPA	Environmental group #16 on Urban Other Principal Arterial in Table 5.1
V2V	Vehicle-to-Vehicle
VDT	Vehicle Distance Traveled

ACKNOWLEDGMENTS

To my advisor Dr. Linda Ng Boyle – I would like to express my most sincere gratitude to you, for giving me the opportunity to be a PhD student at the very beginning, and for all your guidance, support, understanding and trust throughout the years. This valuable chance has opened up a new chapter in my life, where the experience learning and doing research with you would definitely become one of the most precious memories for me.

To each of my committee members – I would like to thank you, Dr. Don MacKenzie, Dr. Youngjun Choe and Dr. Gary Chan, for all your insightful comments and valuable suggestions on my dissertation work.

To all my colleagues in the Human Factors and Statistical Modeling Lab – These are Xingwei Wu, Yilun Xing, Jundi Liu, Mayuree Binjolkar, Haena Kim, Tianshu Feng, Steven Hwang, Ning Li, Klaas Fiete Krutein, Erika Miller, Chun-Cheng Chang, Yuqing Wu, and Yiyun Peng. Thank you for all your help, inspiration and friendship.

Finally, to my beloved parents – I wouldn't be here without all your love and support. You give me the courage to face the challenges and the unknown. Thank you.

Chapter 1

INTRODUCTION

Road traffic safety remains a serious problem in recent years. In 2016, road traffic injury leads to the death of 1.35 million people worldwide, ranked as the 8^{th} leading cause of death for all age groups and the leading cause of death for children and young adults (5 to 29 years) (WHO, 2018). Crash contributing factors are traditionally classified into three categories, driver, vehicle and roadway/environment (FHWA, 2017). Among the three, driver-related factors are considered the most significant component in all crashes (FHWA, 2010).

Driver-related factors include drivers' demographic attributes (e.g., age and gender) and their behavioral choices that are more directly associated with crash occurrence. The National Motor Vehicle Crash Causation Survey examined a sample of 5,470 crashes from 2005 to 2007 and found 94% of crashes due to driver-related reasons (Singh, 2015). The top driver-related reasons were (1) recognition error, such as driver's inattention, (2) decision error, such as driving too fast for conditions, (3) performance error, such as poor directional control, and (4) non-performance error, such as sleep. Recent crash statistics of 2016 in the U.S. show a consistent observation, where speeding, drunk driving, driver distraction, and drowsy driving accounted for 66% of all fatal crashes (NHTSA, 2017a).

Advanced technologies have been introduced to help mitigate the impact of human errors. The Vehicle-to-Vehicle (V2V) communication technology transmits safety information between vehicles and warns drivers in advance of impending crashes (Harding et al., 2014). The autonomous vehicle technology, currently appeared in the form of Advanced Driver Assistance Systems (ADAS) and conditional Automated Driving Systems (ADS), can conduct certain driving tasks and in some situations monitor the driving environment for the driver (NHTSA, 2017b). These technologies provide hope for a driverless future that may finally removes human errors and reduce roadway crashes.

During the meantime, nevertheless, drivers will still be kept in the loop and servers a critical role in traffic safety. It is thus important to learn driver behaviors in different driving contexts and be able to assess how safe or how risky those behaviors are. This information will help deepen our understanding of driver behavior and the underlying relationship of it with crash or near-crash occurrence. It can also benefit real-world applications to help reduce crashes and improve traffic safety, e.g., to tailor education programs that target driver behavior with higher safety concerns. In addition, it could help improve the ADAS or ADS and the V2V communication algorithms by incorporating information on individual driving styles. For example, the ADAS or ADS could be customized to mimic a driver's driving style with modifications on aspects that are considered more risky. The vehicles and roadside infrastructure could be designed to have more frequent communications with vehicles (drivers) that have a higher propensity to behave in a less safe way.

The goal of the dissertation is to assess driver behavior in the context of driving environment. In Chapter 2, previous studies related to this topic, the gap in the literature and the specific research aims are introduced. Chapter 3 describes the data used for the dissertation and the main work is present from Chapters 4 to 6. Chapter 7 summarizes the major findings and contributions, and provides a discussion on limitations and future research opportunities.

Chapter 2
BACKGROUND

This chapter provides a literature review of studies related to driver behavior assessment. Driver behavior is defined as the way in which a person acts in response to the surrounding environment while driving, where the exact process of carrying out the action(s) is the driving performance. The different types of data used to assess driver behavior are firstly introduced in this chapter. Emphasis is given to driver behavior assessment from the perspective of driving style using naturalistic driving data. Gaps in the literature are summarized, followed by the specific research questions and aims for this dissertation.

2.1 Driver Behavior Assessment

2.1.1 Historical Driving Records

Historical driving records, such as crash and moving violation histories, might be the first type of data that comes to mind as for driver behavior assessment. In fact, vehicle insurance companies have a long history using the historical driving records to identify more "risky" drivers and to increase their premiums accordingly (WA OIC, 2016). Das et al. examined the crash data from 2004 to 2011 in Louisiana. They divided drivers into four groups based on whether they were at-fault for crash occurrence and whether they have been involved with multiple crashes. They found that at-fault drivers with multiple crashes were associated with higher fatality and severity rates. They thus suggested more regular targets on this group of drivers through education and regulation (Das et al., 2015). Wang et al. used drivers' two-year crash or moving violation records and demographics to predict their driving risk in the subsequent two years. Four machine learning algorithms were used and the algorithm performs the best for the subset of drivers with recurrent crash involvement or were at-fault in severe crashes. In addition, among the top nine most important features, seven were related to prior crash or moving violation records (Wang et al., 2019).

Historical driving records are direct measures of driver safety. Previous studies validate its potential to predict the riskiness of drivers in a future situation. However, crashes are rare events and moving violations are also infrequent. This type of data could be difficult to obtain in real life and self-reported records may suffer from bias due to, e.g., social desirability or vague memory. Also, behavior assessment based on crash records are more reactive than proactive. Since crashes can have severe consequences, one may not want to wait until a crash happens to take an action. That said, historical driving records are still valuable and can be used as the benchmark for how safe or how risky a driver behaves.

2.1.2 Driver's Socio-demographics and Personality

Drivers' socio-demographics have been used in driver behavior assessment. For example, Massie et al. found elevated rates for drivers aged 16–19 and 75 and over, and for men than women in fatal crashes using crash and survey data in 1990. The author contributed the higher rates for younger and male drivers to an increased propensity to drive in a risky manner (Massie et al., 1995). For marital status, single drivers were shown with higher risk of injury than married people in traffic crashes (Whitlock et al., 2004). Not surprisingly, these socio-demographic factors have all been considered by vehicle insurance companies to rate the potential behavior of their customers (WA OIC, 2016).

Many questionnaires have been developed to measure driver personality, including the NEO five traits inventory (Arthur Jr and Graziano, 1996), the Zuckerman's Sensation Seeking Scale (Arnett, 1990) and the Zuckerman-Kuhlman Personality Questionnaire (Sârbescu et al., 2012). Relationships have been found between driver personality and crash or near-crash rates (Guo and Fang, 2013), as well as between certain personality traits (e.g., sensation seeking, aggression-hostility) and adverse driving styles, such as risky driving and drunk driving (Poó and Ledesma, 2013; McMillen et al., 1992).

Driver's socio-demographics can be easily obtained. Although there are associations between these factors and crash occurrence, the assumption of causality may be implausible. Socio-demographic factors are too general to represent individuals: two drivers with the same socio-demographic information can drive in distinct ways. Driver personality provides more detailed approach to describe individuals. By linking personality traits to specific driving styles, the safety or risky level of a driver can be evaluated. A major limitation of driver personality is that it is static information and does not capture the dynamic nature of driving.

2.1.3 Driving Style

Driving style characterizes the way a driver drives and has long been used to understand the relationship between driver behavior and roadway safety. In a review of driving style and road safety studies, Sagberg et al. defined driving style as a "habitual way of driving, which is characteristic for a driver or a group of drivers" (Sagberg et al., 2015). Previously, driving style was primarily measured by self-reported instruments. Although questionnaires were designed in different ways, they capture some similar aspects of driving styles, such as aggressive driving and high-velocity driving. These studies in general confirm a correlation between certain driving styles and crash or violation involvement. While a significant association is shown in some studies (French et al., 1993), others may only observe a weak correlation, especially for crashes (Dula, 2003). Also, the relative importance of different driving styles and their impacts on crash risk are not yet clear.

Compared to the aforementioned data types, driving style is easy to measure and captures characteristics of an individual's driving behavior. However, the same limitation exists for driving style measured by questionnaires that they don't capture the dynamic nature of driving. Fortunately, with the rapid development of technology in recent years, the new type of data – Naturalistic Driving (ND) data – provide researchers a tool to deal with this limitation. The ND data can be obtained from in-vehicle Controller Area Network (CAN bus) via an On-Board Diagnostics scanner (Kwak et al., 2016) or through a cell phone (Mantouka et al., 2019). The easy availability and the rich information of ND data introduce a new approach to capture driver's driving style. The next section will introduce studies that use this type of data for assessing driver behavior.

2.2 Driving Style Based on ND Data

As we have briefly mentioned in the previous section, driving styles can be categorized into different types. In practice, when using ND data to identify driving style, researchers may focus on only specific types according to their research purpose. In the review by Sagberg et al. (2015), driving styles were categorized into four types:

- *Defensive driving*, a safe driving style.
- *Aggressive driving*, which includes hostile aggression such as road rage, and instrumental aggression to reach a goal faster, such as weaving and tailgating
- *Deviant and risky driving*, which could correlate or overlap the aggressive driving style. Risky driving includes reckless driving while deviant driving refers to unusual behaviors deviated from socially accepted norms (e.g., overtaking on the inside).
- *Concentrated and focused driving*, which focuses on the concentration and attention to driving tasks and includes driving behaviors such as distracted driving.

Another review by Meiring and Myburgh (2015) also proposed four types of driving styles:

- *Normal/safe driving style*
- *Aggressive driving*, which includes risky speeding profiles, improper vehicle position maintenance and inconsistent or excessive acceleration and deceleration. It also includes driver behaviors that intentionally increases the risk of a collision due to driver impatience, annoyance, hostility or an attempt to minimise travelling time.
- *Inattentive driving* due to driver fatigue or driver distraction
- *Drunk driving*, that refers to driver behaviors under influence of intoxication

Categories derived from the two papers are very similar. The *aggressive driving* defined in Meiring and Myburgh (2015) appears to overlap the two categories *aggressive driving* and *deviant and risky driving* defined in Sagberg et al. (2015). *Normal/safe driving* is similar to *defensive driving* and *inattentive driving* is similar to *concentrated and focused driving.* Four types of driving styles were thus considered and each introduced in the remainder of this section. These are aggressive and risky driving, inattentive driving, DUI (Driving Under the Influence) driving, and normal/safe driving.

2.2.1 Aggressive and Risky Driving Styles

Aggressive and risky driving is the most widely examined driving style by previous studies. Most studies use the word "aggressive", but the two words are quite exchangeable. For example, Castignani et al. (2015) proposed an algorithm to detect "risky" driving events which were considered more likely to happen in a so-called "aggressive" run. In a few studies, where both "aggressive" and "risky" driving were examined, the "aggressive" driving was inferred by higher percentage of harsh acceleration or deceleration, while "risky" driving was inferred by speeding (Mantouka et al., 2019). Due to the similarity of the two driving styles, we will use "aggressive" driving to represent both of them unless the literature explicitly differs "risky" driving from "aggressive" driving.

Studies in this category relates driving style to road safety from two different perspectives. The first is to identify aggressive maneuvers as a representation of aggressive driving style. The aggressive maneuver almost always includes speeding, hard acceleration and hard deceleration. Some studies include additional maneuvers such as aggressive lane change and aggressive turning. The maneuver can be detect either by setting a threshold on vehicle kinematics (Castignani et al., 2015) or by referring to stored templates of the specific maneuver (Johnson and Trivedi, 2011; Wu et al., 2013; Júnior et al., 2017). Drivers with more detected maneuvers were considered more risky (Toledo et al., 2008).

The second approach uses vehicle kinematics to directly infer a driver's driving style. Some studies classify drivers into a "normal" or an "aggressive" driving style group by setting thresholds on multiple vehicle kinematics, e.g., longitudinal and lateral speed and acceleration. A Fuzzy Logit model is frequently used in this situation (Aljaafreh et al., 2012;

Castignani et al., 2013). Alternatively, studies classify an individual driving style according to pre-defined aggressive or safe driving style templates (Constantinescu et al., 2010; Eren et al., 2012). In both situations, the standard or threshold defining safe or aggressive driving styles could be based on expert knowledge, historical crash or moving violation records and/or driving style questionnaires (Hong et al., 2014).

2.2.2 Inattentive Driving Style

Bergasa et al. (2014) used sensors and forward cameras of a cell phone to detect driver drowsiness and distraction. Drowsiness was inferred by lane weaving and drifting while distraction was inferred by sudden longitudinal and lateral movements. Lane weaving was confirmed by an unintentional lane change and drifting was based on an indicator called Lanex. For distraction, sudden harsh acceleration or deceleration behaviors were determined by referring thresholds taken from previous literature. A score was given to the driver as a measure of how safe they were based on the frequency and the level of the two inattentive driving types. Mantouka et al. (2019) used data collected from a smartphone application and examined the inattentive driving in addition to aggressive and risky driving. They found a similar small portion of events ($\sim$ 8%) associated with inattentive driving in both the aggressive and non-aggressive groups. Li et al. (2016) also used smartphone to detect four pre-defined dangerous behavior types: three relates to aggressive driving and one relates to inattentive driving, i.e., operating a cell phone. The behavior was detected based on acceleration and angular rates of the smartphone sensor.

2.2.3 DUI Driving Style

DUI driving has not been extensively examined as a driving style. However, many previous studies revealed increased crash or fatality rates under the influence of alcohol or drugs while driving (Blomberg et al., 2005; Brewer et al., 1994; Romano et al., 2017). Efforts have also been successfully made to detect DUI driving using vehicle sensor data on lateral and longitudinal acceleration, lane position and steering angle (Dai et al., 2010; Li et al., 2015).

2.2.4 Normal/Safe Driving Style

Compared to aggressive, inattentive and DUI driving styles, normal/safe driving style usually is not the main focus of a study but serves as the baseline or a comparison group. Different approaches have been used in previous studies to identify this driving style. One approach focuses more on the "safe" aspect of a driving style, oftentimes based on expert knowledge. For example, Eren et al. let two experienced drivers sit on the passenger side of the vehicle and evaluated drivers' driving performance to determine their safety level (Eren et al., 2012). Another approach focuses more on the "normal" aspect of a driving style, usually by considering the average behavior of all observed drivers as the "normal" one. Vlahogianni et al. (2014) used this approach on ND data to detect irregular driving patterns of drivers operating powered-two-wheelers.

2.3 Gaps in the Literature

Different data types and methods used in previous studies to assess driver behavior were reviewed in Section 2.1 and Section 2.2. Historical driving records are direct measures of driver safety but are rare events and reactive in the sense of crash avoidance. Drivers' socio-demographics and personalities associate with driver safety. Nevertheless, they either may be too general to represent individual behaviors or only capture static information thus fail to reflect the dynamic nature of a driving task. Finally, driving style was introduced. When combined with ND data, this concept shows great potential to assess driver behavior from a dynamic perspective and in a more specific way.

That said, previous studies do have have some limitations. First, these studies have primarily focus on aggressive and risky driving, sometimes inattentive driving, but not all driving styles at the same time. This could be due to the difficulty to detect all driving styles at the same time, e.g., when there are too many thresholds for each vehicle kinematic.

Secondly, previous studies usually rely on the detection of risky behaviors to characterize the different driving styles. The definition for a risky behavior involves a threshold determined by either expert knowledge (Aljaafreh et al., 2012) or previous studies (Castignani et al., 2013), thus could vary across studies. Also, discretize a continuous measure (e.g., acceleration) into a categorical measure leads to information loss. For example, if a hard brake is considered when the longitudinal deceleration is smaller than -0.4 g (g-force), then -0.41 g relates to a hard brake but -0.39 g does not. Meanwhile, the actual difference in the magnitude of the two deceleration values is tiny.

Thirdly, previous studies tend to ignore the impact of driving environment. Driving is an interactive process between the driver and the driving environment. An individual driver's driving style – the habitual way of driving – will depend on the driving environment. For example, Bassani et al. (2016) showed that average speed are higher in free-flow condition during night-time than day-time. Furthermore, how safe or how risky a behavior is could also depend on the driving environment. The role of driving environment has gradually been realized in behavior assessment studies in recent years. Castignani et al. (2015) and Hu et al. (2018) used contextual information (weather and time of day; posted speed limit and average speed) to define safety-critical events. Zhu et al. (2017) considered link type, posted speed limit and estimated link average speed when assessing driver behavior using ND data. Bejani and Ghatee (2018) detected and classified driving maneuver as dangerous or normal by considering congestion level. These studies confirm the importance to account for environmental factors in driver behavior assessment. Nevertheless, the current studies appear to choose the set of environmental factors either based on prior knowledge or data availability. A more systematic way to determine the relevant environmental factor would be desired.

2.4 Study Objectives and Specific Aims

The overall goal of the dissertation is to assess driver behavior in the context of driving environment. The assessment will be based on a so-called "abnormal" driving style. An "abnormal" driving style is defined as a complement to the "normal" driving style to account for all potential influences of adverse driving styles reflected on vehicle kinematics.

A "normal" driving style is the average driving performance of all drivers under a specific environmental condition. Since "normal" does not by definition infer "safe", "abnormal" is not necessarily "unsafe". Rather, the dissertation examines the relationship between driver safety and driving "abnormality". It is also of interest to reveal any potential moderating effect of the driving environment on this relationship. Three specific aims and research questions are proposed to achieve the goal of this dissertation.

Research Question 1: What are the most important environmental factors that affect how people "normally" drive?

Aim 1: The first research question aims to determine environmental factors that contribute the most to the formation of "normal" driving styles. This is an essential step to define "abnormal" driving styles. Although various environmental factors have been observed affecting driver behavior in previous studies, the magnitude of impact could be inconsistent and the relative importance among environmental factors is yet clear. Thus, rather than assume *a priori* a set of most influential environmental factors, a clustering and classification analysis was used to identify the environmental factor from the data in a systematic way.

Research Question 2: What's the relationship between driver safety and the level of driving "abnormality"? Do less safe drivers behave more "abnormally"? Will the relationship depend on driving environment?

Aim 2: Once the environmental factors were determined, environmental groups can be set up so as "normal" driving styles. An "abnormal" driving style was then defined and calculated for each traversal under a specific driving environmental condition. The second research question thus aims to examine the relationship between driver safety and the level of driving "abnormality". In this study, driver safety was measured by the crash or near-crash status during the naturalistic driving data collection time. The potential moderating effect of driving environment on this relationship was also examined.

Research question 3: What is the relationship between driver safety and the different aspects of an "abnormal" driving style? Do less safe drivers tend to engage in specific "ab-

normal" behaviors? Will the relationship depend on driving environment?

Aim 3: This research question aims to examine the relationship between driver safety and the different aspects of an "abnormal" driving style. Aim 3 extends the focus of Aim 2 in the hope to identify certain "abnormal" patterns that may be used to indicate certain type of "abnormal" behaviors. It is also of interest if the relationship could be moderated by the driving environment.

Chapter 3
DATA

This chapter introduces the naturalistic driving data used in the study. Certain steps were taken to ensure data used for any forthcoming analyses include all necessary information, were consistently recorded, and free of extreme or unreliable values that could be due to, e.g., measurement errors.

3.1 Data Source

3.1.1 SHRP2 NDS

This dissertation used data from the second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS). The SHRP2 NDS was the largest and most comprehensive study of its kind ever undertaken. Data collected in the SHRP2 NDS were from more than 3,000 drivers over a 3-year time period from 2010 to 2013. Six cities around the United States were involved in data collection, i.e., Tampa, Florida; Bloomington, Indiana; Buffalo, New York; Durham, North Carolina; State College, Pennsylvania, and Seattle, Washington. The SHRP2 NDS produces rich data aims to help researchers investigate and understand the role of driving performance and driver behavior in traffic safety, especially their relationship to the risk of crash or near-crash (Victor et al., 2015; Hankey et al., 2016).

We requested a specific part of the SHRP2 NDS that were previously extracted by the Honda R&D Co. and later became publicly available through the VTTI Dataverse (Sears et al., 2019). This dataset encompassed all crashes, near-crashes, and baselines available at the time of extraction on the InSight website (Transportation Research Board of the National Academies of Science, 2013). This include time-series data on all types of vehicle kinematics and video-coded variables related to the roadway and environmental conditions. The data was thus considered an appropriate dataset to serve the goal of dissertation, i.e., examining driver behavior in the context of driving environment.

3.1.2 Roadway Information Database (RID)

The RID was created in parallel with the SHRP2 NDS. The RID includes roadway data newly collected in the SHRP2 NDS through mobile data collection project, and other existing roadway data and supplemental traffic operations data provided by the departments of transportation. These data include information on horizontal curvature, grade, lanes, signs and etc. The RID can be linked to the SHRP2 NDS database and to provide additionally high-quality roadway information for all the associated traversals (Smadi et al., 2015).

3.2 Data Preparation

All baselines in the requested dataset were considered for the proposed analysis. A baseline event in SHRP2 is defined as an epoch of data not involved in a crash, a near-crash or any other types of traffic conflicts. Baselines in the SHRP2 study were randomly selected with the goal of having a minimum of one baseline per driver. Each of the baseline is 21-second long. Nevertheless, the roadway and environmental information was coded "at the time of the start of the Precipitating Event" that is "1 second prior to the end of the epoch" for all baselines (Transportation Research Board of the National Academies of Science, 2013). Thus, only the second-half (i.e, the last 10 seconds) of baselines were used to ensure a more consistent roadway and environmental information with the actual driving condition.

Four vehicle kinematics were extracted and used: longitudinal speed (km/h), longitudinal acceleration (g), lateral acceleration (g) and yaw rate (deg/sec). The first two kinematics measure drivers' driving performance on the longitudinal direction and the latter two measure their driving performance on the lateral direction. All kinematics have been successfully used to examine driving performance and/or access driver behavior in previous studies (Ali et al., 2020; Eren et al., 2012; Simons-Morton et al., 2013). The two accelerations and yaw rate were all recorded at 10 Hz and each represented by a unique kinematic variable. Nevertheless, there are two variables available for longitudinal speed in the time-series data. The two variables have different sampling rates and recorded the longitudinal speed sometimes with considerable discrepancy. The next section describes the procedure used to recode the longitudinal variable before it was used in any forthcoming analysis. This procedure produced a new variable for longitudinal speed at a constant sampling rate of 1 Hz.

3.2.1 Recode Speed

The two variables of longitudinal speed are *speed_gps* and *speed_network*. By definition, *speed_gps* refers to the speed from the GPS and *speed_network* records the speed indicated on speedometer from the vehicle network. As regards the sampling rate, *speed_gps* is said to be at 1 Hz and *speed_network* at 10 Hz. While 80% of the *speed_gps* data are at 1 Hz, less than 50% of *speed_network* were recorded at 10 Hz, with a considerable amount of them recoded at 1 Hz or 2 Hz. On the other hand, when *speed_gps* and *speed_network* indicate different speed patterns, one does not appear to always outperform the other. Seeing the situation, it is decided that a new variable would be created as a combination of *speed_gps* and *speed_network* to represent longitudinal speed.

Figure 3.1 summarizes the procedure in a flow chart, which can be divided in two parts:

1. The first part segmented all 10-second data into ten 1-second data segments. Any traversals with neither *speed_gps* nor *speed_network* for all 10 data segments were removed. In summary, 82% of traversals have *speed_gps* for all 10 data segments, 78% have *speed_network* for all 10 data segments, and 96% have either *speed_gps* or *speed_network* for all 10 data segments. This is the "Step 1" in Figure 3.1.
2. The second part was to evaluate data quality of *speed_gps* and *speed_network* by referring to the longitudinal acceleration. Records failed the evaluation were removed. Among the retained data, the speed variable with a higher degree of consistency with the longitudinal acceleration was used. These are Steps 2 to 4 in Figure 3.1.

Evaluation of the data quality was through a comparison of the recorded speed values to the calculated speed values based on longitudinal acceleration.

1. Convert the longitudinal acceleration to 1 Hz by taking the mean value within each 1-sec data segment.
2. Set the initial speed v_{t_1} to the recorded speed (*speed_gps* or *speed_network*) value of the first data segment (or time slot).
3. Calculate speed values v_{t_i} for the next 9 time points ($i = 2, 3, ..., 9$). It is assumed that the vehicle was traveling within each 1-sec time slot (1) with a constant longitudinal acceleration at time t_i and (2) in a straight line. That is,

$$v_{t_i} = v_{t_{i-1}} + a_{t_i} \times 1 \ sec \tag{3.1}$$

4. An average speed difference between the recorded speed curve and the calculated speed curve was generated. Examination of multiple speed curves with varied magnitude of speed differences selected a threshold of 8 km/h. That is, any data records with an average speed difference larger than 8 km/h were removed. A variable with a lower speed difference is considered more consistent with the longitudinal acceleration.

3.2.2 Data Merge and Reduction

Functional class and posted speed limit from the RID were merged with the baseline time-series based on *Link_ID* in *ArcMap*. Location information (i.e., urban or rural) were in addition extracted for IN and NC due to their functional classes coded in a different way from the other four states, i.e., without specifying urban or rural. It might be worth mentioning that, *Link_ID* in SHRP2 is available through a so-called Map Matching table. Not all traversals can be identified in the table due to Personally Identifiable Information (PII) concerns. That is, traversals within a pre-defined area of the trip origin or destination were not included in the Map Matching table and thus failed to be merged with RID information. This part of the data were removed for any forthcoming analysis.

