Abstract

JAFARI, REZA. Safety Effects of the Access Points near Signalized Intersections. (Under the direction of Dr. Joseph E. Hummer.)

In the US in 2009, 5.5 million collisions occurred in which over 2.2 million people were injured and over 33,000 people died due to highway collisions. Over half of these total crashes were intersection and access point – related. Most collision reporting systems do not provide the necessary level of information to identify access – related collisions but collision data, where available, indicate a high incidence of access – related collisions.

The objectives of this research were to develop a valid statistical model to estimate the number of access point – related collisions occurring at access points near signalized intersections and providing checklist for site planners and decision-makers to distinguish higher collision sites from lower collision sites and avoid constructing higher collision sites.

Geometric, traffic, and access – related collision data over 5 years, from January 2005 to December 2009 were collected for 108 sites. Out of the 15 independent variables tested, only AADT, driveway width, and Synchro through movement 95% queue at the intersection near the access point were statistically significant in developing the collision prediction statistical model. This model could be used by state DOTs and municipal traffic engineers to address access management requirements and to predict problems likely to result from site traffic impacts.

To provide checklist for site planners to distinguish the higher collision sites from lower collision sites, the data that were previously collected and some new information such as demographic and socio economic data were used. The higher collision sites were investigated one by one. Quantitative, binary, and categorical variables, and demographic and socio economic information, were analyzed and compared between the higher and lower collision sites. Statistical tests were used to find the contributing factors and provide checklist to certify no access points will be constructed before the safety issues are considered. The proposed policy checklist stated that an access point most likely would be a higher collision site if it was operating full movement, had a driveway peak hour volume of over 120 vehicles per hour, had an intersection peak hour through movement Synchro 95% queue of over 230 feet, and had a driveway left turn proportion (100 × left turns from driveway per peak hour /AADT) of over 0.2.
Safety Effects of the Access Points near Signalized Intersections

by
Reza Jafari

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

2011

APPROVED BY:

Dr. Nagui M. Rouphail    Dr. Billy M. Williams    Dr. John F. Monahan

Dr. Joseph E. Hummer
(Chair of Advisory Committee)
Dedication

This work is dedicated to my mother, who loves me unconditionally, to my father, who has taught me how to work hard and is proud of me, and to the color of my life, Zohreh, without whose caring support it would not have been possible. This work is also dedicated to my brothers, who have always been there for me.
Biography

Reza Jafari is a civil engineering alumna from Northeastern University in Boston, Massachusetts with a Master of Science degree. He also got a Bachelor of Science degree from Sharif University of Technology and a Master of Science degree from K.N.T. University, both in Mechanical Engineering from Tehran, Iran.

Reza has over seven years of experience in transportation and traffic engineering. He worked for WSP-SELLS and the Congestion Management Unit of the North Carolina Department of Transportation (NCDOT) and on a few projects at North Carolina State University and Northeastern University. He worked on different projects including safety analysis, signal design, capacity analysis, traffic impact analysis, parking studies, pedestrian access improvements, and roadway improvements for public and private clients.

Reza held teaching assistant and research assistant positions while pursuing his MSc and Ph.D. in Civil Engineering at Northeastern University and North Carolina State University. In North Carolina State University and under the direction of Professor Joseph Hummer, he participated in a research project sponsored by the NCDOT to evaluate a railroad crossing wayside horn and in research to analyze the safety effects of the access points near signalized intersections.

Reza is a registered Professional Engineer (PE) in Civil Engineering in the State of North Carolina and is a Professional Traffic Operations Engineer (PTOE) license holder.
Acknowledgements

Any findings, conclusions, and recommendations expressed in this dissertation research are those of the author and do not necessarily reflect the views of the people and organizations with whom I would like to acknowledge for their gracious support.

I cannot really find the right words to express my sincere gratitude and appreciation to Dr. Joseph Hummer who was not only my advisor throughout the process of this dissertation, but also a good friend. His intelligence, guidance, inspirational encouragement, advice, continuous support, and patience have been of the utmost importance to the completion of my research work.

I would like to thank my friend and former classmate, Dr. Jongdae Baek whose research was very helpful in developing my statistical analysis. His friendship and technical suggestions on this research was invaluable.

I would like to express my gratitude to my committee members, Dr. Nagui Rouphail, Dr. John Monahan, and Dr. Billy Williams and also to Dr. John Stone whose precious suggestions and continuous support helped me to improve the quality of this research document and analyses.

My special thanks go to Dr. Ezra Hauer who always had responded to my questions in a timely manner with comprehensive detailed information. His statistical analyses were used as the basis of this research. Not only me, but the entire society of roadway safety, worldwide appreciate his effort in producing valuable research to be used now and in the future by safety professionals to reduce road collisions and ultimately save more lives.

This research was initially generated by the help of my former supervisor at NCDOT, James Dunlop. He was very supportive and assisted me with obtaining some of the necessary data required for this study. I also extend my appreciation to Kevin Lacey and Jeff Jaeger of the NCDOT for their permission in using the North Carolina traffic collision database.

The current and former directors of the Civil Engineering graduate program, Dr. Vernon Matzen and Dr. David Johnston and the student services managers, Renee Howard and Bri Ryen were helpful in all aspects during my study and work at NC State University. I do appreciate their support.
I would like to express my special appreciation to Phil Demosthenes, Bernie Kalus of WSP-Sells, Lori Cove of the Town of Cary, John Albeck of the Trafficware, Kent Taylor and Byron Sanders of the NCDOT, and Kyle Ward of the North Carolina Capital Area MPO who provided me with turning counts and other valuable traffic data used in this study. I also would like to thank Bowman Kelly of the City of Raleigh who provided me with traffic volume data and supported me since the first day I met him at NC State University.

I would like to thank the North Carolina Section of the ITE for supporting me with their scholarship that encouraged me to stay, work, and complete my PhD degree in Transportation Engineering.

I would like to thank my former advisor, Dr. Honarvar who taught me to “see the cup as half full”. I also deeply appreciated all the assistance from my teacher, Dr. Sadati in helping me finding my path, and discovering better ways to be more helpful to myself and my society. Dr. Sadati was more than a teacher to me. He is, and was, my mentor, older brother, and a blameless friend.

