ABSTRACT

AHMED, ISHTIAK. Lane Changing and Car-following: Safety and Mobility Perspectives. (Under the direction of Dr. Nagui Rouphail and Dr. Billy Williams).

Characteristics of different driving maneuvers such as lane changing and car-following have pronounced effects on safety and mobility of a traffic stream. Both empirical and traffic-flow theory-based analyses of these maneuvers are important for the planning and designing of roadways. From the standpoint of lane changing maneuvers, this research aims to develop lane change detection, characterization, and prediction frameworks for freeway segments. From the standpoint of car-following maneuvers, this research aims to investigate the relationship between driver reaction time required for traffic stream stability and rear-end crashes based on car-following theory.

Five research questions are addressed in this dissertation. The first question was how to detect a lane change maneuver efficiently. A GPS-equipped vehicle’s geographic position and digitized infrastructure data were used to address this question. An algorithm was developed to minimize the effect of GPS errors by constraining the temporal duration and lateral displacement of a lane change detected using preliminary lane positioning. Field experiments showed that a temporal threshold of 1 second and an optimized lateral shift threshold of 6.9 ft. minimize the total error. Out of 46 ground truth lane changes, the algorithm detected 42 correctly with two false positives, resulting in a 4.35% and 8.7% false positive and false negative error, respectively.

The second research question was how to identify extreme lane change maneuvers in a traffic stream. This question was addressed by analyzing a complete trajectory dataset collected by the Next Generation SIMulation (NGSIM) program on US-101 in California. Two metrics were applied to identify extreme lane changing maneuvers, namely lane change rate (per distance
traveled) and time-to-line-crossing (TLC) during lane changes. The top 1% lane change rates exceeded three lane changes per 1,000 ft. traveled. Lane changes associated with the lowest 1% critical TLCs, which varied from 0.71 to 1.63 seconds, exhibited generally high entry angle and/or speed.

The third research question was how to predict discretionary lane change (DLC) intensity at weaving segments. This question was addressed by applying statistical modeling techniques to geometric and traffic data from 19 sites. A regression tree approach outperformed a linear regression model when predicting both DLCs per hour and DLCs per vehicle. A comparison with a widely used weaving model in the Highway Capacity Manual (HCM6) demonstrated the superiority of the proposed statistical models.

The fourth research question was how to develop a discontinuous form of the macroscopic Gazis-Herman-Rothery (GHR) car-following model. It was addressed by introducing an overlapping range defined in terms of density between the uncongested and congested flow regimes of the macroscopic GHR car-following model. The proposed model outperformed both continuous and discontinuous forms of the HCM6 traffic flow model when these models were applied to steady-state traffic flow observations from five sites.

The fifth research question was what is the association between rear-end crashes and driver reaction time required for car-following stability. This question was addressed by leveraging the discontinuous GHR model developed to answer the fourth research question. The discontinuity in the proposed model enables the estimation of the change in the required reaction time during traffic state transition. It was hypothesized that the greater the change in the required reaction time, the greater the rear-end crash risk for a segment. This hypothesis was tested by estimating rear-end crash rate and the changes in required reaction time for 28 basic freeway
segments. A moderate positive correlation coefficient of 0.63 was found between rear-end crash rate and the change in required reaction time at the end of the transition regime.
© Copyright 2020 by Ishtiak Ahmed

All Rights Reserved
Lane Changing and Car-following: Safety and Mobility Perspectives

by
Ishtiak Ahmed

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Civil Engineering

Raleigh, North Carolina
2020

APPROVED BY:

__________________________________________  _________________________________________
Dr. Nagui Rouphail  Dr. Billy Williams
Committee Co-Chair  Committee Co-Chair

__________________________________________  _________________________________________
Dr. George List  Dr. Eleni Bardaka

__________________________________________
Dr. Alan Karr
External Member
DEDICATION

This work is dedicated to my father, late Amin Uddin Ahmed for his support and love toward my family.
BIOGRAPHY

Ishtiak Ahmed was born in Dhaka, Bangladesh to Shirin Amin and Amin Uddin Ahmed. He received his BS in Civil Engineering in 2013 from Bangladesh University of Engineering and Technology (BUET). He worked as a Junior Engineer at Institute of Water Modelling before joining North Carolina State University in Fall 2015. He completed his MS in Civil Engineering with a major in Transportation Engineering in 2017. He worked as a Research Assistant at North Carolina State University from 2015 to 2020.
ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisors, Dr. Nagui M. Roupail and Dr. Billy M. Williams. Dr. Roupail taught me how to utilize my strengths and how to improve my weaknesses. He also helped me to dream big about my future research career. Dr. Williams has always been a great source of inspiration and aid whenever I needed those to overcome any challenge in research. I am grateful to Dr. Alan Karr, Dr. George List, Dr. Eleni Bardaka, and Dr. Yahya Fathi for their support and valuable feedback on this dissertation.

I am thankful to Dr. Shams Tanvir for inspiring me through his innovative research ideas. I am grateful to my classmate Dr. Shoaib Samandar for guiding me through the maze of exams and deadlines throughout my graduate life. I am indebted to Mehedi Hasnat, Tanmay Das, Samiul Haque, Dr. Nabaruna Karmakar, and Gyounghoon Chun; they were always there whenever I needed any help in research. Thanks to Institute for Transportation Research and Education and its many employees: Dr. Behzad Aghdashi, Dr. Daniel Findley, Shannon Warchol, Thomas Chase, Christopher Cunningham, Chris Vaughan, and Robert McCain for providing me with a helpful and friendly work environment.

I would like to acknowledge the North Carolina Department of Transportation, the Southeastern Transportation Research, Innovation, Development, and Education Center, and the North Carolina Section of ITE for supporting my research in various ways.

I am indebted to my mother for the sacrifices she made to shape my entire educational background. I am thankful to my brother as well as to his family for helping me completing my Ph.D. I am grateful to my wife, Mosarrat Samiha Kabir, for always holding my hand through all the ups and downs I faced throughout this journey. I would not be able to reach this stage
without her motivation. Above all, I thank the almighty Allah for giving me the strength and courage to complete this dissertation.
TABLE OF CONTENTS

LIST OF TABLES ......................................................................................................................... ix
LIST OF FIGURES ......................................................................................................................  x
CHAPTER 1 – INTRODUCTION ....................................................................................................  1
  1.1 BACKGROUND AND MOTIVATION .................................................................................  1
  1.2 DRIVING MANEUVERS ....................................................................................................  2
    1.2.1 Lane Changing Maneuvers .......................................................................................  2
    1.2.2 Car-following Maneuver ......................................................................................... 10
  1.3 SCOPE OF THE STUDY ..................................................................................................... 14
  1.4 ORGANIZATION OF THE DISSERTATION ..................................................................... 15
  1.5 REFERENCES .................................................................................................................... 16
CHAPTER 2 – CHARACTERIZING LANE CHANGES VIA DIGITIZED INFRASTRUCTURE AND LOW COST GPS .................................................................................................................. 21
  2.1 INTRODUCTION .............................................................................................................. 21
  2.2 LITERATURE REVIEW .................................................................................................... 23
  2.3 Methodology .................................................................................................................. 25
    2.3.1 Description of the Test Site ....................................................................................... 25
    2.3.2 Description of the Data ............................................................................................ 26
    2.3.3 Data Pre-processing ................................................................................................ 28
    2.3.4 Quality Check for the Data ..................................................................................... 29
    2.3.5 Lane Change Detection Algorithm ......................................................................... 31
    2.3.6 Characterization of Lane Changes .......................................................................... 36
  2.4 RESULTS ........................................................................................................................ 38
    2.4.1 Designed Experiment ............................................................................................... 38
    2.4.2 Application to the Naturalistic Data ......................................................................... 40
  2.5 SUMMARY ...................................................................................................................... 45
  2.6 REFERENCES .................................................................................................................... 48
CHAPTER 3 – MINING LANE CHANGING BEHAVIOR FROM TRAJECTORY DATA: CHARACTERIZATION AND EXTREME DRIVING IDENTIFICATION ................................................................................. 52
  3.1 INTRODUCTION AND RESEARCH OBJECTIVES ............................................................ 52
  3.2 LITERATURE REVIEW .................................................................................................... 54
  3.3 NGSIM US-101 DATASET ................................................................................................ 56
    3.3.1 Site Description ......................................................................................................... 56
    3.3.2 Traffic Characteristics ............................................................................................... 57
5.3.2 Fitting a Two-regime GHR Steady State Model .................................. 136
5.3.3 Requirements for Additional Constraints ........................................... 138
5.3.4 Robust Regression ............................................................................. 142
5.3.5 The HCM6 Model: As a Continuous Fundamental Diagram .................. 144
5.4 DESCRIPTION OF DATA AND STUDY AREA ....................................... 145
5.5 RESULTS ................................................................................................. 145
5.5.1 Application of CST, CDT, and SFDT .................................................. 146
5.5.2 Fitted Models ..................................................................................... 146
5.5.3 Comparison with HCM6 Models .......................................................... 150
5.6 SUMMARY AND CONCLUSIONS ........................................................... 153
5.7 References ............................................................................................... 155

CHAPTER 6 – INVESTIGATING THE RELATIONSHIP BETWEEN FREEWAY REAR-END CRASH RATES AND MACROSCOPICALLY MODELLED REACTION TIME ...... 158
6.1 INTRODUCTION ...................................................................................... 158
6.2 LITERATURE REVIEW .......................................................................... 159
6.3 METHODOLOGY .................................................................................... 162
6.3.1 The Macroscopic Model Equivalent to the GHR Car-following Model ........ 162
6.3.2 Estimating Required Driver Reaction Time for Different Traffic States .......... 164
6.3.3 Change in Required Reaction Time between Two Regimes ....................... 165
6.4 DATA DESCRIPTION ............................................................................. 167
6.4.1 Traffic Data from Sensors .................................................................. 167
6.4.2 Crash Data ......................................................................................... 168
6.5 ANALYSIS AND RESULTS ....................................................................... 169
6.5.1 Fitted Two-Regime Traffic Flow Models .............................................. 169
6.5.2 Analysis of Crash Data ...................................................................... 174
6.5.3 Crash Rates vs. Change in Required Reaction Time ................................ 175
6.6 CONCLUSIONS ...................................................................................... 179
6.7 REFERENCES ......................................................................................... 181

CHAPTER 7 – SUMMARY AND CONCLUSIONS .......................................... 184
7.1 SUMMARY .............................................................................................. 184
7.2 KEY FINDINGS ....................................................................................... 185
7.2.1 Key Findings Related to Lane Changing Behavior ................................. 185
7.2.2 Key Findings Related to Car-following Behavior .................................. 187
7.3 RECOMMENDATIONS FOR FUTURE WORK ....................................... 188
LIST OF TABLES

Table 1-1: Organization of the dissertation ................................................................. 15
Table 2-1: Designed trips by origin-destination and number of lane changes .................. 28
Table 2-2: Total lane change rates by lane origin-destination ...................................... 41
Table 4-1: List of weaving segments constituting the database ..................................... 95
Table 4-2: Linear regression model for DLCs per hour .................................................. 116
Table 4-3: Linear regression model for DLCs per vehicle ............................................. 118
Table 4-4: Comparison of model performance ............................................................. 121
Table 5-1: Parameter values for different sensors obtained from the fitted model ............ 149
Table 5-2: Parameter values for five sensors obtained by the HCM6 model ................. 152
Table 6-1: Fitted parameter values for different sensors and the resulting
reaction time changes ............................................................................................... 171
Table 6-2: Regression model forms and evaluations ..................................................... 176
LIST OF FIGURES

Figure 1-1: Demonstrating different lane change parameters (a) trajectory-level parameters (b) single lane change event........................................................................................................... 4
Figure 1-2: Demonstrating traffic stream stability................................................................................................................................. 12
Figure 2-1: Lane configuration of the weaving segment and its location near Raleigh, NC............................................................................. 26
Figure 2-2: Lane boundary crossing frequency by crossing direction in the designed field experiment.................................................................................................................. 27
Figure 2-3: Distribution of first difference of lateral distances from the median marking when the vehicle stayed on a lane.................................................................................. 30
Figure 2-4: Algorithm for detecting lane changes ........................................................................................................................................ 31
Figure 2-5: Illustrating the lane change detection algorithm using an example trajectory................................................................. 33
Figure 2-6: Sensitivity of lane change outcomes to displacement threshold value ....................................................................................... 39
Figure 2-7: Estimated total and excess lane changes by trips with various routes ..................................................................................... 41
Figure 2-8: Distribution of lateral speed of lane changes.................................................................................................................. 43
Figure 2-9: Distribution of weaving lane change distances from the entry gore.......................................................................................... 44
Figure 3-1: Schematic of the US-101 weaving segment .............................................................................................................................. 56
Figure 3-2: Top view of lane changing trajectories (a) toward the left (b) toward the right ........................................................................... 58
Figure 3-3: Example of lateral drifting by vehicle ID 1446 in period 1.................................................................................................. 60
Figure 3-4: Example of speed gain by changing lanes discretionarily ........................................................................................................ 61
Figure 3-5: Example of changing lanes frequently and with a sharp entry angle..................................................................................... 62
Figure 3-6: (a) ΔX and Δt for a lane change followed by another lane change in the opposite direction (b) ΔX vs. Δt for all lane changes that are either followed or preceded by a lane change in the opposite direction .......................................................... 65
Figure 3-7: Number of vehicles with different lane change frequencies by period ...................................................................................... 66
Figure 3-8: Probability density plot for change in average speed subsequent to changing lanes .......................................................................................................................... 69
Figure 3-9: Cumulative distribution of space mean speed of trajectories classified by lane change flag and period ................................................................. 70
Figure 3-10: Estimating time to line crossing during a lane change........................................................................................................ 73
Figure 3-11: Critical TLC analysis: (a) Distribution of critical TLC (b) Lowest 1% critical TLC in the average angle vs. average speed plot for all lane changes .................................................................................................................. 75
Figure 3-12: Trajectories with the highest frequency of lane changes per distance traveled ........................................................................ 78
Figure 3-13: Trajectories with lane changes with low critical TLC ........................................................................................................ 80
Figure 4-1: Snapshot of the automated video processing tool with vehicle IDs and traffic gates............................................................................................... 96
Figure 4-2: Explanation of site geometric properties and variables .................................................................................................. 97
Figure 4-3: Illustrating the selection of complexity parameter for predicting DLCs per hour............................................................................... 104
Figure 4-4: Regression tree model for DLCs per hour ........................................................................................................ 105
Figure 4-5: Regression tree model for DLCs per vehicle ........................................................................................................ 107
Figure 4-6: Site-specific relative RMSE of regression tree models (a) for DLCs per hour (b) for DLCs per vehicle ...................................................................................................................... 110
Figure 4-7: Partial dependence plot for interaction between length and freeway-to-ramp flow rate in DLCs per hour model ............................................. 114
Figure 4-8: Diagnostic plots for DLCs per hour model........................................................................................................ 116
Figure 4-9: Site-specific relative RMSE of linear regression models (a) for DLCs per hour (b) for DLCs per vehicle................................................................. 120
Figure 5-1: (a) Inconsistent data filtered by CST and CDT (b) Non-stationary data filtered by SFDT................................................................. 136
Figure 5-2: An example of the proposed flow-density model with and without any constraint......................................................................................... 139
Figure 5-3: Percentage of data reduction in different sensors ........................................ 146
Figure 5-4: Flow-density and Speed-flow diagrams for 5 sensors .................................. 148
Figure 5-5: Fundamental diagrams for sensor 29995 WB fitted by HCM model (a) Flow-density (b) Speed-flow ............................................................. 151
Figure 5-6: Comparison of MSE and BIC values obtained from the proposed and HCM model......................................................................................... 152
Figure 6-1: A typical required reaction time vs. density plot ........................................ 166
Figure 6-2: Location of the side-fire radar sensors in the study area.......................... 168
Figure 6-3: Fundamental diagrams for sensor 18W (a) Flow vs. Density, (b) Speed vs. Flow... 170
Figure 6-4: Rear-end crash rates for different sensor stations and the length of the freeway segments................................................................................. 175
Figure 6-5: Rear-end crash rates vs. (a) $\Delta t_e$ (b) $\Delta t_s$............................................. 175
Figure 6-6: Scale-location plot for (a) Linear (b) Exponential and (c) Logarithmic model.... 178
CHAPTER 1 – INTRODUCTION

1.1 BACKGROUND AND MOTIVATION

Characteristics of different driving maneuvers such as lane changing and car-following have pronounced effects on driving comfort, safety, and mobility of a traffic stream (Chen & Ahn, 2018; Daganzo, Laval, & Muñoz, 2002; Laval & Daganzo, 2006). For example, excessive and abrupt lane changing maneuvers—either due to speed differentials among vehicles or because of a poor lane choice—may cause significant lateral turbulence near freeway-ramp junctions. Another example is sudden acceleration and braking—such as those attributed to shockwaves propagating from a downstream bottleneck during traffic state transition—may severely disrupt the traffic stability and often results in rear-end crashes (Chatterjee & Davis, 2016; Davis & Swenson, 2006; Tanaka, Ranjitkar, & Nakatsuji, 2008). Moreover, extreme driving maneuvers increase the probability of driving errors (Godthelp, 1985). Hence, identifying and characterizing these elements of driving behavior and their causal relationships with traffic conditions and crash occurrences are of utmost importance for maintaining a satisfactory quality of service on roads.

The application of various driving behavior metrics is widespread across different domains of the transportation sector. For instance, the planning and designing of critical geometric elements, e.g., acceleration and deceleration lane length, segment length, and lane configuration of a weaving segment, and informative signs and markings are dictated by the expected lane change frequency at those segments (Transportation Research Board, 2016). Operational treatments like variable speed limits have been implemented to mitigate the effects of rear-end crash risk of a road, which is often an aggregated reflection of the car-following behavior of drivers (Chen, Ahn, & Hegyi, 2014; Han, Chen, & Ahn, 2017). For these cases,
transport agencies mostly rely on aggregated or macroscopic data involving all vehicles in a traffic stream. On the other hand, driver behavior analyses using trajectory-based or microscopic data are now becoming more popular. Common application arenas for this line of research include microsimulation model calibration and development (Chen, Laval, Zheng, & Ahn, 2012), assessing in-vehicle warning sensor performance (Das, Khan, & Ahmed, 2020), and evaluating future connected and autonomous vehicle scenarios (Lin, Gong, Li, & Peeta, 2018). Given this breadth of application, both microscopic and macroscopic level examinations of driving behavior are therefore important.

1.2 DRIVING MANEUVERS

On a microscopic level, the outcomes of driving maneuvers include speed, acceleration, jerk (a derivative of acceleration), yaw (rotation with respect to the vertical axis), roll (rotation with respect to the longitudinal axis), and pitch (rotation with respect to the transverse axis) (Jo, You, Joeng, Lee, & Yi, 2008). On freeways, these kinematic characteristics fall into two types of maneuvers: lane changing and car-following. The remainder of this discussion on driving maneuvers is focused on those two types.

1.2.1 Lane Changing Maneuvers

Lane changing is one of the most complex and challenging tasks associated with freeway driving (Fastenmeier & Gstalter, 2007). Extensive research has been conducted to model the different aspects of a lane changing maneuver and to quantify its impacts on road capacity, mobility, and safety. Yet, several critical research gaps persist in these regards, some of which are being addressed in this dissertation. Before detailing those questions, it is important to explain different characteristics of a lane changing maneuver, as presented in the following
1.2.1.1 Characteristics of Lane Changing Maneuvers

When dissecting different characteristics of a lane changing maneuver, the key aspects past researchers have focused on include why, where, and how many times lane changing occurred within a spatiotemporal boundary. In addition, researchers were also interested in different microscopic vehicular kinematics and dynamics during the maneuver such as speed, acceleration, steering angle, time to collision with neighboring vehicles, and time to cross the outer boundary of the destination lane (Godthelp, 1985; Rathbone & Huckabee, 1999; Schmidt, Beggiato, Hoffmann, & Krems, 2014; Tasca, 1999; van Winsum, De Waard, & Brookhuis, 1999). Figure 1-1 below illustrates these characteristics of lane changing maneuvers, further discussed below.
Figure 1-1: Demonstrating different lane change parameters (a) trajectory-level parameters (b) single lane change event.

**Motivations for a lane Change**

Many past studies classified lane-changing maneuvers into two groups based on the driver’s motivation: mandatory (MLC) and discretionary (DLC) (Kesting, Treiber, & Helbing, 2007; Toledo, Koutsopoulos, & Ben-Akiva, 2003). In this dissertation, the fewest number of lane changes that must be made to reach one’s destination are classified as mandatory or minimum lane changes. MLC for each vehicle at a segment is dictated by lane configuration of the segment. For the weaving segment marked by the green rectangle in Figure 1-1 (a), the minimum number of lane changes for a vehicle traversing from the freeway to the exit ramp would be zero if it had entered the segment through Lane 4, while that for a vehicle merging from the on-ramp...
on to freeway is two (from on-ramp to Lane 3). All other lane changes are classified as discretionary. While MLCs are dictated by the driver’s route, several factors may motivate DLCs, such as pre-positioning, speed adjustment, and friction avoidance (Kan, Sun, & Zhang, 2009).

Note that these definitions of mandatory and discretionary lane changes depend on the segment of interest within which a vehicle’s trajectory is observed. To further illustrate, consider the yellow vehicle in Figure 1-1 (a) that merged onto the freeway from an on-ramp. If the segment of interest is the weaving segment marked by the green rectangle in this figure, then its leftward lane change from lane 4 to lane 3—according to the definition stated above—is a discretionary lane change. However, if the segment of interest includes both the on-ramp upstream of the weave and the weave, then according to the definition stated above, that leftward lane change should be considered as mandatory. Without shifting toward left up to lane 3, this merging vehicle cannot stay on the freeway at the end of the weaving segment. Therefore, before defining and analyzing mandatory and discretionary lane changes, first, the segment of interest needs to be specified. To observe lane changing behavior using traditional field data collection tools such as video cameras, the segment of interest is often determined by the maximum coverage capacity of the tool.

The positioning of a vehicle often dictates how many lane changes it has to make to maintain its origin-destination route. For example, consider the blue vehicle in Figure 1-1 (a) that is traversing from the freeway to the exit ramp. Due to poor lane positioning, it entered the segment through the rightmost lane and made three lane changes within the weave, which are discretionary (given that the portion of the trajectory as shown in Figure 1 (a) is what an analyst can observe). If a significant number of vehicles pre-position in such a poor manner during
heavy traffic conditions, they will create driving discomfort and turbulence by accepting narrow gaps and forcing other vehicles to decelerate. Another example is the green vehicle in Figure 1-1 (a), for which both lane changes are discretionary. The motivation behind these might be to change the speed or to pre-position for the exit ramp located further downstream of the weave. While these lane changes may be beneficial for this driver’s own driving comfort or in terms of pre-positioning for a downstream exit ramp, the weaving segment in question may face additional turbulence due to these DLCs during peak hours. Here, once again, the necessity of specifying the segment of interest appears when describing the benefit of pre-positioning. Early pre-positioning may reduce the turbulence within a segment, but also may increase the turbulence at an upstream segment. This compromise may raise the question of whether pre-positioning is beneficial at all. Pre-positioning may be beneficial if the necessary lane changes take place within a segment that is not in a critical condition (i.e., not prone to flow breakdown). For example, if the weaving segment in Figure 1-1 (a) is prone to bottleneck activation due to lane change-induced turbulences and the upstream on-ramp merge area is not, then the driver of the yellow vehicle should make that leftward lane change (from lane 4 to lane 3) before the vehicle entered the (critical) weaving segment.

**Position of a Lane Change**

Near the ramp entry and exit gores, mandatory lane changes are prevalent, especially if the length of the acceleration, deceleration, or auxiliary lane is short (Marczak, Daamen, & Buisson, 2014; van Beinum, Farah, Wegman, & Hoogendoorn, 2018). Hence, even without any discretionary lane changes, by design, these points are likely to experience significant turbulence. Should many drivers attempt to merge or diverge at the same time near these critical points, the capacity and quality of service of that segment may degrade substantially. In addition,
poor lane positioning often leads drivers to change lane beyond the exit gore, such as the last lane change by the blue vehicle in Figure 1-1 (a). Such a delayed MLC may trigger a crash or a near-crash event. Furthermore, lane changes near these critical points may induce turbulence. On the other hand, lane changes away from these critical points to avoid friction free up space at the neighboring lanes and help reduce the turbulence. All in all, the location of lane changes plays a critical role in the operation of a ramp-junction area. Traditional navigation tools (e.g., Google Maps and GPS devices) do alert drivers to change lanes when reaching within a pre-specified distance of the target exit ramp. However, these tools do not account for the traffic condition and the expected lane change turbulence near that ramp. Investigating the relationships among lane change opportunities, expected lane change frequency, and traffic characteristics would help improve those navigation tools so that they can alert the drivers early if needed. Thus, drivers may avoid unwanted events induced by late diverging or by accepting narrow gaps while changing lanes.

**Lane Change Frequency**

Frequency is an important parameter for individual driving behavior analysis and for the planning and design of roads. There are several ways the number of lane changes within a spatiotemporal boundary can be quantified. For instance, lane change frequency over a period is often reported separately by lane change type and by direction, e.g., mandatory or discretionary, toward the left or the right (Gan & Jin, 2015). It is also common to normalize lane change frequency by the spatial and temporal coverage over which it is observed, e.g., lane changes per hour per mile (Fredericksen & Ogden, 1994). One focus of this research is to develop methods for estimating lane change frequency. For example, in order to identify risky driving, lane change frequency for an individual trajectory is typically the focus (Xuan & Coifman, 2006). On
the other hand planning and design level analysis will tend to focus on aggregated lane change frequency for a traffic stream (Transportation Research Board, 2016).

**Other Dynamics of a Lane Change**

As mentioned earlier, other kinematics and dynamics of a lane change that previous researchers focused on include speed, acceleration, entry angle, accepted gaps, Time To Collision (TTC), and Time to Cross outer lane boundary of the destination Lane (TLC). Among these, speed and acceleration are readily understandable metrics but are difficult to visualize. Figure 1-1 (b) depicts those additional metrics during the lane change maneuver by the blue vehicle. TTC and TLC are estimated from the estimated time to cross the distances shown by the green and red arrows, respectively. These times, in turn, were estimated using the entry angle ($\theta$) in Figure 1-1 (b)), speed, and acceleration data. A low value for TTC and/or TLC, and a high value for speed, acceleration, and entry angle have been used as indicators of potentially aggressive lane changes in past studies (Das et al., 2020; van Winsum et al., 1999).

### 1.2.1.2 Research Questions Related to Lane Changing Behavior

The first research question addressed in this dissertation is:

*Research question no. 1: How to efficiently detect lane changes via vehicle position and digitized infrastructure data?*

While a variety of studies related to lane changing have been conducted in controlled environments such as using driving simulators (Li et al., 2018; van Winsum et al., 1999), naturalistic driving data is critical for lane changing behavior analysis. To that end, traffic videos have been extensively used, but their spatiotemporal coverage tends to be limited. Recently, in-vehicle GPS or vision-based sensors have gained popularity to track vehicle trajectories in real-time (Atia et al., 2017; Song et al., 2017). Vision-based sensors are primarily installed for safety
purposes; they also have the highest accuracy because of their proximity to the lane marking (Li et al., 2014; Kim, 2008). This precision is important for an advanced driver assistance system or in an autonomous vehicle. However, such sensors might be too expensive to be used for traffic-monitoring purposes. GPS-based technologies can be used to identify lane changes as well. Nonetheless, the primary challenge in this regard lies in dealing with the significant amount of randomness and bias in the reported GPS positions. All these challenges and the resulting gaps in research motivate the first research question.

The second research question addressed in this dissertation is:

Research question no. 2: How to identify extreme lane change maneuvers in a real-world traffic stream?

Upon accurate detection of lane changes, the next important task is to characterize those to identify potentially aggressive or extreme behavior. Many researchers conducted controlled experiments to determine what attributes of lane changing characteristics make a lane change aggressive (van Winsum et al., 1999). For instance, DLC frequency has been considered as an indicator of aggressive driving in several studies (Bham, 2011; Sun & Elefteriadou, 2012). The combination of steering angle and speed during lane changes has been used to identify extreme driving behavior by past studies (Godthelp, 1985; van Winsum et al., 1999) using driving simulators. Speed, acceleration, and accepted gaps for changing lanes have been used by Das et al. (Das et al., 2020). While these studies established those measures as indicators of aggressive lane changing behavior, their application to identify potential aggressive behavior from a real-world trajectory dataset for all vehicles in a traffic stream has not been done. This gap motivates the second research question of this dissertation.

The third research question addressed in this dissertation is:
Research question no. 3: How to predict discretionary lane change intensity at weaving segments using macroscopic traffic characteristics and infrastructure data?

One of the most challenging issues related to investigating different characteristics of lane changing behavior is that it necessitates high-resolution trajectory data. Unfortunately, high-resolution trajectory data are not readily available to most transportation engineers and researchers. The most common type of traffic data is aggregated traffic characteristics collected from fixed sensors (e.g., side-fire radar or loop detectors) at various points of a segment. Although these data are not suitable for individual driver behavior analysis, these can provide an estimate of aggregated lane changing frequency within a segment (Transportation Research Board, 2016). To that end, a mathematical bridge must be developed between aggregated lane change counts and macroscopic traffic characteristics supplemented by road characteristics data. Lane change frequency estimated in this fashion is expected to be helpful for making informed decisions when planning and designing interchange influence areas such as minimum weave length, necessary road signs and markings, and lane configuration. It motivates the third research question of this dissertation.

1.2.2 Car-following Maneuver

Car-following is defined as the interaction of two vehicles as one follows the other. This maneuver has a significant influence on road capacity and safety. However, one of the most challenging aspects of conducting research on car-following behavior is that collecting naturalistic data on car-following behavior is extremely difficult. Unlike a lane changing maneuver which takes place over a short time and limited space, a car-following maneuver typically continues over a longer span. Moreover, while most lane change related parameters are measured from the subject vehicle dynamics as described previously, characteristics of car-
following maneuvers mostly depend on both following and leader vehicle dynamics. These issues make it extremely difficult to collect naturalistic data on car-following parameters.