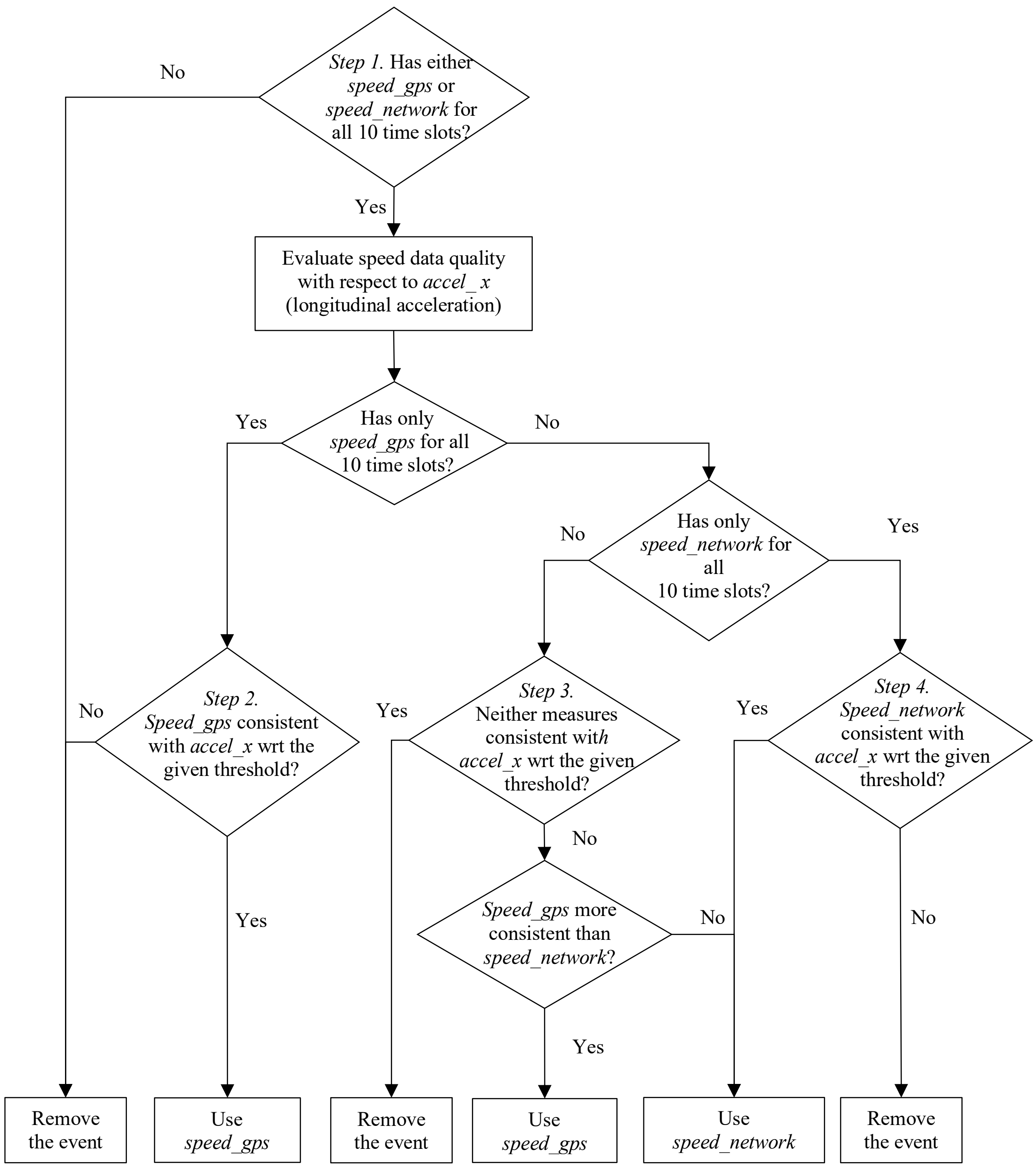

Figure 3.1: Speed recoding procedure

An additional set of data reduction steps have been taken to remove extreme and inconsistently recorded values from the data. More specifically:

- Removed traversals with more than 20% missing data in longitudinal acceleration (∼1% removed), lateral acceleration (∼1% removed) and yaw rate (∼4% removed).
- Removed traversals with no or changed posted speed limit data (∼22% removed).
- Removed traversals with extreme kinematic values. That is, rate of change per second in a 0.1 second time slot is larger than 10 g for longitudinal or lateral acceleration (∼0.06% removed) or 10 deg/sec for yaw rate (∼0.3% removed).
- Removed traversals with inconsistently recorded yaw rate and GPS heading. This was done by calculated and compared (1) the cumulative yaw over the 10-second time period and (2) the difference between the first and the last recorded GSP headings. The comparison was based on cosine values and any data with cosine differences larger than 0.2 were removed (∼1.5% removed).
- Removed potential turning traversals. A turning movement could be problematic since that traversals noted as traveling on a straight road (*alignment* = *"straight"*) may have turned thus had distinct lateral performance compared to those on a straight road but not turned. Initially, an indicator for a turning behavior has been explored based on GPS heading and longitude and latitude information. Nevertheless, since the aforementioned variables are all 1 Hz and with measurement errors, no solid turning behavior indicator can be created. Thus, the difference between the first and the last GPS headings was calculated. Traversals with a cosine of the difference less than 0.9 were removed (∼3.3% removed).
- Finally, removed potential lane-changing traversals. The SHRP2 data has no indicator for a lane-changing behavior. Rather, a vehicle crosses the left marker will have *left_line_right_distance* > 0 and a vehicle crosses the right marker will have *right_line_left_distance* < 0. Values of the two variables return to their normal range when either the vehicle has fully moved to the adjacent lane or has returned back

to the original lane. The former may be considered as a lane-changing behavior and the latter as a lane-departure behavior. Nevertheless, the two behaviors cannot be easily separated solely based on lane position. Therefore, all traversals with either *left_line_right_distance* > 0 or *right_line_left_distance* < 0 were removed ($\sim$28% removed). It is recognized the criteria might somewhat conservative as it also removed any lane-departure traversals.

Chapter 4

CRITICAL ENVIRONMENTAL FACTORS OF "NORMAL" DRIVING STYLES

4.1 Objective

The goal of the dissertation is to assess driver behavior in the context of driving environment based on "abnormal" driving styles, where an "abnormal" driving style is defined as a complement to the "normal" driving style. This chapter thus serves as the first step to identify environmental factors that contribute the most to the formation of "normal" driving styles based on the SHRP2 NDS data. That is:

- *Aim 1: What are the most important environmental factors that affect how people "normally" drive?*

Recall that, in this study, the "normal" driving style is the average driving performance of all drivers under a certain environmental context. That is, the "normal" driving style will depend on environmental factors but not personal characteristics. Many environmental factors have been shown to have an impact on drivers' driving performance. For example, free-flow operational speed was observed the lowest on a local road and the highest on an arterial (Fitzpatrick, 2003); reduction of free-flow speed increases with rain or snow intensity but is on average larger for snow than rain (Rakha et al., 2008). Nevertheless, it is not practical to define a "normal" driving style for every possible environmental context. Furthermore, the different magnitude of impacts suggests the varied importance of environmental factors on "normal" driving styles. The two more specific research questions in Aim 1 thus are:

1. Which environmental factors are the most important to the formation of "normal" driving styles?

2. How do the environment factors form the different "normal" driving styles?

4.2 Methods

A clustering and classification analysis was conducted to address the questions in Aim 1. The analysis includes two steps: (1) the clustering step finds the different types of "normal" driving styles given the data and (2) the classification step identifies the most important environmental factors for all found "normal" driving styles. The "most important" environmental factors in the classification step are those that predict the clusters better than random guess based on the cluster proportion. Details will be introduced later in the section.

A Gaussian Mixture Model (GMM) clustering method was used to cluster time-series kinematic data into homogeneous groups. Since as much as four kinematics were considered, the GMM clustering would be based on the component scores from a Multivariate Functional Principal Component Analysis (MFPCA). The classification step was conducted through a decision tree method. The rest of this section will introduce the analytical methods used in the clustering and classification analysis in detail. Model selection method to determine the "most important" environmental factors will also be introduced.

4.2.1 Clustering Step

Functional Data Analysis

As introduced in Chapter 3, the SHRP 2 naturalistic driving data are time-series data collected continuously over time. A classical way to model time-series data is by an Autoregressive (Integrated) Moving Average model, which predicts a future observation as a linear function of its past observations and past error terms. It is frequently used for forecast and to assess the impact of an intervention. Another way to analyze time-series data is through a Functional Data Analysis (FDA) method, which converts discrete time-series to functions. FDA considers the individual datum, rather than its values at any specific points, as a whole function, and doesn't impose any particular assumptions about the independence of different values within a functional datum (Ramsay and Silverman, 2002). It aims to study the sources of important patterns and to reveal the variation among the data (Ramsay and Silverman, 2005). This exactly fits the goal of the Aim 1 study to characterize and to understand the major trend of driving patterns under different environmental contexts.

The FDA method converts discrete time-series data into functions. For a time-series record x_i, a discretely measured value x_{it} at time t can be represented as a realization of a function with an error term

$$x_{it} = \sum_{k=1}^{K} \theta_{ik} b_k(t) + \epsilon_{it} \tag{4.1}$$

where $b_k, k = 1, ...K$ are the basis functions and $\theta_{ik}, k = 1, ..., K$ are the coefficients of the expansion (Ramsay and Silverman, 2005; Ramsay et al., 2009). Two most popular basis systems are the Fourier basis system and the b-spline basis system, where the former is the usual choice for periodic functions and the latter for nonperiodic functions. Because driving patterns are nonperiodic especially within a short period of time, b-splines were used in this analysis.

Four kinematic measures were included in the analysis, which were (1) delta speed (km/h; 1 Hz), (2) longitudinal acceleration (g; 10 Hz), (3) lateral acceleration (g; 10 Hz), and (4) yaw rate (deg/sec; 10 Hz). Delta speed is defined and calculated as the difference between the longitudinal operational speed and the posted speed limit

$$Delta\ Speed = Operational\ Speed - Posted\ Speed\ Limit \tag{4.2}$$

Since the speed data was collected at a different frequency from the other three kinematics (speed: 1 Hz; others: 10 Hz), different numbers of b-splines were used. More specifically, 8 cubic b-splines were used for delta speed and 13 cubic b-splines were used for the other three kinematics. This is equivalent to 4 interior knot equally splitting every speed record into five 2-second subintervals or 9 interior knot equally splitting every other kinematic record into ten 1-second subintervals.

Multivariate Functional Principal Component Analysis

The method introduced by Happ and Greven was used to conduct the MFPCA in this analysis (Happ and Greven, 2018; Happ-Kurz, 2020). Suppose the data $x_1, ..., x_n$ is from a random process $X = (X^{(1)}, ..., X^{(j)}, ..., X^{(p)})$ with p elements on domains $\mathcal{T}_j \in \mathbb{R}$, the m^{th} multivariate principal component function on the j^{th} element $\hat{\psi}_m^{(j)}$ and the corresponding score for the i^{th} sample $\hat{\rho}_{i,m}$ can be estimated as:

$$
\begin{aligned}
\hat{\psi}_m^{(j)} &= \sum_{n=1}^{M_j} [\hat{c}_m]_n^{(j)} \hat{\phi}_n^{(j)} \\
\hat{\rho}_{i,m} &= \sum_{j=1}^{p} \sum_{n=1}^{M_j} [\hat{c}_m]_n^{(j)} \hat{\xi}_{i,n}^{(j)}
\end{aligned}
\tag{4.3}
$$

where $\hat{\phi}_n^{(j)}$ is the n^{th} univariate principal component of the j^{th} element and $\hat{\xi}_{i,n}^{(j)}$ is the corresponding component score for the i^{th} sample; $\hat{c}_m$ is the eigenvector of the joint covariance matrix $\hat{\boldsymbol{Z}} = \frac{1}{N-1}\boldsymbol{\Xi}^T\boldsymbol{\Xi}$, with $\boldsymbol{\Xi}_{i,\cdot} = (\hat{\xi}_{i,1}^{(1)}, ..., \hat{\xi}_{i,M_1}^{(1)}, ..., \hat{\xi}_{i,1}^{(p)}, ..., \hat{\xi}_{i,M_p}^{(p)})$. The univariate principal component is estimated using the PACE (Principal components Analysis through Conditional Expectation) approach (Yao et al., 2005). According to Equation 4.3, the multivariate principal component for each of the p elements is a weighted sum of their univariate principal components, where the weights are based on the data covariance of all p elements. The multivariate principal components thus take into consideration the correlation among the p elements, where in this study are the four kinematics.

The R packages `funData` (version 1.3-5) and `MFPCA` (version 1.3-6) were used to conduct the analysis. Data were demeaned and an element-wise weight was used to rescale the data so that the amount of variation for the four kinematics are on a similar scale. More specifically, the weight was calculated for each element as an inverse of the integrated pointwise variance $w_j = \left(\int_{\mathcal{T}_j} V\hat{A}R(X^{(j)}(t))dt\right)^{-1}$.

Gaussian Mixture Model Clustering

In GMM clustering, the data $\boldsymbol{X}$ is assumed to be generated by a mixture of multivariate normal distributions with density

$$f(\boldsymbol{X}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \prod_{i=1}^{n} \sum_{k=1}^{G} \tau_k \phi_k(\boldsymbol{x}_i|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \tag{4.4}$$

where $\boldsymbol{x}_i$ is the score vector of a driver i, $i = 1, ..., n$; $(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ are the means and covariance matrix of the multivariate normal distribution associated with cluster k, $k = 1, ..., G$; τ_k is the probability of belonging to the k^{th} cluster. Parameters in the model are estimated by maximum likelihood using the Expectation-Maximization (EM) algorithm (Fraley and Raftery, 2007). GMM clustering was used in preference to other clustering method, such as K-means clustering and hierarchical clustering, due to that it allows clusters to have different sizes and shapes. In addition, it does not require the number of clusters to be determined *a priori*.

The R package `mclust` (version 5.4.7) was used to conduct the GMM clustering. Bayesian Information Criterion (BIC) was used to choose the best parameter setting, which include:

1. Number of clusters, ranging from 1 to 15.
2. Structure of the covariance matrix $\boldsymbol{\Sigma}_k$. The package allows structure of $\boldsymbol{\Sigma}_k$ to be specified with varied volume (equal or variable cluster size), shape (equal or variable cluster shape) and orientation (equal or variable orientation of axes).
3. Data transformation method. The package initializes the EM algorithm using partitions from a model-based hierarchical agglomerative clustering. The initialization step can be applied on data transformed using a number of pre-defined approaches.

For more details regarding the different covariance matrix structures and data transformation methods, see Scrucca and Raftery (2015) and Scrucca et al. (2016).

4.2.2 Classification Step

Once a set of homogeneous groups were extracted from the clustering method, the environmental factors associated with identified clusters were determined using a classification method. This was achieved by applying a decision tree. A decision tree recursively binary splitting the data to grow a large tree and stops when some criteria are reached (Friedman et al., 2009). Compare to Random Forests and Boosted Trees, decision trees provide a single structure of data splits, thus is very interpretable. This is an important advantage to answer the second specific research question "How do environment factors form the 'normal' driving styles?". That is, it allows us to link environmental factors as well as factor levels directly to identified clusters.

The R package `rpart` (version 4.1-15) was used to build the decision tree. The decision tree classifies records into the identified clusters based on the environmental variables listed in Table 4.1. Details regarding the environmental variables and their corresponding categories can be found in "Event Table Detail Information" on the SHRP2 Insight website (Transportation Research Board of the National Academies of Science, 2013). Records associated with categories in any of the environmental variables with no more than 5 observations were removed. The decision tree was applied on a training set of 70% of records, and the remaining 30% of data were retrieved for testing. Data were randomly split into training and testing sets with a similar proportion of records in each cluster. The full decision tree was built based on Gini index and pruned to obtain a subtree to avoid overfitting the data. The complexity parameter `cp` was used to select the best subtree according to

$$R_{cp}(T) = R(T) + cp \times |T| \times R(T_1) \tag{4.5}$$

where $R(T)$ is the risk of the tree T based on zero-one loss, $|T|$ is the number of splits for a tree, and $R(T_1)$ is the tree with no splits. A 10-fold cross validation is used to determine the best value of `cp`. The simplest model within one standard error of the minimum risk was selected (Terry M. Therneau, 2019).

Table 4.1: Environmental variable description

Variable	Description
Lighting	1. Daylight, 2. Dawn, 3. Dusk, 4. Darkness, lighted, 5. Darkness, not lighted
Weather	1. No adverse conditions, 2. Fog, 3. Mist/light rain, 4. Raining
Surface condition	1. Dry, 2. Wet
Traffic flow	1. One-way traffic, 2. Not divided - simple 2-way trafficway, 3. Not divided - center 2-way left turn lane, 4. Divided (median strip or barrier)
Traffic density	1. Level-Of-Service (LOS) A1, 2. LOS A2, 3. LOS B, 4. LOS C, 5. LOS D, 6. LOS E
Relation to junction	1. Non-junction, 2. Driveway, alley access, etc., 3. Entrance/Exit ramp, 4. Interchange area, 5. Intersection, 6. Intersection-related, 7. Parking lot entrance/exit
Alignment	1. Straight, 2. Curve left, 3. Curve right
Grade	1. Level, 2. Grade down, 3. Grade up
Locality	1. Business/Industrial, 2. Bypass/Divided highway with traffic signals, 3. Church, 4. Interstate/Bypass/Divided highway with no traffic signals, 5. Moderate residential, 6. Open country, 7. Open residential, 8. Playground, 9. School, 10. Urban
Construction zone	1. Not construction zone-related, 2. Construction zone, 3. Construction zone-related
Intersection influence	1. No, 2. Yes, interchange, 3. Yes, traffic signal, 4. Yes, uncontrolled

4.2.3 Model Selection Method

The clustering analysis separates data into multiple groups based on MFPCA scores and the classification analysis identifies environmental factors that best explain the data separation. Model selection for the clustering and classification analysis thus includes two questions:

1. How many multivariate principal components should be chosen?
2. How to determine whether an environmental factor should be retained or not?

Principal components selection is not a new problem. For example, the author of the `MFPCA` package suggests finding the appropriate number of principal components based on the decrease of estimated eigenvalues using a scree plot or to set a threshold for the percentage of variance explained (Happ and Greven, 2018), both of which are strategies frequently used in a PCA. The ultimate goal of principal component selection is to determine the number of non-trivial components that summarize meaningful data variation (Peres-Neto et al., 2005). In the context of Aim 1, "meaningful" data variation is considered as data variation that can be explained by some given environmental factors. Thus, the number of principal components to select in the clustering step will need to be determined together with the results from the classification step.

There are two concerns for model selection in the classification step when different numbers of multivariate principal components are used. Firstly, GMM clustering using different number of principal components will very likely generate different number of clusters. Even when the number of clusters is the same, data assignment can be different. In this sense, the classification step are based on different data. To compare and select models based on the *overall accuracy rate* thus is not reasonable. Another concern relates to the *overall accuracy rate* is that, even in an ideal situation, the classification model will not be able to classify data into identified clusters from the clustering step with 100% accuracy. This is because that variation in kinematic data that separates driving records into homogeneous groups might not only be due to environmental factors but also individual factors. For example, a cluster could involve extreme driving patterns that were uniquely defined by aggressive driving tendency. A classification model based on environmental factors will not be able to capture this cluster, which will reduce the overall prediction accuracy. This situation is more likely to happen when higher number of principal components is used.

Due to the two concerns, the *overall accuracy rate* will not be used to compare models. Rather, it will be used together with the *No Information Rate (NIR)* as an initial criterion to ensure an informative model to start with.

Criterion 1. *Overall accuracy rate* of the model is significantly higher than the *NIR*.

The *NIR* equals the largest proportion of the observed classes (or clusters) in the training data. It sets up a lower bound of the *overall accuracy rate* for an informative model. That is, a given model should classify data significantly better than randomly allocate them to the majority class for its associated environmental factors to be considered "informative" in explaining the data variation

Once we confirm a classification model is informative, *precision* and *recall* will be used to assess individual environmental factors. That is, for all informative models consisted of different environmental factors, which factors should be retained and which ones should not? Suppose an environmental factor i is used to classify data into cluster j, the environmental factor will need to satisfy the following criteria to be retained:

Criterion 2. Either the associated *precision* or *recall* for cluster j is high

Criterion 3. Neither the associated *precision* nor *recall* for cluster j is very low

According to Table 4.2, *precision* equals $A/(A+B)$ and *recall* equals $A/(A+C)$. *Precision* of a cluster j measures the proportion of data correctly classified by the given environmental factor i and *recall* measures the proportion of data in cluster j that can be classified by environmental factor i. Let's consider three possibilities of a true relationship between the identified environmental factor i and its associated cluster j:

(a) Cluster j is uniquely defined by environmental factor i

(b) Cluster j is uniquely defined by multiple environmental factors, but only factor i is identified

(c) Environmental factor i is involved in the definition of multiple clusters but only identified for cluster j

For possibility (a), *precision* and *recall* of cluster j. For possibility (b), *precision* might still be high but *recall* will drop down due to increased C in Table 4.2. For possibility (c), *recall* might still be high but *precision* will drop down due to increased B. This motivates model selection criterion #2 that if an environmental factor i is truly involved in separating the data, its associated cluster should have either a high *precision* or a high *recall* (e.g., larger than the *NIR*). Model selection criterion #3 is to avoid the situation that the identified environmental factor contributes too little to explain the variation of a cluster.

Table 4.2: Number of observed and predicted observations in a 2 by 2 table

		Observed	
		Yes	No
Predicted	Yes	A	B
	No	C	D

Finally, if multiple models satisfy the first three criteria and that the identified environmental factors are the same, the most appropriate model will be determined according the the forth criterion.

Criterion 4. Select the model with the largest positive difference between the *overall accuracy rate* and *NIR*

The whole process of selecting the environmental factors that best explain the data separation in Aim is summarized in Algorithm 1.

Algorithm 1: Model selection in the clustering and classification analysis

Result: A selected model

1 Extract the first 10 multivariate principal components based on kinematic data;
2 Determine based on the scree plot the number of n principal components to consider;
3 Run the GMM clustering and decision tree on 1 to n multivariate principal components; record for each model the identified environmental factors and the corresponding *overall accuracy rate, NIR, precision* and *recall* ;
4 Select the best model based on Criteria 1 to 4 ;

4.3 Results

The clustering and classification analysis was each conducted for the three functional class groups with the most data, i.e., Urban Interstate, Urban Other Principal Arterial and Urban Minor Arterial. This section will report the clustering and classification results for each of the functional class type.

4.3.1 Urban Interstate

The MFPCA was applied on $n = 1,542$ 10-second traversals on Urban Interstate. The first 10 multivariate principal components were estimated based on the four kinematics and a scree plot of the 10 PCs is showed in Figure 4.1. The first two PCs explained the most proportion of variance in the data, equaling 30% and 28% respectively. Proportions of variance explained gradually decreased since the third PC and become almost negligible after the sixth one. Thus, the first six principal components were used in the forthcoming clustering and classification analysis to identify homogeneous driving groups and corresponding environmental factors on Urban Interstate.

Table 4.3 shows the results of the clustering and classification analysis using 1 to 6 multivariate principal components. As mentioned earlier, model prediction accuracy is assessed through *overall accuracy rate* and *NIR* and identified environmental factors will be evaluated through *precision* and *recall*.

- Model selection criterion #1 filters out the model based on 6 PCs, whose *overall accu-*

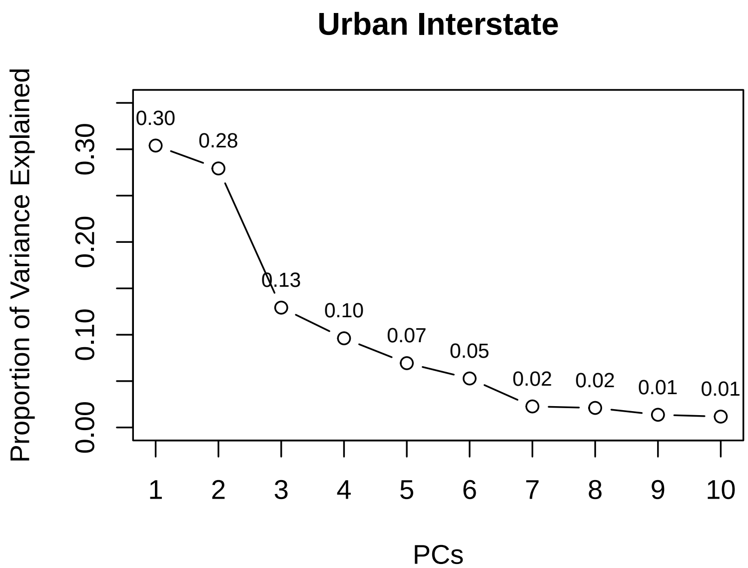

Figure 4.1: Scree plot of the MFPCA for traversals of Urban Interstate

racy rate is not significantly higher than *NIR* at 0.05 confidence level.

- Among the other five models, the model based on 1 PC identifies one environmental factor *traffic density* and all other models identifies an additional environmental factor *alignment*.
- For all models identifies the second environmental factor *alignment*, the model based on 5 PCs has the largest positive difference between *overall accuracy rate* and *NIR* (i.e., ΔAccuracy). In this model, the additional factor *alignment* identifies a new cluster not captured by *traffic density* alone with *precision* = 0.83 and *recall* = 0.42. The *precision* and *recall* are considered satisfy criteria #2 and #3 so that *alignment* is retained.

The model based on 5 PCs was selected for Urban Interstate based on criteria #1 to #4. Table 4.4 summarizes the number of traversals by predicted and observed cluster indexes, along with the *precision* and *recall* for each cluster. The pruned decision tree that specifies the identified environmental factors and their associated clusters of this model is shown in Figure 4.2.

Table 4.3: Performance of the classification model by number of principal components for Urban Interstate

# of PCs	# of clusters	Overall accuracy	NIR	ΔAccuracy	Selected environmental factors
1	2	0.95	0.91	0.04	Traffic density
2	3	0.78	0.64	0.14	Alignment, traffic density
3	3	0.74	0.61	0.13	Alignment, traffic density
4	5	0.68	0.53	0.15	Traffic density, alignment
5	**4**	**0.66**	**0.46**	**0.20**	**Traffic density, alignment**
6	*8*	*0.31*	*0.27*	*0.04*	*Traffic density, alignment, lighting, grade, relation to junction, traffic control, weather*

Note: selected model in bold and non-informative model in italic

Table 4.4: Precision and recall for the model with 5 PCs of Urban Interstate

Prediction	Observation				Precision	Recall
	Cluster 1	Cluster 2	Cluster 3	Cluster 4		
Cluster 1	22	0	0	0	1.00	0.44
Cluster 2	26	208	103	13	0.59	0.98
Cluster 3	2	5	74	8	0.83	0.42
Cluster 4	0	0	0	0	/	0
Total	50	213	177	21		

Let's take a further look at the identified clusters and environmental factors for the selected model. Figure 4.3 depicts the first five multivariate PCs used by this model. The solid black line denotes the mean trend of a specific kinematic and is the same across all panels in the same column. The "+" and "−" signs demonstrate the effect of adding or subtracting a multiple principle component. It can be observed that, the first multivariate PC mostly accounts for variation in delta speed. Traversals with higher scores on the first

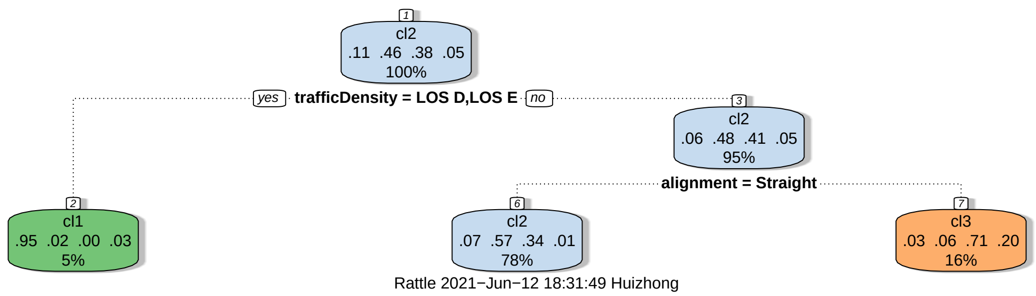

Figure 4.2: Pruned decision tree based on 5 multivariate PCs for Urban Interstate

multivariate PC drove slower than the posted speed limit. The second multivariate PC is associated with large variation on lateral acceleration and yaw rate. Higher scores on the second PC thus denotes steering right. The third multivariate PC represents varied behavior on the longitudinal acceleration metric. Higher scores on the third multivariate PC suggests an accelerating behavior. The fourth and fifth multivariate PCs both account for variations in the lateral direction. Compare to the second PC that steering in one direction, the forth and fifth PCs account for a change in the steering direction.