My special appreciation is extended to my friends and coworkers, Dr. Majed Al-Ghandour, Dr.-to-be Katy Salamati, Anne Holzem, Dr. Elizabeth Harris, Mike Reese, Dr.-to-be Leta Huntsinger, Dr. Bastian Schroeder, Dr. Daniel Findley, Dr. Alix Demers, Erin Harrington, Caroline Kone, Mike Surasky, Gavin Teng, Avani Patel, Dr. Jae Pil Moon, Dr. Hyejun Hu for their continuous support.

I would like to express my heartfelt thanks to my parents and my brothers for their love and great support for my education. Without their endless encouragement, this work could not have been possible.

Last, but certainly not least, I would like to thank my wife, Zohreh. Even though busy with her PhD and office work, she has been unconditionally supportive and helpful.
# Table of Contents

List of Figures ......................................................................................................................... x
List of Tables ........................................................................................................................... xii

1. INTRODUCTION .................................................................................................................. 1
   1.1 Background ..................................................................................................................... 1
   1.2 Objective ......................................................................................................................... 2
   1.3 Scope ............................................................................................................................... 3
   1.4 Organization .................................................................................................................... 5

2. LITERATURE REVIEW ....................................................................................................... 6
   2.1 Driveways ....................................................................................................................... 6
   2.2 Access management ....................................................................................................... 7
   2.3 Corner clearance ............................................................................................................ 11
   2.4 Collision models ............................................................................................................ 12

3. DATA PREPARATION ......................................................................................................... 17
   3.1 Site selection .................................................................................................................. 17
   3.2 Data collection .............................................................................................................. 24
      3.2.1 Geometric data ........................................................................................................ 24
      3.2.2 Traffic data .............................................................................................................. 25
      3.2.3 Collision data .......................................................................................................... 32
   3.3 Data analysis .................................................................................................................. 35
      3.3.1 Corner clearance ..................................................................................................... 35
      3.3.2 Lane configuration at the intersection .................................................................... 36
      3.3.3 Lane configuration at the driveway ...................................................................... 36
      3.3.4 Major road median type ...................................................................................... 37
      3.3.5 Speed limit on the major road ............................................................................... 37
      3.3.6 Grade on the major road ...................................................................................... 38
      3.3.7 Driveway angle ...................................................................................................... 38
      3.3.8 Driveway one-way or two-way operations ............................................................ 40
      3.3.9 Exclusive right turn into driveway ....................................................................... 40
3.3.10 Driveway median status ........................................................................................................ 40
3.3.11 Driveway width ...................................................................................................................... 41
3.3.12 Transition ............................................................................................................................. 41
3.3.13 Driveway radius ................................................................................................................... 42
3.3.14 Second driveway ................................................................................................................ 43
3.4 Summary of data ....................................................................................................................... 43

4. ANALYSIS METHODOLOGY .................................................................................................... 44
4.1 Modeling method ...................................................................................................................... 45
4.1.1 Data traits ............................................................................................................................. 45
4.2 Modeling Process ..................................................................................................................... 47
4.2.1 Candidate functional forms ................................................................................................. 49
4.2.2 Integrate – differentiate method .......................................................................................... 51
4.2.3 Estimation of the model coefficients ................................................................................. 54
4.2.4 Model selection .................................................................................................................... 59
4.3 Goodness of fit ......................................................................................................................... 62

5. MODELING PROCESS ................................................................................................................. 65
5.1 Initial model .............................................................................................................................. 65
5.2 Second model .......................................................................................................................... 67
5.2.1 Driveway volume ................................................................................................................ 68
5.2.2 Corner clearance .................................................................................................................. 71
5.2.3 Synchro queue length ......................................................................................................... 72
5.2.4 Driveway width .................................................................................................................... 76
5.2.5 Lane configuration at the intersection .............................................................................. 78
5.2.6 Lane configuration at the driveway .................................................................................... 79
5.2.7 Grade on the major road ..................................................................................................... 79
5.2.8 Speed limit .......................................................................................................................... 79
5.2.9 Major road median ................................................................................................................ 80
5.2.10 Driveway median ............................................................................................................... 80
5.2.11 Driveway angle .................................................................................................................. 81
5.2.12 Driveway radius ................................................................................................................ 81
6.3 Analysis