Key characteristics of a driver’s car-following behavior include following distance, reaction time, time to collision, and acceleration, jerk, and speed of follower and leader vehicles (Brackstone & McDonald, 1999). Due to the data collection difficulties explained above, transport agencies often find it challenging to employ these microscopic parameters in road planning and design. To solve this issue, researchers had attempted to link these parameters with macroscopic traffic properties — e.g., average traffic speed and density — of which data are more readily available (Lee, Lee, & Kim, 2001). While these macroscopic properties help analyze mobility and capacity of road segments, their applications for safety evaluations are very limited. The research presented in this dissertation related to car-following maneuver aims to utilize a bridge between microscopic and macroscopic car-following properties for crash-risk analysis of freeway segments.

Among various types of crashes, the mechanism of rear-end crashes is closely associated with that of car-following maneuvers. A rear-end crash occurs when a following driver fails to react in time to a change in relative dynamics with its leader and consequently collides with the leader. This reaction time of the following driver is an important parameter that impacts a rear-end crash occurrence. In other words, a reaction time that is less than a certain maximum value is required for the follower to avoid a rear-end crash, i.e., to maintain local car-following stability.

The research presented in this dissertation regarding car-following behavior hypothesizes that the reaction time required for traffic stream stability is associated with the long-term risk of rear-end crash occurrence on a segment. Before describing the key research questions, it is important to explain the physics of traffic stream stability, as presented in the next section.
1.2.2.1  *Traffic Stream Stability*

Consider the stream of vehicles shown in Figure 1–2 moving in the same lane. In response to a change of speed of the frontmost vehicle, its immediate follower must react within a certain time to maintain car-following stability. This stability between a pair of follower-leader vehicles is called the local stability (May, 1990). Despite reacting within the required time, the oscillation generated due to the acceleration or deceleration of these two vehicles will propagate downstream. If the response of any follower driver in the traffic stream is erratic, i.e., slow to react and when reacting does so with extreme acceleration or deceleration, then the amplitude of the oscillation will increase as it propagates downstream, and may lead to a crash between two vehicles further upstream of the traffic stream. The stability of an infinite number of vehicles as discussed here is termed the asymptotic stability (May, 1990).

![Figure 1-2: Demonstrating traffic stream stability.](image)

Increasing oscillations that may lead to an asymptotic instability are common during traffic state shifts from an uncongested to a congested state. During this transitioning state, the required reaction time for asymptotic stability also changes. From a car-following behavior perspective, it implies that drivers must adapt themselves to a lower required reaction time during this state-transition to avoid asymptotic instability. This research, thus, logically hypothesizes that this change in required reaction time has an association with the long-term rear-end crash risk of a freeway segment.
1.2.2.2 Research Questions Related to Car-following Behavior

The first research question related to car-following behavior (fourth overall) is:

*Research question no. 4: How to develop a discontinuous form of the macroscopic GHR car-following model*

As a reminder, the aim of this research is to develop a safety measure related to car-following behavior that can be estimated using macroscopic traffic observations. The seminal research that derived a closed-form equation for determining the required reaction time from a macroscopic traffic flow model parameters is reported by May (May, 1990). This derivation was based on the well-known Gazis-Herman-Rothery macroscopic model that is continuous at the maximum flow point (Gazis, Herman, & Rothery, 1961). However, the major challenge to estimate the *change* in required reaction time during traffic state transition requires one to fit the underlying macroscopic model with a discontinuity at capacity (Edie, 1961). This motivated the fourth research question of this dissertation.

The fifth research question addressed in this dissertation is:

*Research question no. 5: What is the association between rear-end crashes and driver reaction time required for car-following stability?*

Once the above research question is addressed, it would be possible to estimate the change in reaction time required for car-following stability during traffic state transition. After that, what remains is to test the hypothesis that this change is related to the long-term rear-end crash risk of a segment. To identify rear-end crash-prone locations, the state of the practice is documented in the Highway Safety Manual (American Association of State Highway and Transportation Officials, 2010). However, highway safety engineering is founded on the statistical analysis of crash occurrences. Therefore, for the safety evaluation of a road, transport
agencies are obliged to wait for crashes to occur and the crash rate to regress to its mean (Hauer, 1997). This issue can be solved if a relationship between rear-end crash risk and an estimate of the required driver reaction time can be developed so that the latter can be used as a surrogate measure of rear-end crash risk during the transition of traffic states. To that end, the fifth and the last research question addressed in this dissertation was emerged.

1.3 SCOPE OF THE STUDY

The general scope of this dissertation research is noted below. Note that more specific components related to each research question are presented in the corresponding chapter.

- This dissertation deals with driving behavior focused exclusively on freeways. Moreover, it addresses two key driving maneuvers: lane changing and car-following. Other maneuver types, for instance, response to crashes or work zones, are beyond the scope of this study.
- Distracted driving has been more common nowadays with the increase in the use of cellphones, navigation tools, and entertainment features while driving. These activities often have detrimental effects on road safety and mobility, and numerous studies have been going on to address these. However, research on these activities is also beyond the scope of this dissertation.
- Driving behavior data employed in this study apply to either individual driver-level or aggregated traffic-level characteristics. The author acknowledges that much research on driving behavior employed metrics that are associated with the dynamics of a pair of vehicles such as time-to-collision, headway, and following distance. However, these metrics are not considered in this research due to a lack of reliable data. Details about the issue of driving behavior data used in this research are described in the corresponding chapter.
1.4 ORGANIZATION OF THE DISSERTATION

In total, there are seven chapters in this dissertation. Following this introductory chapter, the next five chapters discuss Research question 1 through Research question 5 in that order. Each of these five chapters include its own, more specific research motivations, literature review, methods, findings, summary, and references. The last chapter presents the overall summary and the most critical findings from this dissertation and recommendations for future research. The overall organization of this research and the flow of work are shown in Table 1-1.

Table 1-1: Organization of the dissertation.

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Research question</th>
<th>Chapter title</th>
<th>Research theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>Introduction</td>
<td>Background, motivations, and problem statements</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>Characterizing Lane Changes via Digitized Infrastructure and Low-cost GPS</td>
<td>Presents a method for lane change identification and characterization using in-vehicle sensor and digitized infrastructure</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>Mining Lane Changing Behavior from Trajectory Data: Characterization and Extreme Driving Identification</td>
<td>Presents a method for identifying extreme driving behavior from high-resolution trajectory data</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>Modeling Framework for Predicting Lane Change Intensity at Freeway Weaving Segments</td>
<td>Develop a prediction model for lane changing frequency within weaves using macroscopic traffic and road characteristics data</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>Developing a Discontinuous Form of Macroscopic Gazis-Herman-Rothery Car-following Model</td>
<td>Develop a macroscopic steady state GHR model with a discontinuity at the transition of uncongested and congested flow regimes.</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>Investigating the Relationship between Freeway Rear-end Crash Rates and Macroscopically Modeled Reaction Time</td>
<td>Evaluates the association between rear-end crash rate and the change in required reaction time during traffic state transition</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>Summary, Conclusions, and Recommendations</td>
<td>Highlights major findings and provides guidance for future research</td>
</tr>
</tbody>
</table>
1.5 REFERENCES


CHAPTER 2 – CHARACTERIZING LANE CHANGES VIA DIGITIZED INFRASTRUCTURE AND LOW COST GPS

2.1 INTRODUCTION

Lane changing is one of the most complex and challenging tasks associated with freeway driving. Wide-ranging research has been conducted to model and characterize lane change patterns and their effect on mobility and safety under different traffic and geometric conditions (Bakhit, Osman, & Ishak, 2017; Sparmann, 1979; Toledo & Zohar, 2007). In particular, near ramp-junctions where changing lanes is mandatory, lane positioning and frequency of lane changes can impact the mobility and safety of the roadway significantly (Laval & Daganzo, 2006; Shoaib Samandar, Billy Williams, and Ishtiak Ahmed, in press; Sparmann, 1979; Tanvir, Karmakar, Roupail, & Schroeder, 2016). To study these properties and impacts of lane changes, the first and foremost task is to accurately detect their location and time path and to estimate their frequency.

Lane changes detection processes currently in practice can be categorized into two classes: (a) using video footage, (b) using in-vehicle sensors. With the improvement of imagery analysis, several researchers have demonstrated the successful applications of aerial videography for detecting and analyzing lane changes (Marczak, Daamen, & Buisson, 2014a; USDOT, 2006; van Beinum, Farah, Wegman, & Hoogendoorn, 2018). However, the inherent limitation of this technique is that its spatial coverage is limited, and it requires extensive post-processing. In-vehicle sensors used for detecting lane changes are of two types: vision-based and GPS-based. Vision-based sensors are primarily installed for safety purpose; they also have the highest

---

1 Part of this chapter was published as: Ahmed, I., Karr, A., Roupail, N. M., Chun, G., & Tanvir, S. (2019). Characterizing lane changes via digitized infrastructure and low-cost GPS. Transportation research record, 2673(8), 298-309.
accuracy because of their close proximity to the lane marking (Kim, 2008; Li, Q., Chen, Li, Shaw, & Nüchter, 2014). This precision is important for Advanced Driver Assistance System (ADAS) or in an autonomous vehicle. However, such sensors might be too expensive to be used for traffic monitoring purposes. GPS based technology can be used to identify lane changes as well. Nonetheless, the primary challenge in this regard lies in dealing with the significant amount of randomness and bias in the reported GPS positions.

Nowadays, such in-vehicle sensors have the capability of sharing various information related to a vehicle and the surrounding environment with the provision of connected vehicle (CV) technology. With a simple cellular network equipped with these sensors, a vehicle can act as internet of thing (IoT) (Live drive.2013). Assuming that the privacy issue of data-sharing can be resolved, traffic agencies can monitor vehicle lane position information using such a vehicle-to-infrastructure (V2I) connection.

In this study, an algorithm for detecting lane changes is developed by integrating a high-resolution V2I connection and digitized lane marking data. A low-cost GPS receiver equipped with a cellular connection was used. It transfers vehicle positions at a frequency of 1 Hz. in real-time. Detecting false signal of lane changes attributed to the randomness and bias in the data was minimized by calibrating the algorithm through a designed field experiment. The algorithm was applied to a freeway weaving segment to detect and characterize naturalistic lane change behavior of the connected vehicles by their lane change frequency, excess lane change tendency, lateral speed while changing lanes, and the spatial extent of the mandatory lane changes.

A review of past studies on lane change detection using in-vehicle sensors is presented in the next section. It is followed by a description of the methodological tools for developing the lane change algorithm. A technique for the robust selection of the algorithm thresholds for
filtering the detected lane changes is described as well. Key findings from the application of the algorithm are summarized in the last section.

2.2 LITERATURE REVIEW

Several past studies have used in-vehicle GPS technologies for detecting lane changes (Xuan & Coifman, 2006), and often fused those with vehicle kinematics information (Atia et al., 2017; Li, T., Yang, Xu, Zhou, & Wang, 2016; Song et al., 2017; Toledo-Moreo & Zamora-Izquierdo, 2009). A technique for identifying mandatory and discretionary lane changes was developed (Xuan & Coifman, 2006) using probe vehicles equipped with a high-precision, high-cost Differential GPS unit. A reference trajectory was defined from the median of a set of probe vehicle runs through a predefined origin-destination route. Based on the observation that a mandatory lane change along the route impacts the median of the reference trajectories more abruptly than it does the mean value, the difference between these two values was used to detect the mandatory lane changes. Discretionary lane changes were identified by applying a lateral displacement and a lateral speed threshold selected based on the average lane width.

Toledo-Moreo et al. (2009) fused position data from a low-cost GPS with velocity measurements using an Interactive Multiple Model to predict lane changes on highways. Characterizing lane changes as a Markovian process with known probability, a Keep Lane and a Change Lane model are combined to estimate the position and state of a vehicle. An Extended Kalman Filter algorithm is applied to include the kinematic model of vehicle dynamics assuming constant acceleration. Atia et al. (2017) also used low-cost GPS along with Inertial Measurement Units (IMU) and identified lane changes by a map-matching technique. Upon combining the GPS and IMU data using a Kalman-filter, a Hidden Markov Model and a least square regression model were used to find the lane position and detect lane changes. Song et al. (2017)
demonstrated a low-cost technique for detecting lane changes by combining different features of a smartphone such as camera, accelerometer, and gyroscope. A module was developed that detected the two stages of the maneuver: lane changing and settlement on the new lane. However, post-processing of the recorded image is required to implement this process. Li et al. (2016) used a vehicle equipped with an RTK-GPS and an inertial motion tracking device along with the digitized road map for detecting lane changes. It investigated three features namely the lateral displacement, yaw angle displacement, and lane change duration in the algorithm. Thresholds for these parameters are selected based on the characteristics of the road curvature.

Lane change characterization by their number, motivation, lateral speed, and the location is considered important for highway design and operations. The Highway Capacity Manual (HCM6) (Transportation Research Board, 2016) estimates lane change frequencies in a weaving segment for weaving and non-weaving vehicles separately. Therefore, whether a lane change is discretionary or mandatory is dependent on the origin-destination path of the vehicles. Knoop, Hoogendoorn, Shiomi, and Buisson (2012) modeled the number of lane changes as a function of macroscopic traffic characteristics using video record and loop detector data. Although this study classified the lane position of the lane changes, it did not distinguish the discretionary lane changes based on lane position.

Several studies used NGSIM or similar video recorded trajectory information to characterize lane changes. The duration and longitudinal speed of a lane changing vehicle analyzed by Moridpour, Sarvi, and Rose (2010) were found to be significantly different for different vehicle classes. A moving average smoothing technique was implemented prior to identify the start and end point of the lane change based on their lateral displacement. Location of the lane changes in weaving segments was investigated by a few studies (Marczak et al., 2014;
Using video-recorded traffic data for a single weaving site, it was revealed that all the weaving lane changes occurred within the first 60% of the total weaving length (Marczak et al., 2014). Similarly, it was observed by Vaan Beinum et al. (2018) based on trajectory data from eight weaving sites that about 65%-95% of the weaving lane changes occurring within the first 25% of the total length.

From the review of the literature, it is evident that studies that used only GPS positions to detect lane changes required high precision and expensive devices. Several studies fused low-cost GPS data with vehicle dynamics, but it can make the process complex and introduce additional sources of error. To the analyst’s knowledge, there is a lack of research on detecting lane changes solely based on low-cost GPS position data and open-source digitized maps. Regarding lane change characterization, past studies used full trajectory data to classify lane changes based on their route origin-destination, duration, and location of lane changes. From the perspective of both highway design and operations, these properties of lane changes were deemed important in past studies.

2.3 METHODOLOGY

In this section, the proposed method for detecting lane changes is described. It begins with the description of the selected test site and the data used in this study. Then, the quality of the GPS data is assessed. A step-by-step description of the algorithm developed for detecting lane changes follows next. Finally, the characterization of the identified lane changes is discussed.

2.3.1 Description of the Test Site

In this study, a five-lane freeway weaving segment on EB I-40 in Raleigh, NC was selected as the study site (see Figure 2-1). The selected section is approximately 1.1-mile long.
with a one-lane entry and a two-lane exit ramp. The minimum number of lane changes for the ramp to freeway vehicles through this weaving segment is two, but it is zero for the freeway to ramp vehicles. Being located in the close proximity of the Research Triangle Park, it was tagged as one of the most severe recurring bottlenecks near Raleigh (Ahmed, Rouchail, & Tanvir, 2018).

2.3.2 Description of the Data

2.3.2.1 Digitized Lane Markings at the test-site

The lane, shoulder, and median markings at the test site were digitized using the Google Earth Pro tool. The spatial resolution of the digitized marking (i.e., the length of each straight digitized line) was 15 feet, in order to maintain sufficient accuracy. The numbering scheme shown in Figure 2-1 is used throughout this chapter for designating the lanes.

Figure 2-1: Lane configuration of the weaving segment and its location near Raleigh, NC.

2.3.2.2 Vehicle Position Data

Vehicle position data were collected using a low-cost GPS device manufactured by U-blox (model: MAX-M8Q) (Ublox.2016) and retailed by LiveDrive (LiveDrive, 2013). With the provision of a cellular network connection, the device acted as an IoT and transferred the position data real-time at a frequency of 1 Hz. to a cloud server. Two sets of position data used in this study are described in the following subsections.
**Designed Field Experiment:** A designed experiment was conducted comprising 29 trips with passenger vehicles, where a total of 46 pre-specified lane changes were made. The trips were recorded with an in-vehicle video recorder. The recorder and the GPS device both were placed in the middle of the car. The video feed and the time-stamp of the recording were used to obtain the time series information of the ground truth lane position. These information were used to develop the proposed lane change detection algorithm. The basis of designing this experiment was to cross all lane marking from both directions. Figure 2-2 shows the number of times each lane marking was crossed by crossing direction. The classification of the field trips among different origin-destination routes and their number of lane changes are shown in Table 2-1.

![Bar chart](image)

**Figure 2-2:** Lane boundary crossing frequency by crossing direction in the designed field experiment.
Table 2-1: Designed trips by origin-destination and number of lane changes.

<table>
<thead>
<tr>
<th>Origin-Destination Routes</th>
<th>Number of trips with pre-planned lane changes</th>
<th>Total Trips</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Freeway-Freeway</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Freeway-Ramp</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ramp-Freeway</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ramp-Ramp</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total Trips</strong></td>
<td><strong>5</strong></td>
<td><strong>3</strong></td>
</tr>
</tbody>
</table>

**Total Number of Lane Changes** | 46

Figure 2-2 reveals that all lane markings were crossed from both directions at least once. More lane changes from lane 4 to lane 5 and lane 3 to lane 4 could not be conducted more as it perceived unsafe at that time. Table 2-1 shows that drivers were instructed to remain in their lane during four freeway-freeway and one ramp-ramp trips. The lateral displacement of the vehicle from a reference lane marking during these five trips is to be used to assess the quality of the data.

**Naturalistic Driving Data:** The second dataset contains 637 naturalistic vehicle trajectories through the test site obtained from a fleet of 40 passenger vehicles. These trips were made between 2015 and 2016 by commuter drivers residing near the Triangle Area of Raleigh. In this study, the proposed lane change detection algorithm is applied to this dataset in order to characterize the lane changes.

**2.3.3 Data Pre-processing**

Following the road digitization, each vehicle position was mapped using ArcGIS. A small proportion of the data (0.5% and 5% of the designed and naturalistic dataset, respectively) were located outside the road. The average lateral displacement of these outlying latitude/longitude points from the nearest road edge was 3.5 feet. Since these erroneous points may result in
abnormal lane selection behavior, they were moved to the nearest edge marking using the “Near” tool of ArcGIS. However, although this technique can fix some obvious outliers, there is no guarantee that the observations within the road segment are all accurate. Therefore, a quality check of the position data is also required. For roads with more complex geometry such as at junctions, map-matching algorithm proposed by an earlier study (Zhu, Bao, Wang, Zhou, & Bao, 2012) can be used to plot such vehicle trajectory data.

2.3.4 Quality Check for the Data

Prior to implementation, it is imperative to check whether the GPS data is sufficiently accurate for the purpose of detecting lane changes. In this study, this assessment was made based on the variation of the lateral orthogonal distance of the trajectories from a reference line when a vehicle stayed on the same lane throughout the segment. Lateral distances were calculated using the “Geosphere” package in R (Hijmans, Williams, & Vennes, 2015). It uses spherical trigonometry functions for geographic applications. Its “distance to a polyline” function is used to compute the shortest distance between a set of points and a set of spherical polylines. The coordinates of a trajectory and each lane marking are provided as inputs. For each trajectory position, it generated the perpendicular distance and bearing from each lane marking.

For the five trips listed in Table 2-1 with no lane changes, the first difference of the lateral displacement from the median marking is calculated for a total of 272 observations. Figure 2-3 shows the distribution of these first differences with a fitted distribution density plot.
Figure 2-3: Distribution of first difference of lateral distances from the median marking when the vehicle stayed on a lane.

For the estimated mean and standard deviation shown in Figure 2-3, it was found that the first differences are symmetric around a zero mean ($p$-value= 0.87) and that these are normally distributed ($p$-value=0.53). For the measured lane width at the site of 12 ft. the estimated average of the lateral fluctuation (0.03 ft.) and standard deviation (2.44 ft.) also suggests that the quality of the data might be sufficient for identifying lane changes accurately when assisted by a robust filter.

It should be noted that a quality check of the digitized lane marking was carried out as well. The marking between lane 3 and lane 4 was re-digitized by a different person. The absolute average and 95th percentile lateral distance between this new and previous digitization were found 0.26 ft. and 0.63 ft., respectively. The standard deviation of this lateral distance was 0.73 ft. Such a small magnitude of deviation attests that the precision of the digitization was satisfactory.
2.3.5 Lane Change Detection Algorithm

This section presents the description of the proposed algorithm for detecting valid lane changes using GPS position and digitized lane marking data. Figure 2-4 shows the critical steps of the algorithm for each trip through the weaving segment.

Figure 2-4: Algorithm for detecting lane changes.

Step 1: Determining Preliminary Lane Positions

After obtaining the pre-processed GPS position for each trip and digitized lane markings, the first task is to determine the preliminary lane positions. Since the width of the lanes varies within the weaving segment, the relative position of the GPS positions with respect to the digitized lane markings were used to pinpoint the preliminary lane position.
To explain this process, Figure 2-5 shows the time-series plot of the position of a trajectory with respect to the shoulder line. To determine the lane position of an observation $i$, the relative position of $i$ as a binary sign (positive or negative) from the different lane markings was determined using the Geosphere tool. Consider that $s_{j,i}$ represents the sign of the $i^{th}$ observation of a trajectory from the marking between lane $j$ and lane $j+1$; it is positive if $i$ is toward the left of the marking and negative if toward the right. Here, lanes are numbered according to Figure 2-2-1, and $j=0$ and $j=6$ represent the shoulder and median, respectively. The vertical arrows toward $i=8$ and $i=9$ in Figure 2-5 show the signs of these points from various markings in two colors: green for positive, red for negative. For example, the sign for $i=8$ from the marking between lane 1 and lane 2 is -ve, and that of $i=9$ from the same marking is +ve.

After determining the signed position of all observations of a trajectory from all markings, the lane position of an observation $i$ ($lp_i$) is determined based on the following condition.

\[
\text{For each marking } j = 0 \text{ to } 5\{ \\
\quad s_{j,i} \neq s_{j+1,i} \rightarrow lp_i = j + 1 \}
\]

Intuitively, the above condition is met only once for all $j$. It should be noted that while the Geosphere utility calculates these distances, it does assume that the markings are straight lines. A curved road should be treated as a series of short straight segments to implement this technique.

**Step 2: Detecting the core of a lane change**

A potential lane change is detected by observing the change in preliminary lane positions. Two successive observations having their preliminary positions in two different lanes make up the core of a lane change. With a 1 Hz resolution, the two points associated with a lane change core are always found in two neighboring lanes, even if a vehicle swiftly traverses two or more consecutive lanes. For example, in Figure 2-5, the observations marked by red circles ($i = 8$ and $i = 9$) make up the core of the lane change as the vehicle went from lane 1 to lane 2.
**Step 3: Growing the Lane Change**

In this step, the two points of the lane change core on the origin and destination lane are tracked backward and forward in time, respectively. The intent is to trace the start and end of the lane change maneuver, which is explained below.

**Start and End of a Lane Change Maneuver:** Consider that the two core points for the lane change shown in Figure 2-5 are \( C \) and \( C + 1 \) \((i = 8 \text{ and } i = 9)\). Then, \( i = 1, 2, 3 \ldots C \) and \( i = C + 1, C + 2 \ldots n \) are the sets of observations on lane 1 and lane 2, respectively. Here, \( n \) is the smaller between the start of another lane change core or the last observation of the trajectory. Then, the start point of the lane change maneuver \( k \) is identified when the following condition is met.

\[
\text{For each observation } i = 1 \text{ to } C \{ \Delta(i, C) > \Delta(i - 1, C) \land \Delta(i, C) > \Delta(i - 2, C) \rightarrow k = i \}
\]

The endpoint of the lane change maneuver \( l \) is identified when the following condition is met.

\[
\text{For each observation } i \text{ in } C + 1 \text{ to } n \{ \Delta(i, C + 1) > \Delta(i + 1, C + 1) \land \Delta(i, C + 1) > \Delta(i + 2, C + 1) \rightarrow l = i \}
\]
Here, \( \Delta(.) \) is a function to calculate the lateral displacement between two positions with respect to a reference line.

These conditions state that a lane change maneuver’s starting point is marked by an increasing displacement toward the lane change direction at least over two consecutive observations. The two consecutive periods are included in evaluating the conditions to minimize the effect of the randomness and bias in the GPS data. According to Figure 2-4, the 3\(^{rd}\) observation shows an increasing displacement compared to the two preceding observations, which makes it the starting point of the lane change maneuver. Conversely, the 11\(^{th}\) observation is marked as the end of the lane change maneuver, where the displacement decreased over the next two consecutive time periods. Therefore, the core of lane change between \( i=8 \) and \( i=9 \) is grown up to \( i=3 \) (backward) and \( i=11 \) (forward).

**Step 4: Implementation of Filters**

Because of the randomness and possible bias in the GPS data, not all potential lane changes are valid. To tackle this issue, the lane change detection is constrained by the temporal duration of the vehicle on the target lane and the lateral displacement within a grown lane change core. According to the first filter, a potential lane change is invalidated if the trajectory goes to the target lane only for one second and comes back to the origin lane in the next one. Application of such a single point removal process is common in handling position data (Patterson, McConnell, Fedak, Bravington, & Hindell, 2010).

The second filter implies that the lateral displacement between the start and the end of a lane change maneuver that has passed the single point threshold must be higher than a certain value to declare that as valid. Implementation of this threshold is reasoned based on the fact that a minimum lateral displacement occurs during a lane change.
At this point, it is imperative to discuss the value of the lateral displacement threshold. Past studies on lane change maneuvers reported that the lateral displacement of a vehicle during lane changes is about one lane width (Xuan and Coifman, 2006). However, because of the randomness and bias in the data, it is difficult to just pick a single cut-off value for this parameter. Had the GPS positions and digitized map been sufficiently accurate, a threshold equal to the average lane width would be reasonable. Since the real-life scenario is far noisier, the analyst investigates the sensitivity of the predicted correct and false number of lane changes to different lateral displacement thresholds.

Here, the temporal extent threshold was used as an initial filter so that the apparent false lane change alarms generated from overlapping the GPS position of the designed field trips on the digitized lane marking can be excluded. On the other hand, the lateral displacement threshold was applied to fine-tune the algorithm. There are two reasons why the first threshold was arbitrarily selected, while the second one was optimized. First, a temporal extent threshold that is optimized based on the designed field trips data might reduce the total error of the algorithm. However, the application of that temporal threshold on a naturalistic trajectory dataset (for which ground truth is unknown) may remove some valid lane changes that happened over a short period. For avoiding any unsafe maneuver, no such quick lane changes were made when the designed field trips were conducted. However, in naturalistic driving data, such a quick lane change may be unlikely but not impossible. On the other hand, the lateral displacement threshold has a minimum limit (i.e., a vehicle has to move laterally at least by its own width to move its body from one lane to another). Hence, the researcher can qualitatively assess the optimization of the lateral drift threshold, but cannot do the same for the temporal extent threshold.
The second reason is that the temporal duration of a lane change depends on the speed and heading angle of the subject vehicle when changing lanes. These two variables can be significantly influenced by the surrounding traffic condition. In a naturalistic driving dataset, traffic condition may vary significantly than that when the field trips were made. As a result, a temporal extent threshold that is optimized based on the designed field trips data may not be applicable to the naturalistic driving data. On the other hand, the lateral displacement between the start and end of a lane change maneuver is not likely to be significantly influenced by the traffic condition. For these reasons, the temporal extent threshold is selected arbitrarily as an initial filter to remove the apparent false positive lane changes, while the lateral displacement threshold was optimized to fine-tune the algorithm.

If a potential lane change passes both the single-point and a certain lateral displacement threshold, it is considered to be a correct positive if the video record verifies it. Conversely, a false positive lane change is one that did not happen per the video record but passes the two thresholds. A false negative lane change is one that is verified by the video record but does not pass the thresholds. In this study, the lateral displacement threshold that gives the minimum number of false positives and false negatives is selected for the algorithm. Others may emphasize minimizing false negatives as the core objective, by using a lower threshold value.

It should be noted that according to this algorithm, the first two preliminary lane positions of a trajectory within the study site are accepted as valid. This assumption may create some errors in the lane change detection by the algorithm.

2.3.6 Characterization of Lane Changes

The weaving segment chapter of the HCM6 classified lane changes as weaving and non-weaving vehicles. While this classification is useful, lane change measures based on entry and
exit lanes provide detailed information on the excess lane change pattern more than the weaving and non-weaving based rates. This study defines three classes of lane change types in weaving segments. Those are based on the entry and exit lane, or the origin and destination (freeway or ramp) observed for each vehicle within the weaving section. The three classes are (a) the minimum number of lane changes per vehicle assuming proper pre-positioning, (b) the required number of lane changes based on entry and exit lanes, and (c) the actual number of observed lane changes.

The lane configuration of the test-site and the vehicle shown in Figure 2-2-1 can be referred here to aid the discussion. In this figure, the vehicle in lane 5 is going from the freeway to ramp and it is evident that (a) the minimum number of lane changes is zero, had the subject vehicle in lane 5 was correctly pre-positioned in lane 2 to exit. Next, (b) the required number of lane changes to exit assuming entry in lane 5 is three, and finally, the actual number of lane changes is five, as shown in the figure. Here, the minimum number of lane changes is strictly a function of the segment configuration, while the required number is of the entry and exit lanes. The difference between actual and required lane changes is termed as “excess to required”. In terms of aggregate weaving and non-weaving vehicles, only traffic weaving from the on-ramp to the freeway must execute a minimum of two lane changes in this example. All other lane changes, termed as “excess to minimum” would be considered discretionary.