As noted in Table 4.3, the model based on 5 PCs generates 4 clusters. Figure 4.4 shows the PC scores by cluster. Cluster #1 has the highest score on the first PC and a slightly higher variation on the third PC. It does not show much variation on the second, fourth and fifth PCs that mostly accounts for variation on the lateral position. This cluster appears to be a group of traversals with lower speed and frequent accelerating or decelerating behavior. This observation is consistent with the environmental factor that classifies this cluster: *traffic density* in "LOS D" or "LOS E". Cluster #2 and cluster #3 are the two largest clusters. The two clusters are similar on the first and third PCs of variations on the longitudinal direction and differ on the second and forth PCs on the lateral direction. More specifically, traversals in cluster #3 have more lateral-direction variations, which is consistent with the environmental factor that separates it from cluster #1, i.e., *alignment* in "curve left" or "curve right". Cluster #4 is the smallest cluster with universally higher variation on all used principal

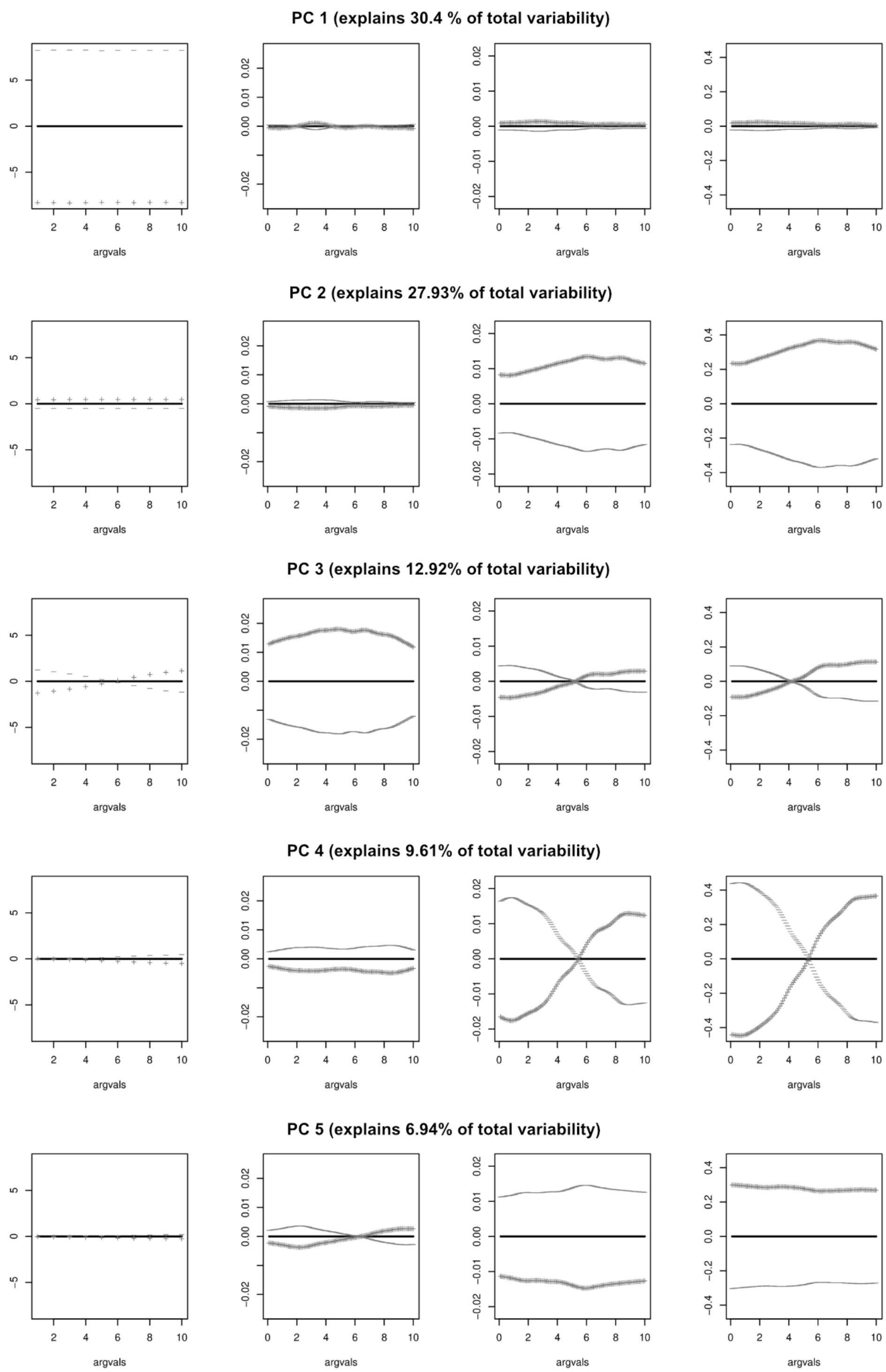

Figure 4.3: The first five principal components of (from left to right) delta speed, longitudinal acceleration, lateral acceleration, and yaw rate of traversals on Urban Interstate.

components. This cluster is not identified by any of the identified environmental factors, indicating that it could be due to individual factors or other unidentified environmental factors.

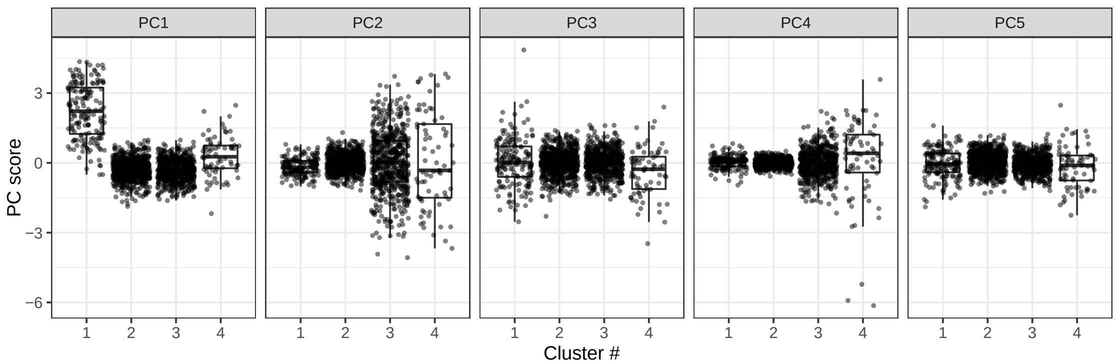

Figure 4.4: Multivariate PC scores by cluster based on 5 PCs for Urban Interstate

4.3.2 Urban Other Principal Arterial

A total of 1,125 traversals on Urban Other Principal Arterial were used in a MFPCA. Figure 4.5 shows the proportion of variance explained for the first 10 principal components. The first multivariate principal component explains 34% of the data variance. The proportion variance explained gradually decreased since the second principal component and becomes negligible after the seventh component. Thus, the forthcoming clustering and analysis were applied on all traversals on UOPA with 1 to 7 multivariate principal components.

Table 4.5 summarizes the clustering and classification results when using 1 to 7 multivariate principal components.

- No model fails the first model selection criterion.
- All models select the environmental factor *intersection influence*.
- The two models based on 4 or 7 PCs identify an additional environmental factor *alignment*. In the model based on 4 PCs (with higher ΔAccuracy than the model based on

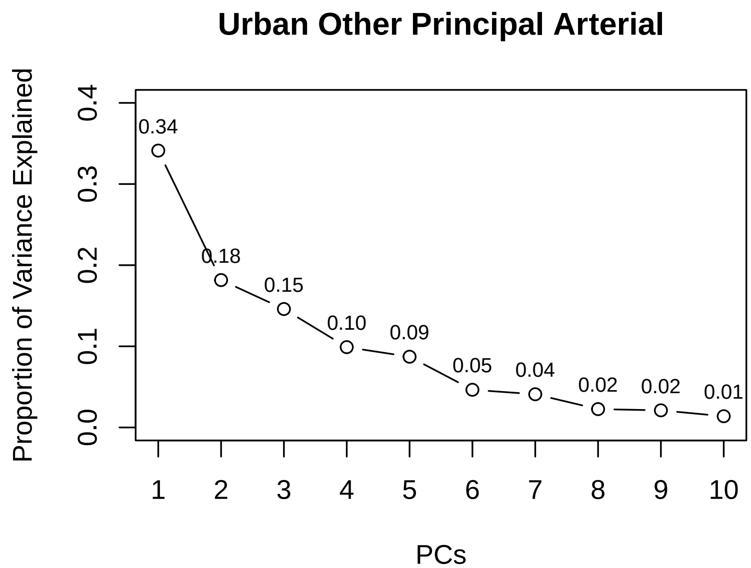

Figure 4.5: Scree plot of the MFPCA for traversals of Urban Other Principal Arterial

7 PCs), *alignment* defines a new cluster not captured by *intersection influence* with *precision* = 0.78 and *recall* = 0.21. The *precision* and *recall* are considered satisfy criteria #2 and #3 and *alignment* was retained.

- The model based on 5 PCs identifies a third environmental factor *traffic density*. This factor reassign traversals with higher traffic density in one cluster (#2) to another (#1). On the testing set, 18 traversals were reassigned from cluster #2 to #1 by including this factor. Among them, 9 traversals were correctly classified while 9 traversals were incorrectly reassigned. Inclusion of this factor increases the *recall* of cluster #1 by 0.11 meanwhile decreases its *precision* by 0.10. For cluster #2, this factor increases the *precision* by 0.02 and decreases the *recall* by 0.04. In general, inclusion of this factor does not appear to explain data variation better thereby the variable was not retained.
- The model based on 6 PCs identifies a forth factor *locality*. This factor defines a new cluster that is not explained by *traffic density*, *alignment* or *traffic density*. *Precision*

of this cluster is 0.6 and *recall* is 0.04. Because that *precision* is not very high and *recall* is extremely low, this factor was not retained.

Table 4.5: Performance of the classification model by number of principal components for Urban Other Principal Arterial

# of PCs	# of clusters	Overall accuracy	NIR	ΔAccuracy	Selected environmental factors
1	2	0.89	0.76	0.13	Intersection influence
2	4	0.80	0.70	0.10	Intersection influence
3	4	0.72	0.61	0.11	Intersection influence
4	**4**	**0.70**	**0.56**	**0.14**	**Intersection influence, alignment**
5	4	0.67	0.49	0.18	Intersection influence, alignment, traffic density
6	4	0.64	0.48	0.16	Intersection influence, alignment, traffic density, locality
7	7	0.40	0.31	0.09	Intersection influence, alignment

Note: selected model in bold

The final selected model for Urban Other Principal Arterial is based on 4 PCs. Table 4.6 summarizes the observed and predicted observations based on this model. Figure 4.6 demonstrates how the identified environmental factors classify data records into the identified clusters. Similarly to the model selected for Urban Interstate, one cluster (#1) with the least number of traversals failed to be captured.

Table 4.6: Precision and recall for the model with 4 PCs of Urban Other Principal Arterial

Prediction	Observation				Precision	Recall
	Cluster 1	Cluster 2	Cluster 3	Cluster 4		
Cluster 1	0	0	0	0	/	0
Cluster 2	8	185	40	37	0.68	0.97
Cluster 3	0	2	11	1	0.78	0.21
Cluster 4	9	3	2	39	0.74	0.51
Total	17	190	53	77		

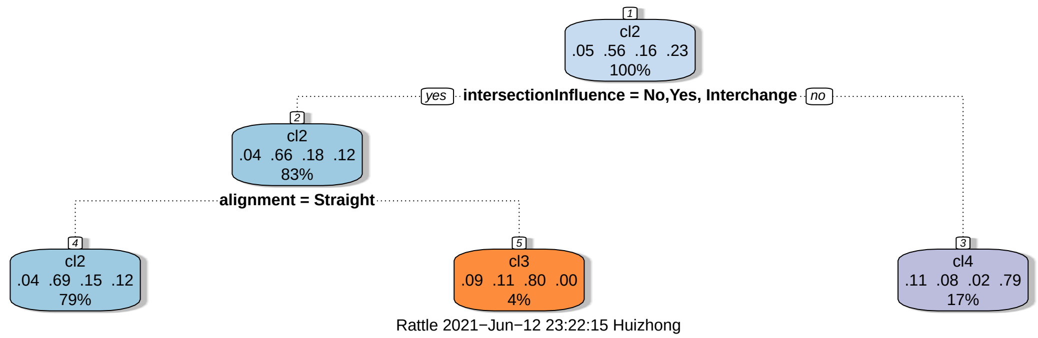

Figure 4.6: Pruned decision tree based on 4 multivariate PCs for Urban Other Principal Arterial

Figure 4.7 depicts the first four multivariate PCs used by the selected model. Again, the solid black line denotes the mean trend and the "+" and "−" signs demonstrate the effect of adding or subtracting a multiple principle component. The first multivariate PC is associated with variations on delta speed and the longitudinal acceleration. That is, traversals with higher scores on the first PC were decelerating while those with lower scores were accelerating. The third multivariate PC captures a similar trend to the first one, nevertheless with a higher delta speed to start with. The second and the forth multivariate PCs both account for variations on the lateral direction. The second one is associated with lateral acceleration and yaw rate in the same direction and the forth one with the two in opposite directions possibly due to, e.g., releasing the steering wheel from a turning position.

Figure 4.8 shows the PC scores by cluster for the selected model of Urban Other Principal Arterial. Traversals in cluster #4 have higher scores on PC1 and a relatively larger variation than clusters #2 and #3 on PC3. As mentioned earlier, both PC1 and PC3 are accounting for variations on the longitudinal direction. This observation is consistent with how the identified environmental factor *intersection influence* classifies this cluster, i.e., traversals were either under the influence of a traffic signal or an uncontrolled intersection ("Yes, traffic signal" or "Yes, uncontrolled"). Traversals in clusters #2 and #3 have similar score distribution on PC1 and PC3, but traversals in cluster #3 are associated with higher scores

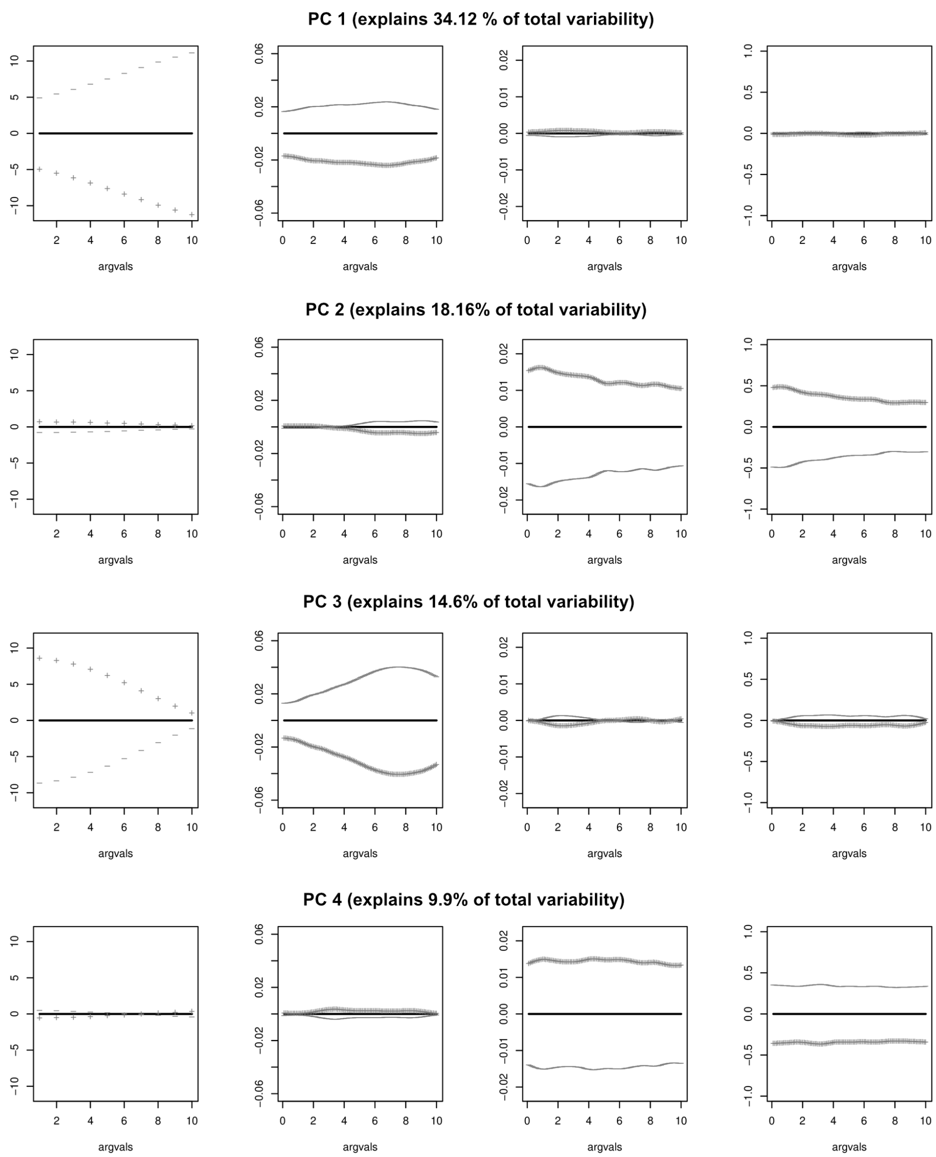

Figure 4.7: The first four principal components of (from left to right) delta speed, longitudinal acceleration, lateral acceleration, and yaw rate of traversals for Urban Other Principal Arterial

on PC4 and larger variations on PC2 than those in cluster #2. This indicates that the two clusters of traversals are similar in the longitudinal direction but differ on the lateral direction. Actually, the two clusters are both under no intersection influence or influence of an interchange area ("No" or "Yes, interchange"). Nevertheless, traversals in cluster #3 are on a curved left or curved right road rather than a straight road. Again, we have a cluster (#1) with higher variation on all the used PCs but failed to be identified by any selected environmental factor or their combinations.

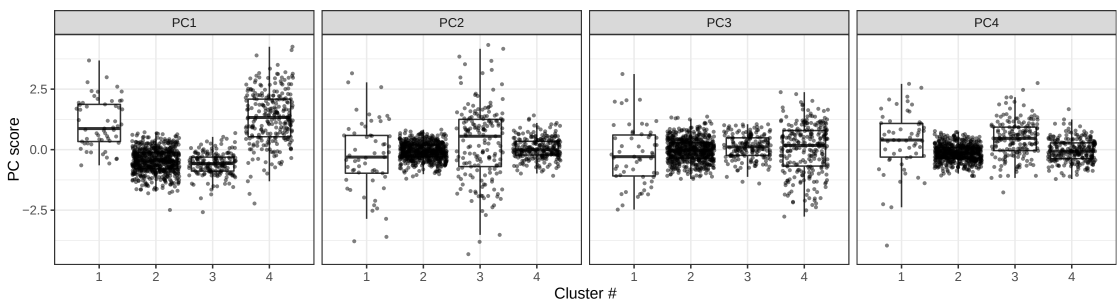

Figure 4.8: Multivariate PC scores by cluster based on 4 PCs for Urban Other Principal Arterial

4.3.3 Urban Minor Arterial

MFPCA was applied to 749 traversals on Urban Minor Arterial. Figure 4.9 shows the proportion of variance explained for the first 10 principal components. Similar to Urban Other Principal Arterial, the first multivariate principal component explains the most data variance (31%). The second to forth principal components each explains approximately 15% data variance. The proportion of explained variance then drop down again and becomes negligible after the seventh component. The forthcoming clustering and analysis were applied to 1 to 7 multivariate principal components for traversals on Urban Minor Arterial.

Clustering and classification results based on 1 to 8 multivariate principal components for UMA are shown in Table 4.7.

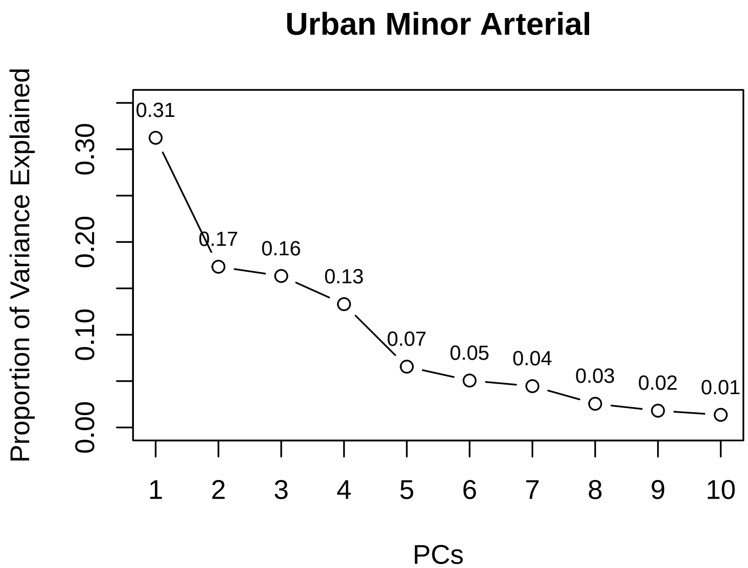

Figure 4.9: Scree plot of the MFPCA for traversals of Urban Minor Arterial

- No model fails Criterion #1.
- All models select environmental factor *intersection influence*.
- Models based on 4, 5 and 6 multivariate PCs identify a second environmental factor *alignment*. In the model based on 6 PCs (with the highest ΔAccuracy among the three models), *alignment* finds a third cluster not captured by solely based on *intersection influence*. This cluster has *precision* $= 0.61$ and *recall* $= 0.27$. The *precision* and *recall* were considered satisfied criteria #2 and #3, thus *alignment* was retained.

The final model selected for Urban Minor Arterial is based on 6 multivariate PCs according to Criteria #1 to #4. Table 4.8 summarizes the distribution of observed and predicted cluster members on a testing set using this model. Figure 4.11 specifies how the two identified environmental factors were used to separate traversals into each cluster. Similar to the situation for the other two functional classes, one cluster (#3) with the least members cannot be captured by the identified environmental factors.

Table 4.7: Performance of the classification model by number of principal components on Urban Minor Arterial

# of PCs	# of clusters	Overall accuracy	NIR	ΔAccuracy	Selected environmental factors
1	2	0.85	0.74	0.11	Intersection influence
2	3	0.64	0.57	0.07	Intersection influence
3	4	0.44	0.36	0.08	Intersection influence
4	5	0.35	0.28	0.07	Intersection influence
5	4	0.63	0.50	0.13	Intersection influence, alignment
6	**4**	**0.67**	**0.49**	**0.18**	**Intersection influence, alignment**
7	7	0.32	0.24	0.08	Intersection influence, alignment

Note: selected model in bold

Table 4.8: Precision and recall for the model with 6 PCs of Urban Minor Arterial

Prediction	Observation				Precision	Recall
	Cluster 1	Cluster 2	Cluster 3	Cluster 4		
Cluster 1	28	0	5	1	0.82	0.62
Cluster 2	0	14	8	1	0.61	0.27
Cluster 3	0	0	0	0	/	0
Cluster 4	17	38	4	106	0.64	0.98
Total	45	52	17	108		

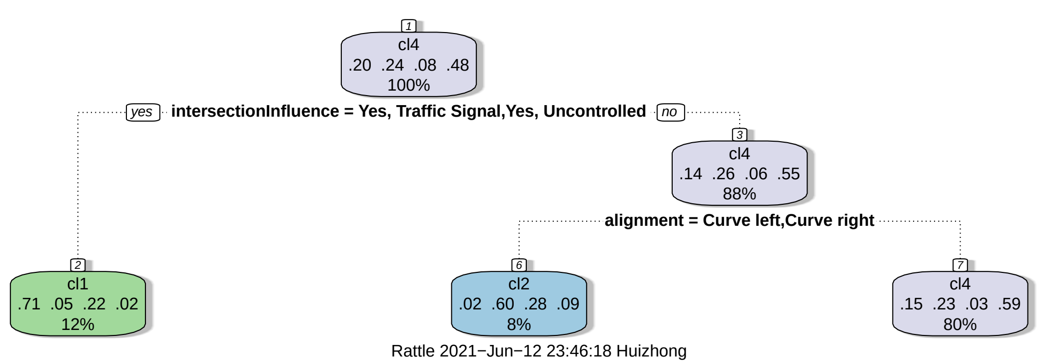

Figure 4.10: Pruned decision tree based on 6 multivariate PCs on Urban Minor Arterial

The first six multivariate PCs used for the selected model are shown in Figure 4.11. It can be observed that the first three multivariate PCs all accounted for variations on the longitudinal direction, with higher scores representing a decelerating behavior and lower scores an accelerating behavior. The three PCs differ on the two kinematics of lateral performance: the first PC accounts for no variation on the lateral direction; the second and third PCs are associated with both variations on the longitudinal and lateral directions but with opposite relationships between the two. The forth multivariate PC shows a change in the lateral direction. The fifth multivariate PC reveals lateral acceleration and yaw rate in opposite directions. The sixth multivariate PC appears to account for fluctuations on both the longitudinal and lateral directions.

PC scores by cluster based on the selected model are shown in Figure 4.12. Cluster #1 has higher scores on the first PC than all other clusters. This is consistent with the environmental factor that classifies it, i.e., *intersection influence* in "Yes, traffic signal" or "Yes, uncontrolled". Recall that "Yes, uncontrolled" refers to under influence of an intersection without traffic control. Clusters #2 and #4 have very similar score distributions on the first and the sixth multivariate PCs. Traversals in cluster #2 were associated with higher variations on the second to fifth PCs than cluster #4, which all explain variations on the lateral direction. As expected, cluster #2 is the cluster consists of traversals traveling on a curved left or curved right road. Once again, the cluster (#3) not captured by any identified environmental factors or their combinations is the one with the least size and relatively higher variation on all used PCs.