6.3.1 Quantitative variables

6.3.2 Binary variables

6.3.3 Driveway left turn proportion

6.3.4 Categorical variables

6.3.5 Hourly distribution of collisions

6.3.6 Socio economics

6.4 Conclusion

7. CONTRIBUTING FACTORS

7.1 Methodology

7.2 Prediction model

7.3 Higher collision site identification

7.4 Conclusion

8. CONCLUSIONS AND RECOMMENDATIONS

8.1 Research summary

8.2 Findings and conclusions

8.3 Future research recommendations

References

Appendices

Appendix A – Distribution of the study sites in Wake County

Appendix B – Google Earth Pro daily traffic counts

Appendix C – Traffic counts data collection method

Appendix D – Crash Report Form DMV-349

Appendix E – Collision IDs

Appendix F – Summary of study sites and collected data in Wake County

Appendix G – SAS® Output for Collision Models

G.1 – Initial Model

G.2 – Second Model

G.3 – Thirds (final) Model
List of Figures

Figure 1 – Intersection physical and functional areas (7) ................................................................. 7
Figure 2 – Driveway upstream of a signalized intersection............................................................. 7
Figure 3 – Corner clearance and radius of the corner on minor street (17) ................................. 11
Figure 4 – Data preparation overview ......................................................................................... 17
Figure 5 – Wake County in the State of North Carolina (38) .......................................................... 18
Figure 6 – (A) Signalized right turn, (B) Free flow right movement ............................................ 19
Figure 7 – Full movement, left-over, and RIRO designs ................................................................. 22
Figure 8 – A snapshot of the AADT map (39) ................................................................................ 25
Figure 9 – Queue on Oneal Road in Raleigh, NC ......................................................................... 30
Figure 10 – Synchro estimated queue versus observed queue ...................................................... 31
Figure 11 – A snapshot of a police report diagram and narrative .................................................. 34
Figure 12 – Corner clearance of the 108 sample sites ................................................................ 35
Figure 13 – Corner clearance measurements .............................................................................. 36
Figure 14 – Driveway angle ............................................................................................................. 39
Figure 15 – Short driveway stem .................................................................................................... 40
Figure 16 – Driveways with smooth and elevated transitions ......................................................... 42
Figure 17 – Driveways with large and small radii ........................................................................... 42
Figure 18 – Modeling methodology overview .............................................................................. 44
Figure 19 – Correlation between access – related collisions per year and AADT/10,000 .............. 47
Figure 20 – Modeling methodology overview (46) ...................................................................... 48
Figure 21 – EIF diagram (46) ......................................................................................................... 52
Figure 22 – EIF for the data in Figure 19 ........................................................................................ 52
Figure 23 – Example of two candidate functions fitting the cumulative function ...................... 53
Figure 24 – Codes and output for GENMOD procedures in SAS® .............................................. 56
Figure 25 – Codes and output for NLMIXED procedures in SAS® ............................................. 57
Figure 26 – Optimum local solution and global solution (48) ....................................................... 58
Figure 27 – Different graphical behavior of quadratic function of  \( f(x) = ax^2 + bx \) ............. 59
Figure 28 – Residual plot between annual collected and predicted collisions ............................... 63
Figure 29 – CURE plot fitted in \( +2\sigma^* \) and \( -2\sigma^* \) for \( f(x) = x^{1.725}e^{-0.749x} \) .............. 64
Figure 30 – Modeling process overview ....................................................................................... 65
Figure 31 – Comparison between annual collected and the first model predicted collisions ...............67
Figure 32 – $f_2(x_2)$ when $x_2$ is the driveway traffic volume .................................................................69
Figure 33 – EIF of $f_2(x_2)$ when $x_2$ is driveway traffic volume ...............................................................69
Figure 34 – $f_2(x_2)$ when $x_2$ is the corner clearance ........................................................................71
Figure 35 – EIF of $f_2(x_2)$ when $x_2$ is the corner clearance .................................................................71
Figure 36 – $f_2(x_2)$ when $x_2$ is the left turn lane Synchro queue .........................................................73
Figure 37 – $f_2(x_2)$ when $x_2$ is the through lane Synchro queue ..........................................................73
Figure 38 – $f_2(x_2)$ when $x_2$ is the right turn lane Synchro queue .........................................................74
Figure 39 – EIF of $f_2(x_2)$ when $x_2$ is the Synchro queue for left movement ....................................74
Figure 40 – EIF of $f_2(x_2)$ when $x_2$ is the Synchro queue for thru movement ...............................75
Figure 41 – EIF of $f_2(x_2)$ when $x_2$ is the Synchro queue for right movement .....................................75
Figure 42 – $f_2(x_2)$ when $x_2$ is the driveway width ...............................................................................77
Figure 43 – EIF of $f_2(x_2)$ when $x_2$ is the driveway width .................................................................77
Figure 44 – CURE plot fitted in $+2\sigma*$ and $-2\sigma*$ for second model ...........................................84
Figure 45 – Comparison between annual collected and the second model predicted collisions ..........85
Figure 46 – $f_3(x_3)$ when $x_3$ is the driveway traffic volume ...............................................................87
Figure 47 – EIF of $f_3(x_3)$ when $x_3$ is the driveway traffic volume .......................................................88
Figure 48 – $f_3(x_3)$ when $x_3$ is the through movement Synchro queue .............................................89
Figure 49 – EIF of $f_3(x_3)$ when $x_3$ is the trough movement Synchro queue ...................................90
Figure 50 – CURE plot fitted in $+2\sigma*$ and $-2\sigma*$ for third model ..................................................92
Figure 51 – Comparison between annual collected and the third model predicted collisions ..............92
Figure 52 – $f_4(x_4)$ when $x_4$ is the corner clearance ..............................................................................95
Figure 53 – EIF of $f_4(x_4)$ when $x_4$ is the corner clearance ...............................................................96
Figure 54 – Prediction model residuals versus predicted annual access – related collisions ...............98
Figure 55 – Predicted versus observed annual access – related collisions .............................................99
Figure 56 – Model validation (predicted versus observed annual access – related collisions) .............103
Figure 57 – Higher collision site identification overview .................................................................105
Figure 58 – Site 1 aerial image ...........................................................................................................107
Figure 59 – Site 1 collision diagram ....................................................................................................108
Figure 60 – Site 2 aerial image ...........................................................................................................109
Figure 61 – Site 2 collision diagram ....................................................................................................110
Figure 62 – Steep hill upstream of the access point at site 2 .........................................................110
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>63</td>
<td>Site 3 aerial image</td>
<td>112</td>
</tr>
<tr>
<td>64</td>
<td>Site 3 collision diagram</td>
<td>113</td>
</tr>
<tr>
<td>65</td>
<td>Site 4 aerial image</td>
<td>114</td>
</tr>
<tr>
<td>66</td>
<td>Site 4 collision diagram</td>
<td>115</td>
</tr>
<tr>
<td>67</td>
<td>Site 4 access point exit issues</td>
<td>115</td>
</tr>
<tr>
<td>68</td>
<td>Site 5 aerial image</td>
<td>116</td>
</tr>
<tr>
<td>69</td>
<td>Site 5 collision diagram</td>
<td>117</td>
</tr>
<tr>
<td>70</td>
<td>Site 5 very steep grade driveway</td>
<td>117</td>
</tr>
<tr>
<td>71</td>
<td>Site 5 driveway Google Earth Pro elevation</td>
<td>118</td>
</tr>
<tr>
<td>72</td>
<td>Site 6 aerial image</td>
<td>118</td>
</tr>
<tr>
<td>73</td>
<td>Site 6 collision diagram</td>
<td>119</td>
</tr>
<tr>
<td>74</td>
<td>Site 7 aerial image</td>
<td>120</td>
</tr>
<tr>
<td>75</td>
<td>Site 7 collision diagram</td>
<td>121</td>
</tr>
<tr>
<td>76</td>
<td>Hourly distribution of collisions</td>
<td>127</td>
</tr>
<tr>
<td>77</td>
<td>Full movement higher and lower collision sites’ hourly distribution</td>
<td>127</td>
</tr>
<tr>
<td>78</td>
<td>Household income</td>
<td>128</td>
</tr>
<tr>
<td>79</td>
<td>Education level</td>
<td>129</td>
</tr>
<tr>
<td>80</td>
<td>Car ownership</td>
<td>129</td>
</tr>
<tr>
<td>81</td>
<td>Contributing factors overview</td>
<td>131</td>
</tr>
<tr>
<td>82</td>
<td>Restricted left turn makes a U-turn at the intersection</td>
<td>140</td>
</tr>
</tbody>
</table>
List of Tables