Although the naturalistic trajectory dataset used in this study did not come from a representative sample of the traffic population in the study area, it represents the behavior of typical commuter drivers in this particular weaving segment. With the provision of infrastructure-to-vehicle (I2V) connection, where local agencies can control the excess lane changes, it is possible to reduce those numbers. Suppose, providing additional guidance to the
vehicles can help them pre-positioning prior to entering the weaving segment. This technique can reduce the “excess to minimum” lane changes. Banning lane changes within a portion of the weaving area can lessen the “excess to required” lane changes.

In addition to the excess number of lane changes, the lateral speed and duration of lane changes is considered imperative for freeway capacity and safety analysis (Moridpour et al., 2010). It is estimated as the ratio of lateral displacement to the duration between the start and end of a lane change. The spatial distribution of mandatory lane changes for the weaving vehicles is another critical factor for weaving segment analysis (Daganzo, Laval, & Muñoz, 2002). For example, changing lanes to take the exit toward the end of the gore at the weaving segment can be an indicator of aggressive or ignorant driving behavior. Hence, the identified lane changes are analyzed to estimate their lateral speed and spatial distribution of the weaving lane changes with respect to the entry gore.

2.4 RESULTS

In this section, results from the lane change detection and characterization are presented. First, findings from the designed experiment are discussed. Then, the application of the algorithm on the naturalistic trajectory data is presented.

2.4.1 Designed Experiment

2.4.1.1 Determining Lane Change Cores

All observations from the designed experiment were mapped onto the digitized road to identify the preliminary lane positions and the core lane changes. In total 265 potential lane changes were signaled after applying the first two steps of the algorithm in Figure 2-4, compared to 46 ground truth lane changes. Out of these 265, 220 were false positives and 45 correct positives. Even before implementing the filters, the GPS data did not enable the detection of one
ground truth lane change. Regardless of this false negative, the sheer number of false positives mandated the use of filters to improve the predictions.

2.4.1.2 Filtering Valid Lane Changes

Initially, the single point threshold was applied to remove the lane changes in which the trajectory stayed on the destination lane for only one second. Implementing this filter introduced another false negative lane change. However, the total estimated number of lane changes dropped from 265 to 76, including 32 false positives and 44 correct positives. To estimate the optimum value of the lateral displacement filter, variations in the total, correct positive, false positive, and false negative lane changes estimated by varying that threshold are plotted in Figure 2-6.

![Figure 2-6: Sensitivity of lane change outcomes to displacement threshold value.](image)

Figure 2-6 shows that as the threshold increases from 0 to about 20 ft, the total estimated number of lane changes reduces drastically. The estimated false positive follows the same trend. The number of false negative lane changes remains only two up to a 3.28 ft. threshold, but then it
starts increasing. At a 6.9 ft. threshold, the total errors (false positive + false negative) is minimized. Above this threshold, the number of false positive decreases but the number of false negative increases sharply. Therefore, the lateral displacement threshold was selected as 6.9 ft. It resulted in 42 correct positive, two false positive, and four false negative lane changes, that is, about 4.35% and 8.7% false positive and false negative errors, respectively. Those compare to 2.2% and 30.4% positive and negative errors had a simple standard lane width of 12 ft. threshold were used.

2.4.2 Application to the Naturalistic Data

The proposed lane change detection algorithm was applied to the naturalistic trajectory data. This application yielded the following origin-destination based counts: 281 freeway-to-freeway (FF), 165 freeway-to-ramp (FR), 143 ramp-to-freeway (RF), and 48 ramp-to-ramp (RR). For the lane configuration at the test-site, the minimum number of lane changes is two for the RF trips and zero for all others. Hence, the total minimum number of lane change is $143 \times 2 = 286$. In the following subsections, the naturalistic trajectories are characterized based on their excessive number of lane changes, lateral speed, and their spatial distribution of weaving lane changes.

2.4.2.1 Excess Lane Changes

The distribution of the estimated total, excess to minimum, and excess to required lane changes from the naturalistic trajectory data are shown in Figure 2-7. The values in parenthesis above each bar show the lane change rate per trip of corresponding origin-destination type.
From Figure 2-7, it is apparent that both excess to minimum and excess to the required lane changes are highest for FF trips. On a per trip basis, RF trips have the highest rate (1.1 per trip) of excess to minimum lane changes. RR trips have the highest rate of excess to required lane changes (0.35 per trips) which could be attributed to the small sample size of this origin-destination type. Table 2-2 shows the distribution of lane change rates by various origin-destination lanes.

Table 2-2: Total lane change rates by lane origin-destination.

<table>
<thead>
<tr>
<th>Orig.\Dest.</th>
<th>Aux- Lane 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aux- Lane 1</td>
<td>0.33</td>
<td>1.33</td>
<td>2.43</td>
<td>3.11</td>
<td>4.00</td>
</tr>
<tr>
<td>2</td>
<td>1.10</td>
<td>0.36</td>
<td>1.00</td>
<td>2.43</td>
<td>3.33</td>
</tr>
<tr>
<td>3</td>
<td>2.00</td>
<td>1.12</td>
<td>0.41</td>
<td>1.86</td>
<td>2.00</td>
</tr>
<tr>
<td>4</td>
<td>3.00</td>
<td>2.00</td>
<td>1.53</td>
<td>0.14</td>
<td>1.13</td>
</tr>
<tr>
<td>5</td>
<td>4.00</td>
<td>3.00</td>
<td>2.33</td>
<td>1.20</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Intuitively, the lane change rates in Table 2-2 are at the minimum for diagonal cells and maximum for cells away from the diagonal. Two types of weaves can be classified by the diagonal of this matrix. Shaded cells below the diagonal are for FR and those above the diagonal.
are for RF. Since the minimum lane change for FR trips is zero, the excess to minimum lane change rate is highest toward the lower left corner of the matrix. For example, origin-destination pairs 1-5 and 5-1 both have a lane change rate of 4 per trip, but the excess to minimum rate is higher for 5-1 (4 per trip) than that for 1-5 (2 per trip). Among all non-weaving pairs, 2-5 exhibited the highest excess to minimum lane change rates (3.33 per trip). Such excess to minimum lane change tendencies can be significantly reduced by providing guidance to the vehicles before reaching the weaving segment to help them pre-position.

Note that the required number of lane changes for any pair of lane origin-destination in Table 2-2 can be calculated simply as the difference of the lane position numbers of each pair. For example, lane origin-destination pair 3-4 has one required lane change. From that, the excess to required lane change rate for this pair is estimated as 0.86 per trip. Interestingly, lane pair 3-4 and 4-3 have a higher rate of excess to required lane change rate (0.86 and 0.53, respectively) than the rest. The reason behind these high lane change rate between lane 3 and lane 4 could be attributed to a number of phenomena like overtaking and cooperative lane change tendency of the vehicles. With I2V connection, banning lane changes between lane 3 and 4 would reduce this high rate of excess to required lane changes.

2.4.2.2 Lateral Speed during Lane Changes

Figure 2-8 shows the fitted probability density function of the lateral speed during lane changes of trajectories on different paths. The average lateral speed varied from 2.60 (for FR trips) to 2.80 fps (FF trips). The average lane change duration was 5.5 second, which is close to an average of about 5 sec reported by several past studies (Toledo & Zohar, 1999; Moridpour et al., 2010).
Figure 2-8: Distribution of lateral speed of lane changes.

The distributions of lane change speeds across paths do not exhibit any significant difference. Nonetheless, there is a group of lane changes by FF and RR trips with noticeably higher speed resulting in a second mode at 7.7 fps. About 31% of all the lane changes had duration more than 6 seconds in contrast to 26% had less than or equal to 3 seconds. The motivations behind the high-speed and short-duration lane changes need further investigation.

2.4.2.3 Position of the Weaving Lane Changes

Locating lane change starts is important for three weaving lane pairs. Lane 1 to 2 and Lane 2 to 3 represent the two mandatory lane changes for the RF trips (labeled as RF 12 and RF 23 in Figure 2-9). Although the minimum (theoretical) number of lane changes for FR trips is zero, many such trips that were initially not on lane 2 must first proceed from lane 3 to 2. For this reason, this lane pair is included in the analysis (labeled as FR 32). Figure 2-9 shows the spatial distribution of the lane change locations by weaving vehicles between these three lane pairs.
Figure 2-9: Distribution of weaving lane change distances from the entry gore.

Figure 2-9 shows that the distribution modes for both FR and RF weaving lane changes are quite close. This situation creates maximum turbulence for the weaving maneuvers at a critical section at about 800 ft. from the entry gore.

The distribution of lane changes for FR trips going from lane 3 to 2 is flatter than the other two with an average value of 2,300 ft. from the entry ramp. Conversely, the first mandatory lane change for RF trips (from 1 to 2) shows a distinct skewness to the right. In addition, all the lane changes for these trips took place within about 3300 ft. from the entry gore (62.5% of the weaving length). Moreover, the mode of the RF 12 distribution (at 500 ft.) indicates that those vehicles quickly exited that first lane much sooner than other vehicles, once they entered the freeway from the ramp. The distribution of the second mandatory lane change by the RF trips (RF 23) is somewhat skewed to the right with a mean of about 2200 ft. However, the tail is extended up to the exit gore, which could be attributed to the fact that some drivers realize very late that lane 2 is an exit lane from the freeway.
The findings from the weaving segment lane change location are very similar to what was obtained by previous studies using the total traffic population data (Marczak et al., 2014; van Beinum et al., 2018). These also suggest that most of the weaving lane changes are occurring within a confined section of the road, where the friction incurred by the weaving lane changes is probably the highest. From the perspective of traffic control in a connected environment, these findings invoke the hypothesis that spreading the spatial extent of the weaving lane changes by providing some control guidance can improve the operational condition of the segment.

2.5 SUMMARY

This study presents an algorithm for detecting lane changes using GPS position and digitized lane marking data. A low-cost GPS receiver equipped with cellular communication capability was used which transferred second-by-second vehicle position in real-time. A freeway weaving segment was selected to characterize the lane changes of 637 naturalistic trajectories through it. An experiment was designed to calibrate the threshold parameters of the proposed algorithm. The identified lane changes were characterized by route-level and lane-level discretionary lane change rate, their lateral speed, and spatial distribution.

A quality assessment of the data showed that the fluctuation of lateral displacement of the vehicle from a reference line while keeping a lane varied symmetrically with a mean approximately zero and a standard deviation of 2.44 ft. It indicated that the quality might be sufficient to detect lane changes with the assist of a robust filter. Mapping all the designed experiment data on the digitized lane marking yielded 220 false positive lane changes, which was reduced to 32 by the application of the single-point, single second filter. By estimating the correct positive, false negative, and false positive lane changes for a range of lateral displacement thresholds, an optimal value of 6.9 ft. was found to minimize the total error. Out of
46 ground truth lane changes, the algorithm detected 42 correctly with two false positives, resulting in a 4.35% and 8.7% false positive and false negative error, respectively.

Application of the algorithm to the naturalistic trajectories estimated that freeway-freeway trips have the highest number of excessive lane changes. Ramp-freeway trips have the highest rate of discretionary lane changes calculated from their route-level minimum rate. Lane position based lane change rates revealed that poor positioning of vehicles can result in a high rate of lane changes. Also, excess lane changes rates were noticeably high between two non-weaving lanes in the study site. These results indicate that strategies based on infrastructure to vehicle connection could be used to reduce excess lane change tendencies of regular commuters.

The estimated average lateral speed and duration of lane changes were found close to norm: 2.7 fps and 5.5 sec., respectively. However, several freeway-freeway and ramp-ramp vehicles changed lanes at lateral speeds of about 7.7 fps. The spatial distribution of the weaving lane changes with respect to the entry gore revealed that ramp-freeway vehicles changed their first mandatory lane within the first 62.5% of the total weaving length. Although the other two weaving lane changes occurred over the entire segment, the distributions showed that there is a critical short section within the weaving segment. Weaving vehicles are likely to generate very high friction at that location.

One limitation of this study is that the proposed lane change detection algorithm should be tested at different highway locations, preferably with differences in curvature and geometry. In addition, the accuracy of the digitization process was not assessed. Nonetheless, this study demonstrates a successful application of Vehicle to Infrastructure connectivity through a low-cost IoT to detect and analyze lane changes. With increasing availability of such a low-sample yet high-resolution connected vehicle data, more information regarding driver behavior at
different road geometries can be obtained with sufficient accuracy. Imminent availability of infrastructure to vehicle connectivity will enable monitoring and controlling aggressive driving behavior such as frequent and high-speed lane changes. In addition, controlling the location of the mandatory lane changes may improve the capacity of weaving bottlenecks. Future research regarding this topic include implementing these control strategies in a simulation environment to investigate their impacts on traffic mobility and safety.
2.6 REFERENCES


Shoaib Samandar, Billy Williams, and Ishtiak Ahmed. (in press). Weigh station impact on truck travel time reliability: Results and findings from a field study and a simulation experiment. *Transportation Research Record: Journal of the Transportation Research Board,*


CHAPTER 3 – MINING LANE CHANGING BEHAVIOR FROM TRAJECTORY DATA: CHARACTERIZATION AND EXTREME DRIVING IDENTIFICATION

3.1 INTRODUCTION AND RESEARCH OBJECTIVES

Extreme lane changing behavior, especially in high-density traffic, is well-known to have detrimental effects on both mobility and safety of freeway sections (Moridpour, Sarvi, & Rose, 2010a; Pande & Abdel-Aty, 2006; Sun & Kondyli, 2010). Such behavior may reduce the quality of service of the roads by activating bottlenecks or causing traffic stream instability, and may lead to crashes.

Among different aspects of lane changing behavior—lane change frequency, speed, acceleration, headway, and steering angle during changing lanes—have been key topics addressed by past studies (Rathbone & Huckabee, 1999; Schmidt, Beggiato, Hoffmann, & Krems, 2014; Sun & Elefteriadou, 2012; Tasca, 1999; van Winsum, De Waard, & Brookhuis, 1999). Discretionary lane change frequency has been considered as an indicator of aggressive driving in several studies (Bham, 2011; Sun & Elefteriadou, 2012; Tasca, 1999). A key motivation for changing lanes discretionarily is to gain speed. However, the extant literature does not address patterns of lane change frequency or the ability to gain speed by discretionary lane changes under varying traffic densities.

Winsum et al. (1999) considered lane changes occurring at high speed and at a sharp angle to be risky because driving error increases with increasing steering wheel rotation and speed of the vehicle. They combined these two measures, along with the vehicle’s relative position in a lane, to define a safety measure called “time-to-line-crossing.” They used this measure in a driving simulator study. While other researchers (Eidehall, Pohl, Gustafsson, &
Ekmark, 2007; Lin & Ulsoy, 1995) used this measure to develop a lane guidance system, it has not been used to characterize lane changing behavior of all drivers in a traffic stream.

With the prospective proliferation of connected vehicles (CV) on the road, it becomes possible to analyze individual driver behavior and to identify extreme lane changing characteristics. However, the current low market penetration rate of CVs enables one to observe the behavior of only a small sample of drivers in a traffic stream, which may not be representative of the behavior of all drivers. In this regard, high-resolution trajectory data from a real-word traffic stream are a useful surrogate for a future 100% CV scenario. Therefore, this chapter develops a framework for analyzing extreme lane changing behavior using archived high-resolution trajectory data.

This study has two primary objectives: i) to assess lane changing behavior of drivers and its variation under different congestion levels using a real-world and complete trajectory dataset, and ii) to define and identify levels of extreme lane changing behavior in that dataset.

To meet those objectives, the researcher explores and analyzes lane changing behavior of drivers to reveal changes in patterns with traffic conditions, to illuminate microscopic characteristics, and to quantify extremeness in lane changing maneuvers. High-resolution trajectory data from the US-101 site of the Next Generation SIMulation or NGSIM effort (USDOT, 2006) were used. A detailed investigation was carried out regarding the quality of the vehicles’ lane positions reported by NGSIM and the data were corrected accordingly. The modified dataset was then used to carry out the remaining analyses.

The following section presents a review of past studies on the topic of lane changing behavior. Trajectory data are explored next. Then, the mechanism to pre-process the data is described. Techniques and key findings regarding driver lane change behavior under different
congestion levels are presented in the following section. Next, extreme lane changing behavior is defined and identified from the data. The last section presents concluding remarks.

3.2 LITERATURE REVIEW

Several microscopic lane change models have been developed considering the differences in driver behavior during lane changes (Moridpour, Sarvi, & Rose, 2010b; Rahman, Chowdhury, Xie, & He, 2013). The ARTemis model (Hidas, 2005) considered the courtesy of the lag vehicle when modeling lane changes. MOBIL model (Kesting, Treiber, & Helbing, 2007) considered politeness factor in driver interaction. The LMRS model (Schakel, Knoop, & van Arem, 2012) considered a driver’s level of desire to make lane changes and level of cooperation by the vehicles in the target lane. The particle hopping model (Chowdhury, Wolf, & Schreckenberg, 1997) considered asymmetric rules of anticipation. While these models are useful for traffic microsimulations, they did not consider extreme driving behavior. Analysis of individual vehicle trajectories in terms of lane change extremeness requires customized methods. Neural Driver Agents (NDAs) is a neural network-based model trained by real driving data (Dumbuya et al., 2009), and can replicate extreme lane-changing behavior. However, NDA’s driver extremeness is limited by the training data. Moreover, NDA’s trajectory formulation is erroneous in extreme speed conditions (Tomar, Verma, & Tomar, 2010). In contrast to microscopic models, macroscopic lane change models mostly focus on overall lane change frequency and its variation with road geometry and traffic conditions (Seraj, Bie, & Qiu, 2017; Transportation Research Board, 2016). Individuals’ lane changing characteristics are not considered in these studies.

Traffic congestion level was found to influence lane changing behavior of drivers (Lv, Song, Fang, & Ma, 2013; Sun & Kondyli, 2010). Lv et al. (2013) investigated the efficiency gains from lane changing at different density conditions in an arterial bottleneck location. They
showed that efficiency gains from lane changes are possible in medium to low density conditions. However, the gains diminish in high density conditions. Lee, Olsen, and Wierwille (2004) investigated the duration, urgency, and severity of naturalistic lane changes for sixteen drivers. They investigated the forward and rearward area of the vehicles and found that a slow-leading vehicle creates an incentive to change lanes. They defined lane change severity using the presence of a vehicle in the passing zone, which is an area in the destination lane. They found that high severity lane changes are related to the presence of a slow-moving vehicle in front. Similar findings came from Oketch (2000), who showed that a driver with extreme maneuvering tendency is more likely to accept a smaller gap for changing lanes when hindered by a slow vehicle. All these studies require an understanding of gaps in the destination lane. Consequently, the extremeness metrics are developed around gap-based measures. Although it is confirmed by these studies that a lane change can be extreme on the basis of other metrics, none of them discussed any method to classify lane change maneuvers based on data coming from the subject vehicle only.

The NGSIM data are a unique source of microscale traffic information providing a comprehensive view of all the vehicles within a stream. However, their quality has been questioned in several past studies. Coifman and Li (2017) showed that for several leader-follower pairs in the I-80 dataset, the reported gaps imply that the two vehicles are colliding. Furthermore, this and other studies have verified that the reported accelerations have unrealistic values in several NGSIM datasets. Montanino and Punzo (2013) reconstructed the trajectories from several NGSIM datasets by constraining extreme values of acceleration, jerk, and the number of sign inversions of jerk. Thiemann, Treiber, and Kesting (2008) proposed a smoothing algorithm for the position, speed, and acceleration data provided by NGSIM. They reported that
the original dataset contained trajectories that crossed multiple lane markings in quick successions. An empirical rule was used to remove lane changes that are either followed or preceded by another lane change within 5 seconds.

### 3.3 NGSIM US-101 DATASET

#### 3.3.1 Site Description

The dataset used in this study were collected using high-resolution videos under the NGSIM initiative (USDOT, 2006). It is a (5+1)-lane ramp weave segment of US-101 in Los Angeles, CA, as shown in Figure 3-1. The weave length gore-to-gore is 698 ft. The trajectory data were collected on a weekday from 7:50-8:35 AM at 10 Hz. frequency over a longer distance of 2,100 ft. The data are divided into three 15-minute periods, which are defined as period 1 to period 3. Individual lanes were labeled as 1 (leftmost lane) to lane 5 followed by the auxiliary lane “Aux”.

Each trajectory observation includes the lane number in which the vehicle was present, vehicle speed, position in terms of two local spatial axes (Local X and Local Y), acceleration, and time and space headway to the nearest lead vehicle. The Local X and Local Y are approximately aligned with perpendicular and parallel to the roadway segment, respectively. Thus, an analyst can track the frequency and position of all lane changes.

![Figure 3-1: Schematic of the US-101 weaving segment.](image)
3.3.2 Traffic Characteristics

The dataset contains records for 4,824 vehicles, of which 4,681 are passenger cars, 105 are heavy vehicles, and 38 are motorcycles. As lane change patterns for motorcycles and heavy vehicles tend to differ from those for passenger cars (Moridpour et al., 2010a), motorcycles and heavy vehicles were excluded from further analyses.

About 90% of the vehicles were non-weaving, and are not required to make any lane change. Hence, most lane changes in the dataset were discretionary. For further context, this section was heavily congested during the 45-minute data collection period, with congestion worsening from period 1 through period 3, and average traffic speeds dropping from 47 to 37 to 32 fps across the three periods. In addition, and because the auxiliary lane had little use due to the paucity of weaving traffic, speed on that lane tended to be higher than that on the mainline lanes in the vicinity of the weave.

Figures 3-2 (a) and 3-2 (b) show a top view of the portion of lane changing vehicle paths immediately before and after each lane change toward the left and the right, respectively. (In these and other figures, the flow of traffic is from left to right.) The overall pattern apparent is that vehicles moved more frequently toward the left (1,206) than toward the right (552), presumably to avoid the turbulence caused by the weaving traffic. From the perspective of relative lane positions, the left lane(s) usually experience a higher speed than the right lanes. Therefore, the pattern in Figure 3-2 is not surprising for a congested freeway segment. These figures also show that many weaving lane changes (i.e., between the auxiliary lane and lane 5) took place beyond the entry and exit gores, which are marked by the two vertical black lines. This observation indicates that many drivers changed lanes over the solid lane markings that
define the gores. This behavior may result from the difference in density between lane 5 and the auxiliary lane, which tempted the drivers to change lanes over the solid lane markings.

Figure 3-2: Top view of lane changing trajectories (a) toward the left (b) toward the right.
3.4 DATASET ASSESSMENT

The researcher employs a two-panel display of each trajectory to gain insights into the lane changing behavior of drivers without having to pore over numerical details. Originally developed by Karr (2019), this display reveals changes in the spatiotemporal location, lane position, and speed during lane changing.

The following sections highlight some trajectories from the dataset in Figures 3 through 6. In each, the top panel shows a top view of the segment with the vehicle’s lateral position on the vertical Local X axis and longitudinal position on the horizontal Local Y axis. A unique color is assigned to each lane, and changes in color indicate a lane change. A downward track of the trajectory in this panel signals one or more lane shift to the right, and vice versa. Recall that lanes 1 through 5 represent mainline lanes (left through right) while Aux represents vehicles on both ramps as well as on the auxiliary lane in the weaving segment. In addition, the two vertical lines at approximately 615 ft. and 1,350 ft. represent the entry and exit gores, respectively, highlighting the location of speed and lane changes relative to the ramp positions. The bottom panel shows a traditional speed vs. clock time trajectory, again with lane changes indicated by a change in the point color. Here, $t = 0$ represents the entry point of the vehicle on the site. The two panels are connected through the lane position color scheme. The vehicle ID and period number were combined to create a unique vehicle identification code.

The power of the two-panel display is the ability to relate multiple spatiotemporal attributes, for example, by exploring speed changes during lane changes and the spatiotemporal intervals between successive lane changes. Note that the horizontal axes (Local Y in the upper panel and time in the lower panel) do not align. The trajectories in Figures 3 through 6 exhibit
some interesting as well as unusual lane changing patterns. The interpretations are presented next.

### 3.4.1 Visualizing Lateral Drifting

Figure 3 shows a trajectory that seemingly moved between lane 3 and lane 4 several times. Note the short time intervals and minimal lateral drifts associated with some of these lane changes marked by the red ovals. The vehicle “moved” between these two lanes five times between $t = 25$ and $t = 28$ seconds, over a space of less than 200 ft. Drivers often move to the neighboring lane and come back to the origin lane to pass a slow-moving vehicle, but doing so multiple times over such short intervals is physically infeasible. This illustration raises the question of whether these pairs of lane changes were executed to complete a passing maneuver or instead, are “false alarms” due to lateral drifting of the vehicle that caused the lane assignment algorithm of NGSIM to fail to identify the correct lane position. Section 5 presents a method to identify (and hence) eliminate this seemingly problematic observation(s).

![Figure 3](image.png)

**Figure 3-3**: Example of lateral drifting by vehicle ID 1446 in period 1.
3.4.2 Speed Gain by Changing Lanes

Figure 3-4 depicts the trajectory of a vehicle entering the segment on the leftmost lane and remaining on the freeway. It shows two discretionary lane changes with the seeming intent to avoid the stop and go condition in lane 1 during the period 25—35 seconds. This speed drop motivated a lane change to the right about halfway between the two gores and another one about 15 seconds later into lane 3. These lane changes yielded a considerable increment (68%) in speed. It would be interesting, although somewhat difficult and quite speculative, to predict the counterfactual speed outcome had the vehicle remained in lane 1 throughout the entire time.

![Figure 3-4: Example of speed gain by changing lanes discretionarily.](image)

3.4.3 Lane Change Frequency, Entry Angle, and Speed

Figure 3-5 shows a trajectory with five discretionary lane changes. After entering the segment through lane 4, the vehicle moved both left and right, but exited through lane 5 at the end. This trajectory was observed in period 3, when the overall congestion level was highest. In addition to the significant number of discretionary lane changes in such a congested condition, the entry angle onto the destination lane when the vehicle went from lane 4 to lane 5 at $t = 100$
seconds is quite sharp (marked by the red oval). This is evidenced by the highlighted steep slope of the line in the top panel. However, this action is tempered by the fact that those entries took place at a low speed, as shown in the bottom panel. Section 6 presents a detailed analysis of extreme driving identification based on lane change frequency and combining the entry angle and speed during lane changes.

![Figure 3-5: Example of changing lanes frequently and with a sharp entry angle.](image)

### 3.4.4 Summary of Dataset Assessment

The trajectory visualization analyses provide the following insights:

- Lane changes in quick succession between two neighboring lanes raise the possibility of false lane changes due to lateral drifts by the vehicle or incorrect lane assignments by NGSIM.
- Gaining speed is one of the primary motivations for discretionary lane changes. However, its causal relationship with lane changing and congestion level needs further investigation.
• Since a lane change combining both high-speed and a sharp entry angle can be an extreme maneuver, a parameter incorporating these two variables while changing a lane should help identify extreme lane changes. Further, some trajectories exhibited multiple lane changes in quick succession—an indicator of extreme driving behavior that also needs to be considered.

The analyses presented in the remainder of this chapter is motivated by the above observations. It is evident that the reported lane changes from the NGSIM dataset need to be scrutinized to remove possible false lane changes, which is discussed in the next section. Following that, the analyses of lane change characteristics, their relationship to traffic congestion level, and extreme lane changing behavior identification are conducted.

3.4.5 Dataset Pre-Processing

Pairs of lane changes that took place in quick succession (Figure 3) warrant an investigation of the NGSIM dataset for false lane change alarms. This issue was initially reported by Thiemann et al. (2010). They removed all lane changes that were either preceded or followed by another lane change within five seconds. However, this use of a temporal threshold was quite subjective—no minimum value for the space gap between two “real” lane changes was reported.

A more precise way of validating the NGSIM reported lane changes is to check the original videotapes manually. However, since there are 1,758 reported lane changes, this process would be excessively time- and labor-intensive. Hence, the researcher first made a list of lane change pairs where following the first lane change, the subject vehicle came back to its previous lane. The researcher investigated these lane changes based on two parameters: i) the temporal duration of the vehicle remaining in the destination lane before coming back to the origin lane (Δt), ii) the distance by which the vehicle moved laterally into the destination lane from the lane
marking, measured based on the reported Local X in the raw data ($\Delta X$). Figure 3-6 (a) explains these two parameters for a pair of lane changes. This pair of lane changes should be tagged as invalid if the value of $\Delta X$ or $\Delta t$ (or both) is too small to be physically feasible for two complete lane changes. One can, therefore, employ a reasonable threshold value for these parameters to decide on the validity of these lane changes. Note that this spatiotemporal threshold-based technique is similar to technique introduced in Chapter 2 to detect false positive lane changes derived from in-vehicle GPS sensor data.

A plot of $\Delta X$ vs. $\Delta t$ for these 340 lane changes is shown in Figure 3-6 (b). It is apparent that there is an accumulation of data points in the region bounded by $\Delta X < 3$ ft. and $\Delta t < 4$ seconds. Note that a vehicle must laterally shift at least half of its width into the destination lane to move its entire body from one lane to another. The dashed horizontal line in this figure at $\Delta X = 3$ ft. represents approximately half of the width of a standard passenger car (Harwood, Torbic, Richard, Glaуз, & Elefterdiadou, 2003). Therefore, observations below this line are not likely to be completed lane changes. Instead, these are lateral drifts by the subject vehicles where a fraction of the vehicle body barely crossed the lane marking due to drifting and came back toward the center of the lane. This reported lane position could also be due to measurement or coding errors from the raw videos. The proposed threshold, however, may remove some lane changes that are valid because the Local X is not uniformly perpendicular to the travel direction of vehicles. In addition, Figure 3-6 (b) shows that some lane changes below this threshold have reasonably long $\Delta t$. However, this plot also guarantees that the number of such cases is not significant since most lane changes below this threshold have unrealistic combinations of $\Delta X$ and $\Delta t$. 
Figure 3-6: (a) $\Delta X$ and $\Delta t$ for a lane change followed by another lane change in the opposite direction (b) $\Delta X$ vs. $\Delta t$ for all lane changes that are either followed or preceded by a lane change in the opposite direction.