4.4 Discussion

The primary reason to find major environmental factors that explain the driving data variation is to set up environmental groups for "normal" driving styles. This has been introduced in the beginning of this chapter. To summarize, the current data identifies multiple "normal" driving styles based on three environmental factors – functional class and two other factors identified given the functional class. More specifically, a "normal" driving style was

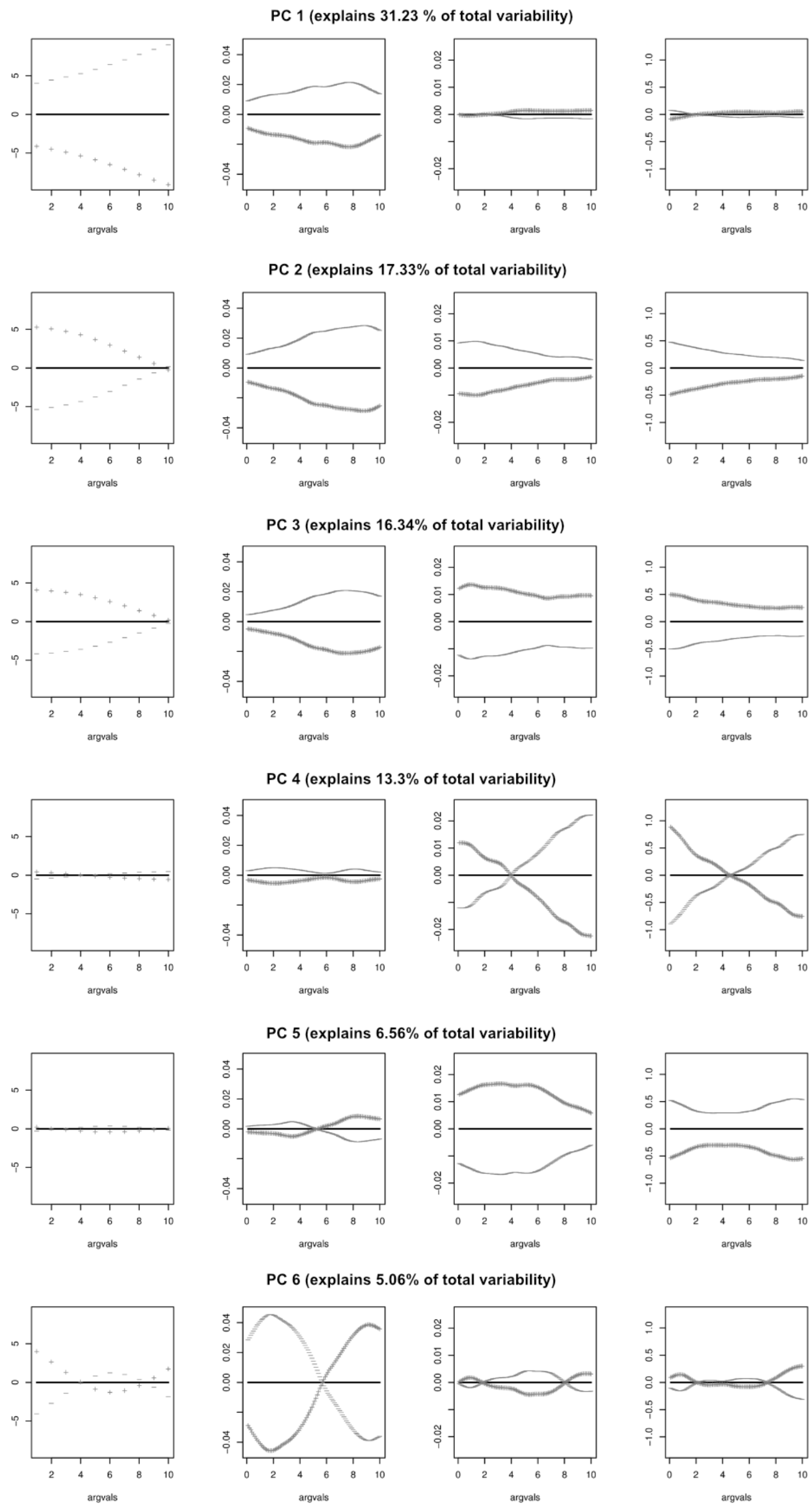

Figure 4.11: The first six principal components of (from left to right) delta speed, longitudinal acceleration, lateral acceleration, and yaw rate of traversals for Urban Minor Arterial

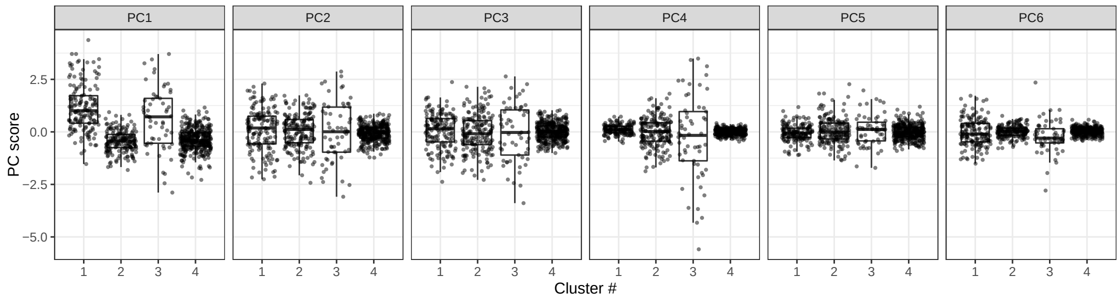

Figure 4.12: Multivariate PC scores by cluster based on 6 PCs on Urban Minor Arterial

mostly influenced by whether the road was "congested" or not on Urban Interstate, where "congested" could be defined as *traffic density* equaling or higher than "LOS D". On Urban Other Principal Arterial or Urban Minor Arterial, the "normal" driving style is mostly influenced by the existence of an intersection (i.e., *intersection influence*) regardless of its control type. *Alignment* is the second important factor for all functional classes when traversals were not on a "congested" road or influenced by an intersection. The result confirms the use of some environmental factors in previous studies when analyzing driver behavior – e.g., "congestion level" in Bejani and Ghatee (2018) – but not others – e.g., "weather" in Hu et al. (2018).

The other benefit of setting up environmental groups is that it helps block potential confounding effects of environmental factors on the relationship between driver safety and driving "abnormality", which is the main focus of the following two chapters. Ideally, drivers' safety level – as a personal characteristic – would be independent of the environmental factor in a specific traversal, e.g., whether the driver is risky or not would not depend on the weather condition or vice versa. Nevertheless, the two might be connected by some common unobserved factors and become correlated. For example, an individual driving mostly during peak-hours might be associated with higher traffic density levels (environmental factor) and higher near-crash rate (safety level). The unobserved factor cannot be controlled for. Rather, by setting up environmental groups (i.e., blocking), all traversals within the same environmental group is assumed to have the same “baseline” driving style in terms of environmental factors – the “normal” driving style. Environmental factors are thus isolated from driving “abnormality”.

Again, practically it is not possible to set up an environmental group for every environmental context. The clustering and classification analysis in Aim 1 identifies the environmental factor that have the most impact on driving performance and these factors were used to set up the environmental group. By doing so, it is assumed that environmental factors not used to set up the environmental groups can be safely ignored. This is considered a plausible assumption at least for the current data given the clustering and classification analysis method. If an additional environmental factor has a considerable impact on the traversals within its environmental group, the clustering and classification analysis should be able to identify this factor. A preliminary analysis was conducted on a path model (will be introduced in the next chapter) with and without environmental factors. Estimated coefficients and standard errors between driver safety and driving “abnormality” were very similar in the two cases, which also support the assumption.

An alternative way to control the potential confounding effect of environmental factors on the relationship between driver safety and driving “abnormality” is to include all environmental factors in an analytical model. This way, the set-up of environmental groups is not needed. This approach is consider inappropriate for two reasons.

The first reason is from a conceptual perspective. If there is no environmental group, then the “normal” driving style will have to be based on all traversals, e.g., by using the grand mean. Mathematically this can be done, but the so-called “normal” driving style will have no practical meaning. This is similar to taking an average of a binary variable, such as gender, but worse than it in the sense that this can not even be considered as a proportion. For example, if we have 50 traversals driving on an Interstate and 50 traversals on a local road, an average of the speed can be calculated but is nonsense – what does it mean to be 50% driving on an Interstate and 50% on a local road? After all, I believe a “normal” driving style will need to be defined under some environmental context to be meaningful.

The second reason is from a practical perspective. The dissertation is about assessing driver behavior in the context of driving environment. This will be done by comparing the relationship between driver safety and driving "abnormality" among environmental groups. If environmental groups are not used, this will need to be achieved by assessing interaction terms in the analytical model. Without knowing which environmental factor is more important, this procedure will require examine all possible two-way interaction terms between the environmental factor and driver safety or even three-way or four-way interaction terms since the environmental groups are determined by two or three environmental factors. The process will be difficult to handle while very likely to generate trivial interactions of less interest or even misleading.

Chapter 5

LEVELS OF DRIVING "ABNORMALITY" AND RELATIONSHIP TO DRIVER SAFETY

5.1 Objective

Once the environmental factors were identified, environmental groups could be set up so as "normal" driving styles. "Abnormal" driving style were defined and calculated accordingly. Nevertheless, it is yet clear how driving "abnormality" relates to driver safety. This chapter will demonstrate the analysis used to examine the research questions in Aim 2.

- *Aim 2: What's the relationship between driver safety and the level of driving "abnormality"? Do less safe drivers behave more "abnormally"? Will the relationship depend on driving environment?*

5.2 Methods

5.2.1 Environmental Groups

Chapter 4 identifies two environmental factors for each of the three functional classes – Irban Interstate, Urban Other Principal Arterial, and Urban Minor Arterial. Twelve environmental groups were established based on the identified environmental factors on three functional classes. Specially, "curve left" and "curve right" of *alignment* were separated when setting up the groups given their distinct impact on lateral behaviors. The twelve groups are shown in Table 5.1. It can be observed that, traversals are not equally distributed in the twelve groups. Rather, three groups (#1, 6 and 10) that are on a straight roadway are associated with the most number of traversals (78%).

Table 5.1: Number of traversals by environmental groups on Urban Interstate, Urban Other Principal Arterial and Urban Minor Arterial

#	Functional class	Traffic density	Alignment	Intersection influence	# of traversals
1	Urban Interstate	A1, A2, B, or C	Straight		1196
2		A1, A2, B, or C	Curve left		121
3		A1, A2, B, or C	Curve right		145
4		D or E	All		80
5	Urban Other Principal Arterial		All	Yes, Traffic Signal; Yes, Uncontrolled	184
6			Straight	No; Yes, Interchange	892
7			Curve left	No; Yes, Interchange	24
8			Curve right	No; Yes, Interchange	25
9	Urban Minor Arterial		All	Yes, Traffic Signal; Yes, Uncontrolled	97
10			Straight	No	586
11			Curve left	No	32
12			Curve right	No	34

5.2.2 Outcome Variables

The outcome of the Aim 2 is the “abnormality” score. This score quantifies the level of “abnormality” associated with a specific “abnormal” driving style within a short a period of time (i.e., 10 seconds). More specifically, it calculates the difference between a specific driving pattern and the “normal” pattern under a certain environmental condition. The “abnormality” will be calculated for all four vehicle kinematics, i.e., delta speed, longitudinal acceleration, lateral acceleration and yaw rate.

For a traversal i of a driver j in an environmental condition l, denotes the driving pattern associated with vehicle kinematic k as x_{ijkl}. The “normal” driving pattern x_{kl} in the environmental condition l is an average of x_{ijkl} over all traversals in that condition. Suppose $\boldsymbol{x}_{ijkl} = \left\{x_{ijkl}(t_1), ..., x_{ijkl}(t_N)\right\}$ and $\boldsymbol{x}_{kl} = \left\{x_{kl}(t_1), ..., x_{kl}(t_N)\right\}$ are two realizations of x_{ijkl} and x_{kl} at time $t_1, ..., t_N$, then the level of “abnormality” a_{ijkl} can be measured by the Euclidean distance as:

$$a_{ijkl}(\text{Euclidean}) = \sqrt{\left(\boldsymbol{x}_{ijkl} - \boldsymbol{x}_{kl}\right)^T \left(\boldsymbol{x}_{ijkl} - \boldsymbol{x}_{kl}\right)} \tag{5.1}$$

The originally proposed Mahalanobis distance is considered inappropriate for the current data. The Mahalanobis distance takes into consideration the correlation structure of the data and rescales the data so that every time point is with equal variance. The method down-weights values at time points with stronger variations and may thus diminish the observed “abnormality” of driving patterns.

To ensure that “abnormality” score of the four kinematics are on the same scale, “abnormality” scores of delta speed was divided by 10 and “abnormality” scores of longitudinal acceleration and lateral acceleration were multiplied by 9.8. For longitudinal and lateral acceleration recorded in unit g (gravity), this is equivalent to convert their units to m/s^2. In addition, all rescaled “abnormality” scores were log-transformed to accommodate the normality assumption of the modeling method.

5.2.3 Explanatory Variables

The explanatory variable of primary interest is level of driver safety. In this study, the safety level is measured by crash rate or near-crash status during the SHRP2 data collection period. In other words, the crashes and near-crashes used in this study are from the SHRP2 NDS records. The crash rate is defined as the number of crashes per 1,000 Vehicle Distance Traveled (VDT; in km) and calculated as

$$\text{Crash rate}_j = \frac{\#\text{ of crashes for driver } j}{\text{Total VDT for driver } j} \times 1,000 \tag{5.2}$$

Similarly, the near-crash rate is defined as the number of near-crashes per 1,000 VDT and calculated as

$$\text{Near-crash rate}_j = \frac{\#\text{ of near-crashes for driver } j}{\text{Total VDT for driver } j} \times 1,000 \tag{5.3}$$

Drivers with no VDT information or driving less than 3,000 km during the SHRP2 NDS period were removed from the analysis. SHRP2 NDS data was collected over a 3-year period and most drivers participated from 1 to 2 years (Victor et al., 2015). Since the exact participating time is not available, 3,000 km (or 1,864 miles; approximately the 10^{th} percentile of VDT) was used as a rough threshold to filter out infrequent drivers, e.g., driving about 1,000 miles per year.

Because that crash or near-crash rates are highly skewed and cannot be easily transformed to satisfy the normality assumption of the modeling method, this variable would enter the model using one of the following form:

1. Crashed or not (binary)
2. Log-transformed crash rates for drivers with crashes (continuous)
3. Near-crashed or not (binary)
4. Log-transformed near-crash rates for drivers with near-crashes (continuous)

The strategy thus analyzed crash or near-crash status in a two-folded manner. It provides us an opportunity to reveal the potentially different underlying reasons for drivers who crashed or near-crashed and for drivers with higher crash or near-crash rates.

Gender (binary, female and male) and age (continuous) are included in the analysis as control variables for their potential confounding effects on drivers safety and driving “abnormality”. Age is recorded as a categorical variable in the SHRP2 data with 16 levels: the youngest age group from 16 to 19 and the rest 15 age groups starting from 20 to 94 with a step of 5 years (i.e., 20–24, 25–29, ..., 90–94). This variable was converted to numeric by taking the mean of the lower and upper bound of each age group, e.g., 17.5 for age group 16-19 and 22 for age group 20-24. A quadratic relationship between age and driver safety level or driving performance has been observed in previous studies (Williams and Carsten, 1989; Guo et al., 2017). Converting the age variable to numeric allows this type of relationship to be examined.

5.2.4 Modeling Method

The analysis of Aim 2 involves four outcome variables, i.e., the “abnormality” scores on the four selected vehicle kinematics. Because “abnormality” scores were calculated for each individual traversal, the four outcomes are very likely not independent of each other due to exposure to the same environmental and vehicular conditions. For example, drivers manipulate a curved road may slow down while steering, leading to simultaneous changes in both longitudinal and lateral kinematic metrics. On the other hand, the explanatory variables are likely correlated too. Although the primary interest is the association between driver safety and the level of driving “abnormality”, both variables are expected to be affected by drivers' gender and age. The specific structure of outcome and explanatory variables requires an analytical method that can simultaneously test the multiple relationships among variables.

Path models was thus used for the analysis of Aim 2. Path models are a multivariate analysis method that test theoretical relationships among observed variables. The path model is an extension of multiple regression models by allowing a number of independent and dependent variables (Schumacker and Lomax, 2012). It is also a special type of Structural Equation Modeling (SEM) that includes only observed variables (i.e., no latent variables).

Figure 5.1 shows the path diagram of the theoretically proposed “abnormality” score model with outcome and explanatory variables introduced in Section 5.2.2 and 5.2.3. More specifically, we assume that in a specific traversal, the associated “abnormality” score would be directly influenced by drivers’ gender, age and the safety level measured by crash or near-crash status. The “abnormality” score could also be indirectly influenced by gender and age through crash or near-crash status. An error term is specified for each endogenous variable to account for additional influential factors that were not included in the current path model. Errors of “abnormality” scores for the four kinematics were allowed to correlate (i.e., e_1 to e_4), indicating that they could be exposed to similar unexplained influences, e.g., environmental and vehicular factors. For simplicity, the correlation are not shown in Figure 5.1. A quadratic term of age was retained if (1) the term is statistically significant at 0.1 confidence level or (2) removal of the term leads to a significantly worse model fit either indicated by a Likelihood-Ratio Test or that any of the examined global statistics exceed the preferred threshold.

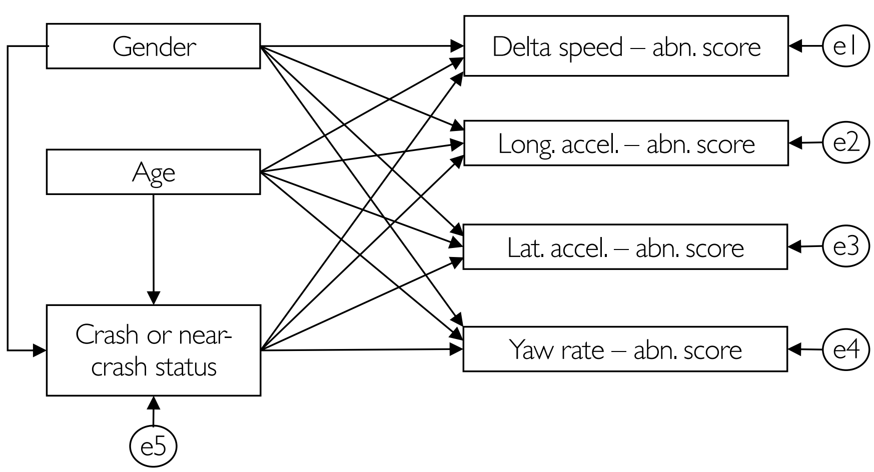

Figure 5.1: “Abnormality” score model

Four global fit statistics were used to help evaluate the model goodness-of-fit as suggested in Kline (2015). A short description and the general accepted threshold of a good model fit are summarized as below:

1. *Model chi-square test statistic*: This statistic is calculated as $\chi^2_M = (N-1)F_{ML}$ or $N(F_{ML})$, where N is the sample size and F_{ML} is the value of the fit function minimized in model estimation. The corresponding χ^2 test examines whether the model predicted covariance equals the observed covariance. Fail to reject the null hypothesis $\chi^2_M = 0$ suggests the two covariance matrices are not significantly different, thus a good fit.
2. *Root Mean Square Error of Approximation (RMSEA)*: This statistics is calculated as $\hat{\epsilon} = \sqrt{\frac{\hat{\Delta}_M}{df_M(N-1)}}$, where N is the sample size and $\hat{\Delta}_M = max(0, \chi^2_M - df_M)$ is the limit of the close fit. RMSEA thus evaluates model departure from a close or approximate fit. RMSEA values ≤ 0.05 can be considered as a good fit and values between 0.05 and 0.08 as an adequate fit (Browne, 1993).
3. *Comparative Fit Index (CFI)*: This statistics is calculated as $\text{CFI} = 1 - \frac{\hat{\Delta}_M}{\hat{\Delta}_B}$, where $\hat{\Delta}_M$ is defined similarly as for RMSEA and $\hat{\Delta}_B = max(0, \chi^2_B - df_B)$ is defined for a baseline model, i.e., an independence model of zero covariance. This statistics thus evaluates the close fit model to an independence model. Models with CFI values ≥ 0.97 can be considered a good fit and values between 0.97 and 0.95 an acceptable fit (Schermelleh-Engel et al., 2003).
4. *Standardized Root Mean Square Residual (SRMR)*: This statistics is calculated as $\text{SRMR} = \sqrt{\frac{\sum_{i=1}^{p}\sum_{j=1}^{i}\left(r_{ij} - \hat{\sigma}_{ij}/(s_{ii}s_{jj})\right)}{p(p+1)/2}}$, where r_{ij} is the observed correlation between variable i and j, $\hat{\sigma}_{ij}$ is the model-implied covariance between variable i and j, s_i is the observed standard deviation of variable i, and p is the number of observed variables. SRMR ≤ 0.05 suggests a good fit (Hu and Bentler, 1995) and values between 0.05 and 0.1 suggests an acceptable fit (Schermelleh-Engel et al., 2003).

Global fit statistics are not available if a model is just-identified. In this study, this happens when the quadratic term of age is discarded for any of the "abnormality" score variables and crash or near-crash status. If this happened, model coefficients and standard errors of the model with one quadratic term of age from the last step of model selection were compared with the just-identified model. If all coefficients are similar, the just-identified model was kept, otherwise the model with one quadratic term of age was kept.

In addition to the four global fit statistics, standardized residual covariances were also examined as a measure of local fit goodness. Kline (2015) suggests to watch for standardized residuals larger than 0.1 as a indicator of model-data disagreement (Kline, 2015).

The path model was applied to three of the twelve environmental conditions with the most number of traversals (see Table 5.1). These three conditions are identified as UI, UOPA, and UMA from this point forward:

- UI: Urban Interstate, with *traffic density* in A1 to C, and straight *alignment* (#1)
- UOPA: Urban Other Principal Arterial, with straight *alignment* and *traffic influence* as "No" or "Yes, Interchange" (#6)
- UMA: Urban Minor Arterial, with straight *alignment* and *traffic influence* as "No" (#10)

It can be observed that the three environmental conditions are all with straight alignment. The reason to focus on only the three selections is primarily due to that the SEM method, including path models, requires large samples. Previous studies reveal a median sample size of SEM studies as about 200 cases and SEM based on a sample with less than 100 observations is considered untenable (Kline, 2015). In addition to the sample size issue, curved alignment is also more complex than straight alignment. For example, two traversals both on curved alignment are likely to be associated with different roadway curvature and relative positions as regards the curve. The more complex roadway situation would require larger samples to be accounted for, which would in turn worsen the sample size issue associated with curved alignment traversals. Since only one environmental group in each functional class was used, the three environmental groups will also be referred to by the abbreviation of their related functional classes (UI, UOPA and UMA) in the rest of the paper.

R package `lavaan` (version 0.6-8) was used to conduct the path analysis in Aim 2.

5.3 Results

5.3.1 Descriptive Statistics

Numbers of traversals and the distribution of drivers' age and gender are summarized in Table 5.2. The column, "# of drivers", refers to the number of unique drivers in each environmental group. The three environmental groups are largely different drivers. Only 34 drivers were in all three conditions. Nevertheless, similar distributions of drivers' gender and age are observed across the three groups. The majority of drivers had only 1 or 2 traversals in the analyzed dataset. More specifically, the percentage of drivers with 1 or 2 traversals is 68% and 20% in UI, 78% and 16% in UOPA and 81% and 15% in UMA. On average, there were 1.52, 1.32 and 1.26 traversals per driver in UI, UOPA and UMA respectively. The extreme sparseness (very few traversals per driver) and the unbalanced structure (different number of traversals per driver) of the data were likely to result in model convergence issues (Clarke, 2008). A previous simulation study also found similar model results from single-level and multi-level regression methods on this type of data (Clarke, 2008). Thus, the data was not examined with repeated measures.

Table 5.2: Driver demographic summary

Env. group	# of traversals	# of drivers	Gender		Age (years)			
			Female	Male	Mean	SD	Min	Max
UI	1,134	738	385	353	39	21.33	17.5	92
UOPA	821	635	347	288	41	22.89	17.5	92
UMA	539	428	214	214	39	22.46	17.5	87
Total	2,494	1,393	725	668	40	22.26	17.5	92

Percentages of drivers with crashes during the SHRP2 NDS data collection period are 36%, 40% and 39% in the UI, UOPA and UMA groups. For near-crashes, the percentage is 76%, 67% and 65% respectively. Crash and near-crash rates for drivers with crashes or near-crashes are shown in Figure 5.2. It should be noted that, crashes defined in SHRP2 include "any contact that the subject vehicle has with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated" (Hankey et al., 2016). That is, the recorded SHRP2 crashes were in general less severe than policed-reported crashes. Actually, only 19% (n = 131) of the recorded crashes were police-reportable, including situations with sufficient property damage or the injury of any involved road users (Transportation Research Board of the National Academies of Science, 2013).

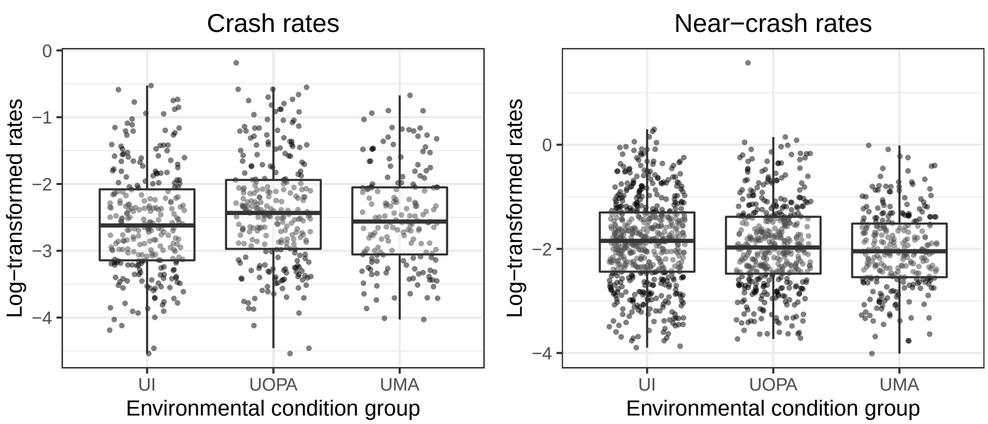

Figure 5.2: Crash and near-crash rates when crash or near-crashed

Figure 5.3 contains four panels of boxplots of the "abnormality" score by kinematic type and environmental group. In general, we observe a consistent distribution of "abnormality" scores across the three environmental groups. That is, the majority of traversals are with a similar "abnormality" score regardless of environmental group. Traversals with higher "abnormality" scores are all on a similar magnitude.