Table 1 – Florida’s corner clearance standards (17) ................................................................. 11
Table 2 – Lane configurations at the driveway (LCD) ................................................................. 20
Table 3 – Sites lane configuration at the driveway ................................................................. 23
Table 4 – K factors (40) .................................................................................................................. 27
Table 5 – Seasonal factors (40) ..................................................................................................... 27
Table 6 – Observed and Synchro estimated maximum queue values ........................................ 31
Table 7 – Lane configuration at the intersection ........................................................................ 37
Table 8 – Full movements and RIRO sites .................................................................................. 37
Table 9 – Posted speed limit on the major roads ........................................................................ 38
Table 10 – Major roads grade at the access point ....................................................................... 38
Table 11 – Driveway angle ........................................................................................................ 39
Table 12 – Driveway median type ............................................................................................. 41
Table 13 – Driveway width ......................................................................................................... 41
Table 14 – Transition between major road and driveway pavements ......................................... 42
Table 15 – Driveway radius ........................................................................................................ 43
Table 16 – Existence of second driveway .................................................................................. 43
Table 17 – Candidate functions to fit the EIF and their corresponding functions ..................... 54
Table 18 – Initial model BIC values and estimated parameters ................................................ 66
Table 19 – Second model estimated parameters ....................................................................... 68
Table 20 – \( f_2(x_2) \) BIC values for the driveway traffic volume .................................................... 70
Table 21 – BIC comparison between driveway entering, exiting, and total traffic ..................... 70
Table 22 – \( f_3(x_2) \) BIC values for the corner clearance ................................................................. 72
Table 23 – Candidate models of the left, through, and right turns’ Synchro queues .................... 76
Table 24 – \( f_3(x_2) \) BIC values for the driveway width ................................................................. 78
Table 25 – Regrouped driveway angle (DA) ............................................................................... 81
Table 26 – Summary of the second variable result ..................................................................... 83
Table 27 – Third model estimated parameters ........................................................................... 86
Table 28 – \( f_3(x_3) \) BIC values for the driveway volume ................................................................. 88
Table 29 – Fourth model estimated parameters ....................................................................... 93
Table 30 – \( f_4(x_4) \) BIC values for the corner clearance ................................................................. 96
Table 31 – Summary of attempts to find important variables

Table 32 – Independent and dependent variables for model validation

Table 33 – 7 higher collision sites

Table 34 – t test for quantitative variables

Table 35 – Chi-square test for binary variables

Table 36 – Categorical variables

Table 37 – Hit and miss statistical method

Table 38 – Prediction model hit and miss

Table 39 – Higher collision sites hit and miss
1. **INTRODUCTION**

1.1 **Background**

Every year 1.3 million people are killed due to road traffic incidents worldwide (1). According to NHTSA (2), every year over five million collisions occur in the US in which over two million people are injured. In 2009 alone over 33,000 people died due to highway collisions. In other words, the rate for injury is 74 and for fatality is 1.13 per 100 million vehicle mile traveled. During the same year, 1,314 people died on the roads in North Carolina; that is 14 fatalities per 100 thousand population. Most collision reporting systems do not provide the necessary level of information to identify access – related collisions but collision data, where available, indicate a high incidence of access – related collisions (3).

Previous efforts have been made in research, education, and other areas to reduce the number of collisions and their severity and ultimately make the roads safer for all users. Researchers have been working on different techniques including access management, traffic calming, road safety audit, and other solutions. To investigate and understand the collisions in recent years numerous studies have been done and collision prediction models have been developed to find the effect of different variables on road collisions.

Overall over half of the total crashes are intersection and access point – related (2). Previous research considered different factors causing collisions within proximity of the intersections. Some models have included terms for the density of access points on a segment of road (4). Although driver confusion, corner clearance, the congestion related to access points, and other related factors have a direct impact on safety and operation of the roadway facilities, no study has considered the parameters causing collisions at access points near intersections. This research attempted to address the issues at access points near intersections by developing a statistical model predicting access – related collisions and providing checklist for site planners and decision-makers to distinguish higher collision sites and prevent construction of those higher collision driveways.
1.2 Objective

Most collision reporting systems do not provide the necessary level of information to identify access – related collisions but collision data, where available, indicate a high incidence of access – related collisions(3). At the end of this research I will provide:

✓ A valid statistical model to estimate the number of access point – related collisions occurring at access points near signalized intersections, and
✓ A checklist for site planners and decision-makers to distinguish the higher collision sites from lower collision sites and eventually prevent constructing higher collision sites.

To achieve these objectives, geometric, traffic, and access – related collision data over a 5-year period were collected for 108 sites in Wake County, North Carolina. Since fatalities and injuries due to access point – related collisions are too infrequent to be analyzed alone, total access – related collisions were considered for analysis and prediction. The NCDOT database TEAAS was used to obtain each individual collision ID. Then the collision information was collected from police crash reports using the NC Department of Motor Vehicles (DMV) website.

In this research 15 independent variables were introduced into the model one by one in a multiplicative form. These variables are listed below and the techniques I used in this research to collect them are shown later in Chapter 3.

1) AADT
2) Driveway volume
3) Corner clearance, i.e., the distance between closest driveway and the intersection
4) Intersection queue length
5) Driveway width
6) Lane configuration at the intersection near the access point on the major road
7) Lane configuration at the driveway on the major road
8) Grade on the major road
9) Speed limit on the major road
10) Major road median type at the access point
11) Driveway median status, i.e., divided or undivided driveways
12) Driveway angle, i.e., the angle between driveway centerline and the edge of travel way
13) Driveway radius
14) Transition between major road and driveway pavements
15) Existence of a second driveway for the same parcel within 150 ft from the main access point

The AADT, driveway volume, corner clearance, intersection queue length, and driveway width are continuous variables while the rest are discrete variables.

To provide checklist for site planners to distinguish higher collision sites from lower collision sites, the data that were previously collected and some new information such as demographic and socio economic were used. The higher collision sites were investigated one by one. Quantitative, binary, and categorical variables, and demographic and socio economic information, were analyzed and compared between the higher and lower collision sites. Statistical tests were used to find factors contributing to the high numbers of collisions and provide checklist to insure that no access points will be constructed before the safety issues are considered.

1.3 Scope

The study sites were access points near signalized intersections. A purely random selection was made from among the 739 signalized intersections in Wake County to collect an unbiased subset of the data for collision prediction. Then I considered a few criteria for site selection that led to 108 sites. A full discussion on these criteria and the reason they were selected are provided later in Section 3.1. These criteria are listed below.