The application of the 3 ft. threshold removed 238 lane changes from the NGSIM US-101 dataset, which constitute about 13% of the total lane changes made by passenger cars. The remaining 102 lane changes that seem to reflect quick passing maneuvers were mostly recorded in the first period when traffic was least congested. The dataset was updated accordingly by correcting the corresponding lane assignments. The remaining analyses presented here were carried out using the revised dataset.

3.5 LANE CHANGING BEHAVIOR AND CONGESTION LEVELS

This section first discusses the variation of overall lane changing frequency of vehicles with varying congestion levels over the three periods. Then, the speed gain after changing lanes at different congestion levels is analyzed to evaluate the motivation for changing lanes. In this context, the researcher contrasts the trajectory-level speeds of trajectories with at least one lane change against those of trajectories with no lane changes.
3.5.1 Lane Change Frequency and Congestion Level

The NGSIM US-101 data are suitable for investigating the variation in lane changing tendency with traffic stream speed since space mean speed (SMS) of traffic dropped from 47 to 37 to 32 fps in the three 15-minute periods. Figure 3-7 shows the numbers of trajectories with different lane change frequencies in each period. A few vehicles changed lanes five or six times within the 2,100 ft. segment, even though only one mandatory lane change is required for weaving vehicles. Overall, the average lane change frequency was only 0.33 per trajectory. A notable observation from Figure 3-7 is the reduction in lane change frequency with increasing congestion level, as the second and third 15-minute periods were more congested than the first one. In fact, the number of lane changes decreased by 33% between the first and the last period. This observation agrees with the results of a past study from a completely different environment (Bonneson, 1998).

![Lane Change Frequencies by Period](image)

Figure 3-7: Number of vehicles with different lane change frequencies by period.
3.5.2 Speed Gain by Changing Lanes and Congestion Level

The reasons for the reduction in lane change frequency with decreasing traffic speed as shown above could be two-fold. First, space is very limited to execute lane changes frequently in a congested condition. Second, the opportunity for gaining speed by individual vehicles also decreases with increasing congestion as traffic density increased from 48 to 60 pc/mi/ln between the first and the last period. The first postulation is very intuitive because indeed, traffic density is inverse to the space headway between successive vehicles. For investigating the second hypothesis, two sets of analyses were conducted. The first addresses individual lane change events: the speed gain or loss ($\Delta v_n$) following the $n$th lane change is estimated. To that end, the average speeds of the trajectory while it was on the origin and destination lane were estimated separately. For the $n$th lane change, the average speed of the trajectory while it was on the origin lane ($\bar{v}_{o,n}$) is estimated from the point in time it entered that lane to the time the $n$th lane change happened. Note that for the first lane change, i.e., $n = 1$, the time at which the trajectory entered the origin lane is the time it appeared within the study area. For other lane changes, i.e., $n > 1$, the time at which the trajectory entered the origin lane is the time instant right after the $(n-1)^{\text{st}}$ lane change. Similarly, the average speed of the trajectory while it was on the destination lane ($\bar{v}_{d,n}$) is estimated from the time instant right after the $n$th lane change to the time it left the destination lane or went outside the coverage area. To further clarify, refer to the second lane change by the trajectory shown in Figure 3-5. For this lane change from lane 3 to lane 2, $\bar{v}_{o,n=2}$ is estimated as the distance traversed between $t = 32$ seconds and $t = 42$ seconds divided by this time duration of 10 seconds. $\bar{v}_{d,n=2}$ is estimated as the distance traversed between $t = 42$ seconds and $t = 66$ seconds divided by this time duration of 24 seconds. Finally, $\Delta v_{n=2}$ is estimated as
(\bar{v}_{d,n=2} - \bar{v}_{o,n=2}). A positive value of \( \Delta v_n \) represents an increase in the average speed following the \( n^{th} \) lane change.

Note that \( \Delta v_n \) represents the change in average speed before and after a lane change. In the event of an observed increase in speed, one cannot be certain that the driver benefited optimally (or even at all) from changing lanes. This is because the vehicle might have maintained or increased its speed had it not changed lanes. Regardless, a positive value of \( \Delta v_n \) is considered here as a speed gain due to lane changing. An alternative way of estimating the speed gain by the subject vehicle through discretionary lane changes is to compare the average speed for several data points before and after a lane change. In that case, the number of data points to select before and after each lane change must be determined through a sensitivity analysis to ensure the stability of the vehicle speed during a lane change. This proposed alternative approach is recommended to be tested in future research.

To gain further insights, the second set of analyses was conducted on individual trajectories classified into two groups: a) trajectories with at least one lane change and b) trajectories with no lane changes. The SMSs of these two trajectory types are contrasted by periods to show the congestion effect.

The average \( \Delta v \) over all lane changes in the dataset was +5.4 fps, which indicates that overall, drivers did increase their average speed after changing lanes. The probability density plot of speed gain and loss (\( \Delta v \)) following each lane change is shown in Figure 3-8, broken down by periods. The mean \( \Delta v \) was +9.1, +3.7, and +1.6 fps, and the standard deviation was 12.3, 14.0, and 15.0 fps, respectively in the three chronological periods. Student \( t \)-tests confirmed that the decreases in mean \( \Delta v \) across these three periods are statistically significant. This plot shows that as overall congestion worsened, gains in speed following a lane change also
decreased. The modes of the distribution for the first and second periods were almost identical; however, the heavy left tail of the distribution in the second period pulled down its mean $\Delta v$.

![Probability density plot for change in average speed subsequent to changing lanes.](image)

Figure 3-8: Probability density plot for change in average speed subsequent to changing lanes.

The cumulative distribution of SMS of the two groups of trajectories—lane changing and no lane changing—for the three periods are shown in Figure 3-9. It is clear that the SMS for lane changing trajectories is significantly higher than that of trajectories with no lane changes. The difference in average speed between the two sets of trajectories is 3.9 fps. As mentioned earlier, most lane changes in this dataset are discretionary, indicating that drivers managed to increase (or maintain) their speed by changing lanes discretionarily. Furthermore, the difference in mean of SMS distribution of these two classes of trajectories is highest (~5.1 fps) in the first period, when the traffic was least congested. In periods 2 and 3, these differences were almost similar (~2.8 fps). Nonetheless, the difference was statistically significant in all three periods.
Figure 3-9: Cumulative distribution of space mean speed of trajectories classified by lane change flag and period.

Figure 3-8 and Figure 3-9 both support the hypothesis that discretionary lane changes are motivated by speed gains: in most cases, the speed of the lane changing vehicles indeed increases. The relationships shown in these figures with different congestion levels indicate that the speed gain diminishes with the increasing congestion level. This observation complies with what was postulated about Figure 3-7—the diminishing chance to gain speed is also responsible for the decreasing lane change frequency as congestion worsens.

3.6 EXTREME LANE CHANGING BEHAVIOR IDENTIFICATION

This section presents the method and results of identifying extreme lane changing behavior from the US-101 data based on two metrics: time-to-line-crossing during lane changes and lane change frequency per unit distance traveled. The importance and estimation method for the former metric needs to be explained first, as presented in the next subsection. In contrast, estimating the lane change frequency per unit distance traveled is a straightforward task using the
corrected lane position data. Here, normalizing the lane change frequency by the distance traveled by the trajectory is important since the distance covered is not the same for vehicles with different origin-destination paths in this dataset. Note that the first metric is estimated for each lane change event, while the second one is a trajectory-level measure.

### 3.6.1 Time-to-line-crossing During Lane Changes

#### 3.6.1.1 Estimation Method

A lane change with a sharp entry angle but at low speed or with a flat angle but at high speed may not represent extreme lane changing behavior. Between these two, a low value of one usually compensates a high value of another—as reported by Winsum et al. (1999) and will be evidenced by this dataset through Figure 3-11 (b). Hence, it is difficult to identify extreme lane changing by considering these two parameters separately. They can, however, be combined into a single metric called time-to-line-crossing (TLC), which has been used in previous studies to develop in-vehicle lane keeping and lane changing warning systems (Eidehall et al., 2007; Lin & Ulsoy, 1995; van Winsum et al., 1999).

TLC is defined as the time required by the subject vehicle to reach a reference line assuming that the vehicle’s dynamics remain unchanged. It is estimated using the speed, travel direction, and position of the subject vehicle at a time-step. The cited research used it to signal lateral drifting within a lane and to warn of a potential aggressive maneuver while changing a lane. In the first case, the reference line is the boundary of the current lane of the vehicle. In the second case, of interest in this research—the reference line is the far lane marking of the destination lane.

The calculation of TLC is illustrated in Figure 3-10. In it, the blue circles show a lane changing trajectory—representing the center of the subject vehicle mass at 0.1-second intervals.
At the red-circled point (the $i^{th}$ observation of the trajectory), the heading angle ($\theta_i$) is the one formed between a straight line connecting the $i^{th}$ and $(i+1)^{st}$ observations and the road alignment. Equation 3-1 shows how TLC with respect to the far lane marking of the destination lane is estimated. It is a function of four parameters namely $\theta_i$, the reported speed at the $i^{th}$ observation ($v_i$), the distance to the near lane marking along Local Y ($d_i$), and the width of the destination lane ($W$). A similar equation was used by Eidehall et al. (2017) for calculating TLC.

$$\text{TLC}_i = \frac{(d_i + W)}{v_i \cdot \sin(|\theta_i|)}.$$  \hspace{1cm} \text{Eq. 3-1}

The next question is how to obtain the critical TLC for a lane change as a measure of extreme behavior. To answer this, one needs to define the start and end point of a lane change maneuver. First, the core of a lane change needs to be identified from the reported lane positions. Two successive observations that have lane positions in two different lanes make up the core of a lane change (see Figure 3-10). Next, and starting with those two points, lane positions are tracked backward and forward over $n$ time steps. The researcher denotes this set of $2n$ TLC values for observations surrounding a lane change as $N$. One may estimate the critical TLC for a lane change as the minimum TLC in $N$. However, estimating TLC based on a single observation may introduce bias because of the random variability of the trajectory position data. This issue is tackled by taking the average of the $n$ lowest values in the set $N$ as an estimate of critical TLC ($\text{TLC}^c$) for a lane change. The last remaining question is what should be an appropriate value for $n$. Since the focus is on identifying lane changes with small $\text{TLC}^c$ values, the researcher determines the ten lane changes with the lowest $\text{TLC}^c$ values using different values of $n$ ranging from 2 to 10. It was found that this list of ten most extreme lane changes in terms of $\text{TLC}^c$ remained virtually unchanged across values of $n=2$ to $n=7$. The value $n=4$ was chosen, which
implies that the critical portion of the lane change encompasses eight NGSIM time steps, i.e., 0.8 seconds. Figure 3-10 shows the start and end point of the lane change for \( n = 4 \).

Note that the lateral distance from the vehicle’s center to the outer lane marking of the destination lane (i.e., \( d_i + W \) in Figure 3-10) perceived by a driver may be different for leftward and rightward lane changes. For the study site where traffic drives on the right side of the road, drivers will have a better estimate of this distance when changing lanes to the left compared to when changing lanes to the right. Hence, the estimated critical TLC\( ^c \)s are ranked separately for leftward and rightward lane changes.

![Figure 3-10: Estimating time-to-line-crossing during a lane change.](image)

3.6.1.2 Critical TLC Analysis

Figure 3-11 (a) shows the probability density plot of TLC\( ^c \) estimated for all 1,520 lane changes in the dataset. One intent of plotting the distribution is to investigate whether lane changes with very small TLC\( ^c \) values form a separate mode and thus could be easily distinguished. However, this distribution exhibits no such characteristics. Hence, to define extreme lane changes, TLC\( ^c \) values below a certain percentile of the distribution were selected. In this study, 1% lane changes (i.e., 15 out of 1,520 lane changes) with the lowest TLC\( ^c \) values are
considered as extreme lane changes. Lowest 1% TLC$^c$ values for lane changes to the right (i.e., for 4 out of 427 rightward lane changes) varied from 0.71 to 1.41 seconds. Lowest 1% TLC$^c$ values for lane changes to the left (i.e., for 11 out of 1084 leftward lane changes) varied from 1.17 to 1.63 seconds. These values suggest that between leftward and rightward lane changes, the lowest TLC$^c$ values belong to the rightward lane changes in this dataset, possibly attributed to the inferior evaluation of the lateral distance ($d_i + W$) by the drivers when changing lanes to the right compared to when changing lanes to the left.

Note that the researcher took necessary caution to ensure that these low TLC$^c$ values are not artifacts of any unrealistic acceleration. To that end, the reported acceleration values during these lane changes were checked against maximum feasible acceleration for different speed ranges, as reported in the Traffic Engineering Handbook (Institute of Transportation Engineers, 1992).
Figure 3-11: Critical TLC analysis: (a) Distribution of critical TLC (b) Lowest 1% critical TLC in the average angle vs. average speed plot for leftward and rightward lane changes.

The researcher next scrutinizes the influence of individual driving maneuver parameters on generating critical TLC’s. Among the parameters in equation 3-1, the absolute value of travel
direction angle ($|\theta|$) and speed ($v$) are direct outcomes of driving maneuvers through the steering and acceleration pedal, respectively. Here, their average values ($|\bar{\theta}|$ and $\bar{v}$, respectively) over the same four points defining the lowest TLCs are considered. Figure 11 (b) shows a scatter plot of $|\bar{\theta}|$ vs. $\bar{v}$ for all lane changes, with the lowest 1% TLC$^c$ values distinguished by a different color.

There is an evident negative correlation between $|\bar{\theta}|$ and $\bar{v}$. It indicates that when changing lanes, drivers tend to compensate for a sharp angle by reducing their speed and for a high speed by flattening the angle. The lowest 1% TLC$^c$ lane changes form an envelope along the outer edge of this scatter plot in Figure 11 (b). All 15 lane changes exhibited a sharp entry angle compared to other lane changes. Among these, the rightward lane change with a reasonably high speed (~50 ft/s) and high average absolute entry angle (~0.28 rad) had the lowest TLC$^c$. Two leftward lane changes occurred with a relatively low speed (<35 ft/s), but with a sharp entry angle (>0.25 rad). The remaining 13 lane changes had reasonably high speed, with three of them having $\bar{v} > 65$ ft/s. All in all, most of these extreme lane changes in terms of TLC$^c$ were executed with a reasonable to extremely high speed and/or entry angle.

3.6.2 Visualization of Extreme Lane Changing Trajectories

In this section, trajectories with the highest lane change frequency per distance traveled and lane change events with the lowest critical TLC are visualized using the same two-panel displayed showed in Section 4.1 through 4.3. Hence, this section is divided into these two criteria.

3.6.2.1 Criterion: Lane Changing Frequency per Distance Traveled

Figure 3-7 showed that within the 2,100 ft. long segment, some drivers executed up to six lane changes. Further exploration of the data revealed additional trajectories with a very high
number of lane changes. Note that the average lane change rate of all lane changing trajectories is only 1.37 per 1,000 ft. A vehicle changing lanes more than three times per 1,000 ft. falls into the top 1% lane changing trajectories in the dataset and indeed represents an extreme lane changing behavior. Figure 3-12 (a) through Figure 3-12 (c) visualize three trajectories with the highest number of lane changes per distance traveled, ranging from 3.1 to 3.6 lane changes per 1,000 ft.
Figure 3-12: Trajectories with the highest frequency of lane changes per distance traveled.
Here, all three trajectories represent weaving maneuvers—the first one was exiting from, while the last two were merging onto the freeway. The exiting vehicle in Figure 3-12 (a) entered the study site through the leftmost lane. Hence, it needed to make five lane changes to exit. Had the vehicle entered through lane 5, four of these five lane changes would have been eliminated. At such a short weave and during a congested condition, poor lane-positioning sense of drivers not only leads to a higher lane change frequency but also forces drivers to accept small gaps when change lanes. The vehicle’s speed increased as it moved from lane 5 to the auxiliary lane, although the primary motivation for this lane change was to exit the freeway.

For trajectories in Figure 3-12 (b) and 12 (c), the first lane change was mandatory, while the following four were discretionary. Seemingly, both drivers targeted to move toward the leftmost lane to avoid the weaving turbulence. In Figure 3-12 (b), the vehicle’s speed increased as it moved toward left. However, its speed was apparently dropping toward the end of the segment when it moved to lane 1. Another interesting observation is that the entry angles onto the destination lane(s) for the second and third lane changes are very sharp (were found to be 0.29 and 0.46 rad., respectively), although the low speed does not signal an unsafe maneuver. The vehicle in Figure 3-12 (c) also managed to increase its speed as it moved to lane 2, but its speed fluctuated as it entered lane 1.

3.6.2.2 Criterion: Lane Changes with Lowest Critical TLC

Among the lane changes with the lowest 1% critical TLC, three trajectories with the lowest critical TLCs are shown in Figure 3-13 (a) through 13 (c). The first two are lane changes to the right while the last one is to the left.
Figure 3-13: Trajectories with lane changes with low critical TLC.
It is visible from these plots that these three trajectories executed the lane changes with a sharp angle (see the top panel) and high speed (see the bottom panel). These trajectories have some additional common features as well: they are all weaving vehicles, all these lane changes were mandatory, these lane changes took place outside the gore points, and the vehicles’ speed while they were on the auxiliary lane is evidently higher than that while they were on lane 5. This opportunity to have a higher speed while on the auxiliary lane presumably motivated the drivers to stay on this lane as long as possible.

Interestingly, these three vehicles did not change lanes very frequently. On the other hand, trajectories with frequent lane changes shown in the previous section did not change lane with a low critical TLC. It appears that these two measures of extreme lane changes represent different forms of driver behavior.

### 3.7 CONCLUSIONS

This chapter explores driver lane changing behavior in the NGSIM US-101 dataset and mines extreme lane changing characteristics of passenger cars. It focuses on the relationship between driver lane changing characteristics and traffic congestion level. Moreover, it identifies extreme lane changing behavior based on two metrics.

Exploring the trajectories in the dataset revealed several thought-provoking aspects of lane changing characteristics in the data. Among these, multiple lane changing events between a pair of lanes in quick succession signaled the presence of invalid lane changes in the NGSIM data—an issue also warned about in past studies. About 13% of the total lane changes were removed by applying a three-foot threshold on the lateral shift into the destination lane for all potential false positive lane changes.
While only one dataset was used to explore lane changing behavior, diversity in the congestion level across the three periods was considered. Congestion level had a perceptible effect on lane changing behavior. As traffic speed dropped on an average from 47 to 32 fps over the three 15-minute periods, lane change frequency dropped by 33%. Analysis of changes in speed following each lane change yielded an average speed gain by changing lanes from 9.1 fps in period 1 to 3.7 to 1.6 fps in periods 2 and 3, respectively. Further, space mean speed of vehicles that changed lanes was on an average 3.9 fps higher than those that did not. This difference in space mean speed between lane changing vehicles and vehicles with no lane changes decreased from +5.1 fps to +2.8 fps as congestion worsened in period 1 through 3. These findings corroborate the hypothesis that the opportunity to gain speed by changing lanes diminishes with increasing congestion level. This diminishing opportunity to gain speed further contributes to the decreasing lane changing tendency of drivers as congestion increases.

Two metrics namely lane change frequency per distance traveled and critical time-to-line-crossing (TLC\textsuperscript{c}) were used to identify extreme lane changing behavior. Lane changes with the lowest 1\% TLC\textsuperscript{c}s are considered to be extreme. TLC\textsuperscript{c} for the extreme rightward lane changes varied from 0.71 to 1.41 seconds, while that for the extreme leftward lane changes varied from 1.17 to 1.63 seconds. Two parameters related to driving maneuvers that contributed to low TLC\textsuperscript{c} are the average speed and average entry angle while changing lanes. Further investigation of these two parameters revealed that most lane changes with low TLC\textsuperscript{c} exhibited a combination of reasonable to extremely high entry angle and speed. The highest lane change rates (top 1\%) are on the order of more than three lane changes per 1,000 ft. traveled, while the average overall rate for lane changing trajectories was 1.37 per 1,000 ft. traveled. Interestingly, the trajectories with lane changes with low TLC\textsuperscript{c} did not exhibit an excessive lane change rate. This chapter
concludes that this dataset does not contain trajectories that had extreme lane changing behavior across both metrics.

The characteristics of lane changing behavior and their variation with congestion level revealed here are useful to model lane change behavior in both macroscopic and microscopic simulation models. The framework of identifying extreme driving using trajectory data can be adapted to future scenarios when connected vehicles are pervasive in the traffic stream. Further, this framework can also be used to assess how driver lane changing behavior evolves with increasing autonomy in driving, for example, with assisted or fully automated lane changing technologies.


3.8 REFERENCES


CHAPTER 4 – MODELING FRAMEWORK FOR PREDICTING LANE CHANGE INTENSITY AT FREEWAY WEAVING SEGMENTS

4.1 INTRODUCTION

Excessive lane changing near weaving segments can often result in increased turbulence, decreased safety, and high levels of acceleration and deceleration cycles with the potential for emission increases (Coifman, Mishalani, Wang, & Krishnamurthy, 2006). Discretionary lane changes—i.e., lane changes in addition to the minimum lane changes required to traverse a weaving segment—may trigger bottleneck activation when the segment is operating in near-capacity conditions (Chen & Ahn, 2018). Characterizing lane change intensity is, therefore, essential for developing treatment strategies aimed at reducing such discretionary lane changes. Furthermore, lane change intensity is critical to improved planning and design of road features, such as segment length, lane configuration, road signs, and markings. Since (at least currently) lane changes are quite difficult to measure in the field using traditional detection systems, agencies must rely mostly upon predictive models to estimate lane change intensity.

In recent years, several studies have modeled individual lane changes based on microscopic traffic data (Kesting, Treiber, & Helbing, 2007; Moridpour, Sarvi, & Rose, 2010). However, such microscopic descriptors are neither easily applicable nor readily available to transportation agency decision-makers. Instead, these decision-makers often will opt for aggregated lane change estimates for planning and designing weaving segments. To that end, this study proposes a modeling framework for predicting discretionary lane change intensity at weaving segments, requiring only generally available geometric and traffic data. The models are based on observations and recording of lane changes in the field from 19 sites. In addition, the performance of the proposed models was compared to the lane change model for weaves
developed by the Highway Capacity Manual (HCM6) (Transportation Research Board, 2016) by applying the latter to the same set of observations.

This chapter is structured as follows: Section 2 reviews past research on macroscopic lane change prediction models. Section 3 presents a description of the data used. Section 4 describes the technical details of the modeling approaches employed in this study and discusses the results. Section 5 compares the performance of the proposed and HCM6 models. Section 6 provides conclusions and recommendations for future work in this area.

4.2 LITERATURE REVIEW

Back in 1970, Worrall, Bullen, and Gur (1970) and Worrall and Bullen (1970) explored the influence of macroscopic traffic properties and ramp distances on lane change frequency using data from 30 sites. However, analyses in both studies were limited to scatterplots of lane change counts against a few parameters, such as the relative speed between a pair of adjacent lanes and the ramp distances. These studies did not reveal any systematic trait of lane change frequency observations. Later studies successfully developed systematic relationships between lane change frequency and macroscopic traffic properties. Two of those by Chang and Kao (1991) and Kang and Chang (2004) modeled lane changes per hour and lane changes per vehicle using Poisson and logistic regression, respectively. These studies showed that discretionary lane changes (DLCs) are strongly associated with average speed, flow, density, and the ratio of these variables between a pair of adjacent lanes. Kang and Chang (2004) also showed that DLCs are influenced more by average density than by average speed during uncongested conditions, while the opposite occurred during congested conditions. Although their models gave a satisfactory fit, the authors do not seem to have considered interactions among the explanatory variables.
Furthermore, they could not incorporate site-specific parameters in the models because data from only one site were employed.

More recently, Ng, Susilawati, Kamal, and Chew (2020) and Seraj, Bie, and Qiu (2017) attempted to predict DLCs at the leftmost lanes of a road segment. The sites considered in both studies were sufficiently far away from any ramp influence. Ng et al. (2020) used the density and speed differences between a pair of lanes to predict DLC frequency toward the left and the right, separately. Seraj et al. (2017) predicted DLC frequency in three separate categories based on the traffic density: free lane changes (for free-flow conditions), cooperative lane changes (for near-capacity conditions), and forced lane changes (for congested conditions). The findings of this research are similar to those of Kang and Chang (2004) since the density difference between a pair of lanes was found to be an important lane change predictor during free-flow conditions, while speed difference played the major role in near-capacity conditions. Despite having logical and expected results, these studies lacked data from multiple sites and testing datasets. These shortcomings raise the possibility that their findings are site-specific.

Roess and Ulerio (2009) employed macroscopic observations from 14 U.S. sites to model lane change frequency at weaving segments using linear regression, leading to the models adopted in the HCM6. Lane changes were categorized into two types reflecting vehicles’ origin and destination: weaving and non-weaving. Although the weaving lane change model gave a reasonable fit to data from those 14 sites, the non-weaving lane change model resulted in counterintuitive coefficients for some parameters, along with a very poor fit for some sites. The main reason behind this unexpected outcome was the lack of data for a significant range of lane change frequencies. Details of the HCM6 lane change models are explained in Section 4.4.1 of this chapter. More recently, Ahmed, Xu, Rophail, and Karr (2019) used empirical trajectories
observed on a congested weave and investigated the HCM6 models by matching the relationship of the estimated lane change frequency with density. Their study showed that according to the HCM6 models, weaving lane change rate exhibits a strong, negative relationship with traffic density. In contrast, the non-weaving lane change rate has a weak but positive relationship with traffic density.

Several studies reported a number of descriptive statistics of empirical observations of macroscopic lane change properties. Knoop, Hoogendoorn, Shiomi, and Buisson (2012) observed that an increase in density in the current lane motivates drivers to move to the neighboring lanes. However, in some cases, the opposite was also found, which renders the research findings inconclusive. Gan and Jin (2013; 2015) used trajectory and loop detector data from multiple sites to explore the relationship between lane change frequency and traffic characteristics. Corroborating the hypothesis by Jin (2010), these researchers showed that the leftward lane changes in a weave tend to increase with the on-ramp flow rate, and that the leftward lane change frequency increases linearly from the leftmost to the rightmost lanes.

Several researchers investigated the characteristics of lane change location in a weave. Marczak, Daamen, and Buisson (2014) showed that the mandatory weaving lane changes are concentrated near the entry gore, but tend to disperse as traffic speed increases. A similar finding was reported by Ahmed, Karr, Rouphail, and Chun (2019). Van Beinum, Farah, Wegman, and Hoogendoorn (2018) and Van Beinum and Wegman (2019) collected trajectory data at 14 sites to explore the effect of the distance to nearby ramps on lane change frequency and location. They categorized lane changes based on their apparent motivation. These two studies revealed that pre-positioning lane changes by exiting vehicles and additional leftward lane changes by merging vehicles might spread as far as 3,000 to 3,300 ft. from the subject ramp gore. According
to this research, the ramp-influence area is longer for merge than for diverge segments, and generally shorter than what is suggested in the HCM6 and Dutch roadway design guidelines (Vos, 2017). Although this study did not develop a lane change prediction model, the analyses presented from such a rich trajectory dataset are important for both macroscopic and microscopic modeling of lane changes.

To summarize, ordinary linear regression, logistic regression, and Poisson regression are the most common statistical tools used to predict lane change intensity in the past. Most past studies lacked data from multiple sites. Consequently, these studies could not investigate the effects of site characteristics on lane change intensity. Most of the proposed models did not consider interactions among various parameters. Moreover, the models have not been applied to test data that were not used in the model development process. Several studies explored critical lane changing characteristics at multiple freeway-ramp junctions. Those characteristics should be considered when predicting lane change intensity at similar segment types.

4.3 DATA DESCRIPTION

This section first describes the sources and locations of the lane change observations along with the data extraction methods. Then it describes the resulting database.

4.3.1 Data Sources

This section describes the traffic and geometric characteristics used for developing the predictive models for discretionary lane change intensity. Data came from three sources: a) archived trajectories collected by the Next Generation Simulation effort (NGSIM) (USDOT, 2006), b) macroscopic traffic observations from NCHRP 03-75 (Roess & Ulerio, 2008), and c) data collected on the freeway network in the Research Triangle region of North Carolina.
High-resolution trajectory data that were collected under the NGSIM effort from two sites in California—one on US-10 in Los Angeles and the second on I-80 in Emeryville—are employed in this research. Several past studies reported that these NGSIM datasets contain a significant number of invalid lane changes (Montanino & Punzo, 2013; Thiemann, Treiber, & Kesting, 2008). This problem was tackled by applying a filtering algorithm to each lane change estimated from the raw data as described in Chapter 3, Section 3.4.5. The algorithm removed 8% and 12% of the reported lane changes from the US-101 and I-80 datasets, respectively.

NCHRP 03-75 collected weaving data from 14 sites across the USA. Among these, one two-sided weave and one collector-distributor road were removed from further analysis because lane changing behavior is considered to be significantly different at these sites compared to a conventional weave. The remaining 12 sites also included datasets from the two NGSIM sites—US-101 and I-80 in California. Since the US-101 dataset in NCHRP 03-75 was identical to the NGSIM dataset, only the remaining 11 NCHRP sites were included in the analysis.

The third dataset comprises macroscopic traffic data gathered at six weaving segments near Raleigh, NC. The research team carried out this data collection effort for a project (RP2019-29) funded by the North Carolina Department of Transportation (Rouphail, Karr, Chase, & Ahmed, 2020). In five of the six sites, traffic observations were video-taped at 4K resolution by a drone hovering at an elevation of about 400 ft. The remaining site was very short in length and two ground cameras were able to cover it. Note that unlike ground cameras, drone-mounted cameras can only record for one or two hours (depending on the number of available extra batteries for the drone) because the power consumption rate of drones is very high.

Table 4-1 shows the list of all sites with their basic geometric characteristics and the number of valid observations that were derived from each.
Table 4-1: List of weaving segments constituting the database.