Aim 2 examines the relationship between crash or near-crash status and driving "abnormality" score. In addition, comparisons would be conducted among the three environmental groups to provide insights on potential environmental effects on this relationship. Summary statistics of this section reveal similar distributions of driver demographics, crash or

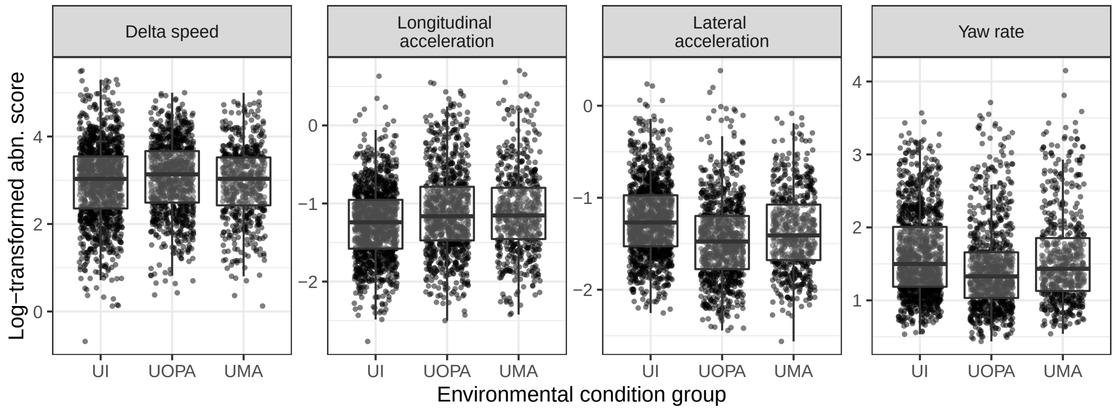

Figure 5.3: "Abnormality" score by kinematic type and environmental group

near-crash status and "abnormality" scores across the three environmental groups. This observation helps to ensure that the comparison across environmental groups is plausible even when different groups of drivers were analyzed in the three conditions.

5.3.2 Model Results

Tables 5.3 to 5.6 summarize the model results with driver safety measured by four different forms of crash or near-crash status. Robust standard errors are reported in case of non-normality not accounted for by transformation. P-values in the table were adjusted p-values using a Benjamini and Hochberg's False Discovery Rate (FDR-BH) method to account for inflated Type I error due to multiple comparisons. This method is considered more powerful than familywise error controlling method such as Bonferroni meanwhile still effective in controlling for Type I error (Smith and Cribbie, 2013). Estimated coefficients significant at 0.05 level are in italic and bold while estimates significant at 0.1 level are in italic only. Robust versions of the four global fit statistics are also provided in the table. All models with global fit statistics satisfy the good-fit model criteria of all four statistics.

Safety level and "abnormality" score

The path models reveal five significant relationships between safety level measured by crash or near-crash status and "abnormality" score on one of the vehicle kinematics. More specifically, drivers that ever crashed during the SHRP2 study were associated with 1.05 times higher "abnormality" score of yaw rate on UI (see Table 5.3). Among drivers who crashed, an $e = 2.72$ times higher crash rates was associated with 1.15 times higher "abnormality" score in delta speed on UI and 1.11 times higher "abnormality" score on lateral acceleration on UMA (see Table 5.4). For near-crash status, the two significant estimates are both observed when drivers traveling on UI. That is, "abnormality" score on the lateral acceleration for drivers with near-crashes were 1.04 times higher than those without near-crashes and every 2.72 times increase in near-crash rates was associated with 5% higher "abnormality" score of longitudinal acceleration (see Tables 5.5 and 5.6).

"Abnormality" score and drivers' age and gender

Estimated coefficients associated with age and gender on "abnormality" scores in Tables 5.3 to 5.6 represent the direct effects of age and gender on "abnormality" score. That is, the effects are not through safety level measured by crash or near-crash status but other possible factors, e.g., driving skills.

Significant influences of drivers' age on their driving "abnormality" are observed for delta speed and lateral acceleration and are mostly on UI than on UOPA or UMA. For delta speed, the influence of age is monotonically negative. That is, older drivers were consistently associated with less deviation from the majority in delta speed – the difference between the operational speed and the posted speed limit. The magnitude of reduction ranges from 8% to 12% for every one standard deviation increase of age (about 21 to 22 years) across the different modeling situations. The influence of age on lateral acceleration, nevertheless, is U-shaped. That is, the "abnormality" in lateral acceleration was higher for both younger and older drivers compared to middle-aged drivers. The lowest points is reached by drivers approximately 0.8×SD older than the mean age, i.e., 56 years.

Gender is less influential on driving "abnormality" compared to age. Female drivers were associated with 6% larger "abnormality" on longitudinal acceleration on UI (see Table 5.5). Among drivers with near-crashes, females were also 12% higher on the "abnormality" score of lateral acceleration (see Table 5.6). Both of the effects are significant only at 0.1 level.

Safety level and drivers' age and gender

Age shows a consistent impact on crash and near-crash status across the three environmental groups. Although due to the different samples of drivers involved in the three groups, the magnitude and significance of impact can be different. Younger and older drivers were more likely to crash than middle-aged drivers and also more likely to have a higher crash rates. Near-crashes, nevertheless, were in general negatively influenced by age. That is, drivers of an older age were associated with lower likelihood of near-crash and lower near-crash rates.

Gender, on the other hand, shows somewhat different impacts on crash or near-crash status. For drivers in all three environmental groups, no significant association between crash occurrence and gender is observed. Among drivers who crashed, female drivers were associated with higher crash rates but only significant in UI. The association between gender and near-crash status also only exists for UI. More specifically, females were less likely to near-crash but more likely to have higher near-crash rates compared to male drivers.

Table 5.3: Path model results with safety level measured by crashed or not

Variable	UI ($N = 1,134$)			UOPA ($N = 821$)			UMA ($N = 539$)		
	Est.	Std. Err.	P-value	Est.	Std. Err.	P-value	Est.	Std. Err.	P-value
"Abnormality" score on delta speed (log-transformed)									
Crash (base: no crash)	0.058	0.035	0.157	0.006	0.036	0.918	-0.039	0.044	0.614
Female (base: male)	0.104	0.054	0.123	0.006	0.058	0.918	-0.049	0.071	0.615
Age (standardized)	*-0.086*	*0.039*	*0.074*	-0.025	0.035	0.863	-0.079	0.047	0.279
"Abnormality" score on longitudinal acceleration (log-transformed)									
Crash	0.030	0.017	0.151	-0.017	0.024	0.863	0.036	0.028	0.441
Female	0.048	0.027	0.151	-0.023	0.039	0.863	0.016	0.047	0.849
Age	-0.017	0.019	0.448	-0.011	0.024	0.863	-0.005	0.030	0.873
"Abnormality" score on lateral acceleration (log-transformed)									
Crash	0.017	0.016	0.356	0.008	0.019	0.863	0.016	0.023	0.615
Female	0.024	0.024	0.407	-0.023	0.032	0.863	0.030	0.038	0.615
Age	***-0.070***	***0.019***	***0.002***	-0.010	0.019	0.863	0.027	0.024	0.497
Age^2	*0.039*	*0.017*	*0.074*						
"Abnormality" score on yaw rate (log-transformed)									
Crash	*0.049*	*0.022*	*0.074*	0.029	0.026	0.863	0.083	0.033	0.169
Female	-0.008	0.035	0.822	-0.004	0.042	0.918	0.010	0.055	0.873
Age	-0.030	0.026	0.353	0.011	0.026	0.863	0.056	0.034	0.279
Crash (probit link)									
Female	0.046	0.078	0.593	0.065	0.090	0.863	0.145	0.111	0.441
Age	***-0.276***	***0.059***	**< 0.001**	***-0.192***	***0.057***	***0.005***	-0.163	0.072	0.169
Age^2	***0.164***	***0.051***	***0.007***	***0.296***	***0.066***	**< 0.001**	0.155	0.083	0.279
Robust Chi-square:	$\chi^2_3 = 2.248$ (p= 0.523)			$\chi^2_4 = 2.171$ (p= 0.704)			$\chi^2_4 = 1.929$ (p= 0.749)		
Robust RMSEA:	0 (90% CI:$[0, 0.045]$, p= 0.972)			0 (90% CI:$[0, 0.039]$, p= 0.984)			0 (90% CI:$[0, 0.046]$, p= 0.965)		
Robust CFI:	1			1			1		
Robust SRMR:	0.002			0.004			0.002		

Table 5.4: Path model results with safety level measured by crash rates for crashed drivers

Variable	UI ($N = 390$)			UOPA ($N = 327$)			UMA ($N = 217$)		
	Est.	Std. Err.	P-value	Est.	Std. Err.	P-value	Est.	Std. Err.	P-value
"Abnormality" score on delta speed (log-transformed)									
Log(crash rate)	***0.142***	***0.052***	***0.025***	0.096	0.054	0.251	0.069	0.074	0.657
Female (base: male)	0.078	0.086	0.525	-0.113	0.087	0.386	0.048	0.111	0.808
Age (standardized)	-0.053	0.048	0.481	0.047	0.055	0.704	***-0.193***	***0.061***	***0.012***
Age^2				-0.123	0.061	0.234			
"Abnormality" score on longitudinal acceleration (log-transformed)									
Log(crash rate)	0.053	0.031	0.199	0.063	0.038	0.262	0.001	0.051	0.982
Female	0.029	0.046	0.568	-0.019	0.061	0.904	0.025	0.079	0.808
Age	-0.019	0.023	0.525	0.002	0.029	0.950	-0.015	0.039	0.808
"Abnormality" score on lateral acceleration (log-transformed)									
Log(crash rate)	0.025	0.026	0.525	-0.010	0.035	0.904	*0.101*	*0.039*	*0.052*
Female	0.026	0.041	0.568	-0.016	0.051	0.904	0.021	0.058	0.808
Age	***-0.124***	***0.030***	***0.001***	-0.035	0.023	0.288	0.016	0.027	0.808
Age^2	0.046	0.024	0.171						
"Abnormality" score on yaw rate (log-transformed)									
Log(crash rate)	0.070	0.038	0.171	-0.008	0.042	0.904	0.105	0.052	0.167
Female	0.018	0.060	0.766	0.015	0.067	0.904	0.029	0.082	0.167
Age	-0.019	0.030	0.568	0.011	0.033	0.904	0.058	0.041	0.167
Log(crash rate)									
Female	***0.252***	***0.078***	***0.006***	0.204	0.083	0.117	0.104	0.097	0.609
Age	-0.097	0.064	0.254	-0.094	0.053	0.251	-0.115	0.069	0.292
Age^2	***0.186***	***0.057***	***0.006***	***0.189***	***0.064***	***0.050***	***0.278***	***0.077***	***0.005***
Robust Chi-square:	$\chi^2_3 = 1.881$ (p= 0.597)			$\chi^2_3 = 3.890$ (p= 0.274)			$\chi^2_4 = 2.380$ (p= 0.666)		
Robust RMSEA:	0 (90% CI:$[0, 0.07]$, p= 0.597)			0 (90% CI:$[0, 0.105]$, p= 0.567)			0 (90% CI:$[0, 0.079]$, p= 0.842)		
Robust CFI:	1			0.991			1		
Robust SRMR:	0.007			0.013			0.011		

Table 5.5: Path model results with safety level measured by near-crashed or not

Variable	UI ($N = 1,134$)			UOPA ($N = 821$)			UMA ($N = 539$)		
	Est.	Std. Err.	P-value	Est.	Std. Err.	P-value	Est.	Std. Err.	P-value
"Abnormality" score on delta speed (log-transformed)									
Near-crash (base: no near-crash)	-0.021	0.036	0.609	-0.019	0.037	0.845	-0.004	0.044	0.928
Female (base: male)	0.101	0.055	0.126	0.004	0.058	0.953	-0.056	0.072	0.598
Age (standardized)	***-0.106***	***0.040***	***0.029***	-0.058	0.030	0.345	-0.073	0.048	0.468
"Abnormality" score on longitudinal acceleration (log-transformed)									
Near-crash	0.021	0.018	0.314	-0.031	0.025	0.816	0.071	0.031	0.107
Female	*0.056*	*0.028*	*0.100*	-0.027	0.039	0.816	0.043	0.048	0.553
Age	-0.021	0.020	0.337	-0.024	0.021	0.816	-0.003	0.031	0.928
"Abnormality" score on lateral acceleration (log-transformed)									
Near-crash	*0.037*	*0.017*	*0.075*	0.014	0.021	0.816	0.028	0.024	0.468
Female	0.036	0.025	0.226	-0.021	0.032	0.816	0.041	0.038	0.468
Age	***-0.067***	***0.018***	***0.004***	-0.004	0.017	0.953	0.027	0.024	0.468
Age^2	***0.042***	***0.017***	***0.039***						
"Abnormality" score on yaw rate (log-transformed)									
Near-crash	0.037	0.023	0.195	-0.003	0.027	0.953	0.021	0.034	0.669
Female	0.005	0.036	0.880	-0.002	0.042	0.953	0.028	0.056	0.710
Age	-0.036	0.026	0.226	0.018	0.022	0.816	0.045	0.034	0.468
Near-crash (probit link)									
Female	***-0.302***	***0.087***	***0.004***	-0.098	0.093	0.816	-0.309	0.114	0.103
Age	***-0.193***	***0.061***	***0.008***	***-0.248***	***0.046***	**< 0.001**	-0.104	0.075	0.468
Age^2							-0.207	0.084	0.103
Robust Chi-square:	$\chi^2_4 = 3.161$ (p= 0.531)			Not available			$\chi^2_4 = 2.120$ (p= 0.714)		
Robust RMSEA:	0 (90% CI:$[0, 0.040]$, p= 0.987)			Not available			0 (90% CI:$[0, 0.048]$, p= 0.957)		
Robust CFI:	1			1			1		
Robust SRMR:	0			0			0.004		

Table 5.6: Path model results with safety level measured by near-crash rates for drivers near-crashed

Variable	UI ($N = 882$) Est.	UI Std. Err.	UI P-value	UOPA ($N = 556$) Est.	UOPA Std. Err.	UOPA P-value	UMA ($N = 355$) Est.	UMA Std. Err.	UMA P-value
"Abnormality" score on delta speed (log-transformed)									
Log(near-crash rate)	0.021	0.036	0.685	-0.084	0.039	0.115	-0.010	0.051	0.927
Female (base: male)	0.098	0.061	0.201	0.003	0.069	0.983	-0.045	0.084	0.829
Age (standardized)	***-0.125***	***0.048***	***0.029***	***-0.107***	***0.037***	***0.019***	*-0.126*	*0.048*	*0.056*
Age^2	0.072	0.040	0.158						
"Abnormality" score on longitudinal acceleration (log-transformed)									
Log(near-crash rate)	***0.051***	***0.017***	***0.015***	0.036	0.027	0.567	-0.054	0.034	0.266
Female	0.043	0.031	0.242	-0.018	0.046	0.803	0.010	0.059	0.927
Age	-0.018	0.016	0.323	-0.023	0.023	0.803	-0.023	0.031	0.729
"Abnormality" score on lateral acceleration (log-transformed)									
Log(near-crash rate)	0.002	0.016	0.937	-0.015	0.024	0.803	0.008	0.030	0.927
Female	0.034	0.028	0.321	-0.015	0.039	0.803	*0.113*	*0.047*	*0.056*
Age	***-0.070***	***0.021***	***0.007***	-0.009	0.020	0.803	0.049	0.025	0.123
Age^2	***0.043***	***0.018***	***0.043***						
"Abnormality" score on yaw rate (log-transformed)									
Log(near-crash rate)	-0.005	0.022	0.925	-0.020	0.032	0.803	-0.041	0.044	0.627
Female	0.002	0.039	0.961	0.021	0.049	0.803	0.089	0.067	0.369
Age	-0.045	0.029	0.204	-0.001	0.027	0.983	*0.095*	*0.038*	*0.056*
Age^2	0.053	0.028	0.138						
Log(near-crash rate)									
Female	***0.156***	***0.058***	***0.029***	-0.048	0.071	0.803	0.002	0.083	0.981
Age	***-0.122***	***0.029***	**< 0.001**	***-0.188***	***0.043***	**< 0.001**	***-0.141***	***0.044***	***0.021***
Age^2				***0.178***	***0.051***	***0.003***			
Robust Chi-square:	$\chi^2_2 = 0.950$ (p= 0.622)			$\chi^2_4 = 1.761$ (p= 0.780)			Not available		
Robust RMSEA:	0 (90% CI:[0, 0.055], p= 0.929)			0 (90% CI:[0, 0.044], p= 0.969)			Not available		
Robust CFI:	1			1			1		
Robust SRMR:	0.004			0.008			0		

5.4 Discussion

This chapter aims to examine the relationship between driver safety and the level of driving "abnormality". The answer to the first research question "Do less safe drivers behave more 'abnormally'?" is "yes". The set of path models suggest five associations between crash or near-crash status and the level of driving "abnormality" as significantly different from a null effect. More specifically, the five associations indicate that drivers with crash or near-crash records also had higher deviation in their driving performance from a "normal" driving style. An interesting finding is that, drivers who crashed or near-crashed were associated with higher level of "abnormality" in the lateral driving direction. On the other hand, among drivers who crashed or near-crashed, those with higher crash or near-crash rates were associated with higher level of "abnormality" in their longitudinal driving performance. While lateral stability is oftentimes used to assess driver skills (Brown et al., 2020; Chen et al., 2021) and is considered more easily affected by impaired driving ability due to e.g., aging (Chen et al., 2021) or inactivity from on-road driving (Upahita et al., 2018), deviations on the longitudinal direction are more frequently used to capture aggressive or risky driving styles, e.g., through the detection of speeding and hard acceleration (Mantouka et al., 2019; Johnson and Trivedi, 2011). The distinct association of longitudinal and lateral "abnormality" with crash or near-crash status may suggest that one is more likely to crash/near-crash due to insufficient driving skills or impaired driving ability, but more likely to have higher crash/near-crash rates due to an aggressive driving tendency.

The "abnormality" score is considered a straight-forward and easily interpreted metric of the level of "abnormality" associated with a specific driving style. Nevertheless, while a relationship has been revealed between driver safety and the "abnormality" score, there are limitations as regards the use of this "abnormality" metric. For example, a composite score does not tell which specific aspects of an "abnormal" driving style contribute to its association with driver safety. More importantly, it is possible that the different aspects of an "abnormal" driving style relate to driver safety in different ways. In this case, the total effect may have canceled out and fails to be captured by the composite "abnormality" score. The next chapter will use a different "abnormality" metric to deal with the limitation.

The answer to the second research question "Will the relationship depend on driving environment?" is also "yes". More specifically, the association between driver safety and the level of driving "abnormality" is more distinguishable in UI than in the other two environmental groups. Nevertheless, it is less clear whether the difference also exists between UOPA and UMA. After all, no significant association was observed on UOPA and only one observed on UMA. The unclearness could be due to the limitation of the composite "abnormality" score as discussed in the previous paragraph. Thus, the discussion for the impact of driving environments will be saved to the next chapter.

The significant direct effect between age and the level of driving "abnormality" on delta speed and lateral acceleration suggests that age affects drivers' driving performance on this two kinematics through approaches other than safety level. The direct effect could be somewhat difficult to interpret because that the actual variable entered into the model – crash or near-crash status – though named by "safety level", could be due to a variety of reasons. Thus, a more straight-forward interpretation of the direct effect would be the effect of age on the level of driving "abnormality" for drivers with the same crash or near-crash status. One reason of the monotonically decreasing "abnormality" in delta speed with age could be that drivers of an increased age are more likely to observe the speed limit (Shinar et al., 2001). Our results that older drivers tend to select a speed more similar to the surrounding vehicles is somewhat different from previous studies that suggest a lower driving speed of older drivers (Shinar et al., 2005). Nevertheless, the underlying rationale could be very similar. Older drivers behave in a certain way to compensate their declining cognitive processing ability and to ensure themselves sufficient time to react to events on the road (Andrews and Westerman, 2012). The compensatory behavior could be driving at a lower speed in the driving simulator with very few surrounding vehicles and potential distraction tasks such as in Shinar et al. (2005). It could also be driving at a similar speed to surrounding vehicles on a highway so that the relative speed difference is minimized. The higher "abnormality" in lateral acceleration for younger and older drivers could be due to different reasons. For example, younger drivers are more prone to distracted driving which is associated with higher deviation on the lateral direction (Dozza et al., 2015; Choudhary and Velaga, 2017). On the other hand, older drivers are more likely with higher "abnormality" on the lateral control performance due to physiological and cognitive performances deteriorate (Chen et al., 2021).

There also exists an effect of drivers' age on their safety level, regardless of the environmental group. The quadratic effect of age on crash status, i.e., younger and older drivers are more prone to crashes than middle-aged drivers, is consistent as in previous studies (Williams and Carsten, 1989; Guo et al., 2017). Although the analysis in Aim 2 include only a subset of drivers with traversals in the three environmental groups, the observed negative relationship between age and near-crash rate is the same with the results based on all SHRP2 drivers (Seacrist et al., 2020). The different impact of age on crashes and near-crashes could be result from the different driver characteristics. Younger drivers have less driving experience therefore detect hazards less quickly and perceive them less holistically than experienced drivers (Deery, 1999). Meanwhile, younger drivers have higher reaction times compared to middle-aged and older drivers (Dozza et al., 2015), which enables them to successfully evade crashes and turn them into near-crashes. Older drivers may drive more cautiously to compensate their reduced physiological and cognitive performances. Nevertheless, if a critical situation happens, older drivers might be less likely than younger and middle-aged drivers to react in a timely manner thus result in a crash.

Chapter 6

ASPECTS OF DRIVING "ABNORMALITY" AND RELATIONSHIP TO DRIVER SAFETY

6.1 Objective

The last chapter examines the relationship between driver safety and the level of driving "abnormality". While the results confirm an association between crash or near-crash status and higher "abnormality" score, it is yet clear what specific aspects of an "abnormal" driving style contribute to the association. Further, it is of interest that whether the different aspects of an "abnormal" driving style might provide insights on specific behaviors engaged by less safe drivers. Aim 3 thus extends the focus of Aim 2 by answering the following questions.

- *Aim 3: What is the relationship between driver safety and the different aspects of an "abnormal" driving style? Do less safe drivers tend to engage in specific "abnormal" behaviors? Will the relationship depend on driving environment?*

6.2 Methods

6.2.1 Outcome Variables

The outcome in Aim 3 is a set of statistical features, or "abnormality" features, that describe the shape of an "abnormality" curve. Recall that in Aim 2, x_{ijkl} was used to denote the driving pattern associated with vehicle kinematic k for a traversal i of a driver j in an environmental condition l and x_{kl} for the "normal" driving pattern of vehicle kinematic k in that environmental condition. Based on that, we defined $\boldsymbol{x}_{ijkl} = \left\{x_{ijkl}(t_1), ..., x_{ijkl}(t_N)\right\}$ and $\boldsymbol{x}_{kl} = \left\{x_{kl}(t_1), ..., x_{kl}(t_N)\right\}$ as two realizations of x_{ijkl} and x_{kl} at time $t_1, ..., t_N$. The "abnormality" curve $\tilde{\boldsymbol{x}}_{ijkl}$ is given as the difference between the individual curve $\boldsymbol{x}_{ijkl}$ and the "normal" curve $\boldsymbol{x}_{kl}$ at time $t_1, ..., t_N$:

$$\tilde{\boldsymbol{x}}_{ijkl} = \boldsymbol{x}_{ijkl} - \boldsymbol{x}_{kl} = \left\{x_{ijkl}(t_1) - x_{kl}(t_1), ..., x_{ijkl}(t_N) - x_{kl}(t_N)\right\} \tag{6.1}$$

Six statistical features were originally proposed to build up an "abnormality" feature set for each "abnormality" curve, i.e., mean, median, minimum, maximum, standard deviation, and the sum of the absolute gradient. The first two features were expected to capture the mean trend of an "abnormality" curve, the third and the forth to capture extreme values and the last two features to account for the variation in the curve. Preliminary analysis revealed high correlations between the mean and the median features for all three environmental conditions (correlations ranging from 0.94 to 0.99). Median was therefore excluded from the feature set. The remaining five statistical features are defined and calculated as below:

1. Mean:

$$mean_{ijkl} = \frac{1}{N}\sum_{i=1}^{N} \tilde{x}_{ijkl}(t_i) \tag{6.2}$$

2. Minimum:

$$min_{ijkl} = \min\left(\tilde{x}_{ijkl}(t_1), ..., \tilde{x}_{ijkl}(t_N)\right) \tag{6.3}$$

3. Maximum:

$$max_{ijkl} = \max\left(\tilde{x}_{ijkl}(t_1), ..., \tilde{x}_{ijkl}(t_N)\right) \tag{6.4}$$

4. Standard deviation:

$$sd_{ijkl} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\tilde{x}_{ijkl}(t_i) - mean_{ijkl}\right)^2} \tag{6.5}$$

5. Sum of absolute gradient:

$$grad_{ijkl} = \sum_{i=1}^{N-1}\left|\frac{\tilde{x}_{ijkl}(t_{i+1}) - \tilde{x}_{ijkl}(t_i)}{t_{i+1} - t_i}\right| \tag{6.6}$$

Specially, the minimum and the maximum features were further excluded for delta speed since the two are also highly correlated with the mean delta speed in all three environmental conditions (correlations ranging from 0.96 to 0.99). The final outcome set of "abnormality" features thus includes 18 variables: 3 feature variables for delta speed, 5 feature variables for each of the other three kinematics.

Similar to Aim 2, to ensure that variables of the four kinematics are on the same scale, all variables of delta speed were divided by 10 and all variables of longitudinal acceleration and lateral acceleration were multiplied by 9.8 (i.e., converting units from gravity to m/s^2). Transformation was applied for some outcome variables to accommodate the normality assumption of a SEM. SEM assumes multivariate normal distribution for all continuous outcome variables. Nevertheless, to examine a joint distribution of multiple variables is often impractical and existed significance tests for multivariate normality can be too conservative to reject. Therefore, it is recommended that normality diagnosis be based on an inspection of skew and kurtosis of univariate variable distributions. Skew larger than 3 or kurtosis larger than 8 may suggest severe non-normality issues (Kline, 2015).