✓ Urban and suburban sites
✓ No pedestrian and animal related collisions
✓ More than one lane on the major road at the access point
✓ No more than two through lanes at the access point
✓ Access points upstream of the intersections
✓ No free flow right movement at the intersection
✓ Closest access point to the intersection
✓ Access points within 800 ft from intersection
✓ Major roads grades less than 4%
✓ No driveway associated with a different parcel within 150 ft of the main access point
✓ Common lane configurations at the driveway that occur at least 10 times in study sample (shown later in Table 3)
✓ Four-leg signalized intersections
✓ No sites with geometry changes within the 5-year study period
✓ No odd geometries
✓ Only public (state or city) owned roads
✓ No small (single-family house) driveways
✓ Full movement and right-in/ right-out (RIRO) driveway movement type

In this study the dependent variable was the number of access – related collisions. A collision needed to satisfy at least one of the criteria below to be included in the access – related collision analysis.

✓ The narrative or diagram of the collision indicates that at least one of the vehicles involving the collision is clearly headed to or from the driveway, or
✓ Any word like "driveway" or "access point" is in the narrative section of the police report, or
✓ Indication number 49 of the police report states that one of the vehicles was making a turn to or from the driveway.

As a part of identifying access – related collision effort (later in Section 3.2), a few criteria were followed to delete collisions from the analysis. These criteria are listed below.

✓ Collisions involving pedestrian or animal
✓ Collisions involving vehicles travelling on different major intersection approaches
✓ Collisions that happened in the physical area of the intersection
1.4 Organization

This dissertation consists of eight chapters. In the next chapter a literature review on related topics is provided. In Chapter Three a data preparation strategy to find the variables of interest is discussed. In Chapter Four the methodology for developing a valid statistical model is provided. Chapter Five describes the actual modeling process and the result. In Chapter Six the higher collision sites are discussed and in Chapter Seven different methodologies are used to find the contributing factors and provide recommendations for decision-makers to distinguish higher collision sites from lower collision sites. Finally a summary of the research results, including findings and conclusions and future recommendations, is provided in Chapter Eight.
2. LITERATURE REVIEW

This chapter reviews the previous related research. The definition of “driveway” is presented from AASHTO (5) and other sources. Then, access management (AM) research is reviewed because this research objective is linked to the focus of AM. Engineers and planners use AM techniques to produce access to land uses while maintaining roadway safety and mobility. An important issue considered in this research is the distance from the driveway to the intersection, so corner clearance standards in North Carolina and other states are shown. Finally, a review of the previous similar statistical modeling efforts is provided.

2.1 Driveways

Driveways provide the transition between a site and the adjacent roadway. Driveway designs should minimize the impact on traffic and provide safe movement. In the State of North Carolina, new and expanded driveways on state roads are permitted by the North Carolina Department of Transportation through a permit application process. This permit application is based on standards in the “Policy on Street and Driveway Access to North Carolina Highways” (6) known as the “Driveway Manual”. Corner clearance, i.e., the distance between the closest driveway and the intersection, is a critical dimension that should be carefully considered in permitting the driveway.

AASHTO (5) indicates that, “Driveways are, in fact, at-grade intersections and should be designed consistent with the intended use. The number of crashes is disproportionately higher at driveways than at other intersections; thus their design and location merit special consideration. Ideally, driveways should not be situated within the functional area of an intersection or in the influence area of an adjacent driveway”. As shown in Figure 1, the intersection functional area includes its upstream and downstream areas (7). This figure also shows that the physical area of the intersection that is a fixed area representing the space within the corners of the intersection. Unlike the physical area, the functional area is variable.
A question here is how the functional area can be specified. It’s also important to find out how far downstream and upstream access points can be located from an intersection. Figure 2 shows an upstream access point near a signalized intersection.

2.2 Access management

Streets and highways are valuable and their safe operation requires appropriate access management (AM) to and from the adjacent properties, businesses, and land developments. The common effective techniques of AM are to minimize or limit left turns, use raised medians, encourage shared driveways for adjacent land developments, provide adequately designed turn lanes, and create service roads for direct land access parallel to major arterial. In fact, AM involves managing traffic movements into and out of driveways. Drivers and road users expect...
to move smoothly and safely and residents and business owners have a right to access the roads. Thus the need for AM is essential as a balance between access and mobility. Prohibiting left turns, channelization, and other appropriate techniques significantly decrease access point crashes (8).

Researchers have been working on different techniques of access management (AM) for more than 30 years. An early classification of AM techniques was introduced by Stover et al. (9) in 1970. This was improved by Glennon et al. in 1975 (10). They provided checklists for the control of direct access to major roads and classified access techniques according to a) highway design and operation, b) driveway location, and c) driveway design and operation. In 1982, Flora (11) classified AM techniques by the following functional objectives: a) limit number of conflict points, b) separate basic conflict points, c) limit deceleration requirements, and d) remove turning vehicles from through lanes. In 1992, Koepke et al. (12) described different policy, planning, and design approaches to AM in various categories such as interchanges, frontage roads, medians, left turns, right turns; and driveway arrangements.

Bellomo-McGee (13) in 1993 included management elements with the access techniques and grouped the AM techniques as a) management, b) facility design, c) access driveway/design, and d) traffic control elements. NCHRP Report 420 in 1999, by Gluck et al., (3) emphasized policy (strategic) and design/operation (tactical) decisions in providing access to properties. They investigated safety, operation, environmental impacts, and economic impacts of the techniques such as intersections spacing, speed, corner clearance, median type, left turn lanes, U-turns, alternatives to direct left turns, and access separation. McCoy (14) and McCoy and Heimann (15) evaluated operational and safety impacts of driveway traffic volume on saturation flow rates at two signalized intersections. They found that driveway traffic can reduce the saturation flow rate on signalized intersection approaches. The amount of this reduction was found to depend on the corner clearance of the driveway and the proportions of volume that enter and exit the driveway.