<table>
<thead>
<tr>
<th>Site No.</th>
<th>Site Location</th>
<th>Source</th>
<th>Weave length (ft.)*</th>
<th>Number of Lanes</th>
<th>Number of 5-minute observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>US-101 SB @ Ventura Blvd. &amp; Cahuenga Blvd., Los Angeles, CA.</td>
<td>NGSIM</td>
<td>694</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>I-80 WB @ Powell St. &amp; Asby Ave., Emeryville, CA.</td>
<td></td>
<td>1,572</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>I-80 WB @ Powell St. &amp; Asby Ave., Emeryville, CA.</td>
<td>NCHRP 03-75</td>
<td>1,572</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>I-405 NB @ SW 6th Ave. &amp; SW 12th Ave., Portland, OR</td>
<td></td>
<td>693</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>I-95 SB @ NW 135th St. &amp; N. Miami Blvd., Miami, FL</td>
<td></td>
<td>1025</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>MD-100 EB @ I-95 SB &amp; I-95 NB, Baltimore, MD</td>
<td></td>
<td>360</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>7</td>
<td>MD-100 EB @ I-95 NB &amp; US-1 SB, Baltimore, MD</td>
<td></td>
<td>1025</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>SR-202 EB @ 32nd St. &amp; 40th St., Phoenix, AZ</td>
<td></td>
<td>2110</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>9</td>
<td>SR-101 EB @ SR-51 &amp; Tatum Blvd., Phoenix, AZ</td>
<td></td>
<td>1930</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>SR-102 WB @ Tatum Blvd. &amp; SR-51, Phoenix, AZ</td>
<td></td>
<td>1720</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>11</td>
<td>SR 217 SB @ SW Pacific Hwy &amp; SW 72nd Ave., Portland, OR</td>
<td></td>
<td>2820</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>12</td>
<td>I-5 SB SW Nyberg Rd. &amp; I-205, Portland, OR</td>
<td></td>
<td>1310</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>I-5 SB SR-217 &amp; Upper Boones Ferry Rd., Portland, OR</td>
<td></td>
<td>2060</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>14</td>
<td>I-40 EB @ S Saunders St. &amp; Hammond Rd., Raleigh, NC</td>
<td>RP2019-29</td>
<td>740</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>15</td>
<td>I-40 WB @ Hammond Rd. &amp; S Saunders St., Raleigh, NC</td>
<td></td>
<td>1285</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>16</td>
<td>I-440 EB @ US-64 &amp; Poole Rd., Raleigh, NC</td>
<td></td>
<td>1000</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>17</td>
<td>I-440 WB @ Poole Rd., &amp; US-64, Raleigh, NC</td>
<td></td>
<td>1610</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>18</td>
<td>I-440 EB @ Ridge Rd., &amp; US-70, Raleigh, NC</td>
<td></td>
<td>270</td>
<td>4</td>
<td>95</td>
</tr>
<tr>
<td>19</td>
<td>Wade Ave. WB @ I-440 &amp; Blue Ridge Rd., Raleigh, NC</td>
<td></td>
<td>1,135</td>
<td>3</td>
<td>25</td>
</tr>
</tbody>
</table>

*Distance between the entry and exit ramp gore

Total number of 5-min. periods = 294
Geometric characteristics for these sites were obtained from Google Maps. Macroscopic traffic data deemed important for the models included discretionary lane change frequency, flow rate by vehicle class and by origin-destination path, average speed, and average density. Although extracting these traffic characteristics from the archived NGSIM or NCHRP 03-75 datasets was straightforward, doing so was quite complex for the raw videos from the RP2019-29 sites. To that end, a combination of manual counting and automated video processing techniques was employed. A web-based service (DataFromSky) and its offline software assisted the automated video processor that tracked individual vehicles. Figure 4-1 depicts a snapshot of the interface of the offline software, showing a unique ID for each vehicle. In the drone-video sites, traffic gates along and across the lane markings were implemented to analyze traffic properties, as shown for the directional segment on the right side of the frame in Figure 4-1. The estimated vehicular flow rate and lane change count for 10% of each video file were checked manually to ensure that the performance of the automated tool was satisfactory.

Figure 4-1: Snapshot of the automated video processing tool with vehicle IDs and traffic gates.
4.3.2 Database Description

All key predictor variables were organized in a database. That database contains 294 observations and 17 explanatory variables, of which 9 are site properties and 8 are traffic characteristics. Variables’ descriptions and their expected roles in the models are presented next.

4.3.2.1 Site Geometric Properties

Figure 4-2 illustrates the key geometric properties for a hypothetical site.

**Segment Length, \( L \) (ft.):** It is the length over which traffic data were collected for a site. Ideally, it should be the distance between the entry and exit gore of a weaving segment. However, the data coverage length for some sites was slightly shorter than that distance. Segment length is expected to have a positive relationship with DLC intensity.

![Figure 4-2: Explanation of site geometric properties and variables.](image)

Lane Configuration Parameters: Four lane configuration parameters are considered: total number of lanes \( (N) \), the minimum number of lane changes required by a freeway-to-ramp
vehicle (LC\textsubscript{FR}) and a ramp-to-freeway vehicle LC\textsubscript{RF}, and the presence of any high occupancy vehicle (HOV) lane. The last one is a binary (Yes/No) variable.

The total number of lanes is presumed to have a positive effect on DLC intensity. The minimum number of lane changes in a weave (LC\textsubscript{FR} and LC\textsubscript{RF}) may influence DLC intensity in a complex way. It is assumed that if these values are high, weaving vehicles may find very limited space for lane changes within the segment in addition to the minimum lane changes. The presence of HOV lanes is supposed to have a negative effect on DLC intensity since it imposes a restriction on lane changes to and from the HOV lanes.

Note that segment length, number of lanes, and minimum lane change requirements by the weaving vehicles are site characteristic variables. Hence, they may act as categorical variables in the models. However, since these variables have potential physical relationships with lane change intensity, they are input into the models as continuous variables. Further, the categorical forms of number of lanes and minimum lane change requirements by the weaving vehicles were tested, but the performance of the resulting models was found to be inferior to that of the models which have these predictors as continuous variables. Hence, these site characteristic predictors are treated as continuous variables.

**Nearby Ramp Distances**: Four variables related to nearby ramp distances are included: interchange density (ID), distance to the nearest upstream off-ramp ($D_{\text{up,off}}$ in feet), presence of an on-ramp within 3,000 ft. upstream of a weave ONR\textsubscript{up}, and the presence of an off-ramp within 3,000 ft. downstream of a weave OFR\textsubscript{down}.

The interchange density parameter is borrowed from HCM6 to represent the difference in driver behavior between urban and rural road networks. It is defined to be the number of
interchanges 3 miles upstream and downstream from the center of the subject weaving segment, divided by 6.

The distance to the nearest upstream off-ramp ($D_{up,off}$) represents the influence of an upstream off-ramp on DLC intensity. Presumably, a lower value of $D_{up,off}$ would lead to more freeway-to-freeway vehicles to be on the leftmost lanes, consequently reducing discretionary lane changes within the subject weave. HCM6 also made a similar proposition in its merge and diverge segment analysis in Chapter 14. This distance, as shown in Figure 4-2, is measured from the exit gore of the upstream off-ramp to the entry gore of the subject weave.

The last two parameters, $ONR_{up}$ and $OFR_{down}$, are binary, indicating the presence (or absence) of an upstream on-ramp and downstream off-ramp within the specified distance, respectively. As discussed in the literature review, a previous study (van Beinum et al., 2018) showed that pre-positioning lane changes toward the right by exiting vehicles may begin as far as approximately 3,000 ft. upstream of an off-ramp. Similarly, leftward lane changes by merging vehicles may extend as far as 3,000 ft. downstream of an on-ramp. These rightward and leftward lane changes triggered by such off-ramps and on-ramps, respectively, may take place within the subject weave if the subject weave is near those ramps. As shown in Figure 4-2, $ONR_{up}$ is measured from the upstream on-ramp’s gore to the subject weave’s entry gore. Conversely, $OFR_{down}$ is measured from the subject weave’s exit gore to the downstream off-ramp’s gore.

4.3.2.2 Traffic Flow Characteristics

Traffic flow variables included in the analysis are: hourly flow rate by vehicle class and by origin-destination route ($v/hr/\text{lane}$), space mean speed (SMS) (mph), average density ($k$) ($v/\text{mi}/\text{lane}$), and discretionary lane change counts (per hour).
There are four hourly flow rate variables by origin-destination (OD) route: freeway-to-freeway ($V_{FF}$), freeway-to-ramp ($V_{FR}$), ramp-to-freeway ($V_{RF}$), and ramp-to-ramp flow rate ($V_{RR}$). In HCM6, these four variables are combined into one predictor called the weaving ratio. However, it is assumed that drivers along these four ODs have different purposes and characteristics in terms of their lane changing requirements. The heavy vehicle flow rate ($V_{HV}$), is included in the analysis for similar reasons. The total flow rate, which is the sum of the four OD-based flow rates, was not included as it exhibited strong collinearity with the freeway-to-freeway flow rate.

Average density and SMS are included as explanatory variables because past research showed that driver lane changing behavior varies significantly at different congestion levels (Ahmed et al., 2020). Average density is estimated using the fundamental equation of traffic flow, i.e., dividing the total flow rate by SMS.

All lane changes in excess of the minimum lane changes are considered discretionary.

The minimum lane change frequency in a weave is estimated as shown in Equation 4-1.

Minimum lane change count per hour = $V_{RF} \cdot LC_{RF} + V_{FR} \cdot LC_{FR}$. Eq. 4-1

The role of DLC intensity as the response variable of the models is discussed next.

**4.4 PREDICTIVE MODELS**

In literature, lane change intensity had been expressed in several ways, such as lane change frequency per hour (Transportation Research Board, 2016), lane changes per vehicle (Kang & Chang, 2004; Ahmed et al., 2018) and lane changes per vehicle-mile traveled (Knoop et al., 2012). The first two forms of intensity for DLCs were considered. DLCs per vehicle are estimated as shown in Equation 4-2.
DLCs per vehicle = \frac{DLC frequency per hour}{total hourly flow rate} \quad \text{Eq. 4-2}

For predicting DLCs per vehicle, the five predictors expressing different flow rates, i.e., $V_{FF}$, $V_{FR}$, $V_{RF}$, $V_{RR}$, and $V_{HV}$ are expressed as percentages of the total flow rate and represented by the symbols $V_{FF\%}$, $V_{FR\%}$, $V_{RF\%}$, $V_{RR\%}$, and $V_{HV\%}$ respectively.

The major challenge for predicting DLC intensity is that no theoretical model form of this variable has been established. Although the predictor variables described in the preceding section are expected to have some physical relationships with DLC intensity, their effects can be of numerous forms and may be intertwined. Modeling of DLC intensity, therefore, requires the testing of the statistical significance of several nonlinear and interaction terms. However, the small sample size of observations from several sites (see Table 4-1) restricts the testing to only a few forms of predictors. To overcome this problem, robust statistical modeling techniques were applied that can efficiently handle nonlinearity and interactions between predictor variables.

Two modeling techniques were employed for predicting DLC intensity: regression trees and linear regression. Both techniques have advantages and drawbacks, which will be discussed in subsequent sections. Since it was proposed to compare the performance of these models against the HCM6 lane change model for weaves, the HCM6 modeling technique is explained as well. The technical discussion on the predictive models may seem complicated to the readers since the following sections will discuss three modeling techniques for each of the two response variables, and hence, some guidance on the should precede. First, the HCM model is explained briefly. Next, the two models developed using regression trees—one for each response variable—are discussed, including the findings in terms of model performance and the interpretation of the outputs. The following subsection presents the model testing results for each site. A similar treatment is given regarding the linear regression models.
### 4.4.1 HCM6 Lane Change Models for Weaving Segments

HCM6 is a widely used guideline for traffic analysts around the globe, and the models proposed in this study also used the NCHRP 03-75 database, which has a substantial overlap with the database used in this research. In the HCM6 framework, lane changes are categorized into those for weaving and non-weaving vehicles. Separate models are developed for each category. Discretionary lane changes per hour for weaving vehicles ($DLC_W$) are estimated according to Equation 4-3.

$$DLC_W = 0.39[(L - 300)^{0.5}N^2(1 + ID)^{0.8}]. \quad \text{Eq. 4-3}$$

The lane change count for non-weaving vehicles—all discretionary, is contingent on an index value called non-weaving index ($I_{NW}$). Equation 4-4 shows how to estimate this index.

Note that the flow rates are in pc/hour unit, so the heavy vehicle percentage is already incorporated in the flow rates.

$$I_{NW} = L \cdot ID \left(\frac{V_{FF} + V_{RR}}{10,000}\right). \quad \text{Eq. 4-4}$$

Non-weaving lane changes per hour ($LC_{NW}$) are then estimated per Equation 4-5.

$$LC_{NW} = \begin{cases} 0.206(V_{FF} + V_{RR}) + 0.542L - 192.6N & \text{if } I_{NW} \leq 1,300 \\ 2,135 + 0.223(V_{FF} + V_{RR} - 2,000) & \text{if } I_{NW} \geq 1,950. \end{cases} \quad \text{Eq. 4-5}$$

Note that the model in Equation 4-5 is discontinuous between $I_{NW} > 1,300$ and $I_{NW} < 1,950$. HCM6 suggests that if $I_{NW}$ is within this region, $LC_{NW}$ should be interpolated between the values obtained for $I_{NW} \leq 1,350$ and $I_{NW} \geq 1,950$. Also of note is that the HCM6 model is applicable to observations with segment length ($L$) $> 300$ ft. or density ($k$) $< 43$ pc/mi/lane.

### 4.4.2 Regression Tree Models

Compared to linear regression and related techniques, regression trees can more readily handle both categorical and numerical variables as predictors, nonlinear relationships between
predictors and the response, and interactions between and among predictors. The only modeling
decision in this context is the choice of a “complexity” parameter that controls the number of
terminal nodes (also called leaves) in the final tree. This choice is the canonical one between
“model fit” and complexity.

This research uses the CART algorithm (Breiman, Friedman, Olshen, & Stone, 1984) for
predicting DLC intensity. The process starts with a root node, which contains all observations.
The root node is split sequentially into child nodes. Splits are made in order of decreasing
statistical significance, which can be interpreted heuristically as sequentially the difference of the
means in the two child nodes.\textsuperscript{2} Splitting ceases when the complexity parameter (CP)—which
decreases as the number of nodes increases—falls below a model-determined threshold
calculated by means of 10-fold cross-validation. Figure 4-3 shows graphically the process of
selecting the appropriate CP threshold for predicting DLCs per hour. The number of nodes is on
the upper x-axis and the (relative) cross-validation error on the y-axis. Two CP thresholds are
widely used: one is associated with the error given by the dotted horizontal line, which represents
the minimum relative error plus one standard deviation of the 10-fold cross-validation process. It
is called the local minimum CP. Another one is associated with the minimum cross-validation
error, termed as global minimum CP. According to the upper horizontal axis in Figure 4-3, the
global and local minimum CPs generate regression trees with 17 and 6 nodes, respectively. The
tradeoff of using the global and local minimum CPs is that the first one will generate a larger
tree, with lower error, but with a higher chance of overfitting the model than the second tree.

In addition to optimizing the CP parameter, in this research, a restriction is imposed so
that no split can create a node containing fewer than eight data points, which represents

\textsuperscript{2} Taking associated variances into account.
approximately 3% of the data. This condition reduces the effect of any erroneous observation on
the prediction. Once the choice of the CP threshold is made, the entire process is algorithmic,
and therefore fully reproducible.

![Figure 4-3: Illustrating the selection of complexity parameter for predicting DLCs per hour.](image)

**4.4.2.1 Regression Tree for DLCs per Hour**

Figure 4-4 shows the regression tree for predicting DLCs per hour that was pruned based on the
global minimum CP as shown in Figure 4-3. The upper portion of the red dashed line is the tree
that the local minimum CP would generate, which has only four predictors. Two of those four
predictors are site characteristics and the many splits on site-specific variables hints that the
model generated by the local minimum CP fits each site separately. Further, the model relative
root-mean-squared error (RMSE) for the local minimum tree is 0.35 (not shown in the figure),
while that of the global minimum tree is 0.29 of the mean DLCs per hour. Hence, the global
minimum tree is the selected model for predicting DLCs per hour.
Figure 4-4: Regression tree model for DLCs per hour.

Ultimately, the selected model places each of the 294 data points in one of the 17 terminal nodes. The standard deviations (sd) for each node are shown beneath the corresponding node in Figure 4-4. Key observations from this tree are noted below.

- The mean number of DLCs over the 17 terminal nodes ranges from 145 to 5,142 per hour, corresponding to the colors in Figure 4-4. The standard deviation increases with the predicted value, but its relative value is higher for the nodes with lower predicted values.

- The first split was made based on segment length, which is also used in two further splits. Freeway-to-ramp and ramp-to-freeway flow rates were also used in multiple splits. These imply that segment length and weaving flow rates have a strong influence on the DLCs in a weaving segment. However, it is cautioned that because segment length is a site...
characteristic, in this relatively small dataset, splits based on it may effectively be splits based on site.

- The branches generated by most traffic variables are consistent with their expected effects on DLCs per hour. For instance, freeway-to-freeway, ramp-to-freeway, freeway-to-ramp, and heavy vehicle flow rates show an increasing effect on DLCs per hour.
- Site-specific variables also show anticipated effects on DLCs. For instance, segment length, interchange density, and the distance to the nearest upstream off-ramp show an increasing effect on DLCs per hour.

4.4.2.2 Regression Tree for DLCs per Vehicle

Figure 4-5 is the companion version of Figure 4-4 for predicting DLCs per vehicle. Unlike the earlier model, the relative RMSEs for the local and global minimum CP trees shown in Figure 4-5 were quite close to each other—0.20 and 0.18, respectively. However, because half of the predictors in the local minimum tree are site-specific, the global minimum tree is recommended for predicting DLCs per vehicle as well. Key observations from this tree are noted below.

- Among the 15 nodes, the predicted DLCs per vehicle vary from 0.029 to 0.96. The relative standard deviation shows a similar trait to that of the DLCs per hour model.
- Similar to the DLCs per hour model, the first split as well as three other splits at the lower branches of the DLCs per vehicle model are based on segment length. Other predictors that are common between the two trees are interchange density, SMS, and distance to the nearest upstream off-ramp. Unexpectedly, the effect of the nearest upstream off-ramp distance is in the opposite direction in the DLCs per vehicle model when compared to the DLCs per hour model.
- As a reminder, for this model, flow rates are input as percentages of the total flow rate in this model, and hence, their interpretations are different as well. For instance, several splits in this model imply that freeway-to-freeway vehicle fraction has a negative effect on DLCs per vehicle. The effect of freeway-to-ramp vehicle percentage on the response variable varies at different splits. Ramp-to-freeway vehicle percentage exhibits an increasing effect on DLCs per vehicle.

![Regression tree model for DLCs per vehicle.](image)

**Figure 4-5: Regression tree model for DLCs per vehicle.**

### 4.4.2.3 Site-Specific Validation of Regression Tree Models

Inside the CART algorithm, the size of a regression tree is determined by minimizing the cross-validation error of a 10-fold cross-validation algorithm. The resulting model RMSE and $R$-squared values are satisfactory—varying between 0.18–0.29 and 0.93–0.96, respectively. Note that $R$-squared values may be misleadingly high, because they do not account for bias. On the other hand, RMSE accounts for both bias and variability.
To further evaluate the validity of the models, their prediction strength and
generalizability for each site need to be tested. To that end, both regression tree models are
evaluated by successively omitting each of the 19 sites and constructing the associated model (or
tree) using data from the remaining 18 sites. Then, the predicted DLC intensity is compared
using that truncated model for data from the omitted (or out-of-bag) site with the actual DLC
intensity. Figure 4-6 shows the results, which require some additional explanation:

- There are three sets of predictions for each site
  - “No Model” corresponds to simply using the observed mean of DLCs as the
    predictor for the site. If the No Model performs better than the regression tree at a
    site, this implies that the site differs so much from the remaining sites that
    “borrowing of strength” is not possible by the regression tree.
  - “Full Model” corresponds to the regression trees in Figure 4-4 and Figure 4-5,
    which uses the data from all sites. If the Full Model performs better than the No
    Model, then it implies that there is benefit from cross-site modeling.
  - “Omitted Model” corresponds to the truncated model described above, based on
    data from all sites other than the omitted site. If the Omitted Model performs
    reasonably, then it suggests that there is scientific generalizability of the modeling
    approach.

- The performance of a model for a site is evaluated in terms of its relative RMSE, which is the
  RMSE divided by the mean response variable for that site.

Site-specific performance for the models in Figure 4-4 and Figure 4-5 is discussed below
with the help of Figure 4-6. Note that the vertical axis in Figure 4-6 (a) and Figure 4-6 (b) is
capped at 1.2 for improved display of the low values. For reference purposes, site numbers are the same as in Table 4-1.
Figure 4-6: Site-specific relative RMSE of regression tree models (a) for DLCs per hour (b) for DLCs per vehicle.

- With DLCs per hour as the response variable, Figure 4-6 (a) shows that relative RMSE is uniformly lower for Full Model than for No Model with the exception of Sites 2, 3, 6, 8, 9.
and 17. This finding confirms that there is benefit from cross-site modeling for most sites.

With DLCs per vehicle as the response, ten sites have relative RMSE higher for Full Model than for No Model, although for many of those the values were approximately equal.

- For both response variables, several sites exhibited a very high relative RMSE for Omitted Model—more than five times than that for the Full model. These are Site 9 for DLCs per hour, Sites 2 and 8 for DLCs per vehicle, and Sites 11 and 13 for both models. Evidently, these sites have certain unaccounted for features that are significantly different from what the other sites have. For the remaining sites, the performance of the Omitted Model was acceptable. This finding indicates that there is scientific generalizability for the modeling approach, which would be stronger if there were more observations for each site.

- The DLCs per vehicle model has a lower overall relative RMSE (0.18) than what the DLCs per hour model has (0.29), but the site-specific validation of the former model is not as good as the later one. The total flow rate plays an important role in the variation of DLC count. Since DLCs per vehicle is estimated by dividing the DLC count by total flow rate, the variation of DLCs per vehicle becomes significantly less than that of DLCs per hour for a site. Consequently, the No Model relative RMSE is low compared to the Full Model relative RMSE for many sites when predicting DLCs per vehicle.

**4.4.3 Linear Regression Models**

The major advantage of linear regression is that the response variable is mathematically connected to the predictors through the coefficient values (including signs), hence the marginal effect of one predictor can be assessed, assuming all other predictors are fixed at their mean values. However, the application of linear regression has some caveats, requiring, for example, that the response variable and the resulting residuals be normally distributed. Besides, linear
regression cannot automatically handle critical issues like multicollinearity, outliers or influential observations, nonlinear relationships, and interactions among predictors. Many past studies that modeled lane change intensity using linear regression overlooked most, if not all, of these limitations of linear regression.

In this research, significant efforts were made to overcome these constraints before applying this technique for predicting DLC intensity. Before delving into the model development, necessary measures taken to minimize their effects on the final models are discussed below.

**Normality of the response variable:** The Box-Cox transformation (Sakia, 1992) is used to transform the response variables to have distributions that more closely approximate normality. By maximizing a log-likelihood function, the conversion factors obtained for DLCs per hour and DLCs per vehicle were 0.091 and 0.151, respectively. The resulting t-test indicated that even after transformation, neither transformed variable is normally distributed, giving caution about the models’ performance.

**Detecting outliers:** An added variable plot (Cook & Weisberg, 1994) is used to detect outliers. It isolates the relationship between the response and each predictor from the remaining predictors. Three outlier observations were removed from each model based on the added variable plots of more than two predictors.

**Detecting nonlinearity of predictors:** Partial residual plots (Larsen & McCleary, 1972), which show the relationship between a predictor and the response assuming that all other predictors are in the model, were used to check for nonlinearities. For this dataset, none were found.

**Detecting collinearity:** Significant collinearities were found between two pairs of variables: SMS and average density and number of lanes and freeway-to-freeway flow rate. SMS is not
very sensitive to traffic flow rate in most of the free flow regime, and hence, was removed.

Number of lanes was excluded from the DLCs per hour model.

**Checking for interaction terms:** Two-way interactions between pairs of variables are tested for statistical significance using their partial dependence plots (PDPs) (Greenwell, 2017) and $p$-values of their coefficients. Because of the limited number of observations, each pair of interactions was tested iteratively. PDPs show the marginal effect of one predictor on another in estimating the response variable. For instance, Figure 4-7 shows the PDP of segment length ($L$) and freeway-to-ramp flow rate ($V_{FR}$) for predicting DLCs per hour. The color indicates the variation of the response variable. The key observation is that the variation of the response variable with one of the predictors, say $V_{FR}$, is opposite for different values of the other predictor, $L$, and vice versa. Note that had the database included (many) more observations, all possible two-way interactions could have been accounted for to test their significance using a variable selection algorithm.
Figure 4-7: Partial dependence plot for interaction between length and freeway-to-ramp flow rate in DLCs per hour model.

**Modeling Form Selection:** A stepwise selection algorithm was opted for that uses the Akaike information criterion (AIC) to include or exclude predictors in a stepwise fashion (Venables & Ripley, 2002). Overfitting the models was avoided by supplementing the stepwise selection algorithm by a bootstrap procedure that randomly generates 100 bootstrap samples and runs the stepwise selection algorithm iteratively on each sample. A predictor was included in (or excluded from) the model by recording the fraction of bootstrapped samples where it was statistically significant, and whether its coefficient sign was consistent across those samples.

Despite its widespread application, the stepwise selection algorithm often selects variables that are not statistically significant (Zhang, 2016). Hence, an alternative approach, called the best subset selection, was also employed (Lumley, 2013). It seeks the best subset of
predictors from all possible subsets in terms of $R$-squared value or AIC. The algorithm was applied to the database by iteratively taking a random sample from it, developing the model, and applying the model to the out-of-sample data points. The final model was selected based on the minimum cross-validation error. In the end, both approaches yielded similar estimates. Hence, the findings reported here are based on the stepwise selection algorithm.

4.4.3.1 \textit{Linear Regression for DLCs per Hour}

Table 4-2 shows the predictor variables, their coefficients, and statistical significance at a significance level of 0.05 for the DLCs per hour model. With a relative RMSE of 0.30 and an $R$-squared of 0.95, the overall goodness of fit of this model seems satisfactory. Figure 4-8 shows a group of diagnostic plots for this model. Key observations from Table 4-2 and Figure 4-8 are noted below.

- There are 17 terms in addition to the intercept—12 are main effects and 5 are interactions. All coefficients except one are statistically significant.

- $LC_{FR}$, heavy vehicle flow rate ($V_{HV}$), and presence of an HOV lane, upstream on-ramp (ONR$_{up}$) and downstream off-ramp (OFR$_{down}$) have only their main effect terms in the model. Their coefficients—assuming all other predictors are held constant—imply that $LC_{FR}$ and presence of HOV lane have negative effects, while $V_{HV}$ and the presence of upstream on-ramp have positive effects on DLCs per hour. All these effects are expected as explained in Section 3.2.
Table 4-2: Linear regression model for DLCs per hour.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Coefficients and their statistical significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>8.592*</td>
</tr>
<tr>
<td>$L$</td>
<td>0.002974*</td>
</tr>
<tr>
<td>$L_{CFR}$</td>
<td>-0.6193*</td>
</tr>
<tr>
<td>$L_{CRF}$</td>
<td>-2.761*</td>
</tr>
<tr>
<td>$V_{FF}$</td>
<td>-0.0008754*</td>
</tr>
<tr>
<td>$V_{FR}$</td>
<td>-0.0006433*</td>
</tr>
<tr>
<td>$V_{RF}$</td>
<td>0.0012234*</td>
</tr>
<tr>
<td>$V_{HV}$</td>
<td>0.0009521*</td>
</tr>
<tr>
<td>$k$</td>
<td>0.02862*</td>
</tr>
<tr>
<td>ID</td>
<td>-0.8652*</td>
</tr>
<tr>
<td>Presence of HOV lane = $Y$</td>
<td>-4.982*</td>
</tr>
<tr>
<td>OFR$_{down}$ = $Y$</td>
<td>-2.653*</td>
</tr>
<tr>
<td>ONR$_{up}$ = $Y$</td>
<td>3.616*</td>
</tr>
<tr>
<td>$L \cdot V_{FR}$</td>
<td>0.000000755*</td>
</tr>
<tr>
<td>$L \cdot V_{RF}$</td>
<td>-0.0000005732*</td>
</tr>
<tr>
<td>$L_{CFR} \cdot V_{FF}$</td>
<td>0.0008844*</td>
</tr>
<tr>
<td>$L_{CRF} \cdot k$</td>
<td>-0.0248</td>
</tr>
<tr>
<td>$V_{FF} \cdot ID$</td>
<td>0.0002361*</td>
</tr>
</tbody>
</table>

* indicates statistical significance at a level of 0.05

Figure 4-8: Diagnostic plots for DLCs per hour model.
Figure 4-8 (a), 4-8 (c), and 4-8 (d), respectively, show that there is little to no nonlinearity, heteroscedasticity, or strongly influential points. Conversely, Figure 4-8 (b) reveals that the model residual distribution is light-tailed (Vittinghoff, Glidden, Shiboski, & McCulloch, 2012), as it clearly deviates from the theoretical normal distribution at the two tails. This gives a caution about the model performance.