Outcome variables with either skew larger than 3 or kurtosis larger than 8 were transformed. The goal of the transformation is to bring the skew and kurtosis of these variables below the threshold. Two transformation methods were used:

- *Natural logarithm (log)*: Positive variables (i.e., *sd* and *gradient* features) were transformed by taking the natural log on the original data. Variables with negative values (i.e., *mean*, *min*, and *max* features) were firstly subtracted by one plus the minimum value and then log-transformed.
- *Inverse hyperbolic sine (asinh)*: The inverse hyperbolic sine is defined for real values x as $asinh(x) = log(x + \sqrt{x^2 + 1})$. Variables were transformed by taking the transformation method on the original data. This method appears to work better than log-transformation for variables with high kurtosis, nevertheless more difficult to interpret. Thus, this method was used only when log-transformation failed the goal.

Variables and the corresponding transformation method are summarized as below for each environmental condition.

- UI, UOPA and UMA - All variables related to *sd* and *gradient* are log-transformed. It might be worth mentioning that, not all *sd* and *gradient* features are with high skew and/or kurtosis. Nevertheless, since a large proportion of variables related to the two features needed to be transformed, it is decided that all would be transformed for an easier comparison of the two features across kinematics and groups in later analysis.
- UOPA - Maximum of longitudinal acceleration was log-transformed.
- UOPA - Mean, minimum and maximum of yaw rate was asinh-transformed.
- UMA - Minimum and maximum of yaw rate was asinh-transformed.

6.2.2 Explanatory Variables

Explanatory variables were the same as in Aim 2. That is, the explanatory variables include driver safety measured by crash or near-crash status, age of the driver and gender of the driver. Please refer to Section 5.2.3 for details of the three variables.

6.2.3 Modeling Method

The Aim 3 analysis utilizes a set of statistical features to capture the different aspects of an "abnormal" driving style. The statistical features, although depict the shape of an "abnormality" curve from different perspectives, are highly likely to be correlated. For example, a traversal with higher mean longitudinal acceleration may also be observed with higher maximum value of this kinematic. Similarly, a traversal with more extreme minimum and maximum values may also be more likely to have a higher standard deviation and/or sum of absolute gradient. That is, the statistical features may be considered as measuring higher-level factors related to the driving performance that cannot be directly observed, e.g., the mean trend and variation of a driving curve.

The SEM was therefore used in the Aim 3 analysis. SEM is a statistical technique to quantitatively test a theory based on multiple dependent and independent variables that are observed or latent (Schumacker and Lomax, 2012). The SEM contains a measurement model and a structure model. The measurement part of the SEM – a Confirmatory Factor Analysis (CFA) model – connects multiple dependent variables through a set of latent variables or factors. The structure part of the SEM – a path model – then relates the set of factors to multiple independent variables. In addition to its ability to deal with multiple dependent and independent variables, the SEM explicitly takes into consideration the measurement error of any endogenous variables (Kline, 2015). It is thus considered a more realistic and robust analytical method for a real-world problem than multiple regression and/or other multivariate procedures (e.g., MANOVA).

The SEM method introduced in the last paragraph is also referred to as a covariance-based SEM (CB-SEM), to be distinguished from the other composite-based SEM estimation approach (e.g., PLS-SEM). While both methods estimate "a series of structural equations that represent causal processes" (Rigdon et al., 2017), CB-SEM is a factor-based latent variable method while PLS-SEM is component-based. That is, latent variables in PLS-SEM are exact linear combinations of their associated manifest variables without errors, similar to components in a Principal Component Analysis. Furthermore, any manifest variables in PLS-SEM can be linked to only one latent variable (Monecke and Leisch, 2012). In the context of this study, the "abnormality" feature is expected to capture different aspects in driving performance but not without errors. It is also expected that some "abnormality" features could be associated with multiple latent variable, i.e., capturing multiple aspects of driving. Thus, the CB-SEM approach is considered more appropriate than the PLS-SEM approach for the purpose of this study.

Measurement model

The measurement part of the SEM was explored through an Exploratory Factor Analysis (EFA). The EFA suggested a division of statistical features for every kinematics into two groups: one relates to the mean trend and one to the variation. The two groups were consistent with the two perspectives of an "abnormality" curve that are expected to be captured by the set of statistical features. The EFA-suggested specification was used as an initial input for a CFA and revised based on global fit statistics and theoretical knowledge to derive a final measurement model.

The final specification of the measurement model is depicted in Figure 6.1. There are a total of seven factors in two groups: factors #1 to #3 can be considered as measuring the mean trend of "abnormality" curves for longitudinal acceleration, lateral acceleration and yaw rate while factors #4 to #7 measuring the variation of the curve for the three kinematics and delta speed.

Mean of the delta speed is not included in the measurement model. This variable was not grouped into any identified factors by the initial EFA. This is not unexpected. Among the three statistical features associated with a mean trend factor – mean, minimum and maximum – only mean was used for delta speed while minimum and maximum have been excluded in a previous step due to their high correlations with the mean value. A factor with only one indicator variable is not defined. Rather, this variable would be directly included in the structure model.

Correlations were specified among factors in the following situations. These correlations were included to account for common exposure to potentially influential factors on driving "abnormality" that were not included in the SEM, e.g., environmental factors.

- Between mean trend and variation factors of the same kinematic: f1 and f5, f2 and f6, and f3 and f7
- Among mean trend factors: f1 and f2/f3, and f2 and f3
- Among variation factors: f4 and f5/f6/f7, f5 and f6/f7, and f6 and f7

As a common practice of SEM, every manifest variable (i.e., statistical features) has an error term to account for variance not explained by its associated factor(s). An additional set of correlations were added among error terms of *sd* for longitudinal acceleration, lateral acceleration and yaw rate, as well as among error terms of *gradient* for the three kinematics.

This strategy is to account for potential vehicular influences on the variation of vehicle kinematic. For readability, these correlations are not shown in Figure 6.1.

Structure model

Figure 6.2 shows the structure part of the SEM. The structure model has a very similar form to the path model introduced in Aim 2. The difference is that the four outcome variables in the path model of Aim 2 are now replaced by a measurement model and the variable of mean delta speed. By setting up a SEM in this form, we assume that in a specific traversal, driving "abnormality" captured by the set of statistical features would be directly influenced by drivers' gender, age and safety level measured with crash or near-crash status through the latent variables. In addition, gender and age are assumed to have indirect effects on "abnormality" features through crash or near-crash status.

The model fitting procedure was the same as in Aim 2. More specifically, for each of the three environmental groups, four separate models were applied with driver safety level entering the model either as a binary indicator of crash or near-crash occurrence or a continuous measure of crash or near-crash rates. In the latter situation, only traversals associated with drivers ever crashed or near-crashed were used and the variable was log-transformed to accommodate the normality assumption. The quadratic term of age was retained following the same criteria specified for the path model in Aim 2. Model performance was accessed based on the four global fit statistics (i.e., model chi-square, RMSEA, CFI and SRMR) as well as the standardized residual covariance. See Section 5.2.4 for more details.

R package `lavaan` (version 0.6-8) was used to conduct the SEM in Aim 3.

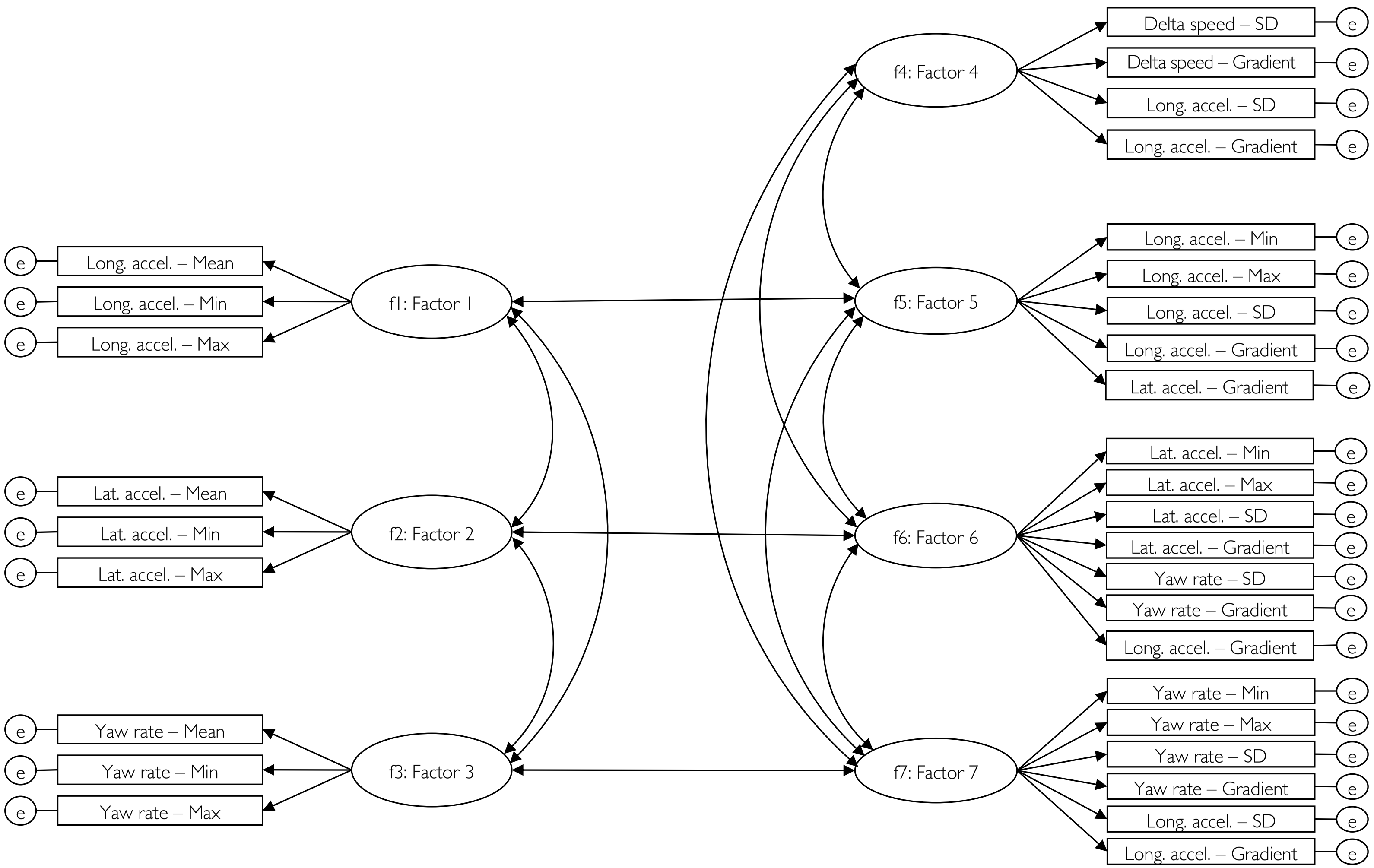

Figure 6.1: Measurement part of the SEM model

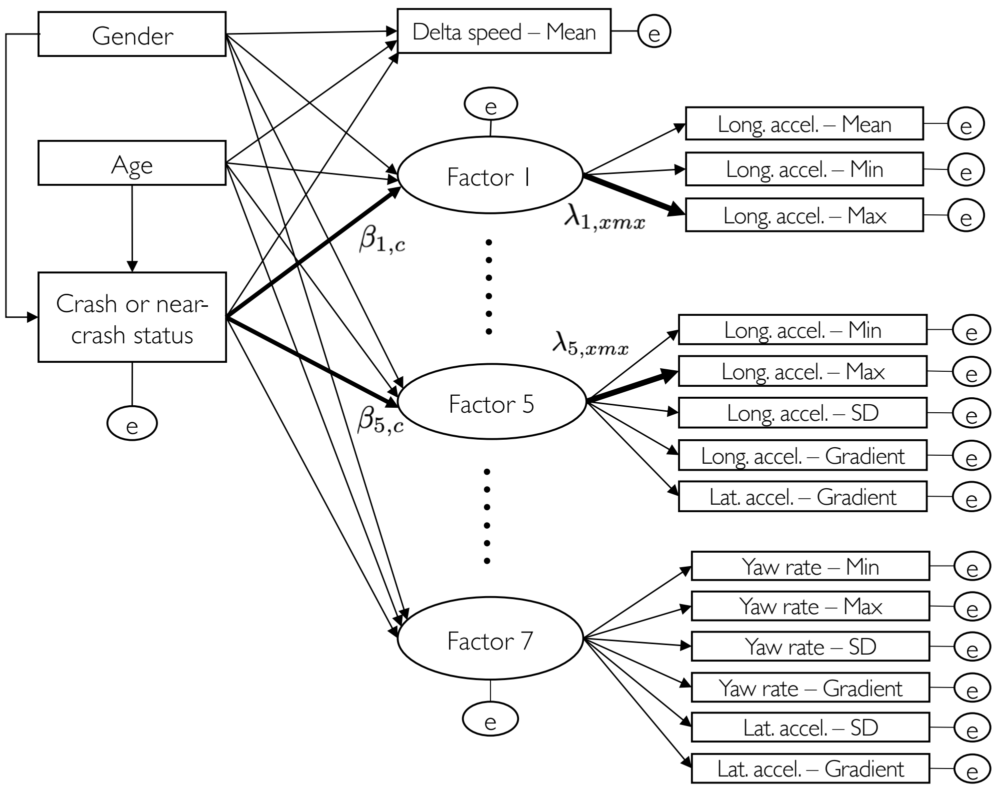

Figure 6.2: Structure part of the SEM model

6.3 Results

6.3.1 Descriptive Statistics

A set of boxplots of the "abnormality" features by environmental group and kinematic type are shown in Figure 6.3. It should be noted that this figure is based on raw feature values, i.e., neither rescaled nor transformed. Also, although the minimum and maximum of delta speed were not used for modeling, the two features are included to provide a complete comparison with the other three kinematics. For readability, scale of the y-axis is set individually.

In general, distribution of all "abnormality" features are fairly consistent across the three environmental groups. The *gradient* feature may stand out for its distinct higher magnitude than all other statistical features. The *gradient* feature sums up the total absolute unit time changes over time (see Equation 6.6) so this difference is expected. The *gradient* of delta speed is an exception for two reasons: (1) speed was recorded at 1 Hz while other three kinematics were all recorded at 10 Hz; (2) speed of traversals in all three environmental groups were quite stable over the analyzed 10-second period of time. This can be observed by the very similar distribution of the mean, the minimum and the maximum of delta speed in Figure 6.3. It is also consistent with the fact that the minimum and the maximum of delta speed are highly correlated with the mean.

Since the Aim 3 analysis was applied on the same set of traversals in Aim 2, summary statistics for drivers' age, gender and crash or near-crash status are also the same. See Section 5.3.1 to have a refresh on this information.

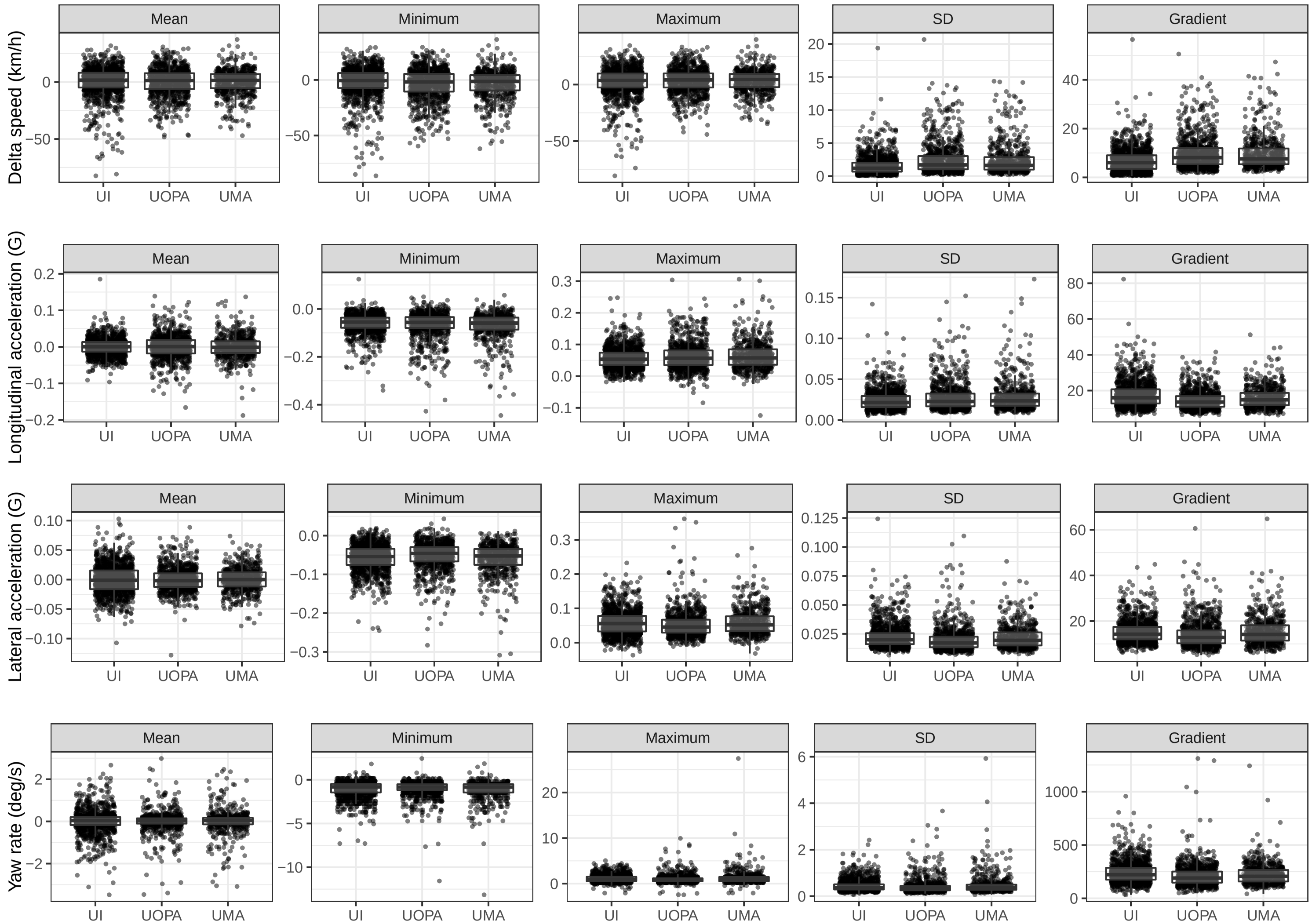

Figure 6.3: Distribution of “abnormality” features by environmental group

6.3.2 Model Results

This section will focus on the model results as regards the relationship between driver safety measured by crash or near-crash status and driving "abnormality" measured by a set of statistical features. Results between crash or near-crash status and drivers' age or gender are similar to those in the path model and can be found in Section 5.3.2. The path model and the SEM were based on the same data but different "abnormality" representations (i.e., scores versus features). The construct between crash or near-crash status and drivers' age or gender are the same in the two models and consistent estimates among them are expected.

Since many "abnormality" features are associated with multiple factors, the total effect of crash or near-crash status on each "abnormality" feature was calculated and will be used for evaluation. For a specific "abnormality" feature u, the total effect of e.g., crashed or not on it was calculated as

$$\sum_{i=1}^{n_u} \beta_{f_{u_i}} \times \lambda_{f_{u_i},u} \tag{6.7}$$

where f_{u_i} is the i^{th} of n_u factors associated with u, $\beta_{f_{u_i}}$ is the estimated coefficient between crashed or not and f_{u_i}, and $\lambda_{f_{u_i},u}$ is the loading of f_{u_i} on u. Using symbols notated in Figure 6.2 as an example, the total effect of crashed or not on the maximum longitudinal acceleration is

$$\begin{aligned} &\sum_{i=1}^{n_u} \beta_{f_{u_i}} \times \lambda_{f_{u_i},u}\Big|_{u=xmx} \\ =&\beta_{f_{xmx_1}} \times \lambda_{f_{xmx_1},xmx} + \beta_{f_{xmx_2}} \times \lambda_{f_{xmx_2},xmx} \\ =&\beta_{1,c} \times \lambda_{1,xmx} + \beta_{5,c} \times \lambda_{5,xmx} \end{aligned} \tag{6.8}$$

Tables 6.1 to 6.4 summarize the total effect of crash or near-crash status on each "abnormality" feature. Column "Std. Est." refers to the estimated total effect standardized by the observed standard deviation of the related feature. The standardized estimate will not be used directly in interpretation but offer a way to assess the magnitude of effect while taking into consideration the data variation. I will discuss this in more detail in the Discussion section. P-values were based on robust standard errors and were FDR-BH adjusted. Estimated coefficients significant at 0.05 level are in italic and bold while estimates significant at 0.1 level are in italic only.

Robust versions of the four global fit statistics are included in the table. All models reject the chi-square test at 0.05 confidence level, which means that they fail the criterion of a good-fit model based on the chi-square test. The chi-square test statistic, although among one of the most frequently used global fit statistics of SEM, has long been criticized for its sensitivity to violation of the multivariate normality assumption, favor of complex models and being difficult to reject when sample size is large (Kline, 2015; Schumacker and Lomax, 2012; Schermelleh-Engel et al., 2003). In practical situations, a model "will almost certainly be rejected if sample size is sufficiently large" (Schermelleh-Engel et al., 2003). Actually, except for the chi-square fit statistic, all models of UI and UOPA satisfy the criteria of a good-fit model measured by all other global fit statistics (i.e., RMSEA, CFI and SRMR; see Section 5.2.4 for details of the criteria). Not all global fit statistics of UMA satisfy the good-fit criteria but all of them pass the acceptable-fit threshold with values close to the good-fit threshold. In general, all models are considered to fit the data well.

Safety level and "abnormality" features

Crashed or not (Table 6.1): Compared to drivers not crashed, those with crashes were associated with higher variation in delta speed in UMA. More specifically, standard deviation and sum of absolute gradient in delta speed of drivers with crashes were 14% and 10% higher than those without crashes respectively. This is consistent with the higher mean-trend in longitudinal acceleration for crashed drivers in UMA, where their mean and maximum longitudinal acceleration were 0.051 m/s^2 and 0.078 m/s^2 higher.

In UMA, all features entirely or partially considered as measuring the variation on lateral control are significantly associated with crash occurrence, most at 0.05 confidence level. More specifically, drivers with crashes were associated with 7% higher *sd* and 3% higher *gradient* of lateral acceleration and 13% and 6% higher *sd* and *gradient* for yaw rate. Crashed drivers also had higher *max* lateral acceleration or yaw rate meanwhile lower *min* of the two kinematics. It should be noted, a positive *mean*, *max* or *min* of lateral acceleration or yaw rate indicating accelerating or steering to the right of the sensor position.

Variable transformed with inverse hyperbolic sine (i.e., asinh-transformed) can be interpreted as introduced by Bellemare and Wichman (2020). That is, for an outcome variable y and its asinh-transformation $\tilde{y} = asinh(y) = log(y + \sqrt{y^2 + 1})$, suppose an estimated coefficient between x and $\tilde{y}$ as $\tilde{\beta}$ (e.g., $\tilde{y} = \alpha + \tilde{\beta} \cdot x + \epsilon$), then the association between x and y can be derived as $\beta = \tilde{\beta}\sqrt{1 + \frac{1}{y^2}}$. Using the maximum yaw rate as an example, crashed drivers had 0.106 higher asinh-transformed *max* yaw rate, equivalently $0.106 \times \sqrt{1 + \frac{1}{1.207^2}} = 0.138$ higher yaw rate on the original scale, where 1.207 is the mean yaw rate of traversals in UMA.

In UI, *sd* and *gradient* of yaw rate were also positively associated with crash occurrence. Drivers with crashes on UI were more likely to steer to the left. While in the UOPA group, drivers with crashes were more likely to accelerate towards the right.

Crash rates (Table 6.2): Crash rates for crashed drivers were only significantly associated with driving “abnormality” in the UI group. More specifically, although crashed or not-crashed drivers were not distinguishable by their variation on delta speed on UI, among drivers who crashed, higher crash rates were associated with higher variation in delta speed (20% in *sd* and 17% higher in *gradient*). Drivers with higher crash rates were also associated with higher variation on the lateral direction. They were more likely to accelerate or steer to the left and had on average 8% higher *sd* and 3% higher *gradient* in lateral acceleration and 9% higher *gradient* in yaw rate.

Near-crashed or not (Table 6.3): Similar to models for crashed or not, we observe more significant associations between near-crash occurrence and “abnormality” features in the UMA group. Again, drivers with near-crashes were associated with higher variation in delta speed (12% higher *sd* and 9% higher *gradient*) and higher maximum longitudinal acceleration (0.074 m/s^2 higher). In addition, near-crashed drivers also had 7% higher *sd* of longitudinal acceleration.

Max and *sd* of lateral acceleration were both higher for drivers with near-crashes. The fact that only *sd* but not *gradient* is significant may suggest that near-crashed drivers had a variation of the mean trend towards the right (given the positive *mean* and *maximum*) rather than a frequent fluctuation around the mean trend. For yaw rate, drivers with near-crashes were observed with both higher *sd* (9%) and higher *gradient* (4%), though with lower magnitude of influences compared to crashed or not.

On UI, drivers with near-crashes were associated with positive *mean* and *max* lateral acceleration, indicating they were accelerating to the right. These drivers were also associated with higher *sd* and *gradient* of yaw rate, 6% and 4% respectively, and a negative *min* yaw rate, indicating a larger steering towards the left. The negative *min* yaw rate does not necessarily contradict the positive *mean* and *max* lateral acceleration. This will be explained in the Discussion section.

Near-crash rates (Table 6.4): Also similar to models for crash rates, significant association between near-crash rates and "abnormality" features for near-crashed drivers are only observed in the UI group. The difference is that, nevertheless, drivers with higher near-crash rates were more distinguishable on their longitudinal driving performance than drivers with higher crash rates, who were more distinguishable on their lateral driving performance. Still, near-crashed drivers were associated with 9% higher *sd* and 8% higher *gradient* of delta speed. They also had on average 0.033 m/s^2 higher *max* and 5% and 2% higher *sd* and *gradient* in longitudinal acceleration. On the lateral direction, drivers with higher near-crash rates were more likely to accelerate to the right.