According to the Access Management Manual (8), 10 principles should be applied to reach a complete AM plan. These principles are:
1. Provide a specialized roadway system – it is important to design and manage roadways according to the primary functions that they are expected to serve;
2. Limit direct access to major roadways – roadways that serve higher volumes of regional through traffic need more access control to preserve their traffic function;
3. Promote intersection hierarchy – an efficient transportation network provides appropriate transitions from one classification of roadway to another;
4. Locate signals to favor through movements – long, uniform spacing of intersections and signals on major roadways enhances the ability to coordinate signals and ensure continuous movement of traffic at the desired speed;
5. Preserve the functional area of intersections and interchanges – the functional area is where motorists are responding to the intersection (i.e., decelerating, maneuvering into the appropriate lane to stop or complete a turn);
6. Limit the number of conflict points – drivers make more mistakes and are more likely to have collisions when they are presented with the complex driving situations created by numerous conflicts. Traffic conflicts occur when the paths of vehicles intersect and may involve merging, diverging, stopping, weaving, or crossing movements;
7. Separate conflict areas – drivers need sufficient time to address one potential set of conflicts before facing another;
8. Remove turning vehicles from through-traffic lanes – turning lanes allow drivers to decelerate gradually out of the through lane and wait in a protected area for an opportunity to complete a turn, thereby reducing the severity and duration of conflict between turning vehicles and through traffic;
9. Use non traversable medians to manage turn movements – they minimize left turns or reduce driver workload and can be especially effective in improving roadway safety; and
10. Provide a supporting street and circulation system – a supporting network of local and collector streets accommodates development and unifies property access and circulation systems. Interconnected streets provide alternate routes for bicyclists, pedestrians, and drivers.

The Pennsylvania AM Handbook (16) summarizes the benefits of a good AM as follows:

**Community and neighborhoods:**
- Safer transportation system
• More attractive roadway corridors
• Lower taxes for future roadway investment
• Preservation of property values
• Safer pedestrian and bicycle travel
• Improved appearance of the built environment
• Reduced fuel consumption and air emissions

Business community:
• Stable property values
• More consistent development environment
• Reduced transportation and delivery costs

Pedestrians:
• Safer walking routes due to fewer conflicts with traffic
• Refuge areas created by medians

Bicyclists:
• Fewer conflicts with traffic
• More predictable traffic patterns
• Greater choice of alternative travel routes

Transit riders:
• Reduced delay and travel times
• Safer walking environment for access to stations

Motorists:
• Fewer traffic conflicts which increases driver safety
• Fewer traffic delays

Governmental agencies:
• Lower cost of providing a safe and efficient roadway
• Improved internal and intergovernmental coordination
• More success in accomplishing transportation goals
• Lowered collision and collision response costs
2.3 Corner clearance

Corner clearance is the distance between an intersection and the next driveway. The recommended corner clearance varies by state. Some of the states such as Florida provide different standards for different circumstances. Other states like Pennsylvania and North Carolina simply define a standard value regardless of the various circumstances.

The Florida Driveway Handbook (17) specifies the corner clearance based on access class and the roadway speed limit as shown in Table 1.

<table>
<thead>
<tr>
<th>Access class</th>
<th>Speed &gt; 45 mph</th>
<th>Speed ≤ 45 mph</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>N/A - freeway</td>
<td>N/A - freeway</td>
</tr>
<tr>
<td>2</td>
<td>1,320</td>
<td>660</td>
</tr>
<tr>
<td>3</td>
<td>660</td>
<td>440</td>
</tr>
<tr>
<td>4</td>
<td>660</td>
<td>440</td>
</tr>
<tr>
<td>5</td>
<td>440</td>
<td>245</td>
</tr>
<tr>
<td>6</td>
<td>440</td>
<td>245</td>
</tr>
<tr>
<td>7</td>
<td>125</td>
<td>125</td>
</tr>
</tbody>
</table>

This corner clearance could vary with the radius of the corner on the minor road as shown in Figure 3.

The Pennsylvania Municipalities Handbook (16) recommends using a corner clearance of 600 ft for a principal arterial, 400 ft for a minor arterial, and 200 ft for a major collector. In North Carolina the standards are based on the Driveway Manual (6). This manual recommends that the
corner clearance should be at least 100 ft from the point of tangency of the curb curvature of the intersecting streets and no distance less than 50 ft is allowed. This distance for a full movement driveway next to a signalized intersection should be more than 100 ft. Per Mr. James Dunlop, the North Carolina Congestion Management Engineer, “The district engineers evaluate each permit individually, however developers have mostly learned that the closer to an intersection, the less likely full movement access will be granted. If a parcel’s frontage is only 100’, we cannot deny a driveway connection, however we can restrict it to right-in/ right-out (RIRO). If the parcel has direct access to another state road, we can restrict access to that other road in some situations”.

2.4 Collision models

Collision prediction models show the relation between the dependent variable of the number of collisions and the independent variables like the roadway speed, number of lanes, and other factors. Collision models usually are provided based on the historical data from the same or similar roadways. A simple linear form of the model, assuming a normal probability distribution of collisions is:

$$\mu_i = \beta_0 + \beta_1 X_{i1} + \cdots + \beta_m X_{im} = \beta_0 + \sum_{j=1}^{m} \beta_j X_{ij} \quad \text{(Equation 1)}$$

Where,

- $\mu_i$ is the expected number of crashes at the $i$th segment, $i = 1, \ldots, n$
- $X_{ij}$ is the $j$th variable at the $i$th segment, $i = 1, \ldots, n$ and $j = 1, \ldots, m$
- $\beta_j$ is the $j$th regression coefficient, $j = 1, \ldots, m$
- $\beta_0, \beta_1, \ldots, \beta_m$ are the regression coefficients that need to be estimated.

Since the number of collisions is typically not normally distributed, a linear model doesn’t fit the crash data but the natural logarithm distribution has very similar form. The number of crashes can’t be a negative value, thus a lognormal regression model or a log-linear regression model could be applied to the data. This is based on the assumption that collision counts follow Poisson or negative binomial distribution. Previous research (18) shows that the Poisson or negative binomial distributions are acceptable for modeling discrete rare events like road collisions.
In a Poisson distribution it is assumed that the variance of the number of collisions at the segment i is the same as the mean value at this segment. The relation between expected number of collisions and the variables is expressed as:

\[ \ln(\mu_i) = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_m x_{im} = \beta_0 + \sum_{j=1}^{m} \beta_j x_{ij} \]  
(Equation 2)

Where \( \ln() \) is the natural logarithm of the expected number of collisions and the Poisson model is:

\[ P(y_i) = \frac{\exp(-\mu_i)(\mu_i)^{y_i}}{y_i!} \]  
(Equation 3)

Where \( P(y_i) \) is the probability of \( y_i \) collisions at the ith segment. If the variance of a collision data set exceeds the mean value of this data, it’s said that the data are overdispersed. In this case, the variance won’t be the same as the mean, like the Poisson model. The negative binomial distribution is an alternative to treat the overdispersion problem by adding a quadratic term to the variance representing overdispersion. The relationship between the expected number of crashes at the segment i and the variables is the same as for Poisson model but the negative binomial model is:

\[ P(y_i) = \frac{r^{y_i}(\frac{r}{r+1})^{y_i} K_i^{y_i} (\frac{1}{r+K_i})^{\frac{1}{r}}} {y_i! (\frac{r}{r+1})^r} \]  
(Equation 4)

Where, \( P(y_i) \) is the probability of the collisions at the segment i and \( K \) is the dispersion parameter. The variance is \( \mu_i + K \mu_i^2 \). As \( K \to 0 \), the variance will be the same as the mean and so the negative binomial model becomes the Poisson model.