4.4.3.2 Linear Regression for DLCs per Vehicle

Table 4-3 is the companion of Table 4-2 for DLCs per vehicle. Both relative RMSE (0.29) and $R^2$ values were very similar to those of the DLCs per hour model. Key observations from this model are noted below. Note that the diagnostic plots for this model are very similar to those for the DLCs per hour model, and hence, the same remarks apply.
Table 4-3: Linear regression model for DLCs per vehicle.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Coefficients and their statistical significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-12.28*</td>
</tr>
<tr>
<td>$L$</td>
<td>0.007922*</td>
</tr>
<tr>
<td>$N$</td>
<td>0.09941*</td>
</tr>
<tr>
<td>$LC_{FR}$</td>
<td>-0.5349*</td>
</tr>
<tr>
<td>$LC_{RF}$</td>
<td>0.09355</td>
</tr>
<tr>
<td>$V_{FF%}$</td>
<td>0.09374*</td>
</tr>
<tr>
<td>$V_{FR%}$</td>
<td>0.06996*</td>
</tr>
<tr>
<td>$V_{RF%}$</td>
<td>0.1874*</td>
</tr>
<tr>
<td>$V_{HV%}$</td>
<td>-0.008998*</td>
</tr>
<tr>
<td>$k$</td>
<td>-0.01182</td>
</tr>
<tr>
<td>Presence of an HOV lane = Y</td>
<td>0.3097</td>
</tr>
<tr>
<td>$OFR_{down} = Y$</td>
<td>0.2281</td>
</tr>
<tr>
<td>$L \cdot V_{FF%}$</td>
<td>-0.0000689*</td>
</tr>
<tr>
<td>$L \cdot V_{FR%}$</td>
<td>-0.00005931*</td>
</tr>
<tr>
<td>$L \cdot V_{RF%}$</td>
<td>-0.0001054*</td>
</tr>
<tr>
<td>$LC_{RF} \cdot k$</td>
<td>0.01432*</td>
</tr>
<tr>
<td>$LC_{RF} \cdot V_{RF}$</td>
<td>-0.03999*</td>
</tr>
<tr>
<td>$V_{HV} \cdot k$</td>
<td>0.0011</td>
</tr>
<tr>
<td>$k \cdot (Presence$ of HOV lane = Y)</td>
<td>-0.01618*</td>
</tr>
</tbody>
</table>

* indicates statistical significance at a level of 0.05

- This model has 18 terms—11 are main effects and 7 are interactions. Many main effect terms are not statistically significant, but those are kept in the model because interaction terms involving them are significant.

- Number of lanes, $LC_{FR}$, and the presence of downstream off-ramp have only main effect terms in the model. The positive coefficient of number of lanes and presence of downstream off-ramp are expected, so is the negative coefficient of $LC_{FR}$.

- The coefficient sign for some predictors in the model may seem counterintuitive. For example, $LC_{FR}$ and the indicator for the presence of an HOV lane have a positive coefficient. However, their overall effect on the model is dominated by their interaction term’s coefficients.
4.4.3.3 Site-Specific Validation of Linear Regression Models

Similar to the regression trees, relative RMSEs for No Model, Full Model, and Omitted Model are estimated for each site using the linear regression technique, as shown in Figure 4-9.

- With DLCs per hour as the response variable, Figure 4-9 (a) shows that only 8 of 19 sites have lower relative RMSE for Full Model compared to No Model. The performance of the DLCs per vehicle model is worse, as only three sites exhibit such a low relative RMSE for Full Model (Sites 5, 13, 16 in Figure 4-9 (b)).

- For both response variables, relative RMSE for Omitted Model is high for most sites. However, for DLCs per hour (Figure 4-9 (a)), only two sites exhibited a very high relative RMSE for Omitted Model (more than five times that of Full Model). These are Sites 7 and 11. For DLCs per vehicle (Figure 4-9 (b)), four sites exhibited such a high relative RMSE for Omitted Model (Sites 2, 8, 17, and 18).

- Among the four site-specific modeling results shown in Figures 6 and 9, several sites exhibited very high relative RMSE for all of them. For instance, Sites 3 and 6 exhibited high relative RMSE for Full Model in all four models. Sites 8 and 9 also exhibited high relative RMSE for Full Model in three of the four models. Omitted Model relative RMSE was very high for Site 11 in three of the four models. Note that many sites did not exhibit significant variability in terms of congestion level because most data were collected during off-peak periods and the collection duration was short. Consequently, the observed discretionary lane change intensity did not vary much within a site, resulting in poor performance for Full Model relative to the corresponding No Model.
Figure 4-9: Site-specific relative RMSE of linear regression models (a) for DLCs per hour (b) for DLCs per vehicle.
4.5 COMPARISON OF MODEL PERFORMANCE

In this section, the performance of regression trees and linear regression for predicting DLC intensity is compared using several measures. The performance of the weaving segment lane change model in HCM6 is included in this comparison. Table 4-4 shows the results. Six criteria were used, of which the number of unique predictors in a model represents the data required for applying the model. Overall relative RMSE represents the model error combining all sites. For each site, the comparison of relative RMSE between Full Model vs. No Model and Omitted Model vs. Full Model was discussed earlier. The remaining two criteria that evaluate the number of sites with Full Model and Omitted Model relative RMSE less than 0.5 are somewhat arbitrary. Note that for HCM6, the Full Model is applicable to the 11 NCHRP sites only. Omitted Model relative RMSE for HCM6 was estimated by applying the Full Model on five sites that are not NCHRP sites and have $L < 300$ ft. and $k < 43$ pc/mi/ln.

Table 4-4: Comparison of model performance.

<table>
<thead>
<tr>
<th>Performance Measure</th>
<th>Regression tree</th>
<th>Linear regression</th>
<th>HCM6 model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of unique predictors</td>
<td>9</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Overall relative RMSE</td>
<td>0.29</td>
<td>0.30</td>
<td>0.57</td>
</tr>
<tr>
<td>Number of sites with Full Model relative RMSE &lt; 0.5</td>
<td>18*</td>
<td>18</td>
<td>5 (out of 11)</td>
</tr>
<tr>
<td>Number of sites with Full Model relative RMSE &lt; No Model relative RMSE</td>
<td>13</td>
<td>8</td>
<td>2 (out of 11)</td>
</tr>
<tr>
<td>Number of sites with Omitted Model relative RMSE &lt; 0.5</td>
<td>11</td>
<td>10</td>
<td>2 (out of 5)</td>
</tr>
<tr>
<td>Number of sites with Omitted Model relative RMSE &lt; 5 times Full Model relative RMSE</td>
<td>16</td>
<td>17</td>
<td>5 (out of 5)</td>
</tr>
</tbody>
</table>

*There are total 19 sites

Table 4-4 shows that the HCM6 model has fewer predictors than the remaining models. However, its overall relative RMSE is almost twice that for the other two models. In fact, its Full Model relative RMSE is higher than No Model RMSE for most sites. It only performed better
than the competitors in terms of Omitted Model relative RMSE, which is less than five times the Full Model relative RMSE for all five sites. The regression tree outperforms the linear regression model across most measures. While the differences are marginal, trees are superior in terms of the number of sites with the Full Model relative RMSE lower than No Model relative RMSE. In only 8 of 19 sites, the prediction improved through linear regression relative to No Model. Several factors contribute to these differences, among which the regression tree’s ability to handle interactions and outlier observations and the issues related to the linear regression, such as non-normality of the response variable and the distribution of residuals are evident.

4.6 SUMMARY AND CONCLUSIONS

This study proposed and tested a new modeling framework for predicting lane change intensity at freeway weaving segments at the aggregate level. Data from 19 sites—obtained from two archived sources and field surveys using drone cameras—were used to develop and test the models. Predictor variables included both traffic stream and geometric features. Two modeling techniques, using regression trees and linear regression, were employed to predict discretionary lane changes (DLCs) per hour and DLCs per vehicle. Transportation agencies can readily use this model to improve the planning and design of weaving segments.

The models revealed that segment length, weaving flow rates, and interchange density are important variables for predicting DLC intensity in a weaving segment. The minimum lane change requirements and traffic density are also important according to some models. Most splits in the regression trees and coefficient signs in the linear regression models are logical. The performance of the models in terms of overall relative RMSE and R-squared values were satisfactory, ranging from 0.18 to 0.30 and from 0.93 to 0.96, respectively. However, site-
specific validation revealed that for some sites, the performance of the models, particularly the linear regression models, should be viewed with caution.

Site-specific validation also showed that the overall performance of DLCs per hour model was superior to that of DLCs per vehicle model. The regression tree for DLCs per hour was a definite improvement of the prediction than No Model (i.e., taking the mean DLCs per hour for a site as its predicted value) for more than 2/3 of the sites. Applicability of this modeling technique to a site that was not included in the model development process (omitted site) was satisfactory except for three sites. A comparison of the DLCs per hour models showed that the regression tree outperforms the linear regression model as the later one generated a higher error than a No Model for more sites. The weaving lane change model of HCM6 also gave a satisfactory performance given that it uses only six unique predictors and performed better than the other two on the omitted sites. However, its overall relative RMSE was very high (0.57), underscoring the value of advanced modeling tools with additional predictors.

This research represents a substantial effort of lane change intensity modeling, particularly in terms of the number and diversity of sites included in the model development database. However, the modeling process was hampered by a lack of sufficient observations from a number of sites. Various robust statistical techniques were incorporated or newly included in both modeling approaches, such as automatic interaction detection, bootstrapping, and k-fold cross-validation, to overcome the data limitation problem. In addition, due to the lack of variation in congestion levels and consequently in lane change behavior at each site, the proposed models did not show improvements in the prediction of lane change intensity compared to the No Model case at several sites. This limitation emphasizes the need for additional data from each site in future studies. As mentioned earlier, the number of data points were few for the
sites at which drone-mounted cameras were used to observe traffic. Future studies should overcome this limitation due to the short battery life of drones by using tethered drones (Desai, Faivre, & Zerillo, 2016). In addition, other predictor variables such as the difference in free flow speed between the on-ramp and freeway of a site should be tested for statistical significance for predicting discretionary lane change intensity at weaves.
4.7 REFERENCES


CHAPTER 5 – DEVELOPING A DISCONTINUOUS FORM OF MACROSCOPIC GAZIS-HERMAN-ROThERY MODEL TO STEADY STATE FREeway TRAFFIC STREAM OBSERVATIONS

5.1 INTRODUCTION

The transition of traffic state between uncongested and congested flow regimes has been a key topic of discussion in many past studies (Edie, 1961; Tanaka, Mitsuru, Prakash Ranjitkar, and Takashi Nakatsuji, 2008). Several important phenomena such as flow breakdown, capacity drop, and the evolution of oscillations in a traffic stream are explicitly associated with the transition of traffic state between those two regimes (May, 1990; Ahn, Laval, and Cassidy, 2010). Moreover, during this transition period, drivers need to adjust their car-following behavior (of interest in this dissertation) according to the changing traffic condition to avoid crashes or near-crash situations.

Traffic models representing the so-called fundamental diagram of traffic flow have been widely used by researchers and practitioners for analyzing traffic characteristics at various flow-density conditions. However, both traffic and driver behavior characteristics associated with the transition of traffic state are difficult to explain using a fundamental model that is continuous between uncongested and congested flow regimes. This continuity of the fundamental diagram has been under debate. Although the literature includes studies that acknowledge a ubiquitously observed discontinuity between uncongested and congested flow (Edie, 1961; Kerner, 1998; Koshi, Iwasaki, & Ohkura, 1983; Zhang & Kim, 2005), others have argued against discontinuity on empirical and theoretical grounds as mentioned in the literature review that follows. The

---

Highway Capacity Manual (HCM6) (Transportation Research Board, 2016) does not directly speak to traffic state discontinuity in its basic discussion of freeway segment analysis methodology but does acknowledge and allow for a capacity drop at queue discharge in its freeway facilities methodology.

In this study, a steady-state macroscopic model is developed for basic freeway segments with a discontinuity between uncongested and congested flow regimes. To that end, the macroscopic model form derived from the microscopic Gazis-Herman-Rothery (GHR) car-following model was adopted (Gazis, Herman, & Rothery, 1961; May, 1990). Various versions of the GHR model have been developed and proposed based on investigation and analysis of empirical traffic stream data in the past (Ceder & May, 1976; Edie, 1961; May & Keller, 1967). From the perspective of this dissertation, the major advantage of the macroscopic version of the GHR car-following model is that it provides a rational means to estimate the reaction time necessary for car-following stability. Thus, the proposed discontinuous GHR model will expectedly enable the estimation of change in required reaction time during traffic state transition—a measure that is researched in the next chapter for its association with rear-end crashes.

The proposed discontinuous model is applied to field observations of aggregated traffic flow. In addition, the performance of the proposed model was compared to the macroscopic model for basic segments developed by HCM6 by applying the latter to the same set of observations. One of the practical challenges that lies behind empirically fitting a steady state macroscopic model is identifying non-stationary and mixed state observations in addition to identifying outliers (Xu, Williams, Routhall, & Chase, 2013). If not removed prior to or during model fitting, such observations will introduce bias in the model parameters and resulting
estimates. Although many outlier detection algorithms are available in the literature (Li, X., Li, Han, & Lee, 2009; Xu et al., 2013), the selection of algorithms should be specific to the modeling objective and data characteristics. This study explicitly addresses the issue of outliers and mixed-state observations in field data when developing the proposed macroscopic model.

A review of past studies on the relevant subject is presented in the next section. Following that, the methodology for fitting the two regimes independently is described. The model is then applied to the data collected from five different sensors operated by Traffic.com (ITIP Stakeholder, 2010). Next, a comparison of outcomes from the proposed and HCM6 models is made based on the statistical performance measure. The key findings from fitting the models are summarized in the last section.

5.2 LITERATURE REVIEW

Edie (Edie, 1961) developed a complementary theory for steady state conditions by combining the models proposed by Greenberg (1959) and Underwood (1960). It was the first study to observe a sharp speed drop in a small density range in some observed speed–density phase plots and proposed a two-regime phase diagram to model it. The frequent appearance of this discontinuity in several cases was mentioned in this study, which evidenced that the discontinuity did not appear because of random factors or by circumstances upstream from a bottleneck. The Lincoln Tunnel data used by Greenberg (1959) and Gazis et al. (1961) was fitted in this study discontinuously to the proposed model and showed that the regression lines have better fits than the continuous model.

Similarly, Drake et al. (Drake, Schofer, & May, 1965) revealed that the two-regime model produced the best fit with the lowest standard error compared to the single-regime model. The term “inverted Lambda shaped flow-density curve” was first used by Koshi et al. (1983)
who analyzed traffic data from the Tokyo Expressway and acknowledged the presence of the discontinuity.

More recently, Zhang and Kim (2005) proposed a modified car-following theory to model capacity drop and traffic hysteresis phenomena. The Pipe’s theory of car-following model (Pipes, 1966) was modified and four different models were proposed. One of the models produced a fundamental diagram that resembles the mirrored image of the reversed lambda hypothesized by Koshi et al. (1983). The model was developed by incorporating the concept varying “gap-time” which is defined as the time required for the follower to travel a gap-distance.

However, several past studies modeled the relationship between traffic stream parameters continuously. The speed-flow relationship for base conditions described in HCM6 can be interpreted as a three-regime continuous model (although as mentioned above, the HCM6 freeway facilities methodology incorporates a capacity drop discontinuity). According to Chapter 10 and Chapter 25 of HCM6, the flow-density curve for the oversaturated regime assumes a simplified linear relationship. Others have argued against the existence of a discontinuity in the fundamental diagrams as a permanent characteristic and advocated the continuous flow-density relationship (Cassidy, 1998; Li, J. & Zhang, 2013). Cassidy’s (1998) empirical observation from the re-scaled curves of cumulative arrival number vs. time and cumulative occupancy vs. time led to the assumption that the flow-density relationship is continuous under steady-state traffic flow conditions. The existence of a gap that appeared in the observed loop detector data is explained in this study as an absence of density values in some ranges. Li and Zhang (2013) theoretically showed that a discontinuous fundamental diagram is non-differentiable at the discontinuous point and results in infinite characteristic wave speeds.
Past outlier detection algorithms based on various traffic data are also instructive. To obtain only valid data points from PeMS (Caltrans,) archived data, a study adopted an empirical rule of excluding observation days for which PeMS reported a functionality of 80% or less (Li et al., 2009). In addition, this study detected outliers even from the valid observation days by plotting box plots of observed flow rate over multiple days and imposing a threshold of 1.5 times the interquartile range.

Xu et al. (2013) developed a method for fitting the HCM6 oversaturated model to fixed sensor data by filtering steady state observations only. Mixed state and inconsistent data points were removed by imposing three thresholds. Detailed about these thresholds are described in the Methodology section.

The use of robust regression technique in dealing with erroneous data has gained popularity in recent times (Bary, 2017; Rousseeuw & Leroy, 2005). This technique iteratively dampens the influence of outlying observations to provide a better fit to the majority of the data. However, the application of robust regression or similar technique in handling traffic related data is limited in the literature.

From the above review, it can be said that the discontinuity in the steady state fundamental diagrams has been recognized from as early as 1960s. Nonetheless, modern computational tools can improve the development of such models by using data collected over an extended time period. In contrast, a few studies challenged the existence of fundamental diagram discontinuity. Finally, there are opportunities for including state-of-art techniques of outlier detection in empirical traffic data modeling.
5.3 METHODOLOGY

5.3.1 Initial Cleaning of Data

For developing the proposed discontinuous steady-state model, traffic stream data from fixed-location sensors is needed for a long span of time (preferably one year). Since the model development is data driven, inconsistent and mixed state data were removed in two stages of the process. This section presents the first stage, which involves applying three thresholds that are introduced in a prior study (Xu et al., 2013). In this study, the thresholds are modified and applied in a slightly different manner to conform to the study objectives.

The first two thresholds applied to detect inconsistent observations are known as the critical speed threshold (CST) and critical density threshold (CDT). The combined application of these two thresholds removes the low-speed observations associated with low volume. Such observations exist due to inclement weather, work zones, incidents, or any other sort of capacity drop phenomena as well as from observations that include both congested and uncongested flow. CST is used to identify low speed (congested) observations. Analysis of traffic data reveals that 10 mph below the speed limit is a reasonable threshold to determine congested observations. CDT is used to identify low volume observations. A density of 35 pc/mi/ln, the threshold that separates LOS of D and E for urban freeways in the HCM6 is defined as the CDT threshold. Observations that fall below both CST and CDT thresholds are tagged as inconsistent (low-speed and low-volume) data points. Figure 5-1 shows speed-flow data obtained from Traffic.com for calendar year 2013 and for sensor ID 23774 WB. The inconsistent data points are shown in Figure 5-1 (a) bounded by the CST and CDT threshold lines (green points).
Time interval (in this case, 5 minutes) data points that represent non-stationary traffic conditions were detected by applying the third threshold termed as speed first difference threshold (SFDT). SFDT excludes observations whose speeds differ from the previous 5-minute observation by more than 10 mph. These non-stationary observations for sensor 23774 WB is shown in Figure 5-1 (b) by the green points. It should be noted that outliers generating from system measurement error may still exist even after applying these thresholds.

5.3.2 Fitting a Two-regime GHR Steady State Model

The process of fitting the proposed steady state model is explained in this section. The starting point is the flow density relationship for the single regime full GHR model (Gazis et al., 1961) as shown in Equation 5-1.

\[ q_{M,i} = k_i \times u_F \times \left(1 - \left(\frac{k_i}{k_f}\right)^{l-1}\right)^\frac{1}{1-m} \]  

Eq. 5-1

Here,

\( i = 1, 2, 3 \ldots \) Observation index;

\( q_{M,i} = \) Model flow for observation \( i \) (pc/hr/ln);
\( u_r = \) Free flow speed (mph);
\( k_i = \) Observed density (pc/mi/ln);
\( k_j = \) Jam density (pc/mi/ln);
\( l = \) Distance headway exponent;
\( m = \) Speed difference exponent.

In a two regime (uncongested and congested regime) GHR model, the same form of Equation 5-1 is used for each regime, and the regime boundary is determined by a density break point \((k_B)\). In the proposed model, a transition range is introduced by allowing the two regimes to overlap each other which produces the recognized inverted lambda shape when plotted in the flow-density domain. The major challenge is to model this transition range. Empirical observations of traffic stream data depict that the data points in the transition range follow either the uncongested or the congested regime’s characteristics. In other words, the transition range (defined in terms of density) include a mixture of observations from both regimes. In this study, it is proposed to model each data point within the overlap by the regime model, either uncongested and congested, that results in the smallest absolute error.

The algorithm for modeling the uncongested and congested regimes along with the overlap range is defined by Equation 5-2, 5-3, and 5-4.
Observed Density

Formula for model flow

\[ k_i \leq k_{B,r=2} \]

\[ q_{M,i,r=1} = k_i \times u_{F,r=1} \times \left( 1 - \left( \frac{k_i}{k_{l,r=1}} \right)^{l_{r=1}-1} \right) \frac{1}{1-m_{r=1}}. \quad \text{Eq. 5-2} \]

\[ k_i \geq k_{B,r=1} \]

\[ q_{M,i,r=2} = k_i \times u_{F,r=2} \times \left( 1 - \left( \frac{k_i}{k_{l,r=2}} \right)^{l_{r=2}-1} \right) \frac{1}{1-m_{r=2}}. \quad \text{Eq. 5-3} \]

\[ k_{B,r=2} < k_i < k_{B,r=1} \]

\[ q_{M,i,r} = \begin{cases} q_{M,i,r=1} & \text{if } |q_{M,i,r=1} - q_i| < |q_{M,i,r=2} - q_i| \\ q_{M,i,r=2}, & \text{otherwise} \end{cases} \quad \text{Eq. 5-4} \]

Here, \( r = 1, 2 \) for uncongested and congested regime respectively.

In Equation 5-2 through 5-4, the two density breakpoints \((k_{B,r=1} \text{ and } k_{B,r=2})\) define the upper and lower limit of the transition range, respectively.

### 5.3.3 Requirements for Additional Constraints

Before discussing the constraints added to fit the proposed model, it is important to explain the limitations of distinguishing the two regimes and defining the overlap based on empirical data. All data in this study are from HERE roadside sensors (formerly Traffic.com) in the Raleigh, North Carolina urban area. Figure (5-2) shows an example application of the algorithm described above to clarify the limitations. In this figure, the observed data are for HERE sensor ID 23771 WB for the calendar year of 2013 aggregated at a 5-minute interval. The model was fitted in MATLAB software using a nonlinear optimization tool with an objective function of minimizing the sum of squared error of flow. The red points show the fitted model with no additional constraints, while the green ones show the same but with the inclusion of several constraints described later.
Figure 5-2: An example of the proposed flow-density model with and without any constraint.

Several important aspects of unconstrained fitted model shown in Figure 5-2 are noted below:

- The uncongested curve of the unconstrained model (red line) do not capture a group of data points with high flow (>2,000 pc/hr/ln). Instead, the slope of the uncongested regime curve flattens out at a flow value near 2,000 pc/hr/ln. It is apparent that this model is poorly fitting several important and valid observations. Assuming a free flow speed of 65 mph (posted speed limit at this location was 60 mph), the capacity of the road section should be 2350 pc/hr/ln according to HCM6. Therefore, the fit of the uncongested curve can be improved by constraining the slope at capacity using this information.

- The queue discharge flow rate (or post-breakdown flow rate) is appeared to be 1,700 pc/hr/ln from Figure 5-2. Approximating the pre-breakdown flow rate from the observed data as 2,350 pc/hr/ln, it is found that the drop in post-breakdown flow rate is about 28% of the pre-breakdown flow rate. This percentage is significantly higher than the usual range mentioned in HCM6 (2%-20%).
• The jam density for the congested regime estimated by the model is 397 pc/mi/ln. The physical interpretation of this value is that the average length of the vehicles in the traffic stream must be unrealistically small (about 13 feet).

The observations highlighted above are also present in the datasets for the other traffic monitoring stations used in this study. The essence of the above discussion is that fitting the proposed empirical model without additional constraints may result in unusual and infeasible values that deviate from the findings of past research cited in the HCM. To resolve these issues, three constraints were added to properly fit the model. Although these constraints are selected empirically, they are justified by the guidelines provided in HCM. The resulting model after adding these constraints are shown in Figure 5-2 by the green points.

5.3.3.1 Constraint on Queue Discharge Flow Rate

The post-breakdown flow rate is usually lower than the pre-breakdown flow rate, resulting in a significant loss of freeway throughput during congestion. According to a past study cited in Chapter 10 and Chapter 26 of HCM6 (23), the average difference between the post-breakdown and the pre-breakdown flow rates vary widely from as little as 2% to as much as 20%. In light of this information, the limit of post-breakdown flowrate is set as 80%-98% of the pre-breakdown flowrate. The mathematical expressions for pre-breakdown and post-breakdown flow rates are shown in Eq. 5-5 and Eq. 5-6.

\[ q_{\text{PRE}} = k_{B,r=1} * u_{F,r=1} * \left( 1 - \left( \frac{k_{B,r=1}}{k_{J,r=1}} \right)^{l_{r=1}-1} \right)^{\frac{1}{1-m_{r=1}}} \]  \hspace{1cm} \text{Eq. 5-5} \\

\[ q_{\text{POST}} = k_{B,r=2} * u_{F,r=2} * \left( 1 - \left( \frac{k_{B,r=2}}{k_{J,r=2}} \right)^{l_{r=2}-1} \right)^{\frac{1}{1-m_{r=2}}} \]  \hspace{1cm} \text{Eq. 5-6}
Constraint: \( q_{\text{PRE}} \ast 0.98 \geq q_{\text{POST}} \geq q_{\text{PRE}} \ast 0.8 \).

5.3.3.2 Constraint on Slope at Capacity

It is shown earlier in Figure 5-2 that the capacity is underestimated by the unconstrained model as the slope of the uncongested regime flattens out, and the regime continues to a high-density value. The slope of the flow-density curve at any point of the uncongested regime can be obtained by differentiating the model flow equation with respect to density.

\[
\frac{dq_{M,i,r=1}}{dk_i} = u_{F,r=1} \left( 1 - \left( \frac{k_i}{k_{J,r=1}} \right)^{l_{r=1} - 1} \right) \frac{1}{1-m_{r=1}}
- k_i \left[ \frac{1}{1-m_{r=1}} \ast \left( 1 - \left( \frac{k_i}{k_{J,r=1}} \right)^{l_{r=1} - 1} \right) \frac{m_{r=1}}{1-m_{r=1}} \ast \frac{l_{r=1} - 1}{k_{J,r=1}} \ast \left( \frac{k_i}{k_{J,r=1}} \right)^{l_{r=1} - 2} \right]
\]

Eq. 5-7

Substituting \( k_i = k_{B,r=1} \) in Equation 5-7 yields the slope at capacity. To resolve the issue of decreasing slope at capacity, the slope at capacity value in HCM6 is used as a guideline. According to Chapter 10 and Chapter 25 of HCM6, the equations for flow in a basic freeway segment at different density regimes are shown in Equation 5-8.

\[
q_{\text{HCM},i} = \begin{cases} 
  k_i \ast u_F & \forall k_i \in [0, k_B], \\
  q_B + \sqrt{1 - 4 \ast \beta \ast q_B \ast k_i + 4 \ast \beta \ast u_F \ast k_i^2 - 1} & \forall k_i \in (k_B, 45], \\
  q_C \ast \left( 1 - \frac{k_i - 45}{k_j - 45} \right) & \forall k \in (45, k_j]. 
\end{cases}
\]

Eq. 5-8

Here,

\( q_{\text{HCM},i} \)=HCM6 model flow for observation \( i \);

\( \beta \)= A coefficient defined as a function of free flow speed;

\( q_B \)= Flow breakdown (pc/hr/ln) as a function of free flow speed;

\( q_C \)= Capacity of the segment (pc/hr/ln) as a function of free flow speed and
\( k_B \) = Density at breakpoint (pc/mi/ln).

To obtain the slope at capacity equation, Equation 5-8 for density range 45 to \( k_B \) is differentiated with respect to the observed density and the following expression is obtained.

\[
\frac{dq_{HCM,i}}{dk_i} = \frac{1 + \frac{k_i(-4q_B\beta + 8\beta u_F k_i)}{2\sqrt{1 - 4q_B\beta k_i + 4\beta u_F k_i^2}} - \sqrt{1 - 4q_B\beta k_i + 4\beta u_F k_i^2}}{2\beta k_i^2}. \quad \text{Eq. 5-9}
\]

In Equation 5-9, plugging \( k_i = 45 \) gives the slope at capacity value according to the HCM6 model. The proposed constraint here is that the slope at capacity according to the proposed model must be greater than or equal to the HCM6 slope at capacity value.

### 5.3.3.3 Constraint on Jam Density

Jam density is defined by the HCM6 as “the maximum density that can be achieved on a segment.” Jam density occurs under congested flow conditions as speed approaches zero. According to the HCM, the default value for jam density is 190 pc/mi/ln. However, is the HCM6 recommends that jam density be based on site specific data if such data is available. The recommended range of site specific jam density is given in the HCM6 as 150 to 270 pc/mi/ln. Therefore, the fitted jam density for the congested regime in the proposed model is capped at 270 pc/mi/ln.

### 5.3.4 Robust Regression

Although the initial stage of model development process involves removing clearly inconsistent and non-stationary data from the raw datasets, outlying observations are likely to remain in the data after the stage one filters. Since the method described above is data driven, it is imperative to fit the model through valid observations only. A customized application of
robust regression method is proposed in this study. According to this approach, the standard error for each data point is estimated by fitting the model. Data points with a standard error higher than a certain threshold are removed from the original dataset. Then, the model is fitted again with the updated dataset. The process is continued until the maximum standard error becomes lower than the threshold.

In this regard, the selection of the threshold is critical. First, there is a significant difference in the number of data points in the uncongested and congested regime. Therefore, it is important to distinguish the two regimes and estimate the standard error for each of them separately. The estimation of the standard error for each regime is shown in Eq. 5-10

\[ SE_{i,r} = \frac{q_{M,i,r} - q_i}{\text{Std}_r}, \]  

Eq. 5-10

where \( \text{Std}_r \) = Standard deviation of flow error for regime \( r \).

The second issue with robust regression is that there are two types of observations that need to be removed: mixed state observations and observations with extreme measurement error. The mixed state data points are valid observations that represent time intervals in which there was a switching between congested and uncongested regimes. Following the initial cleaning described above, such observations that remain are expected to be more prevalent in the congested regime than in the uncongested regime as illustrated in Figure 5-1 (b). The distribution of the mixed state observations will be asymmetric around the flow-density curve of congested regime (lying on the left side of the curve). Because these observations do not result from measurement error, they are likely to be more prevalent than measurement errors. For example, for every flow breakdown, there is likely to be two pronounced mixed state observations, one for the time interval when the queue formation shock wave passed by and one for the time interval
when the queue clearance shock wave passed by. Considering these facts, a symmetric threshold for the standard error of ±3.5 is applied for the uncongested regime. This symmetric threshold is expected to remove the observations with high measurement error in the uncongested regime. For the congested regime, an asymmetric threshold of +2 is applied for all but the final step of the robust regression. This threshold is expected to remove the remaining mixed state data points and the measurement error outliers on the left side of the flow-density curve. In the final step of robust regression, a threshold of -3.5 is applied to exclude any remaining outliers from the congested regime on the right side of the flow-density curve.