Table 6.1: Total effects of crashed or not on "abnormality" features

Variable	UI ($N = 1,134$) Est.	UI Std. Est.	UI P-value	UOPA ($N = 821$) Est.	UOPA Std. Est.	UOPA P-value	UMA ($N = 539$) Est.	UMA Std. Est.	UMA P-value
"Abnormality" features on delta speed (km/h)									
Mean	-0.103	-0.008	0.959	1.070	0.092	0.210	-0.839	-0.076	0.209
Log(sd)	0.028	0.031	0.901	0.013	0.016	0.741	***0.127***	***0.160***	***0.002***
Log(grad)	0.024	0.032	0.901	0.010	0.017	0.741	***0.096***	***0.168***	***0.002***
"Abnormality" features on longitudinal acceleration (m/s^2)									
Mean	-0.001	-0.003	0.959	-0.011	-0.035	0.551	***0.051***	***0.164***	***0.002***
Max	0.005	0.016	0.901	-0.027*	-0.064*	0.389*	***0.078***	***0.178***	***0.002***
Min	-0.008	-0.020	0.901	0.002	0.004	0.923	0.033	0.059	0.292
Log(sd)	0.012	0.028	0.901	-0.018	-0.037	0.551	*0.046*	*0.092*	*0.098*
Log(grad)	0.003	0.010	0.901	-0.016	-0.048	0.221	0.006	0.018	0.668
"Abnormality" features on lateral acceleration (m/s^2)									
Mean	0.004	0.015	0.901	***0.025***	***0.130***	***0.011***	0.007	0.035	0.526
Max	0.007	0.021	0.901	0.011	0.055	0.439	***0.063***	***0.168***	***0.004***
Min	0.001	0.002	0.959	***0.046***	***0.135***	***0.011***	*-0.034*	*-0.094*	*0.097*
Log(sd)	0.008	0.022	0.901	-0.022	-0.053	0.439	***0.070***	***0.174***	***0.002***
Log(grad)	-0.005	-0.019	0.901	-0.023	-0.065	0.210	*0.034*	*0.096*	*0.062*
"Abnormality" features on yaw rate (deg/sec)									
Mean	*-0.048*	*-0.086*	*0.078*	-0.009†	-0.024†	0.741†	0.022	0.039	0.513
Max	-0.003	-0.004	0.959	0.009†	0.018†	0.741†	***0.106***†	***0.192***†	***0.001***†
Min	***-0.109***	***-0.126***	***0.007***	-0.024†	-0.051†	0.450†	*-0.058*†	*-0.114*†	*0.070*†
Log(sd)	***0.045***	***0.099***	***0.040***	0.023	0.046	0.453	***0.121***	***0.223***	**< 0.001**
Log(grad)	***0.032***	***0.084***	***0.016***	0.027	0.066	0.210	***0.057***	***0.152***	***0.001***
Robust Chi-square:	$\chi^2_{142} = 243.706$ (p< 0.001)			$\chi^2_{143} = 264.928$ (p< 0.001)			$\chi^2_{144} = 292.377$ (p< 0.001)		
Robust RMSEA:	0.025 (90% CI:[0.020, 0.030], p= 1)			0.032 (90% CI:[0.026, 0.038], p= 1)			0.044 (90% CI:[0.037, 0.051], p= 0.922)		
Robust CFI:	0.986			0.980			0.960		
Robust SRMR:	0.026			0.028			0.039		

*: related variable log-transformed
†: related variable asinh-transformed

Table 6.2: Total effects of crash rate on “abnormality” features for crashed drivers

Variable	UI ($N = 390$)			UOPA ($N = 327$)			UMA ($N = 217$)		
	Est.	Std. Est.	P-value	Est.	Std. Est.	P-value	Est.	Std. Est.	P-value
“Abnormality” features on delta speed (km/h)									
Mean	-0.623	-0.037	0.619	-1.349	-0.095	0.218	0.112	0.008	0.891
Log(sd)	***0.180***	***0.152***	***0.013***	0.116	0.113	0.167	0.060	0.054	0.544
Log(grad)	***0.154***	***0.159***	***0.013***	0.092	0.123	0.167	0.044	0.058	0.544
“Abnormality” features on longitudinal acceleration (m/s^2)									
Mean	-0.002	-0.007	0.886	-0.038	-0.094	0.218	0.024	0.052	0.544
Max	0.005	0.012	0.867	-0.012*	-0.024*	0.851*	-0.011	-0.019	0.863
Min	-0.012	-0.022	0.754	-0.088	-0.116	0.192	0.075	0.087	0.486
Log(sd)	0.028	0.049	0.526	0.060	0.091	0.218	-0.037	-0.053	0.544
Log(grad)	-0.009	-0.021	0.642	0.015	0.036	0.598	-0.005	-0.012	0.863
“Abnormality” features on lateral acceleration (m/s^2)									
Mean	-0.017	-0.054	0.472	-0.018	-0.070	0.598	0.014	0.057	0.544
Max	0.026	0.062	0.430	-0.011	-0.044	0.652	0.065	0.128	0.314
Min	***-0.067***	***-0.152***	***0.013***	-0.026	-0.060	0.636	-0.029	-0.057	0.544
Log(sd)	***0.079***	***0.165***	***0.010***	0.006	0.011	0.977	0.066	0.120	0.314
Log(grad)	***0.026***	***0.080***	***0.029***	0.003	0.005	0.977	0.023	0.048	0.544
“Abnormality” features on yaw rate (deg/sec)									
Mean	-0.042	-0.056	0.472	0.028†	0.058†	0.598†	0.051	0.068	0.544
Max	0.044	0.041	0.564	0.032†	0.050†	0.636†	0.098†	0.139†	0.314†
Min	***-0.143***	***-0.122***	***0.043***	0.024†	0.038†	0.652†	-0.024†	-0.034†	0.701†
Log(sd)	***0.085***	***0.138***	***0.013***	0.001	0.002	0.977	0.081	0.115	0.314
Log(grad)	0.026	0.051	0.419	-0.002	-0.004	0.977	0.029	0.058	0.486
Robust Chi-square:	$\chi^2_{145} = 260.131$ (p< 0.001)			$\chi^2_{142} = 220.292$ (p< 0.001)			$\chi^2_{145} = 216.522$ (p< 0.001)		
Robust RMSEA:	0.045 (90% CI:[0.037, 0.053], p= 0.836)			0.041 (90% CI:[0.031, 0.051], p= 0.930)			0.048 (90% CI:[0.034, 0.060], p= 0.606)		
Robust CFI:	0.975			0.980			0.970		
Robust SRMR:	0.033			0.032			0.038		

*: related variable log-transformed
†: related variable asinh-transformed

Table 6.3: Total effects of near-crashed or not on "abnormality" features

Variable	UI ($N = 1,134$)			UOPA ($N = 821$)			UMA ($N = 539$)		
	Est.	Std. Est.	P-value	Est.	Std. Est.	P-value	Est.	Std. Est.	P-value
"Abnormality" features on delta speed (km/h)									
Mean	-0.441	-0.034	0.565	1.330	0.116	0.215	0.116	0.011	0.873
Log(sd)	0.018	0.020	0.657	-0.048	-0.060	0.278	***0.112***	***0.145***	***0.019***
Log(grad)	0.015	0.021	0.657	-0.038	-0.065	0.278	***0.084***	***0.152***	***0.019***
"Abnormality" features on longitudinal acceleration (m/s^2)									
Mean	0.003	0.013	0.744	0.019	0.061	0.278	0.027	0.089	0.219
Max	0.017	0.054	0.337	$0.028^{\star}$	$0.068^{\star}$	$0.278^{\star}$	***0.074***	***0.175***	***0.018***
Min	-0.012	-0.032	0.565	0.017	0.033	0.579	-0.016	-0.030	0.665
Log(sd)	0.024	0.056	0.287	0.001	0.003	0.946	***0.070***	***0.145***	***0.020***
Log(grad)	0.016	0.046	0.216	0.018	0.055	0.246	0.017	0.051	0.322
"Abnormality" features on lateral acceleration (m/s^2)									
Mean	***0.030***	***0.127***	***0.014***	0.011	0.057	0.285	*0.022*	*0.116*	*0.074*
Max	***0.048***	***0.140***	***0.007***	0.019	0.100	0.246	***0.060***	***0.163***	***0.018***
Min	0.016	0.048	0.402	-0.005	-0.016	0.769	0.003	0.009	0.873
Log(sd)	0.030	0.081	0.165	0.032	0.079	0.278	*0.040*	*0.104*	*0.086*
Log(grad)	0.007	0.025	0.565	0.019	0.053	0.278	0.018	0.054	0.322
"Abnormality" features on yaw rate (deg/sec)									
Mean	-0.011	-0.021	0.657	$0.029^{\dagger}$	$0.080^{\dagger}$	$0.268^{\dagger}$	0.034	0.061	0.322
Max	0.056	0.071	0.213	$0.051^{\dagger}$	$0.107^{\dagger}$	$0.215^{\dagger}$	***0.092***†	***0.171***†	***0.018***†
Min	***-0.087***	***-0.100***	***0.029***	$0.008^{\dagger}$	$0.018^{\dagger}$	$0.769^{\dagger}$	$-0.031^{\dagger}$	$-0.062^{\dagger}$	$0.322^{\dagger}$
Log(sd)	***0.061***	***0.134***	***0.007***	0.028	0.057	0.285	***0.088***	***0.167***	***0.018***
Log(grad)	***0.037***	***0.095***	***0.007***	0.011	0.028	0.579	***0.043***	***0.118***	***0.018***
Robust Chi-square:	$\chi^2_{142} = 240.230$ (p< 0.001)			$\chi^2_{144} = 264.118$ (p< 0.001)			$\chi^2_{145} = 315.567$ (p< 0.001)		
Robust RMSEA:	0.025 (90% CI:[0.019, 0.030], p= 1)			0.032 (90% CI:[0.026, 0.038], p= 1)			0.047 (90% CI:[0.040, 0.054], p= 0.769)		
Robust CFI:	0.987			0.980			0.955		
Robust SRMR:	0.025			0.029			0.040		

*: related variable log-transformed

†: related variable asinh-transformed

Table 6.4: Total effects of near-crash rate on "abnormality" features for near-crashed drivers

Variable	UI ($N = 882$)			UOPA ($N = 556$)			UMA ($N = 355$)		
	Est.	Std. Est.	P-value	Est.	Std. Est.	P-value	Est.	Std. Est.	P-value
"Abnormality" features on delta speed (km/h)									
Mean	-0.106	-0.007	0.926	-0.226	-0.016	0.926	0.523	0.038	0.974
Log(sd)	***0.089***	***0.085***	***0.020***	0.043	0.045	0.702	-0.052	-0.050	0.974
Log(grad)	***0.075***	***0.086***	***0.020***	0.034	0.048	0.702	-0.041	-0.053	0.974
"Abnormality" features on longitudinal acceleration (m/s^2)									
Mean	0.007	0.026	0.589	-0.024	-0.063	0.702	0.025	0.059	0.974
Max	***0.033***	***0.092***	***0.020***	0.010*	0.021†	0.926†	0.005	0.008	0.974
Min	-0.024	-0.054	0.143	-0.069	-0.110	0.164	0.065	0.083	0.974
Log(sd)	***0.054***	***0.108***	***0.003***	0.056	0.093	0.164	-0.041	-0.062	0.974
Log(grad)	***0.024***	***0.057***	***0.023***	0.014	0.035	0.702	-0.005	-0.011	0.974
"Abnormality" features on lateral acceleration (m/s^2)									
Mean	***0.035***	***0.124***	***0.002***	-0.004	-0.019	0.926	-0.009	-0.038	0.974
Max	***0.049***	***0.123***	***0.003***	-0.003	-0.012	0.926	-0.003	-0.005	0.974
Min	*0.026*	*0.064*	*0.076*	-0.006	-0.014	0.926	-0.020	-0.044	0.974
Log(sd)	0.021	0.048	0.221	0.000	0.000	0.997	0.012	0.023	0.974
Log(grad)	*0.015*	*0.047*	*0.076*	0.010	0.025	0.926	0.004	0.010	0.974
"Abnormality" features on yaw rate (deg/sec)									
Mean	0.008	0.013	0.815	-0.003†	-0.006†	0.933†	-0.002	-0.003	0.974
Max	0.022	0.023	0.613	-0.011†	-0.020†	0.926†	0.003†	0.005†	0.974†
Min	-0.004	-0.004	0.926	0.005†	0.009†	0.926†	-0.006†	-0.009†	0.974†
Log(sd)	0.012	0.023	0.613	-0.012	-0.021	0.926	0.007	0.010	0.974
Log(grad)	0.001	0.002	0.926	-0.011	-0.022	0.926	0.001	0.001	0.974
Robust Chi-square:	$\chi^2_{141} = 380.051$ ($p < 0.001$)			$\chi^2_{142} = 336.944$ ($p < 0.001$)			$\chi^2_{144} = 282.800$ ($p < 0.001$)		
Robust RMSEA:	0.044 (90% CI:[0.039, 0.049], p= 0.978)			0.050 (90% CI:[0.043, 0.056], p= 0.520)			0.052 (90% CI:[0.045, 0.060], p= 0.329)		
Robust CFI:	0.979			0.971			0.963		
Robust SRMR:	0.030			0.031			0.052		

*: related variable log-transformed

†: related variable asinh-transformed

6.4 Discussion

Models of Aim 3 examine the relationship between driver safety and specific aspects of driving "abnormality". Consistent results were observed in Aim 3 as in Aim 2 but in a more informative way thanks to the use of "abnormality" features. In summary, there exist multiple relationships between the crash or near-crash status of drivers and different aspects of their driving "abnormality". The associations is more likely observed for driving performance on the lateral direction than the longitudinal direction and for features measuring the variation than the mean trend of an "abnormality" curve.

One might be suspicious of the revealed association because of the small effect size. Nevertheless, I would argue that the effect size is small but not trivial. First of all, all significant estimated associations – although small – are different from 0 based on p-values already adjusted for inflated Type I error. Secondly, the standardized estimate of significant associations, which can be treated as a Cohen's d, are from 0.1 to 0.2 in most cases. An effect size of $d = 0.2$ is smaller than a medium effect that is "visible to the naked eye of a careful observer" but not "so small as to be trivial" (Sullivan and Feinn, 2012). In fact, the small effect size may be expected. After all, the traversals were only 10-second long to ensure that environmental conditions were consistently captured. Also, all traversals were not involved in any type of traffic conflicts, were on a straight road with low traffic density (on UI) or no intersection influence (on UOPA and UMA), and were not turning or changing lane.

Impacts of the driving environment

The association between driver safety level and their driving "abnormality" depends on the driving environment. More specifically:

- Drivers who crashed or near-crashed were most distinguishable by their driving "abnormality" on UMA.
- Among drivers who crashed or near-crashed, those with higher crash or near-crash rates were most distinguishable by their driving "abnormality" on UI.

- Drivers with varied crash or near-crash status were almost not distinguishable by their driving "abnormality" on UOPA.

The distinct observations may stem from the different characteristics of the three functional classes. Urban Interstates are limited access, divided highways that are designed and constructed to maximize the mobility of users. The purpose of the Interstate is not to serve abutting land uses, so it rarely has at-grade intersections (FHWA, 2017). The driving environment on Urban Interstate is relatively simple in the sense that it consists only motor vehicles driving in the same direction and with similar speeds. Drivers in this driving environment thus have less pressure to deal with the surrounding traffic and are allowed to have larger variations on the longitudinal direction. This could be especially true given the lower traffic density in our analyzed situation ("LOS A1" to "LOS C" of the UI group).

On the other hand, Urban Minor Arterial serves to connect smaller geographic areas to the higher Arterial systems. They are undivided, uncontrolled and have at-grade intersections to access local land uses (FHWA, 2017). The driving environment on Urban Minor Arterial may consist of different types of road users moving in different directions and various traffic rules to follow. The environment itself is likely more complex than Urban Interstate with trees, buildings and parked cars on the roadside. Drivers on the Urban Minor Arterial thus need to deal with a more complex driving environment, meantime have less variability on their longitudinal driving performance than those on the Urban Interstate.

In the last chapter, it has been discussed that the different associations as regards the longitudinal and lateral performance may indicate different underlying reasons for those who crashed or near-crashed and those with higher crash or near-crash rates. Interestingly, the different associations between UI and UMA observed in this chapter suggest a similar conclusion. That is, drivers crashed or near-crashed are more likely due to their inexperience, less skillful and/or impaired ability to deal with a complex driving environment such as in UMA. Meanwhile, drivers with higher crash or near-crash rates are more likely due to their adverse driving tendencies, e.g., risky or distracted driving, which can be more easily observed on Urban Interstate.

Urban Other Principle Arterial is in between. On the one hand, it also aims to provide high degree of mobility and does not serve local land uses as much as Urban Minor Arterial. On the other hand, it is not necessarily access-controlled, could be undivided, and may have driveways or at-grade intersections. Since the association between driver safety level and driving "abnormality" shows different trend on Urban Interstate and Urban Minor Arterial, their effects may have canceled out on Urban Other Principle Arterial, resulted in the observed null effect.

Possible implied "abnormal" behaviors

This subsection will discuss the possible implied driving behaviors based on the identified significant "abnormality" features. The set of statistical features for a specific kinematic are correlated and features of different kinematics may also correlate, especially when both measuring the longitudinal or lateral performance. The various significant associations, when observed simultaneously within the same model, may thus indicate certain types of driving behavior. The discussion is to provide some concrete examples of how the "abnormality" features could be connected to explain the association between drivers safety and driving "abnormality". It should be noted that, the implied behaviors are just possible behaviors. It is not and should not be interpreted as the exact or the unique "abnormal" behavior observed in the data.

Drivers with crashes on UI steered to the left and right more frequently than those without crashes (higher *sd* and *gradient* of yaw rate; see Table 6.1) and were in general more likely to steer to the left. The magnitude of steering may be low since we observe no significant difference between the two types of drivers on their lateral acceleration. This difference, nevertheless, was observed for drivers with higher crash rates on UI. These drivers had significantly higher *sd* but not *gradient* on yaw rate (see Table 6.2). Noted that, although both *sd* and *gradient* capture the variation of a kinematic curve, the *sd* primarily accounts for the variation in the mean trend of the curve while the *gradient* accounts for the variation, or more precisely the fluctuation, around the mean trend of the curve. The observation thus indicates an intentional change of the yaw rate back and forth. The change was more sharply to the left than to the right (higher absolute value of *min* than *max* yaw rate; see Table 6.2) and sufficiently large to reflect on the lateral acceleration with a similar trend. This behavior of the group of drivers with higher crash rates may be depicted as trying to overtake the lead vehicle. The higher *sd* and *gradient* on delta speed could be due to speeding up while trying to overtake. But since no lane-changing traversals were included in the analysis, the magnitude of acceleration might not be large enough to show a difference. Back to the crashed or not situation, the reduced effect on *sd* and *gradient* but not *min* of yaw rate might suggests that drivers with crashes were consisted of two types: one group of drivers who more frequently try to overtake and another group of drivers who were incapable to maintain a stable steering wheel control.

Drivers with near-crashes on UI reveal an interesting trend on their lateral performance. More specifically, similar to drivers with crashes or higher crash rates on UI, we observe higher variation of yaw rate and a sharper steering towards the left. Unlike the crash status though, there is no significant difference in the variation but a significant positive mean trend of the lateral acceleration, indicating that the drivers were accelerating towards the right. The possible behavior thus cannot be trying to overtake: it does not explain why steering and lateral accelerating were in opposite directions. Rather, it is more likely that drivers steered to the left noticing their vehicles were drifting towards the right, e.g., due to being distracted from the driving tasks. Driver's seat mostly locates on the left in the U.S., so when distracted they would more likely lean towards the right. For example, Bao et al. found that drivers engaged in visual-manually tasks deviated more to the right side than the left side of the center of the lane (Bao et al., 2015). A similar positive lateral acceleration in the mean trend was observed for drivers with higher near-crash rates on UI. Nevertheless, there is no significant difference in terms of yaw rate. Since traversals analyzed in the analysis are only 10-second long, it is possible that a steering behavior is yet captured. If this is the case, the longer reaction time in steering vehicle back to a center position may explain why the group of drivers were with higher near-crash rates. Drivers with higher near-crash rates also had higher variations of the longitudinal acceleration and the delta speed. Both means are no different from zero, which might suggest an attempt to stay close to a "normal" driving speed.

Drivers with crashes on UMA steered to the left and right more frequently than those without crashes (higher *sd* and *gradient* in yaw rate; see Table 6.1). Magnitude of the effect of *sd* is noticeably larger than *gradient*, which may suggest that the variation caused not only by the incapability to hold the steering wheel stably but also an intentional rotate of the steering wheel. That said, the two causes were not necessarily due to the same group of drivers. The intentional steering was more sharply towards the right than to the left, though the mean yaw rate remained no significantly different. The similar trend was also observed for lateral acceleration. Drivers with crashes on UMA also had higher variation on delta speed. The longitudinal acceleration features are all positive, suggesting the higher variation of delta speed was due to speeding up. It should be noted that, though insignificant, the mean delta speed is negative. This might indicate that the accelerating behavior was to catch up the "normal" driving speed of surrounding vehicles rather than speeding. Similar relationships on driving "abnormality" were also observed for drivers with near-crashes on UMA. The relationship between these implied "abnormal" behaviors and driver safety on UMA is less clear than those on UI, especially for the intentional steering to the right. One possible explanation for the tendency to be closer to the right is to compensate an increased perceived risk e.g., due to engaging in a secondary task, when driving on an undivided road with oncoming traffic in the adjacent lane.

One might be concerned that the significant associations between "abnormality" features and driver safety were due to environmental factors rather than "abnormal" behaviors. For example, drivers with near-crashes on UI steered to the left given their vehicles drifted to the right not because they were distracted but because of the pavement cross slope. However, this associations may be more likely due to drivers' behavioral factors rather than environmental factors. The reason is twofold. Firstly, all drivers on the Urban Interstate regardless of their safety status should experience similar pavement cross slopes, which were 1.5 to 2% on a paved road (AASHTO, 2018) . Thus, the impact of cross slope on driving performance would be accounted for by the “normal” driving style under a specific environmental context rather than the “abnormal” driving style. Secondly, even the different traversals were associated with different pavement cross slopes, the impact of this factor on the relationship between driving “abnormality” and driver safety could be safely ignored, at least when analyzing within the same environmental group in the context of this study. This has been discussed in the discussion section in Chapter 4, where more details are provided on why the set-up of environmental groups helps block the potential confounding effects of roadway and environmental factors on the relationship between driver safety and driving “abnormality”. That said, if an environmental factor is of particular interest, it can be included in the model to evaluate its impact on driving “abnormality” as well as its potential moderating effect on other relationships.

Chapter 7

GENERAL CONCLUSIONS

This chapter summarizes the overall findings of the dissertation, the theoretical implications of the findings and contribution and publications. Limitations and future research will also be discussed.

7.1 Overall Findings

The overall objective of this dissertation was to assess driver behavior in the context of driving environment. The SHRP2 NDS data was used to achieve the goal of this dissertation. In order to quantify and assess driver behavior presented in a time-series manner, an "abnormal" driving style was proposed as a complement to a "normal" driving style. The level of "abnormality" thus measures how much an individual driver behavior deviates from the average of all drivers under a specific environmental condition. A clustering and classification method was conducted to identify major environmental factors that contribute to the formation of "normal" driving styles. The selected environmental factors were used to set up environmental groups, where within each group a unique "normal" driving style was assumed. The level of "abnormality" for each individual traversal was quantified by an "abnormality" score. A set of statistical features were also established to capture the different aspects of an "abnormal" driving style. The relationship between "abnormality" and safety was studied based on drivers' crash or near-crash status. Potential effects of the driving environment on the relationship were also examined. The key findings of the dissertation are summarized below:

- *Environmental factors and the "normal" driving style*: *Traffic density* and *alignment* were identified as the most important environmental factors for "normal" driving styles on Urban Interstate. *Traffic density* was split at "LOS D" given that this level and higher was considered "congested". *Intersection influence* and *alignment* was important for Urban Other Principal Arterial and Urban Minor Arterial. For *intersection influence*, "Yes, traffic signal" and "Yes, uncontrolled" were combined and examined separately from "No" (and "Yes, interchange" for UOPA), suggesting an influence of the intersection on driver behavior regardless of its control type.

- *Levels of "abnormality" and driver safety*: Drivers who had a crash or near-crash were more likely to be associated with higher "abnormality" score, mostly for their driving performance in the UI group. These drivers were more distinguishable by their level of "abnormality" on the lateral driving performance when compared to those who did not have any safety critical events (crash, near crash). Among those who have had safety critical events, those with higher crash rates or near-crash rates were more distinguishable by their longitudinal driving performance.

- *Driver demographic and driving "abnormality"*: Age shows a larger and more consistent effect on the level of driving "abnormality" than gender across all environmental groups. More specifically, level of "abnormality" on delta speed decreases monotonically with increased age, i.e., older drivers tend to choose a speed more similar to the surrounding vehicles. On the other hand, level of "abnormality" on lateral acceleration were higher for both younger and older drivers compared to middle-aged drivers.

- *Driver demographic and driver safety*: Younger and older drivers were more likely to crash and have higher crash rates compared to middle-aged drivers. On the contrary, drivers of an older age were less likely to have a near-crash and had lower near-crash rates.

- *Aspects of "abnormality" and driver safety*: The association between driving "abnormality" and driver safety is more likely observed on the lateral than the longitudinal direction and for features that capture variation in the kinematic measures than the mean trend.

- *Possible "abnormal" behaviors*: The different aspects of "abnormality" associated with driver safety, when examined in groups, reveal possible "abnormal" behaviors. For example, the results depicted a possible behavior of drivers with higher crash rates on UI as frequently trying to overtake while drivers who crashed were likely just incapable to maintain a stable steering wheel control. A possible behavior of drivers who near-crashed on UI is distracted driving. This behavior is also probable for drivers with higher near-crash rates on UI, who seems to engage in the distracted tasks more intensely and for a longer time.

- *Impact of the driving environment*: The association between levels of "abnormality" (i.e., score) and driver safety were most likely observed on UI, followed by UMA and not at all on UOPA. As regards the aspects of "abnormality" (i.e., statistical features), the association with crashed/near-crashed or not were more likely observed on UMA while the association with higher crash/near-crash rates were more likely observed on UI. Still, almost no associations were observed on UOPA.

7.2 Theoretical Implications

This dissertation assess driver behavior based on "abnormal" driving style. The "abnormal" driving style does not assume *a priori* what is a "risky" behavior. Rather, the revealed associations between driver safety and driving "abnormality", either measured by a composite score or by a set of statistical features, appear to suggest two types of reasons for crash or near-crash. One type of reason relates to the insufficient driving ability or relatively poor driver skill, which could be due to insufficient driving experience or impaired ability as a consequence of e.g., aging. This type of reason more likely explain why a driver crashed or near-crashed. The other type of reason relates to the tendency towards adverse driving behaviors, such as frequently tying to overtake or distracted driving. This type of reason more likely explain the higher crash or near-crash rates among those drivers who crashed or near-crashed.