In recent years numerous studies have been done and collision prediction models, with a limited number of variables, have been developed to find the effect of different variables on road collisions. Even though researchers have been using different methods such as the Empirical Bayes Method (19), fuzzy logic (20), and neural network (21) for collision prediction models, multivariable analysis is the most successfully applied methodology.
Apparently one of the first safety collision models was developed by McDonald in 1966 (22). He developed a model to relate crash frequency to traffic flows at divided highway intersections. His model is shown as:

$$N = 0.000783 V_m^{0.455} V_c^{0.633}$$  \hspace{1cm} (Equation 5)

Where $N$ was the number of crashes per year and $V_m$ and $V_c$ were major road and cross-road ADTs.

In 1986 Persaud (23) and Lovell et al. (24) used traffic control type as an explanatory variable in their intersection crash models. They showed that the safety effect of converting to all-way stop was contradictory. Lovell and Hauer affirmed the benefit of converting to four-way stop, while Persaud rejected its effectiveness.

In 1975 King et al. (25) concluded that signalization reduces right-angle crashes but increases rear-end crashes, with no significant change in total crashes – related disutility. Lau and May (26,27) in 1988 and 1989, and Naclerio et al. (28) in 1989 looked at crashes at signalized and unsignalized intersections. These models were developed to identify locations where collision experience was more frequent or more severe than normal, and to evaluate the safety consequences of alternative improvements. Three types of crash severity were modeled separately: fatal, injury, and property damage only. They used a nonparametric statistical modeling method known as the Classification Regression Tree. This method has particular applicability to categorical and discontinuous variables.

In 1988, Hauer et al. (29) developed collision prediction models in negative binomial form for signalized intersections by maneuver pattern before the occurrence of crashes. Each pattern involved two conflicting flows. These models have limited application for designers because they are based on the traffic flow patterns that are not subject to designer control. Unlike traffic flow patterns, physical elements such as channelization and alignment are controllable to improve safety.
In 1995, Fridstrom et al. (30) developed a statistical model to predict collisions using traffic flow, climate, lighting condition, and speed limit. They considered negative binomial regression and a new methodology for goodness-of-fit measures.

In the same year Hadi et al. (31) proposed several collision prediction models for multilane roads and two-lane roads. They used Poisson and negative binomial regression models and related crashes to AADT and road environmental factors.

In 2000, Abdel-Aty et al. (32) developed a statistical model using the negative binomial distribution and related the collisions to AADT, degree of horizontal curvature, section length, lane, shoulder width, median widths, driver sex, driver age, and urban/rural designation.

In 2003 Golob et al. (33) used both linear and non-linear multivariate statistical analysis to relate collisions to traffic flow, weather, and lighting conditions. Continuing this research in 2004 they developed another model using crash type, location, and severity (34).

In 2004 Hauer et al. used a negative a multinomial likelihood distribution for collision counts to model them on urban four-lane undivided roads (35). In another effort Hauer (36) presented a derivation of the negative multinomial likelihood function for collision models. He used a ratio of recorded collisions and collisions predicted by the current model. The introduced variable was first grouped into bins and then the ratios were calculated for each bin. This identifies the functional form for each variable. The independent variables were geometric characteristics and traffic flow. Hauer suggested checklists for the functional form to each variable in the model and observed that the model function may have both a multiplicative and an additive component.

Indeed, the multiplicative part is for the influence of variables that have a continuous role along a road (such as lane width or shoulder type) but the additive component is for the presence of hazardous points (such as driveways or narrow bridges). He introduced the cumulative residuals (CURE) method for measuring the goodness-of-fit of the models. CURE involves plotting the cumulative residuals as a function of the independent variable of interest. An acceptable goodness-of-fit using a CURE plot oscillates around zero. The model he suggested for continuous variables along the road has a multiplicative form as:
\[ y = f(x_1) \times g(x_2) \times h(x_3) \quad \text{... (Equation 6)} \]

Where \( y \) is the estimated number of collisions, \( x_1, x_2, x_3, \text{...} \) are the variables, and \( f(), g(), h() \text{...} \) are the functions of the variables to be estimated.

Previous research considered different factors causing collisions on the road and within the proximity of the intersections. Some models have included terms for the density of access points on a segment of road (4). However, no study has considered the parameters causing collisions at access points near intersections.
3. Data Preparation

I used historical data from a 5-year monitoring period extending from January 2005 to December 2009 to develop a statistical collision prediction model. To obtain proper data I first chose the appropriate study locations and then collected geometry, traffic, and crash data. In the end, to choose an accurate modeling method the collected data was analyzed. Figure 4 illustrates the tasks performed to prepare data for this research.

![Data Preparation Overview](image)

3.1 Site selection

I had access to a list of the intersections in North Carolina that was provided by NCDOT. This list contained 7,236 signalized and 8,736 unsignalized intersections in 100 counties. Wake County had 739 signalized and 503 unsignalized intersections in the list or almost 8% of the total. The total collision rate per 100 population is 2.44 and 2.24 in Wake County and other counties in North Carolina, respectively (37). Most of the roads in Wake County, like other counties of North Carolina, are in flat terrain. Climate condition, roadway geometry, road traffic volumes, drivers' behavior, and other factors are also very similar between Wake County and other counties in North Carolina. Also most of the roads in North Carolina are state owned and
are built and maintained to the same standards. Therefore Wake County is a good representative of North Carolina counties and was chose for data collection. Figure 5 illustrates Wake County among the other North Carolina counties.

![Figure 5 – Wake County in the State of North Carolina (38)](image)

For the purpose of this research a few criteria were taken into account for site selection. The criteria are listed and described in details below.

1- Random selections: Collection of an unbiased and random subset of the data is a key process to yield collision predictions. Each one of the 739 signalized intersections in Wake County was assigned a random number. Then these random numbers were sorted and the first 200 signalized intersections were chosen. For this study a site was defined as a leg of the intersection that fulfills the minimum requirements as described below. Each intersection may consist of 3, 4, or 5 sites depending on the number of the legs at the intersection. The 200 signalized intersections consisted of 674 sites. This means most of the intersections had 4 legs.