The symmetric threshold of ±3.5 standard error applied on the uncongested regime observations represents a confidence interval of approximately 99.95% (assuming that the errors have Gaussian distribution). On the other hand, a less conservative threshold is applied to the congested regime due to the potential presence of mixed state observations. With these set of thresholds, the robust regression process is expected to converge without removing excessive data points.

5.3.5 The HCM6 Model: As a Continuous Fundamental Diagram

As shown earlier in Equation 5-6, if queue discharge capacity reduction is ignored, the HCM6 models the freeway fundamental diagrams as a three-regime continuous model.

According to this model, the breakpoint, capacity, and the coefficient $\beta$ is expressed as a function of free flow speed. Jam density has a default value of 190 pc/mi/ln, but the HCM6 recommends calibration when field data are available.

Since the comparison between the proposed and HCM6 model is an integral part of this study, the parameters need to be calibrated in a manner that provides a best fit model. Here, it is proposed to calibrate the capacity and jam density using the observed flow by minimizing the
sum of squared error. Free flow speed is estimated as the average speed of observations with $\text{Flow} \leq 1,000 \text{ pc/hr/ln}$ as per guidance provided in HCM6.

In this study, the proposed discontinuous steady state model described above is to be compared with the HCM6 model based on two statistical parameters: mean standard error (MSE) of fitted flow rates and Bayesian information criterion (BIC) (Schwarz, 1978). The MSE value represents the average of squared error for each observed flow. However, the MSE value can be misleading in judging two models if the parameterization of the two models varies greatly. BIC is an effective tool to resolve address issue since it penalizes models for having additional parameters.

5.4 DESCRIPTION OF DATA AND STUDY AREA

To illustrate the application of the proposed model, traffic stream data was used from five HERE side-fire radar stations located near the Triangle Area of Raleigh, North Carolina for the calendar year of 2013. The downloadable data is aggregated in 5-minute intervals. Each sensor is located adjacent to a basic freeway segment according to the HCM6 segmentation rules.

5.5 RESULTS

To fit the traffic stream model by imposing the constraints described above, a nonlinear optimization tool available in MATLAB was used. The run time for each model fitting varied from 1-3 minutes. Results from data filtering, fitted models, interpretation, and sensitivity of the parameters obtained are described in the following subsections.
5.5.1 Application of CST, CDT, and SFDT

This section summarizes the two stages of cleaning inconsistent, mixed state, and outlying observations. Figure 5-3 shows the percentage of data reduced by the initial thresholds (CST, CDT, and SFDT) and by robust regression.

![Percentage of data reduction in different sensors.](image)

The most striking observation from Figure 5-3 is the higher percentage of data reduced by the initial thresholds for sensor 23786 EB (7%) compared to that for other sensors. Data exclusion by robust regression is lower than that by the initial filters across all the sensors. The highest proportion of data reduced by it is only about 3%. The proportion of outlier detection may also depend upon the health of the detectors. Furthermore, in no case was the number of outliers more than 10% in total.

5.5.2 Fitted Models

The final parameter values used to plot the flow-density and speed-flow fundamental diagrams as shown in Figure 5-4 are the parameters resulting from convergence of the robust
regression algorithm. The estimated values along with their standard deviations are listed in Table 5-1. Standard deviation for such non-linear optimization models is calculated using a method described in a past study (Smith, 2013).
Figure 5-4: Flow-density and Speed-flow diagrams for 5 sensors.
Table 5-1: Parameter values for different sensors obtained from the fitted model.

<table>
<thead>
<tr>
<th>Sensor ID</th>
<th>$u_{F,r=1}$ (mph)</th>
<th>$k_{B,r=1}$ (pc/mi/ln)</th>
<th>$k_{J,r=1}$ (pc/mi/ln)</th>
<th>$l_{r=1}$</th>
<th>$m_{r=1}$</th>
<th>$u_{F,r=1}$ (mph)</th>
<th>$k_{B,r=2}$ (pc/mi/ln)</th>
<th>$k_{J,r=1}$ (pc/mi/ln)</th>
<th>$l_{r=2}$</th>
<th>$m_{r=2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>23771 WB</td>
<td>63</td>
<td>39.9</td>
<td>231.2</td>
<td>4.26</td>
<td>0.98</td>
<td>1.00E+05</td>
<td>37.4</td>
<td>270.0</td>
<td>1.03</td>
<td>0.62</td>
</tr>
<tr>
<td>23774 WB</td>
<td>63</td>
<td>36.4</td>
<td>78.8</td>
<td>11.82</td>
<td>1.00</td>
<td>1.00E+05</td>
<td>33.4</td>
<td>270.0</td>
<td>1.05</td>
<td>0.68</td>
</tr>
<tr>
<td>23786 EB</td>
<td>64</td>
<td>44.0</td>
<td>411.8</td>
<td>5.16</td>
<td>1.00</td>
<td>4.80E+05</td>
<td>38.0</td>
<td>270.0</td>
<td>1.02</td>
<td>0.67</td>
</tr>
<tr>
<td>29953 WB</td>
<td>65</td>
<td>48.0</td>
<td>521.5</td>
<td>4.78</td>
<td>1.00</td>
<td>2.45E+06</td>
<td>40.0</td>
<td>270.0</td>
<td>1.03</td>
<td>0.73</td>
</tr>
<tr>
<td>29995 WB</td>
<td>62</td>
<td>42.0</td>
<td>208.4</td>
<td>5.50</td>
<td>0.99</td>
<td>4.69E+05</td>
<td>38.5</td>
<td>269.4</td>
<td>1.02</td>
<td>0.63</td>
</tr>
</tbody>
</table>

In general, the plots in Figure 5-4 illustrate that the fitted models reasonably follow through the steady-state observations. Here, the value of the parameters that have physical interpretation needs to be discussed. The distance head exponent ($l$) and speed exponent ($m$) for both regimes do not have any straightforward physical meaning. Same statement applies to the jam density and free flow speed for the uncongested and congested regime respectively. From the fitted speed-flow plots in Figure 5-4, the free flow speed for the uncongested regime ($u_{F,r=1}$) is evident. The density breakpoint for the uncongested regime ($k_{B,r=1}$) is less than HCM6 default of 45 pc/mi/ln for all but sensor 29953 WB. This parameter along with $u_{F,r=1}$ are of importance for estimating the capacity of the segment at the sensor locations. According to the HCM6 default method, the capacity of a basic roadway segment varies between 2,320 to 2,350 pc/hr/ln for the given range of $u_{F,r=1}$ (62-65 mph). However, the site-specific capacities appear from the plots in Figure 5-4 are apparently less for sensor 23771 WB and 23774 WB and more for sensor 23786 EB and 29953 WB than the HCM6 defaults. The close agreement of HCM6 default capacity to only one out of five sensors data (29995 WB) underscores that the national average capacity values provided by HCM6 need to be calibrated with field data if high fidelity analysis is desired.
The difference between the two density break points represents the overlap range. This range appears to be a unique characteristic for the five locations based on the proposed approach modeled. The fitted overlap in density varies from 2.5 to 8 pc/mi/ln across the five sensors. This site-specific modeling of overlap along with the drop-in capacity have many potential applications such as estimating pre-breakdown and queue discharge flow rate, identifying recurring freeway bottlenecks, etc.

As described earlier, the jam density for the congested regime \((k_{J, r=1})\) is capped at the maximum feasible value mentioned in HCM. Since it was observed in Figure 5-1 that this value tends to go beyond the limit of feasibility, all the fitted values of \(k_{J, r=2}\) are suppressed to this maximum limit by the constraint. However, it was inspected that the overall fit and the key parameter values were not significantly affected after introducing this constraint.

The standard deviations of the estimates shown within parenthesis in Table 5-1 are of small magnitude. It reveals that most of these estimates are significant and the fitted models are sensitive to these parameters. It should be noted that the standard deviations for the two breakpoints, \(k_{B, r=1}\) and \(k_{B, r=2}\) cannot be estimated since these are only classifiers of the regimes. Standard deviation for \(u_{F, r=1}\) is very high because the fitted models are insensitive to this parameter.

5.5.3 Comparison with HCM6 Models

Two forms of steady state HCM6 traffic stream model, with and without capacity drop was applied to the data from the five sensors discussed above. A capacity drop between 2% to 20% is acknowledged by HCM6 when traffic state transfers from uncongested to congested regime. As an example, the fundamental diagrams for sensor 29995 WB obtained by fitting both forms of HCM6 model is shown in Figure 5-5
The free flow speed, density breakpoint, capacity, jam density, and capacity drop obtained from the HCM6 models for five sensors are shown in Table 5-2 below. One important observation from Figure 5-5 which is also found for other sensors is that the linear flow-density relationship for the congested regime is unable to capture most of the data points. In the continuous model, this lack of fidelity is due both to the linear shape and to the constraint imposed by the continuity with the uncongested regime. The poor fit results in an underestimation of capacity for three sensors compared to the HCM6 defaults (top three sensors in Table 5-2). The discontinuous model balances this underestimation of capacity to some extent. The jam density is bounded to its minimum reasonable value (150 pc/mi/ln) for three sensors for both forms of model. The capacity drop is constrained at its maximum limit for four sensors. As this discontinuous model fitting process is applied to more freeway data sets, factors related to capacity drop can be studied in more depth. The free flow speed values estimated by HCM6 do not differ much from what was obtained from the proposed model.
Table 5-2: Parameter values for five sensors obtained by the HCM6 model.

<table>
<thead>
<tr>
<th>Sensor ID</th>
<th>$u_F$ (mi/h)</th>
<th>$k_B$ (pc/mi/ln)</th>
<th>Capacity (pc/hr/ln)</th>
<th>$k_C$ (pc/mi/ln)</th>
</tr>
</thead>
<tbody>
<tr>
<td>23771 WB</td>
<td>65</td>
<td>21.5</td>
<td>1860</td>
<td>165.3</td>
</tr>
<tr>
<td>23774 WB</td>
<td>60</td>
<td>26.7</td>
<td>1808</td>
<td>118.8</td>
</tr>
<tr>
<td>23786 EB</td>
<td>65</td>
<td>21.5</td>
<td>2155</td>
<td>105.9</td>
</tr>
<tr>
<td>29953 WB</td>
<td>65</td>
<td>21.5</td>
<td>2400</td>
<td>127.2</td>
</tr>
<tr>
<td>29995 WB</td>
<td>60</td>
<td>26.7</td>
<td>2348</td>
<td>178.8</td>
</tr>
</tbody>
</table>

To compare the goodness of fit and information criteria of the HCM6 models with the proposed one, Figure 5-6 illustrates the mean squared error (MSE) of flow and Bayesian information criteria (BIC) for the three models in question.

![Figure 5-6: Comparison of MSE and BIC values obtained from the proposed and HCM model.](image)

It is apparent from Figure 5-6 that the MSE value is highest for the continuous HCM6 model and lowest for the proposed model across all five sensors. Similar trend exists in terms of the BIC values of the models. Despite the proposed model is penalized because of the additional fitted parameters which is reflected by the BIC values, the BIC values across all the sensors are lowest for the proposed model. The HCM6 discontinuous model is better than the continuous
one. Nonetheless, the proposed model would be considered preferable to the HCM6 models in terms of both information criterion and squared error.

5.6 SUMMARY AND CONCLUSIONS

This study has proposed an empirical process for fitting a discontinuous two-regime steady state model to data from basic freeway segment sensors. The transition range, also known as the overlap range is modeled by comparing the deviation of the two regimes from the observed data. Three HCM6-justified constraints were added to reasonably fit the model. Non-stationary, mixed-state, and outlying data points were addressed explicitly in two stages. Three initial thresholds and a tailored application of robust regression were applied to filter the invalid observations.

The proposed model was applied to five basic freeway segments near Raleigh, North Carolina. Fundamental diagrams for each of the sensors show acceptable fits for the proposed model. Discussions on the interpretation of the parameters revealed that all the parameters obtained from the fitted models represent reasonable values. The site-specific capacity values apparent from the observed data and echoed in the fitted models show deviation from the national average default values provided by the HCM6.

Two forms of HCM6 model, continuous (without capacity drop) and discontinuous (with capacity drop) are fitted to the same five sensors data in the best possible way and the fitted fundamental diagrams are shown. It appears that the linear flow-density relationship in the congested regime by the continuous HCM6 model could not capture the observed data points properly with this poor fit compounded by the constraint of continuity between the two regimes. Moreover, the capacity resulting from most of the continuous models are substantially lower compared to both the national defaults and to what is obtained from the proposed model. The
insertion of discontinuity in the HCM6 model minimizes these deviations to some extent. Statistical comparison of the results showed that the MSE of flow is highest for HCM6 continuous and lowest for the proposed model across all five sensors. Although the BIC penalizes the proposed models for additional parameters, it nonetheless indicates a preference over all of the HCM6 models.

The next chapter (Chapter 6) of this dissertation leverages this proposed discontinuous macroscopic GHR model to estimate the change in driver reaction time required for car-following stability. Moreover, it is expected that the proposed model will be useful to researchers and practitioners for estimating capacity, identifying bottlenecks, and other freeway operational analyses. Moreover, observed phenomena related to traffic state transition such as queue discharge capacity drop can be modeled with the proposed technique as well.

However, the study does not come without limitations. The nonlinear optimization technique used in this study may require extensive period of time for computation if traditional license-based (e.g. Excel) or open source (e.g. R) software is used. Also, it is recognized by the authors that the first-stage removal of unwanted data points recommended in this study may remove some valid data points. In future, attempts will be taken to combine the two-stage data reduction process into a single robust process that is more efficient and consistent.
5.7 REFERENCES


CHAPTER 6 – INVESTIGATING THE RELATIONSHIP BETWEEN FREEWAY REAR-END CRASH RATES AND MACROSCOPICALLY MODELLED REACTION TIME

6.1 INTRODUCTION

Understanding traffic characteristics that may lead to rear-end crashes is not only important for improving safety, but also important to alleviate non-recurrent congestion on the roads. Past studies showed that rear-end crashes are strongly associated with traffic instability and oscillations (Tanaka, Ranjitkar, & Nakatsuji, 2008; Touran, Brackstone, & McDonald, 1999). Due to traffic oscillations, drivers often fail to react in time and collide with the vehicle in front. The state of the practice as documented in the Highway Safety Manual (American Association of State Highway and Transportation Officials, 2010) is used to identify rear-end crash-prone locations.

However, highway safety engineering is founded on statistical analyses of crash occurrences. Therefore, the identification of rear-end crash-prone locations would be improved if the crash risk could be predicted using non-crash data, for instance, the macroscopic traffic characteristics of a site.

The steady state macroscopic flow model equivalent to the so-called Gazis-Herman-Rothery (GHR) (Gazis, Herman, & Rothery, 1961) car-following model provides a rational means to estimate the reaction time necessary for asymptotic stability (May, 1990). This model can be fitted to observed macroscopic traffic data and subsequently, to analytically estimate the driver reaction time necessary for asymptotic stability. Here, it is hypothesized that this analytically estimated driver reaction time required for asymptotic stability can serve as an effective indicator of the impact of traffic oscillations on rear-end crashes.
This study investigates the relationship between freeway rear-end crash rates and macroscopically derived reaction time required for asymptotic stability. Since rear-end crashes tend to occur more during traffic state transitions, it specifically focuses on the change in required reaction time when traffic state shifts from the uncongested to the congested regime. The discontinuous form of macroscopic GHR model that is developed in Chapter 5 of this dissertation is leveraged to estimate this change in reaction time. The relationship between long-term rear-end crash rates of a freeway segment and change in required reaction time at that location is investigated using regression analysis.

This chapter is organized as follows. A review of past studies on the topic of this chapter is presented in the next section. The following section briefly recaps the discontinuous form of GHR macroscopic model that was described in Chapter 5 of this dissertation. Next, the mathematical formulation for estimating the change in driver reaction time required for stability is explained. Rear-end crash data collection and crash rate estimation method is presented in the following section. Results from the fitted models, rear-end crash analysis, and statistical regression analysis are discussed next. Finally, the summary of the results and limitations of the study is presented in the last section.

6.2 LITERATURE REVIEW

May (1990) defined an unstable driver behaviour as when a following driver responds slowly to the change in the speed of the leading vehicle, but when responds, exerts large acceleration or deceleration rates. Therefore, the stability of car-following behaviour of two vehicles, or in a broader term, of a traffic stream is a function of the driver reaction time and sensitivity. The General Motor researchers developed an analytical formulation of the sensitivity term and showed that the product of this sensitivity and reaction time dictates the stability of car-
following behaviour both on a local and asymptotic scale. Details on the mathematical condition of asymptotic stability are discussed later in this chapter. In recent decades, several studies investigated the relationships among these critical car-following parameters, namely, reaction time, sensitivity, headway (Ahn, Laval, & Cassidy, 2010; Kim & Zhang, 2011; Xu & Laval, 2019; Zielke, Bertini, & Treiber, 2008). Kim and Zhang investigated the relationship among these parameters based on the GHR and other simpler car-following models using car-following behaviour data extracted from the NGSIM (USDOT, 2006) dataset. Another study (Zielke et al., 2008) conducted a comparative investigation of traffic oscillation in terms of the amplitude, propagation velocity, and frequency of shockwave across three countries. Ahn et al. (2010) analytically proved with a triangular-shaped fundamental diagram that the amplitude of an oscillation through a queued traffic stream dampens as it passes an on-ramp and magnifies as it passes an off-ramp.

In addition to car-following behaviour and traffic stability analysis, a few studies extended their focus on these parameters’ effects on crash risk (Chatterjee & Davis, 2016; Davis & Swenson, 2006; Misener, Tsao, Song, & Steinfeld, 2000; Tanaka et al., 2008). Tanaka et al. (2008) used a microscopic dataset of ten trajectories to investigate the effect of reaction time and sensitivity on asymptotic stability using different safety indicators. Although several safety indicators showed that traffic oscillation propagates downstream when the product of sensitivity and reaction time exceeds the threshold proposed by Chandler, Herman, and Montroll (1958), other safety indicators yielded an inconclusive result. Nonetheless, it successfully demonstrated the effect of reaction time on traffic stability. To explore the causal relationship between crash occurrence and following headway along with reaction time, Davis and Swenson used video recorded microscopic traffic data and the kinematic theory developed by Brill (1972) and
simulated three real-world rear-end crashes. It revealed that had the colliding vehicles or a few vehicles further downstream in the sequence maintained a higher following headway than their reaction time, the collisions would probably have been avoided. Chatterjee and Davis extended the analytical formulation of crash occurrence using car-following theory by Brill for a series of vehicles in a platoon. It perceived the stopping distance for the braking of the first car in a platoon as a shared resource and this resource is either consumed or contributed by the following vehicles. If the sum of the consumption of this resource by the following vehicles exceeds a threshold, rear-end crash occurs. This theory was verified by observing 41 shockwaves in which 15 rear-end crashes or swerving events to avoid crashes were occurred. To develop a probabilistic model for rear-end crash occurrence using trajectory data, Oh and Kim (2010) used two probability measures, namely probability of changing lanes and probability of the following vehicle hitting the leading one for a given “Time to Collision”. However, the accuracy of the proposed approach was not tested against field crash data. While the availability of microscopic information on driver reaction time level is increasing with the increasing research using instrumented vehicles (Tanvir, Chase, & Roupahil, 2019), microscopic traffic data like acceleration, following headway, and instantaneous reaction time is still difficult to obtain by transport agencies. Therefore, the practical application of the studies discussed above is limited since these demand several micro-level inputs like acceleration, following headway, and reaction time.

Several studies attempted to predict the rear-end crash risk of a roadway using macroscopic traffic data (Abdel-Aty, Uddin, Pande, Abdalla, & Hsia, 2004; Lord, Manar, & Vizioli, 2005). Among these, Lord et al. used traffic density and volume-to-capacity ratio as explanatory variables to predict freeway crashes. Abdel-Aty et al. showed that a combination of
high coefficient of variation of speed and high occupancy at a downstream segment is a potential crash pre-cursor in the upstream segment of a roadway. Pande and Abdel-Aty (2006) divided rear-end crashes into two groups based on whether they occur before or during congestion. Average and coefficient of variation of speed, average occupancy, and presence of a downstream ramp were the statistically significant variables to predict rear-end crashes. However, the discrepancies in the findings of these studies indicate that the findings are mostly site-specific.

The above survey of literature shows that while many studies analyzed the car-following model parameters related to driver reaction time, only a few focused on its relation to rear-end crash rates. These few studies used microscopic trajectory level data that are difficult to obtain on a network-level. Although several studies investigated the crash precursor potential of various macroscopic traffic characteristics, those mostly focused on a single site and the findings are mostly site-specific.

6.3 METHODOLOGY

In this section, first, the macroscopic model equivalent to the GHR car-following model as developed in Chapter 5 of this dissertation is recapped. Next, the derivation of driver reaction time required for asymptotic stability is described. Finally, the description of the study site and crash data collection and analysis method are presented.

6.3.1 The Macroscopic Model Equivalent to the GHR Car-following Model

The method described in Chapter 5—starting from the initial cleaning of traffic flow data as presented in Section 5.3.1 to modeling the transition regime with an overlap as shown in Section 5.3.2— is adopted by this chapter to develop the macroscopic version of GHR model. The first two HCM-based constraints—one related to the slope of the uncongested regime curve at capacity and the second one associated with the queue discharge flow rate—are also adopted
in this study. However, the application of the third constraint on jam density resulted in a very poor fit of the congested regime curve for some sites. Consequently, the estimated driver reaction time required for asymptotic stability (shown later in Equation 6-5 of this chapter) exhibited unrealistic values for those sites. Therefore, the jam density constraint is not included in the model fitting process. The recursive fitting of the proposed model based on a modified robust regression approach as shown in the previous chapter is also adopted by this chapter.

For better readability, the equations for fitting the two-regime GHR macroscopic models proposed in Chapter 5 are rewritten below.

**Observed Density**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Formula for model speed</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(k_i \leq k_{B,r=2})</td>
<td>[q_{M,i,r=1} = k_i \cdot u_{F,r=1} \cdot \left(1 - \left(\frac{k_i}{k_{J,r=1}}\right)^{l_{r=1}^{-1}}\right)^{1-m_{r=1}}]</td>
<td>Eq. 5-2</td>
</tr>
<tr>
<td>(k_i \geq k_{B,r=1})</td>
<td>[q_{M,i,r=2} = k_i \cdot u_{F,r=2} \cdot \left(1 - \left(\frac{k_i}{k_{J,r=2}}\right)^{l_{r=2}^{-1}}\right)^{1-m_{r=2}}]</td>
<td>Eq. 5-3</td>
</tr>
<tr>
<td>(k_{B,r=2} &lt; k_i &lt; k_{B,r=1})</td>
<td>[q_{M,i,r} = \begin{cases} q_{M,i,r=1} &amp; \text{if }</td>
<td>q_{M,i,r=1} - q_i</td>
</tr>
</tbody>
</table>

Here,

\(i=1,2,3\ldots\) observation index,

\(q_{M,i}\) = Model flow for observation \(i\) (pc/hr/ln),

\(r=1, 2\) for uncongested and congested regime respectively,

\(u_i\) = free flow speed (mi/h),

\(k_i\) = jam density (pc/mi/ln),

\(k\) = Observed density (pc/mi/ln),

\(l\) = Distance headway exponent, and

\(m\) = Speed difference exponent.
6.3.2 Estimating Required Driver Reaction Time for Different Traffic States

The formula for estimating the required driver reaction time for asymptotic stability needs to be derived from the microscopic form of the fifth and final GHR car following model, which is expressed in Equation 6-1. Here, the acceleration of the \((n+1)\)st vehicle in a traffic stream at time \((t + \Delta t)\) (termed as \(x''_{n+1}(t + \Delta t)\)) in response to the relative speed between the \(n\)th and \((n+1)\)st vehicle at time \(t\) is expressed as the product of the sensitivity term and the relative speed between the two vehicles.

\[
x''_{n+1}(t + \Delta t) = \alpha \left[ \frac{x'_{n+1}(t + \Delta t)^m}{[x_n(t) - x_{n+1}(t)]^l} \right] * [x'_n(t) - x'_{n+1}(t)].
\]
Eq. 6-1

Here,

\(n\) = Position of a driver in a traffic stream, \((n=0\) is the most downstream driver),

\(x_n\) = Location of the \(n\)th driver with respect to a reference point,

\(x'_n(t)\) = Speed of the \(n\)th driver at time \(t\).

According to May (1990), the parameter \(\alpha\) can be expressed as shown in Equation 6-2.

\[
\alpha = \frac{(l - 1)u_F^{l-m}}{(1 - m)k^{l-1}_f}.
\]
Eq. 6-2

Thus, the sensitivity factor is equivalent to what is shown in Equation 6-3

\[
\text{Sensitivity factor} = \frac{(l - 1)u_F^{l-m}}{(1 - m)k^{l-1}_f} * \left[ \frac{x'_{n+1}(t + \Delta t)^m}{[x_n(t) - x_{n+1}(t)]^l} \right] * [x'_n(t) - x'_{n+1}(t)].
\]
Eq. 6-3

For steady state observations, individual vehicle speed represents the average speed of the traffic stream \((u)\) and the spacing between two successive vehicles represents the inverse of the average traffic density \((k)\). Thus, Equation 6-4 can be written as:
Sensitivity factor \[= \frac{(l - 1)u_F^{1-m}}{(1 - m)k_J^{l-1}} \frac{u^m}{(1/k)^l}. \quad \text{Eq. 6-4} \]

According to May (1990), for a traffic stream to be asymptotically stable, the product of the reaction time and sensitivity must be less than or equal to 0.5. Hence, the expression for the reaction time \((t)\) required for asymptotic stability for the \(i\)th observation can be derived as:

\[t_i = \frac{(1 - m)k_j^{l-1}}{2(l - 1)u_F^{1-m}} \frac{(1/k_j)^l}{u^m}. \quad \text{Eq. 6-5} \]

Equation 6-5 gives the formula to estimate the driver reaction time required for stability for each observation of flow, speed, and density.

**6.3.3 Change in Required Reaction Time between Two Regimes**

Fitting the proposed discontinuous flow-density model as shown in Equation 5-2, 5-3, and 5-4, and with the empirical constraints and using the robust regression technique enabled the research team to investigate the change in driver reaction time when the traffic state moves from uncongested to the congested regime. If the reaction time described in Equation 6-5 is estimated using the fitted parameters \((u_F, k_J, l, m, \text{ and } k_B)\) for each regime) for a site for a series of density values, the following curves are obtained.
Figure 6-1: A typical required reaction time vs. density plot.

Upon scrutinizing the same plot of reaction time vs. density for different locations, it was found that the reaction time for the congested regime does not vary significantly. Moreover, the reaction time for a very low density value (such as less than 20) may not have any significance as traffic stream barely follows the car-following model. On the other hand, if the transition regime is focused here (bounded by the two vertical arrows in Figure 6-1), there are two reaction times for each density point within this regime. As the traffic state transfers from the uncongested to the congested regime, the required driver reaction time also reduces to get adapted to the change in traffic state. The research team hypothesizes that the higher the change (i.e., reduction) in reaction time, the higher the risk of a rear-end crash to occur.

The two changes in reaction time shown in Figure 6-1 are of particular interest here. These two changes termed as $\Delta t_s$ and $\Delta t_e$ are the changes in reaction time at the start and the end of the transition regime, respectively. $\Delta t_s$ is the highest change in reaction time within the overlap. It is the value by which the minimum driver reaction time needs to be changed to maintain asymptotic stability when traffic state transfers from the uncongested to the congested regime at a density of $k_{B,r=2}$. On the other hand, $\Delta t_e$ is the change in minimum driver reaction
time when traffic state transfers the regime upon reaching the capacity and at a density of $k_{B,r=1}$.

The exploration of the association between rear-end crash rates and change in driver reaction time in this study revolves around these two extreme changes in driver reaction time.

6.4 DATA DESCRIPTION

6.4.1 Traffic Data from Sensors

In this study, the discontinuous macroscopic model described above is proposed to be fitted with field data collected from side-fire radar sensors located on different locations of the freeway system of the Triangle Region of North Carolina. Flow, speed, and lane occupancy data from 28 directional sensors are collected for the calendar year of 2013 in a time resolution of 5 minutes. These sensors are located on three interstates namely I-40, I-440, and I-540. Since this study primarily hinges on driver car-following model, only basic freeway segments (see HCM6 for definition) are selected to ensure that the lane changing activities are minimum near the sensors. Figure 6-2 shows the location of these sensors in the study area. Past studies showed that several recurring bottlenecks exist in the proximity of some of these locations (Ahmed, Routhail, & Tanvir, 2018). The numbers show the tag of each station and the alphabets attached to them indicate the travel direction.
Figure 6-2: Location of the side-fire radar sensors in the study area.

6.4.2 Crash Data

Selecting an appropriate freeway segment surrounding the sensor location was the most critical task when estimating the rear-end crash rate associated with a sensor. As explained in the Highway Safety Manual, the crash rate at a segment can be attributed to several geometric features including the vertical and horizontal curvature, lane width, number of lanes, and presence of ramps. Here, each segment was selected in such a way that it is away from any ramp, does not have any tight curve, and the number of lanes and lane width is consistent throughout the segment.

Four years of police-reported crash data from the selected segments are collected for the period from 2011 to 2014. A tool called “Traffic Engineering Accident Analysis System (TEAAS)” (North Carolina Department of Transportation) was used to extract the police reports. Each crash report was carefully investigated to decide if that crash occurred within the corresponding segment. The average AADT of the selected years is obtained from the archive of
Crash rate = \( \frac{100,000,000 \times C}{(365 \times N \times \overline{AADT} \times 0.5 \times L)} \),

where \( N \) = number of years over which crash data were collected (\( N = 4 \)),

\( C \) = total frequency of crashes in \( N \) years,

\( \overline{AADT} \) = average AADT over \( N \) years, and

\( L \) = segment length in miles.