The dissertation also emphasizes the important role of driving environment in driver behavior assessment. The impact is twofold: it not only affects how people normally drive thus what defines an "abnormal" driving style, but also the specific qualities of drivers when interacting with an environment. The former is straight-forward, e.g., assessment of speed selection may want to consider the functional class, the posted speed limit as well as roadway alignment and traffic density. The latter is not so obvious, but as the dissertation shows, the driving environment affects our ability to identify accident-prone drivers. The different reasons for crash or near-crash mentioned in the last paragraph are also relevant. For example, driving environment on UMA is considered more complex than on UI thus may require more sophisticated driver skill to deal with. Drivers with insufficient driver skill or driving ability could be more easily identified on UMA. Meanwhile, the simpler driving environment on UI shrinks the difference due to driver skill and may thus become more responsive to the difference caused by adverse behaviors.

7.3 Practical Implications

This dissertation evaluates the approach to assess driver behavior based on "abnormal" driving styles. When considered from different perspectives – either a between-subject perspective to compare the level of "abnormality" among different drivers, or a within-subject perspective to compare the level of "abnormality" for a single driver – the driving "abnormality" could provide insights to different stakeholders.

The between-subject difference in driving "abnormality" identifies drivers of higher safety concerns. The information could be used by state or local transportation agencies to pick out accident-prone drivers even before they crash and to tailor education programs accordingly to improve their safety. The insurance company can use the driving performance data (in addition to the demographics and historical driving records) to setup or update the insurance premium based on a customer's level of driving riskiness. For example, higher deviations on the lateral direction usually concerns a risky driver with higher likelihood to crash or near-crash. When combined with information on the longitudinal direction, a more complete picture of the driver can be obtained such that higher deviations in the longitudinal direction would indicate an even higher risk level. The detection can be applied to everyday driving data but should account for the driving environment, which would make it less biased and more cost effective. For example, the algorithm can be selected to examine driving performance on Urban Interstate or Urban Minor Arterial but would not need to be activated for Urban Other Principal Arterial.

From a within-subject perspective, the driving “abnormality” describes an individual’s driver behavior. This information can be of benefit to Original Equipment Manufacturers (OEMs) who can improve the design of an ADAS or ensure that their ADS is more user-centric. For example, an ADS could better mimic an individual driving style but also be modified to increase overall safety. There are also practical implications in terms of the within-subject context. For example, if a driver exhibits higher “abnormality” of driving on UI but not on UMA, the driver may be more likely to crash or near-crash due to adverse behaviors rather than insufficient driving skills. Thus, the in-vehicle system can be designed to monitor the individual’s driving performance and to warn them of risky behavior engagement. Alternatively, for drivers with insufficient or impaired driving skills, the system may provide more instructional feedback and assist in monitoring the surrounding areas that might be less likely to be attended to by an inattentive driver.

The within-subject difference could also account for the different trips within the same driver. For example, a normally safe driver may exhibit higher driving “abnormality” in a specific trip. This could be due to an unusual condition of the driver or an unauthorized driver (a theft). An in-vehicle system could use the information to either assist or warn the driver in an unusual situation or in the latter case, to inform the car owner of a potential theft event. This application, nevertheless, will require repeated-measures of the driving performance so the system can learn the usual behavior of a target driver, including the average and the variation of the performance. A threshold on the variation can be set up based on the specific driver's data (to better represent an individual driving style) or on all drivers' data (to represent the "normal" behavior of a larger population). The current conceptual framework and analytical methods provide the groundwork to address these issues.

7.4 Contributions and Publications

Driving, as an interactive process between the driver and the environment, is contextual-dependent. Nevertheless, traditional ways to assess the driver behavior oftentimes didn't take into account the environmental context or the environment factors were selected arbitrarily. The capability to identify environmental factors that contribute to a driver's "normal" driving style will not only ensure a less biased evaluation of their driving performance but also help the findings to be applied more precisely in a real-world setting. The increasingly accessible naturalistic driving data provides researchers an opportunity to analyze driver behavior in the context of driving environment. Meanwhile, the nature of naturalistic data requires specific analytical methods to deal with the auto-correlation in a data series and the correlation among vehicle kinematics. The dissertation demonstrates a new way to analyze the impact of driving environment on driver behavior based on functional data analysis method. More specifically, a GMM clustering was applied on MFPCA scores of four vehicle kinematics followed by a decision tree to identify the environmental factors that contributes to "normal" driving styles of drivers. The method finds *traffic density* and *alignment* the two most influential environmental factors on Urban Interstate, and *intersection influence* and *alignment* on Urban Other Principal Arterial and Urban Minor Arterial. Additional data will allow additional environmental factors to be identified. This work will be submitted to *Transportation Research Part B: Methodological.*

A real-world driving behavior involves the driver, the environment and the vehicle in a closed-loop system. On the one hand, it suggests the important role of a driving environment when assessing a driver behavior. On the other hand, it indicates that a driver's behavior would ultimately, at least to a large extent, reflect on the vehicle's kinematics. This dissertation serves the goal to demonstrate an approach to assess driver behavior based on vehicle kinematics in the context of driving environment. Although naturalistic driving data enlarges the number of crashes or near-crashes captured, the majority of times drivers are not involved in any types of traffic conflicts. A driver's habitual driving behavior in the everyday driving scenario – an individual driving style – could be considered as an "abnormal" driving style to a certain degree when compared against the average driving behavior of all drivers – the "normal" driving style. This dissertation hypothesized that an individual driver's "abnormal" driving style would have an association with the driver's safety level measured by crash or near-crash status. The dissertation thus demonstrates the ability to access driver behavior and to infer the safety level using everyday driving data. This approach is expected to broaden the data source for driver behavior assessment and be able to capture more realistic representations of individual driving styles. This paper will be submitted to *Accident Analysis & Prevention*.

Previous studies assessing driver behavior with naturalistic driving data usually involve the detection of "risky" behaviors. Nevertheless, a pre-determined set of "risky" behaviors is very likely incomplete and any thresholds that define "risky" behaviors are subjective and oftentimes broadly applied regardless of the driving environment. This dissertation assessed driver behavior based on the "abnormality" associated with an "abnormal" driving style. Through this means, driver behavior was assessed directly based on the vehicle kinematics without having to detect "risky" behaviors beforehand. The "abnormality" metric is expected to capture all types of adverse behaviors reflected by the vehicle kinematics. This is considered an advantage because of its broader coverage and potentially less biased evaluation results due to the removal of subjectivity as regards the "risky" behavior. Revealed associations in the dissertation suggest multiple possible behaviors that are frequently used as "risky" behaviors in previous studies. Meanwhile, certain so-called "risky" behaviors are not necessarily as risky as we may expect. For example, although speeding is frequently considered a "risky" behavior, our results suggest that how much a driver exceeds the posted speed limit on average (i.e., mean delta speed) is not significantly associated with the driver's safety level in any analyzed cases. Variation on this excess could indicate a more risky driver but need to consider the driving environment. This paper will be submitted to *Transportation Research Part F: Traffic Psychology and Behaviour*.

7.5 Limitations and Future Research

One limitation of the study is that each analyzed traversal is only 10-second long. The baselines identified in the SHRP2 NDS database are originally short (21-second long) and were further shortened to ensure that the roadway and environmental conditions were coded consistently with the actual driving environment. Nevertheless, the short duration might not capture the typical driving style of an individual driver. This is the reason why results in the dissertation were in general concluded for group of drivers, e.g., with or without crashes, rather than for individual drivers. In addition, a 10-second time period may not be able capture the entire process of a driver behavior. For example, Chapter 6 observed that drivers with higher near-crash rates were associated with significantly higher lateral acceleration (i.e., to the right) but no significant difference on the yaw rate, which might be due to the steering behavior was yet observed. Longer traversals will allow this type of behavior to be fully observed and more precisely analyzed.

The second limitation relates to the baselines in SHRP2. Baselines in SHRP2 were selected to represent "normal driving" and "typical driver behavior" (Hankey et al., 2016). It rules out any crashes or near-crashes but also less severe events that exceeded pre-determined thresholds on vehicle kinematics yet failed the criteria of being considered as crashes or near-crashes. Baselines in SHRP2 thus have naturally exclude, at least to some extent, many "risky" behaviors, such as hard brake. This may partly explain the small effect size of identified associations. Inclusion of potential "risky" behaviors will allow larger effects on an "abnormal" driving style and additional "abnormal" behaviors to be examined.

Small sample size also limited more driving environment to be examined. Although twelve driving environmental groups were set up based on identified environmental factors in Aim 1, only three of them have sufficient data to be analyzed in the subsequent analysis. Because of this, comparison of the relationship between driver safety and driving "abnormality" is mostly a comparison across functional classes. Additional data will allow this relationship to be examined across different traffic density levels, alignment levels, or even other less important environmental factors that were yet identified using this data.

Future research may want to test the current approach on longer traversals and to cover a broader types of driving environments. Repeated measures on drivers could also be considered so an individual driving style is more precisely captured and the variation within the same driver can be accounted for. This dissertation measures driver safety level by crash records regardless of the severity level. However, focus could be given to drivers of different safety levels and/or crashed due to different reasons. This will shed light on the range of effect size for various aspects of driving “abnormality”. Moreover, it would help to establish a more comprehensive understanding of the different types of driver-related factors for crashes and near-crashes, which could be used to target different goals for driver safety.

The use of driving “abnormality” is to bypass “risky” behavior so that a more general approach can be established to assess driver behavior and to evaluate how risky a driver is. Thus, any implied “abnormal” behaviors suggested by the current approach are indicative of a “risky” driver but may or may not themselves be “risky” (in the sense that the behavior will lead to a crash or near-crash). Future research, on the one hand, may want to apply the method on everyday driving data with more variations on vehicle kinematics than the currently used baseline data to allow more types of “abnormal” behaviors to be identified. On the other hand, it is useful to know how the various aspects of driving “abnormality” relates to the different types of “risky” behaviors. This knowledge could serve to detect and to reduce the engagement of certain behaviors and help “risky” drivers drive more safely.

BIBLIOGRAPHY

AASHTO (2018). American Association of State Highway and Transportation Officials, Washington, DC.

Ali, E. M., Ahmed, M. M., and Yang, G. (2020). Normal and risky driving patterns identification in clear and rainy weather on freeway segments using vehicle kinematics trajectories and time series cluster analysis. *IATSS Research*, 45(1):137–152.

Aljaafreh, A., Alshabatat, N., and Najim Al-Din, M. S. (2012). Driving style recognition using fuzzy logic. *2012 IEEE International Conference on Vehicular Electronics and Safety, ICVES 2012*, pages 460–463.

Andrews, E. C. and Westerman, S. J. (2012). Age differences in simulated driving performance: Compensatory processes. *Accident Analysis & Prevention*, 45:660–668.

Arnett, J. (1990). Drunk driving, sensation seeking, and egocentrism among adolescents. *Personality and Individual Differences*, 11(6):541–546.

Arthur Jr, W. and Graziano, W. G. (1996). The five-factor model, conscientiousness, and driving accident involvement. *Journal of Personality*, 64(3):593–618.

Bao, S., Guo, Z., Flannagan, C., Sullivan, J., Sayer, J. R., and LeBlanc, D. (2015). Distracted driving performance measures: Spectral power analysis. *Transportation Research Record*, 2518(1):68–72.

Bassani, M., Catani, L., Cirillo, C., and Mutani, G. (2016). Night-time and daytime operating speed distribution in urban arterials. *Transportation Research Part F: Traffic Psychology and Behaviour*, 42:56–69.

Bejani, M. M. and Ghatee, M. (2018). A context aware system for driving style evaluation by an ensemble learning on smartphone sensors data. *Transportation Research Part C: Emerging Technologies*, 89:303–320.

Bellemare, M. F. and Wichman, C. J. (2020). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, 82(1):50–61.

Bergasa, L. M., Almeria, D., Almazan, J., Yebes, J. J., and Arroyo, R. (2014). DriveSafe: An app for alerting inattentive drivers and scoring driving behaviors. In *IEEE Intelligent Vehicles Symposium, Proceedings*, pages 240–245.

Blomberg, R. D., Peck, R. C., Moskowitz, H., Burns, M., and Fiorentino, D. (2005). Crash risk of alcohol involved driving: A case-control study. Technical report.

Brewer, R. D., Morris, P. D., Cole, T. B., Watkins, S., Patetta, M. J., and Popkin, C. (1994). The risk of dying in alcohol-related automobile crashes among habitual drunk drivers. *New England Journal of Medicine*, 331(8):513–517.

Brown, J., He, Y., and Lang, H. (2020). Quantifying drivers' driving skills using closed-loop dynamic simulations of articulated heavy vehicles. *Simulation Modelling Practice and Theory*, 99:102014.

Browne, M. W. (1993). Alternative ways of assessing model fit. *Testing structural equation models*, 21:230–258.

Castignani, G., Derrmann, T., Frank, R., and Engel, T. (2015). Driver behavior profiling using smartphones: A low-cost platform for driver monitoring. *IEEE Intelligent Transportation Systems Magazine*, 7(1):91–102.

Castignani, G., Frank, R., and Engel, T. (2013). An evaluation study of driver profiling fuzzy algorithms using smartphones. In *2013 21st IEEE International Conference on Network Protocols (ICNP)*, pages 1–6.

Chen, T., Sze, N., Newnam, S., and Bai, L. (2021). Effectiveness of the compensatory strategy adopted by older drivers: Difference between professional and non-professional drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, 77:168–180.

Choudhary, P. and Velaga, N. R. (2017). Analysis of vehicle-based lateral performance measures during distracted driving due to phone use. *Transportation Research Part F: Traffic Psychology and Behaviour*, 44:120–133.

Clarke, P. (2008). When can group level clustering be ignored? Multilevel models versus single-level models with sparse data. *Journal of Epidemiology & Community Health*, 62(8):752–758.

Constantinescu, Z., Marinoiu, C., and Vladoiu, M. (2010). Driving style analysis using data mining techniques the gipix system for vehicle tracking. *International Journal of Computers, Communications & Control*, V(5):654–663.

Dai, J., Teng, J., Bai, X., Shen, Z., and Xuan, D. (2010). Mobile phone based drunk driving detection. In *2010 4th International Conference on Pervasive Computing Technologies for Healthcare*, pages 1–8. IEEE.

Das, S., Sun, X., Wang, F., and Leboeuf, C. (2015). Estimating likelihood of future crashes for crash-prone drivers. *Journal of Traffic and Transportation Engineering (English edition)*, 2(3):145–157.

Deery, H. A. (1999). Hazard and risk perception among young novice drivers. *Journal of Safety Research*, 30(4):225–236.

Dozza, M., Flannagan, C. A., and Sayer, J. R. (2015). Real-world effects of using a phone while driving on lateral and longitudinal control of vehicles. *Journal of Safety Research*, 55:81–87.

Dula, C. S. (2003). *Validity and reliability assessment of a dangerous driving self-report measure*. PhD thesis, Virginia Tech.

Eren, H., Makinist, S., Akin, E., and Yilmaz, A. (2012). Estimating driving behavior by a smartphone. In *2012 IEEE Intelligent Vehicles Symposium*, pages 234–239.

FHWA (2010). Highway safety improvement program (hsip) manual. Technical report, Federal Highway Administration. FHWA-SA-09-029.

FHWA (2017). Road safety fundamentals: Concepts, strategies, and practices that reduce fatalities and injuries on the road. Technical report, Federal Highway Administration. FHWA-SA-18-003.

FHWA (2017). Section 3. criteria - highway functional classifications - statewide transportation planning - planning processes - fhwa. Available at: `https://www.fhwa.dot.gov/planning/processes/statewide/related/highway_functional_classifications/section03.cfm`. Accessed on July 7, 2021.

Fitzpatrick, K. (2003). Design speed, operating speed, and posted speed practices. Technical report. NCHRP REPORT 504.

Fraley, C. and Raftery, A. E. (2007). Bayesian regularization for normal mixture estimation. *Journal of Classification*, 24:155–181.

French, D. J., West, R. J., Elander, J., and Wilding, J. M. (1993). Decision-making style, driving style, and self-reported involvement in road traffic accidents. *Ergonomics*, 36(6):627–644.

Friedman, J., Hastie, T., and Tibshirani, R. (2009). *The Elements of Statistical Learning. Second Edition.* Springer Series in Statistics, New York.

Guo, F. and Fang, Y. (2013). Individual driver risk assessment using naturalistic driving data. *Accident Analysis & Prevention*, 61:3–9.

Guo, F., Klauer, S. G., Fang, Y., Hankey, J. M., Antin, J. F., Perez, M. A., Lee, S. E., and

Dingus, T. A. (2017). The effects of age on crash risk associated with driver distraction. *International Journal of Epidemiology*, 46(1):258–265.

Hankey, J. M., Perez, M. A., and McClafferty, J. A. (2016). Description of the SHRP 2 naturalistic database and the crash, near-crash, and baseline data sets. Technical report.

Happ, C. and Greven, S. (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *Journal of the American Statistical Association*, 113(522):649–659.

Happ-Kurz, C. (2020). Object-oriented software for functional data. *Journal of Statistical Software*, 93(5).

Harding, J., Powell, G., Yoon, R., Fikentscher, J., Doyle, C., Sade, D., Lukuc, M., Simons, J., Wang, J., et al. (2014). Vehicle-to-vehicle communications: readiness of V2V technology for application. Technical report, National Highway Traffic Safety Administration. DOT HS 812 014.

Hong, J.-H., Margines, B., and Dey, A. K. (2014). A smartphone-based sensing platform to model aggressive driving behaviors. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, pages 4047–4056.

Hu, L.-T. and Bentler, P. M. (1995). Evaluating model fit. In *R. H. Hoyle (Ed.), Structural Equation Modeling: Concepts, Issues, and Applications*, pages 76–99. Sage Publications, Inc.

Hu, X., Zhu, X., Ma, Y.-L., Chiu, Y.-C., and Tang, Q. (2018). Advancing usage-based insurance – A contextual driving risk modelling and analysis approach. *IET Intelligent Transport Systems*, 13(3):453–460.

Johnson, D. A. and Trivedi, M. M. (2011). Driving style recognition using a smartphone as a sensor platform. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pages 1609–1615.

Júnior, J. F., Carvalho, E., Ferreira, B. V., De Souza, C., Suhara, Y., Pentland, A., and Pessin, G. (2017). Driver behavior profiling: An investigation with different smartphone sensors and machine learning. *PLoS ONE*, 12(4):e0174959.

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling, Fourth Edition*. Guilford publications.

Kwak, B. I., Woo, J., and Kim, H. K. (2016). Know your master: Driver profiling-based anti-theft method. In *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, pages 211–218.

Li, F., Zhang, H., Che, H., and Qiu, X. (2016). Dangerous driving behavior detection using smartphone sensors. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, pages 1902–1907.

Li, Z., Jin, X., and Zhao, X. (2015). Drunk driving detection based on classification of multivariate time series. *Journal of Safety Research*, 54:61.e29–64.

Mantouka, E. G., Barmpounakis, E. N., and Vlahogianni, E. I. (2019). Identifying driving safety profiles from smartphone data using unsupervised learning. *Safety Science*, 119:84–90.

Massie, D. L., Campbell, K. L., and Williams, A. F. (1995). Traffic accident involvement rates by driver age and gender. *Accident Analysis & Prevention*, 27(1):73–87.

McMillen, D. L., Adams, M. S., Wells-Parker, E., Pang, M. G., and Anderson, B. J. (1992). Personality traits and behaviors of alcohol-impaired drivers: A comparison of first and multiple offenders. *Addictive Behaviors*, 17(5):407–414.

Meiring, G. A. M. and Myburgh, H. C. (2015). A review of intelligent driving style analysis systems and related artificial intelligence algorithms. *Sensors (Switzerland)*, 15(12):30653–30682.

Monecke, A. and Leisch, F. (2012). SemPLS: Structural equation modeling using partial least squares. *Journal of Statistical Software*, 48.

NHTSA (2017a). 2016 fatal motor vehicle crashes: Overview. Technical report, National Highway Traffic Safety Administration. DOT HS 812 456.

NHTSA (2017b). Automated driving systems: A vision for safety 2.0. Technical report, National Highway Traffic Safety Administration. DOT HS 812 442.

Peres-Neto, P. R., Jackson, D. A., and Somers, K. M. (2005). How many principal components? stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics and Data Analysis*, 49(4):974–997.

Poó, F. M. and Ledesma, R. D. (2013). A study on the relationship between personality and driving styles. *Traffic Injury Prevention*, 14(4):346–352.

Rakha, H., Farzaneh, M., Arafeh, M., and Sterzin, E. (2008). Inclement weather impacts on freeway traffic stream behavior. *Transportation Research Record*, 2071(1):8–18.

Ramsay, J., Hooker, G., and Graves, S. (2009). *Functional Data Analysis with R and MATLAB*. Springer Science & Business Media.

Ramsay, J. O. and Silverman, B. (2005). *Functional Data Analysis*. Springer Science & Business Media.

Ramsay, J. O. and Silverman, B. W. (2002). *Applied Functional Data Analysis: Methods and Case Studies*. Springer.

Rigdon, E. E., Sarstedt, M., and Ringle, C. M. (2017). On comparing results from CB-SEM and PLS-SEM: Five perspectives and five recommendations. *Marketing ZFP*, 39(3):4–16.

Romano, E., Voas, R. B., and Camp, B. (2017). Cannabis and crash responsibility while driving below the alcohol per se legal limit. *Accident Analysis & Prevention*, 108:37–43.

Sagberg, F., Selpi, Bianchi Piccinini, G. F., and Engström, J. (2015). A review of research on driving styles and road safety. *Human Factors*, 57(7):1248–1275.

Sârbescu, P., Costea, I., and Rusu, S. (2012). Using the alternative five factor personality model to explain driving anger expression. *Procedia-Social and Behavioral Sciences*, 33:273–277.

Schermelleh-Engel, K., Moosbrugger, H., and Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *MPR-online*, 8(2):23–74.

Schumacker, R. and Lomax, R. (2012). *A Beginner's Guide to Structural Equation Modeling: Third Edition*. Taylor & Francis.

Scrucca, L., Fop, M., Murphy, T. B., and Raftery, A. E. (2016). Mclust 5: Clustering, classification and density estimation using gaussian finite mixture models. *R Journal*, 8(1):289–317.

Scrucca, L. and Raftery, A. E. (2015). Improved initialisation of model-based clustering using gaussian hierarchical partitions. *Advances in Data Analysis and Classification*, 9(4):447–460.

Seacrist, T., Douglas, E. C., Hannan, C., Rogers, R., Belwadi, A., and Loeb, H. (2020). Near crash characteristics among risky drivers using the shrp2 naturalistic driving study. *Journal of Safety Research*, 73:263–269.

Sears, E., Perez, M. A., Dan, K., Shimamiya, T., Hashimoto, T., Kimura, M., Yamada, S., and Seo, T. (2019). A study on the factors that affect the occurrence of crashes and near-crashes.

Shinar, D., Schechtman, E., and Compton, R. (2001). Self-reports of safe driving behaviors in relationship to sex, age, education and income in the us adult driving population. *Accident Analysis & Prevention*, 33(1):111–116.

Shinar, D., Tractinsky, N., and Compton, R. (2005). Effects of practice, age, and task demands, on interference from a phone task while driving. *Accident Analysis & Prevention*, 37(2):315–326.

Simons-Morton, B. G., Cheon, K., Guo, F., and Albert, P. (2013). Trajectories of kinematic risky driving among novice teenagers. *Accident Analysis & Prevention*, 51:27–32.

Singh, S. (2015). Critical reasons for crashes investigated in the national motor vehicle crash causation survey. (Traffic Safety Facts Crash Stats. Report No. DOT HS 812 115). Washington, DC: National Highway Traffic Safety Administration.

Smadi, O., Hawkins, N., Hans, Z., Aldemir Bektas, B., Knickerbocker, S., Nlenanya, I., Souleyrette, R., and Hallmark, S. (2015). Naturalistic driving study: Development of the roadway information database. Technical report. REPORT S2-S04A-RW-1.

Smith, C. E. and Cribbie, R. A. (2013). Multiplicity control in structural equation modeling: Incorporating parameter dependencies. *Structural Equation Modeling*, 20(1):79–85.

Sullivan, G. M. and Feinn, R. (2012). Using effect size—or why the p value is not enough. *Journal of Graduate Medical Education*, 4(3):279–282.

Terry M. Therneau, E. J. A. (2019). An introduction to recursive partitioning using the rpart routines.

Toledo, T., Musicant, O., and Lotan, T. (2008). In-vehicle data recorders for monitoring and feedback on drivers' behavior. *Transportation Research Part C: Emerging Technologies*, 16(3):320–331.

Transportation Research Board of the National Academies of Science (2013). The 2nd Strategic Highway Research Program Naturalistic Driving Study Dataset. Available from the SHRP2 NDS InSight Data Dissemination web site: `https://insight.shrp2nds.us`.

Upahita, D., Wong, Y., and Lum, K. (2018). Effect of driving inactivity on driver's lateral positioning control: A driving simulator study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 58:893–905.

Victor, T., Dozza, M., Bärgman, J., Boda, C.-N., Engström, J., Flannagan, C., Lee, J. D., and Markkula, G. (2015). Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk. Technical Report Report S2-S08A-RW-1.

Vlahogianni, E. I., Yannis, G., and Golias, J. C. (2014). Detecting powered-two-wheeler incidents from high resolution naturalistic data. *Transportation Research Part F: Traffic Psychology and Behaviour*, 22:86–95.

WA OIC (2016). A consumer's guide to auto insurance: Choosing and using your auto insurance coverage. Retrived from https://www.insurance.wa.gov/sites/default/files/documents/auto-insurance-guide_1.pdf. Accessed: 2019-07-27.

Wang, C., Liu, L., Xu, C., and Lv, W. (2019). Predicting future driving risk of crash-involved drivers based on a systematic machine learning framework. *International Journal of Environmental Research and Public Health*, 16(3):334.

Whitlock, G., Norton, R., Clark, T., Jackson, R., and MacMahon, S. (2004). Motor vehicle driver injury and marital status: A cohort study with prospective and retrospective driver injuries. *Injury Prevention*, 10(1):33–36.

WHO (2018). *Global Status Report on Road Safety 2018*. World Health Organization.

Williams, A. F. and Carsten, O. (1989). Driver age and crash involvement. *American Journal of Public Health*, 79(3):326–327.

Wu, B.-F., Yeh, C.-H., and Chen, Y.-H. (2013). Driving behaviour-based event data recorder. *IET Intelligent Transport Systems*, 8(4):361–367.

Yao, F., Müller, H. G., and Wang, J. L. (2005). Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association*, 100(470):577–590.

Zhu, X., Yuan, Y., Hu, X., Chiu, Y. C., and Ma, Y. L. (2017). A bayesian network model for contextual versus non-contextual driving behavior assessment. *Transportation Research Part C: Emerging Technologies*, 81:172–187.