After estimating the crash rate for each segment, its association with the two changes in required reaction time (\( \Delta t_s \) and \( \Delta t_e \)) are investigated using statistical modeling through regression analysis.

6.5 ANALYSIS AND RESULTS

This section is divided into three major parts. In the first part, results from fitting the proposed two-regime traffic flow model is presented. The fitted parameter values and estimated changes in required reaction time are also discussed here. In the second part, results from the rear-end crash rate analysis are described. Finally, the relationship between the changes in driver reaction time and rear-end crash rates are assessed.

6.5.1 Fitted Two-Regime Traffic Flow Models

To fit the traffic stream model by imposing the constraints described earlier, a nonlinear optimization tool available in MATLAB was used. The fundamental diagrams for the sensor 18W are shown Figure 6-3. The parameter values used to plot these diagrams are the fitted parameters.
obtained from the convergence of the proposed robust regression algorithm. Table 6-1 presents the estimated parameter values, their standard deviations, and the resulting required reaction time changes (Δt_s, Δt_e). Standard deviation for such nonlinear optimization models is calculated using a method described in a past study (Smith, 2013).

Figure 6-3: Fundamental diagrams for sensor 18W (a) Flow vs. Density, (b) Speed vs. Flow.
Table 6-1: Fitted parameter values for different sensors and the resulting reaction time changes.

<table>
<thead>
<tr>
<th>ID</th>
<th>$u_f$ (mi/h)</th>
<th>$k_b$ (pc/mi/ln)</th>
<th>$k_j$ (pc/m/ln)</th>
<th>$l$</th>
<th>$m$</th>
<th>$u_f$ (mi/h)</th>
<th>$k_b$ (pc/mi/ln)</th>
<th>$k_j$ (pc/m/ln)</th>
<th>$l$</th>
<th>$m$</th>
<th>$\Delta t_s$ (sec)</th>
<th>$\Delta t_e$ (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1W</td>
<td>65 (5.9E-3)</td>
<td>40 (9.1E-2)</td>
<td>88.9 (2.8E-3)</td>
<td>4.418</td>
<td>0.609</td>
<td>4.7E5</td>
<td>466 (3.5E-1)</td>
<td>1.022</td>
<td>0.684</td>
<td>1.51</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>2E</td>
<td>62 (8.0E-3)</td>
<td>39 (4.6E-1)</td>
<td>81.5 (5.7E-2)</td>
<td>10.37</td>
<td>0.985</td>
<td>4.5E5</td>
<td>467 (3.3E-1)</td>
<td>1.022</td>
<td>0.677</td>
<td>2.82</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>3E</td>
<td>56.4 (6.9E-3)</td>
<td>42 (5.3E+0)</td>
<td>129.8 (2.5E-1)</td>
<td>9.333</td>
<td>0.99</td>
<td>1.2E7</td>
<td>337 (1.7E-1)</td>
<td>1.003</td>
<td>0.594</td>
<td>2.63</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>4W</td>
<td>59.2 (6.0E-3)</td>
<td>39 (3.1E+1)</td>
<td>237.5 (3.1E+1)</td>
<td>8.215</td>
<td>0.99</td>
<td>4.4E5</td>
<td>641 (7.3E-1)</td>
<td>1.04</td>
<td>0.757</td>
<td>2.67</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>5W</td>
<td>61.7 (5.0E-3)</td>
<td>44 (2.2E-1)</td>
<td>209.6 (2.5E-3)</td>
<td>5.47</td>
<td>0.993</td>
<td>1E5</td>
<td>2133 (2.1E+0)</td>
<td>1.134</td>
<td>0.884</td>
<td>1.8</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>6E</td>
<td>57.7 (1.1E-2)</td>
<td>50 (9.4E-1)</td>
<td>614.2 (1.6E-3)</td>
<td>3.964</td>
<td>0.997</td>
<td>3.3E5</td>
<td>1077 (6.2E-1)</td>
<td>1.062</td>
<td>0.808</td>
<td>1.51</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>7E</td>
<td>63.3 (1.0E-2)</td>
<td>39 (5.9E-1)</td>
<td>76.7 (9.0E-2)</td>
<td>11.016</td>
<td>0.981</td>
<td>5.7E6</td>
<td>918 (9.9E-1)</td>
<td>1.025</td>
<td>0.781</td>
<td>2.25</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>8W</td>
<td>61.7 (9.5E-3)</td>
<td>42 (3.2E-1)</td>
<td>146 (8.5E-3)</td>
<td>6.716</td>
<td>0.993</td>
<td>6.2E6</td>
<td>229 (3.4E-1)</td>
<td>1.002</td>
<td>0.538</td>
<td>1.56</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>9W</td>
<td>65.3 (6.5E-3)</td>
<td>48 (1.8E+1)</td>
<td>572 (4.2E-2)</td>
<td>4.862</td>
<td>0.99</td>
<td>1.4E6</td>
<td>1800 (8.1E-1)</td>
<td>1.11</td>
<td>0.894</td>
<td>2.19</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>10E</td>
<td>62.9 (1.2E-2)</td>
<td>42 (5.7E-1)</td>
<td>399.9 (1.0E-3)</td>
<td>2.973</td>
<td>0.955</td>
<td>3.4E5</td>
<td>1307 (1.8E+0)</td>
<td>1.116</td>
<td>0.877</td>
<td>1.67</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>11E</td>
<td>63.6 (1.3E-2)</td>
<td>37 (4.9E-1)</td>
<td>366.4 (1.6E-3)</td>
<td>4.149</td>
<td>0.996</td>
<td>3.4E5</td>
<td>505 (6.8E-1)</td>
<td>1.029</td>
<td>0.702</td>
<td>0.89</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>12W</td>
<td>65.1 (4.8E-3)</td>
<td>40 (5.1E-1)</td>
<td>451.7 (1.3E-3)</td>
<td>4.352</td>
<td>0.998</td>
<td>3.4E5</td>
<td>447 (4.5E-1)</td>
<td>1.119</td>
<td>0.846</td>
<td>2.14</td>
<td>1.63</td>
<td></td>
</tr>
<tr>
<td>13W</td>
<td>61.9 (6.6E-3)</td>
<td>44.6 (6.0E+0)</td>
<td>342.1 (3.6E-2)</td>
<td>5.609</td>
<td>0.99</td>
<td>1E5</td>
<td>1043 (7.3E-1)</td>
<td>1.119</td>
<td>0.849</td>
<td>1.27</td>
<td>0.88</td>
<td></td>
</tr>
</tbody>
</table>
### Table 6-1 (continued).

| 14W  | 61.4 (8.1E-3) | 45 | 729.5 (6.6E-1) | 3.395 (7.1E-4) | 0.994 (1.3E-5) | 3.4E5 (2.3E2) | 45 | 340 (1.8E-1) | 1.007 (2.3E-6) | 0.519 (3.6E-5) | 0.74 | 0.74 |
| 15E  | 62.4 (9.1E-3) | 40 | 1099.7 (5.9E+0) | 1.973 (1.4E-3) | 0.205 (4.2E-3) | 1E5 (5.3E1) | 36 | 1812 (6.9E-1) | 1.091 (9.3E-6) | 0.839 (1.1E-5) | 2.17 | 0.95 |
| 16W  | 67.1 (1.6E-2) | 42 | 183 (2.8E-1) | 3.056 (1.7E-3) | 0.792 (6.7E-4) | 2.9E6 (5.9E3) | 38 | 4207 (5.3E+0) | 1.066 (4.7E-7) | 0.879 (2.2E-5) | 1.33 | 0.96 |
| 17E  | 67.8 (6.1E-3) | 40 | 196 (1.0E+0) | 6.09 (1.4E-2) | 0.99 (2.6E-4) | 1E5 (1.3E2) | 35 | 1808 (1.1E+0) | 1.174 (2.9E-5) | 0.907 (1.5E-5) | 2.01 | 0.84 |
| 18W  | 62.8 (4.8E-3) | 37 | 69.4 (3.1E-1) | 11.207 (5.6E-2) | 0.971 (1.3E-3) | 5.4E5 (3.5E2) | 33 | 345 (1.3E-1) | 1.035 (6.6E-6) | 0.722 (1.9E-5) | 2.95 | 0.78 |
| 19E  | 58.8 (4.9E-3) | 44 | 246 (1.1E+0) | 5.546 (9.7E-3) | 0.99 (2.0E-4) | 1E5 (2.1E2) | 39 | 609 (7.0E-1) | 1.091 (4.3E-5) | 0.802 (5.2E-5) | 2.82 | 1.47 |
| 20W  | 63.3 (6.8E-3) | 40 | 606.2 (3.3E+0) | 3.846 (5.0E-3) | 0.99 (1.5E-4) | 4.7E5 (2.7E2) | 36 | 510 (2.0E-1) | 1.037 (5.8E-6) | 0.742 (1.6E-5) | 2.46 | 1.68 |
| 21E  | 63.3 (3.9E-3) | 38 | 531.1 (9.8E-1) | 5.26 (2.7E-3) | 1 (7.8E-7) | 3.4E5 (7.6E2) | 37 | 269 (3.7E-1) | 1.023 (1.8E-5) | 0.652 (8.3E-5) | 0.98 | 0.98 |
| 22E  | 66 (6.3E-3) | 39 | 79.7 (8.2E-2) | 5.83 (5.2E-3) | 0.752 (1.2E-3) | 4.6E5 (3.6E2) | 36 | 417 (2.2E-1) | 1.028 (6.8E-6) | 0.702 (2.5E-5) | 1.22 | 0.73 |
| 23W  | 69.9 (6.3E-3) | 42 | 497.6 (3.8E-1) | 4.559 (1.0E-3) | 0.999 (2.7E-6) | 3.4E5 (2.4E2) | 39 | 852 (3.7E-1) | 1.074 (1.1E-5) | 0.817 (1.5E-5) | 1.2 | 0.75 |
| 24E  | 63.8 (8.4E-3) | 44 | 427.7 (3.3E+0) | 4.512 (1.0E-2) | 0.99 (2.7E-4) | 4.3E5 (5.7E2) | 38 | 716 (5.4E-1) | 1.062 (1.8E-5) | 0.802 (2.8E-5) | 1.8 | 0.70 |
| 25E  | 61.8 (9.8E-3) | 42 | 289.3 (5.4E+1) | 6.95 (5.2E-1) | 0.99 (1.1E-2) | 1E5 (1.1E2) | 39 | 1045 (7.1E-1) | 1.112 (2.4E-5) | 0.845 (2.3E-5) | 1.47 | 0.90 |
| 26W  | 62.7 (8.7E-3) | 46 | 477.5 (9.3E-1) | 3.718 (2.0E-3) | 0.989 (5.8E-5) | 3.4E5 (1.5E2) | 43 | 555 (1.8E-1) | 1.024 (3.6E-6) | 0.679 (1.6E-5) | 1.35 | 1.07 |
| 27E  | 67.7 (1.2E-2) | 41 | 473.1 (7.7E+0) | 4.403 (1.9E-2) | 0.99 (5.5E-4) | 1.6E6 (1.5E3) | 36 | 1426 (8.6E-1) | 1.055 (9.5E-6) | 0.834 (1.5E-5) | 1.34 | 0.68 |
| 28W  | 68.1 (8.2E-3) | 41 | 447.7 (3.4E+0) | 4.188 (8.6E-3) | 0.99 (2.4E-4) | 6.8E6 (4.4E3) | 36 | 1289 (5.0E-1) | 1.043 (5.0E-6) | 0.833 (9.0E-6) | 1.72 | 0.81 |
The fundamental diagrams shown in Figure 6-3 illustrate that the fitted models reasonably follow through the steady-state observations. Here, the value of the parameters that have physical interpretation needs to be discussed. The distance headway \((l)\) and speed exponent \((m)\) for both regimes, free flow speed for the congested regime, and jam density for the uncongested regime do not have any physical interpretations. The free flow speed of the uncongested regime varies from about 56 to 70 mph across these sites. The jam density of the congested regime varies over a very wide and somewhat unrealistic range of 229 to 4207 pc/mi/ln. This is due to the fact that observations near jam density condition is very scarce. Hence, the congested regime curve extrapolates to a very high density value when flow=0 in that regime. Unlike the previous chapter (Chapter 5), this chapter is not artificially capping the jam density value which caused the congested regime to have a very poor fit to the observed data for some sensors.

The two density breakpoint values \((k_{B,r=1} \text{ and } k_{B,r=2})\) are of particular interests in this study since these two constitute the overlap and consequently, the required reaction time changes. Here, \(k_{B,r=1}\) ranges from 37 to 50 pc/mi/ln. However, in most sensors, it is less than the density at capacity (45 pc/mi/ln) for basic freeway segment specified by HCM. This wide-range variation of density breakpoint underscores that the national average value provided by HCM needs to be calibrated with field data if high fidelity analysis is desired.

The difference between the two density break points represents the overlap range which appears to be a unique characteristic for these sensors. The fitted overlap in density varies from 0 to 8 pc/mi/ln. The value for the first and second change in driver reaction time vary from 0.74 to 2.95 and from 0.55 to 1.68, respectively, indicating that the sites have significant variations in the estimated changes in required reaction time.
The standard deviations of the estimates for the parameters to which the model is sensitive are of small magnitude, as shown within parenthesis in Table 6-1. This reveals that most of these estimates are significant and the error in the estimates of $\Delta t_s$ and $\Delta t_e$ should be within an acceptable range. It should be noted that the standard deviations for the two breakpoints cannot be estimated because these are only classifiers of the regimes. Standard deviation for the congested regime free flow speed is very high because the fitted models are insensitive to this parameter. For the same reason, the standard deviation of the uncongested regime jam density for a few sensors are high (e.g., for 4W and 25E).

**6.5.2 Analysis of Crash Data**

Figure 6-4 shows the estimated rear-end crash rates (per 100 million VMT) for all the 28 sensors. The segment length, which ranges from about 0.15 to 2 miles, is showed in this figure by color-coding the bars. The figure shows that the selected sites have significant variability in terms of crash frequency since the crash rate and segment length appear to be uncorrelated. The crash rate spans over a wide range of about 10 to 65 rear-end crashes per 100 million VMT.
6.5.3 Crash Rates vs. Change in Required Reaction Time

Figure 6-5 (a) and (b) show the scatterplots of rear-end crash rates and the two changes in required reaction time—Δte and Δts, respectively, as demonstrated in Figure 6-1. In order to investigate their relationship, three regression models were fitted to the data: an exponential model, linear model, and logarithmic model. The model forms are described in Table 6-2.
It is apparent from Figure 6-5 (a) that no association exists between the rear-end crash rates and $\Delta t_e$. It could be due to the fact that $\Delta t_e$ is very sensitive to the selection of break point density for the uncongested regime, which makes this predictor less robust to any potential outliers. All three fitted models demonstrated a very poor fit to the data, as the $R$-squared values were found between 0.004 and 0.01.

On the other hand, a strong positive trend is apparent in the rear-end crash rates vs. $\Delta t_s$ plot as shown in Figure 6-5 (b). The correlation coefficient between these two variables are $+0.63$. Since this evidence is mostly empirical, it is difficult to pre-determine about what should best explain the relationship. Nonetheless, all three forms of model showed a reasonable fit. Table 6-2 shows the form of the fitted models, the fitted parameter values along with several statistical measures to evaluate the relationship.

**Table 6-2: Regression model forms and evaluations.**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential</td>
<td>$y = e^{mx+c}$</td>
<td>11.63</td>
<td>0.37</td>
<td>17.04</td>
<td>m</td>
<td>0.48</td>
<td>0.12</td>
<td>$3\times10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>c</td>
<td>2.43</td>
<td>0.22</td>
<td>$3\times10^{-11}$</td>
</tr>
<tr>
<td>Linear</td>
<td>$y = mx+c$</td>
<td>11.53</td>
<td>0.38</td>
<td>17.34</td>
<td>m</td>
<td>14.79</td>
<td>3.55</td>
<td>$3\times10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>c</td>
<td>3.74</td>
<td>6.77</td>
<td>0.59</td>
</tr>
<tr>
<td>Logarithmic</td>
<td>$y = mln(x)+c$</td>
<td>11.86</td>
<td>0.34</td>
<td>14.98</td>
<td>m</td>
<td>24.20</td>
<td>6.25</td>
<td>$6\times10^{-4}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>c</td>
<td>17.63</td>
<td>3.99</td>
<td>$2\times10^{-4}$</td>
</tr>
</tbody>
</table>

Here, the response variable ($y$) is the rear-end crash rate (in crashes per 100 mill. VMT), the explanatory variable ($x$) is $\Delta t_1$ (in seconds), and $m$ and $c$ are model coefficients. The exponential and linear model exhibited similar performance measures in terms of the residual standard error (SE), adjusted R-square, and F-statistics. On the other hand, the logarithmic
model, while showing a satisfactory fit, generated a slightly higher residual error and a lower $R$-squared and $F$-statistic. The p-value column shows that except for the Y-intercept of the linear model, other coefficients of all three models are statistically significant at least at a level of 0.001.

Here, it is essential to explain the interpretation of the model evaluations from the perspective of this study. For instance, $R$-squared value of 0.38 generated by the linear model interprets that the model explains 38% of the variation in the rear-end crash rates. Note that while $\Delta t_s$ is associated only with the transition from uncongested to the congested regime, not all rear-end crash rates happened during the transition regime. Moreover, a lot of such crashes are not directly related to the driver reaction time required to maintain asymptotic stability. Rather, several other factors including distracted driving, driving under the influence, animal crossing, and abrupt lane changing maneuver may lead to rear-end crashes, which are not reflected in the long-term traffic stream characteristics of a segment. Nonetheless, a lot of rear-end crashes are associated with the onset of congestion and the transition between the two regimes. Hence, only a significant portion of all rear-end crashes are expected to be explained by the fitted models. Having said that, an $R$-squared value within the range of 0.34 to 0.38 appears to be satisfactory for the purpose of this study.

The variance of the residuals generated by a statistical model is another important criterion to evaluate statistical models. Despite having a satisfactory $R$-squared, a model might be heteroscedastic if the residual shows a trend with the fitted response variable. To check for heteroscedasticity, the square root of standardized residuals vs. the fitted response variable plot, also known as the scale-location plot, is commonly used (University of Virginia Library,
Charlottesville. Figure 6-6 shows such plots for all three fitted models along with a smoothed trend line and the confidence interval.

![Scale-Location: Linear Model](image)

(a)

![Scale-Location: Exp. Model](image)

(b)

![Scale-Location: Log Model](image)

(c)

Figure 6-6: Scale-location plot for (a) Linear (b) Exponential and (c) Logarithmic model.

Figure 6-6 shows that the linear and the logarithmic model have a very mild increasing trend with the increase of the fitted values. However, the residuals are more uniformly distributed across the fitted values in the case of the linear model than the logarithmic model. The scale-location plot for the exponential model has a stronger trend than the other two models which infers that the model has heteroscedasticity.
It is apparent that the linear model performed best among the three models both in terms of $R$-squared value and homoscedasticity. It also has the simplest model form. The slope of the model interprets that for each second increment in $\Delta t_s$, the estimated rear-end crash rate increases by 14.79 crashes per 100 million VMT. Note that the high $p$-value of the intercept term of the linear model infers that the null hypothesis that the intercept is zero cannot be rejected. However, the intercept term is a trait obtained by extrapolating the linear model; the change in driver reaction time cannot be zero in reality.

6.6 CONCLUSIONS

The study investigates the relationship between rear-end crash rates and macroscopically derived driver reaction time required for asymptotic stability. Separate GHR models are fitted discontinuously for the uncongested and congested flow regimes using traffic sensor data collected for one year. The changes in reaction time at the beginning and end of the transition regime are mainly focused here. Data from 28 sensors located in basic freeway segments of three interstates are used in this study. A freeway segment surrounding each sensor is selected for crash data analysis such that the road geometric characteristics remain constant throughout the segment. Rear-end crash rate for each segment is then estimated by examining police-reported crash data.

Results from the fitted models and crash data analysis showed that the selected sites have significant variability in terms of both traffic characteristics and rear-end crash risks. Standard deviations of the estimates were of small magnitude for the parameters that have a physical interpretation and to which the model is sensitive. Three model forms, namely linear, exponential, and logarithmic were used to assess the association between crash rates and the changes in required reaction time. The second change in required reaction time exhibited no
significant association with rear-end crash rates. The first change in required reaction time exhibited a strong positive correlation with rear-end crash rates. The correlation coefficient was found +0.63. The three models exhibited $R^2$-squared values ranging from 0.34 to 0.38, with the linear model performing best both in terms of $R^2$-squared and standard error. Analysis of the residuals for each model revealed that the linear and logarithmic model have reasonable homoscedasticity.

While an $R^2$-square within this range might seem low, it should be noted that not all rear-end crashes are associated with the transition of traffic regimes and car-following behavior. Hence, all rear-end crash occurrences at a location cannot be explained by the change in required reaction time. Having said that, the statistical correlation revealed here can be considered strong enough to motivate follow-on research to incorporate macroscopically derived reaction time in safety and planning level studies.

More generally, this study demonstrates a useful application of a discontinuous macroscopic traffic model. However, a few limitations need to be highlighted here to be addressed in future research. A sample size of 28 falls slightly short of the generally accepted sample size of 30 for assumption of normality. As much as it is difficult to extract accurate crash data, it is important to increase the number of study sites. Further, if detailed data are available, it is recommended to filter the crashes that are directly associated with the transition of traffic regimes and car-following behavior.
6.7 REFERENCES


CHAPTER 7 – SUMMARY AND CONCLUSIONS

7.1 SUMMARY

This research presents empirical and traffic flow theory-based analyses of driver lane changing and car-following maneuvers on freeways from the perspective of traffic mobility and safety. Regarding lane changing maneuvers, this research aims to develop lane change detection, characterization, and prediction frameworks for freeway segments. From the perspective of car-following maneuvers, this research aims to investigate the relationship between driver reaction time required for traffic stream stability and rear-end crashes based on a car-following theory.

The first research question addressed in this dissertation was how to efficiently detect a lane change maneuver. A GPS-equipped vehicle’s geographic position and digitized infrastructure data were used to address this question. An algorithm using both sets of data defined and verified the start and end of lane change maneuvers. Tests on the validity of the algorithm were successful in predicting pre-planned lane changes.

The second research question was how to identify extreme lane change maneuvers in a traffic stream. This question was addressed by using a trajectory dataset archived by the Next Generation SIMulation (NGSIM) program as the testbed, with data collected at US-101 in California. Two metrics were applied to identify such maneuvers, namely lane change rate per distance traveled and time-to-line-crossing (TLC)—a measure integrating speed and heading angle during a lane change.

The third research question was how to predict discretionary lane change intensity at freeway weaving segments. Various statistical modeling approaches using site and traffic data from 19 weaving segments were developed and tested. A comparison with a widely used
weaving model in the Highway Capacity Manual demonstrated the superiority of the proposed statistical models.

The fourth research question was how to develop a discontinuous form of the macroscopic Gazis-Herman-Rothery (GHR) car-following model. It was addressed by introducing an overlapping range defined in terms of density between the uncongested and congested flow regimes of the macroscopic GHR car-following model.

The fifth research question was what is the association between rear-end crashes and driver reaction time required for car-following stability. The research indicated that the changes in required reaction time exhibited moderate positive correlations with the long term rear-end crash rates when applied to 28 freeway segments.

7.2 KEY FINDINGS

Key findings related to the posed research questions are summarized in the following two subsections, for lane changing and car following maneuvers, respectively.

7.2.1 Key Findings Related to Lane Changing Behavior

- The occurrence of a significant number of false alarms is a concern when detecting lane changes using vehicle position and digitized lane marking data. The effect of GPS errors on a vehicle’s lane position estimate can be minimized by constraining the magnitude of the lateral shift and the temporal span of the vehicle presence in the destination lane for each lane change. Field experiments showed that an arbitrarily selected temporal threshold of 1 second and an optimized lateral shift threshold of 6.9 ft. minimize the total error. Out of 46 ground truth lane changes, the algorithm detected 42 correctly with two false positives, resulting in a 4.35% and 8.7% false positive and false negative error, respectively.
• Application of the above thresholds to 637 naturalistic driving trajectories through a weaving segment revealed that merging traffic had the highest rate of discretionary lane changes. Merging traffic also tended to execute its first mandatory lane change near the entry gore, where the weaving-induced turbulence is typically at its highest.

• Total lane change frequency of a traffic stream tended to drop with increasing congestion due to two contributing factors: decreasing opportunity to change lane and decreasing chance of speed gain by changing lanes. In the NGSIM US-101 site, as density increased from 48 to 60 pc/mi/lane, total lane change frequency decreased by 33%, and the speed gain by changing lanes decreased from 6.2 to 1.1 mph.

• In the US-101 dataset, the highest lane change rates (top one percentile) exceeded three lane changes per 1,000 ft. traveled. The average overall rate for lane changing vehicles was 1.37 per 1,000 ft. traveled. In this dataset, TLCc values for the extreme rightward lane changes varied from 0.71 to 1.41 seconds. TLCc values for the extreme leftward lane changes varied from 1.17 to 1.63 seconds. These low critical TLC values resulted from a combination of high entry angle and speed during lane changes.

• According to the predictive models developed in this study, weaving flow rates, segment length, and interchange density are the most important predictors for modeling discretionary lane change (DLC) intensity in a weave. The proposed predictive models outperformed the HCM6 lane change model and underscored
the necessity of using advanced statistical modeling tools and the inclusion of additional predictors for predicting DLC intensity.

- A regression tree approach, which can handle nonlinearity and interactions among predictors, outperformed a conventional linear regression model in predicting DLC intensity in a weave. The tree model significantly improved the prediction of DLCs per hour for more than 2/3 of the sites, when compared to the mean value at each site. It also performed well for most cases when applied to a site that was omitted from the model development database.

### 7.2.2 Key Findings Related to Car-following Behavior

- The occurrences of inconsistent and mixed-state observation are a concern, when fitting a macroscopic GHR model to traffic flow data. These occurrences can be minimized—as was shown by a previous research—by constraining the speed differential between consecutive observations and by excluding low-density, low-speed observations. A modified version of these constraints were applied to this dissertation, along with three empirical constraints on several model parameters, so that the impact of the outliers is further minimized.

- The proposed discontinuous GHR model outperformed both continuous and discontinuous forms of a widely used traffic flow model in HCM6. The discontinuity in the proposed model also played a critical role in estimating the changes in required driver reaction time for maintaining car-following stability during traffic state transitions.

- The change in reaction time required for maintaining car-following stability at the end of the transition regime exhibited moderate positive correlations with rear-end
crash rates when applied to 28 freeway segments (*correlation coefficient* = +0.63). Three model forms that were tested to predict rear-end crash rates as a function of the change in required reaction time. The linear form exhibited the highest R-squared value (0.38) and the least heteroscedasticity.

### 7.3 RECOMMENDATIONS FOR FUTURE WORK

Insights gained in the conduct of research on lane changing and car-following maneuver have revealed additional opportunities for future research advances in those areas. Following are some recommendations for future research in the field of driving behavior analysis.

**Lane change detection at sites with varying geometric features and GPS Data Accuracy**

The proposed lane change detection algorithm, which relies heavily on GPS positioning was tested at only one site. Additional applications at highway locations with different horizontal and vertical curvature, and in the presence of sight distance obstacles are recommended. GPS accuracy is likely to degrade at segments with sharp curvature, dense canopy cover, or with high-rise buildings near the road. For such locations, additional filtering algorithms such as the Kalman filter (Kalman, 1960) should be incorporated.

**Testing additional metrics for identifying extreme lane changing behavior**

In the US-101 dataset (collected in 2006), gap time between vehicles, and calculated accelerations were artificially capped. Consequently, performance measures based on these parameters could not be applied to this set in order to identify extreme lane changing maneuvers. More recent high-resolution trajectory datasets have been made publicly available that include data from sites with diverse characteristics (Van Beinum et al., 2018). This and similar datasets
will enable researchers to develop gap and acceleration-based metrics to supplement the ones proposed in this research.

**Alternative Method to Estimate Speed Gain by Lane Changes**

An alternative way to estimate the speed gain by the subject vehicle through discretionary lane changes is to compare the average speed for several data points before and after a lane change. In that case, the number of data points to select before and after each lane change must be determined through a sensitivity analysis to ensure the stability of speed during a lane change. This proposed alternative approach is recommended to be tested in future research.

**Testing candidate treatments for early merging tendency at weaves**

At weaving segments, ramp-to-freeway vehicles exhibited an early merging tendency, often over the solid lane marking near the entry gore. During peak hours, such a trait may result in concentrated turbulence near the entry gore of a weave. Possible treatments to avoid such a flow-disruptive trait may include both geometric (e.g., changing segment length and minimum lane change requirements) and operational (e.g., ramp metering) modifications. Further research—possibly using both observational and simulated data—is necessary to shortlist and test these treatments.

**Improved DLC predictions by sampling across traffic density spectra**

Of the 19 sites used to develop the DLC intensity models, many did not exhibit significant variability in lane change frequency. This may be due to small data samples at some sites and time-invariant traffic flow conditions during data collection. As revealed in this dissertation, lane change intensity is strongly influenced by traffic density. Since most field observations were collected during off-peak periods and over a narrow span of time, they were not able to capture the variability in congestion levels. Consequently, the recorded DLC intensity did not vary
within each site. In addition, the number of data points were few for the sites at which drone-mounted cameras were used to observe traffic. Future studies should overcome this limitation due to the short battery life of drones by using tethered drones.

**Cleaning traffic data to fit a macroscopic traffic flow model**

When fitting the macroscopic GHR car-following model, the first-stage filters to remove unwanted traffic flow data points recommended in this research may remove some valid observations. In the future, attempts should be taken to combine the two-stage data-reduction process into a single, robust process that is more efficient and consistent.

**Incorporating driver reaction time into safety evaluation guidelines**

Data from 28 locations were used to explore the relationship between rear-end crash rates and reaction time required for maintaining car-following stability. To incorporate this macroscopically modeled reaction time in the current safety evaluation guidelines, additional sites—preferably from diverse geographic areas—must be investigated. Further, if detailed data are available, it is recommended to filter the crashes that are directly associated with the transition of traffic regimes and car-following behavior.