2- More than one lane on the major road at the access point: The number of lanes on the major road at the access point is a key issue. Most of the roads near intersections in urban and suburban areas operate with two or more lanes on each direction. Safety and operational issues related to driveways are also more critical on these types of roads than roads with one lane in each direction, so the researcher decided to filter the sites in this study for this characteristic.
3- Access points upstream of the intersections: Driver behavior, signing, and other factors depend on whether an access point is upstream or downstream of the intersection. At the upstream access point, as shown in Figure 2, drivers approaching the intersection should be aware of the traffic leaving or entering the access point. I decided to focus on these upstream access points as their issues are more serious than the downstream access points, in that the intersection queue may back up and block or limit the access point movements.

4- No free flow right movement at the intersection: One of the reasons that the researcher decided to study the access points near intersections is the queue back up. In case of a free flow right turn, queues may not form and the issue of the access point being next to the intersection wouldn't be as interesting. Thus for site selection I decided to focus on sites with no free flow right turn. Figure 6 illustrates regular and free flow right turns.

5- Closest access point to the intersection: Many driveway issues are the ones near intersections, so the researcher decided to focus on the closest access point upstream of the intersection.

6- Access points within 800 ft from intersection: Sites with long corner clearances should have few effects on the operations or safety of the intersection. For the purpose of this research I considered only access points located within 800 ft of the intersection. This is because queuing, turn bays, and other issues typically don't cause problems beyond this point.
7- Major roads grades less than 4%: Usually operational and safety problems increase on intersection approaches with grades of 4% or higher. Because of these negative impacts, and since in Wake County most of the roads are in flat terrain, I decided to concentrate on roads with grades of less than 4%.

All the sites in the sample of 200 intersections were scanned and filtered by implementing the above limitations. Table 2 lists all the lane configurations at the driveway (LCD) in the sample of intersections that emerged after I applied the seven criteria listed above. Up to this point I had 37 configurations at the driveway and a total of 260 sites.

<table>
<thead>
<tr>
<th>Site</th>
<th>6</th>
<th>1</th>
<th>3</th>
<th>1</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sites</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>10</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>LCD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Site</td>
<td>3</td>
<td>2</td>
<td>36</td>
<td>16</td>
<td>11</td>
<td>10</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>LCD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Site</td>
<td>12</td>
<td>6</td>
<td>2</td>
<td>5</td>
<td>31</td>
<td>11</td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>LCD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Site</td>
<td>1</td>
<td>5</td>
<td>9</td>
<td>13</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

With Table 2 in hand I established additional site selection criteria, including:
8- No more than two through lanes at the access point: Operations at the access point would be much more complicated when vehicles exiting the access point needed to cross more than two lanes to be positioned at the desired direction. In the other hand, as shown in Table 2, the number of sites with configurations of more than two through lanes is small compared to the sites with two through lanes and would therefore be hard to model statistically. As a result the researcher decided to scope the research down to sites with one or two through lanes at the access point.

9- No driveway associated with a different parcel within 150 ft of the main access point: If a different parcel has a driveway within 150 ft from the main access point, this site was eliminated from the candidate site list. Operation and safety issues at the second access point would significantly affect the first access point in ways that cannot be documented properly. In particular, collisions could not be attributed to the correct access point. For the cases when one parcel has a second access point within 150 ft from the first one, a variable will be introduced to distinguish them from the other sites, as discussed later.

10- Lane configurations at the driveway with 10 sites or more: Results from the statistical modeling might be unconvincing if the number of sites of each type was not sufficient. For the purpose of this research I decided to model the lane configurations at the driveway with 10 sites or more as shown later in Table 3.

11- Four-leg signalized intersections: I noticed that less than 5% of the study intersections surviving to this point had three legs and no intersections had more than four legs. Thus, the researcher focused on only four-legged intersections.

12- No sites with geometry changes within the last 5 years: Since the collision data were to be collected for a period of 5 years starting January 2005, sites with any geometry changes at or near the access point were eliminated. This is because the collision data should be unbiased from any external change or revisions.

13- No odd geometries: Sight distance is a length of a roadway that is visible to the driver. Sharp horizontal curvature reduces this visibility, reduces safety, and increases the risk of crashes.
Because of these complications, sharply-curved and odd geometry sites were eliminated from the site list. There were only a few cases with these characteristics.

14- Only public (state or city) owned roads: Some law enforcement agencies report the crashes only on public-owned roads. Also non-public roads, such as private roads in shopping centers, are not of much policy interest. As a result if the main study roadway was a private road, such as an entrance to a shopping center, it was eliminated from the database.

15- No small driveways: A site was not considered for data collection if its access point was unpaved or it served only a single-unit residence. The low traffic volumes of these access points provide only a slight risk of causing a collision.

16- Driveway movement type: The median type at the access point on the major road guide the drivers how and where to drive. Channelization, signing, and pavement marking help them understand the roadway configuration and follow the rules. Full movement, left-over (directional cross-over), and RIRO are the main three types of driveways, with the latter two designs being generally more desirable from a safety point of view where an access point is located within the functional area of an intersection. Figure 7 illustrates full movement, left-over, and RIRO designs.

I noticed that there was only 1 site with a left-over design in the sample. The reason may be that since the access points in this sample are near intersections, city and state traffic engineers likely choose to close the median in most cases. Channelizing an access point for a left-over design requires more signing and pavement marking to communicate a complex message to drivers. Due to the small number of left-over designs, the researcher decided to scope the study objective
down to only full movement and RIRO sites and delete the site with left-over design from the list. After applying above criteria, 103 sites were left in the study list.

The next chapter will discuss in detail how I collected the collision data. After a careful review of the sites’ collisions it came to my attention that only 4 out of the 103 sites have relatively higher collision frequencies with 15, 16, 21, and 32 crashes over the study’s 5 years. The rest of the sites had less than 7 collisions over the studied 5 years. As a result I decided to look further and find a few more sites with high number of crashes. I scanned the sites in Wake County randomly again and found 10 sites with some recent median changes at access points that meet all criteria discussed previously. The reason behind this step is that I presumed the improvements to the median took place because of complaints received by residents or investigations done by traffic engineers as a result of the high collisions at those sites. Indeed 5 of these 10 sites were already considered in our study but the crash data between 2005 and 2009 applied to the time when the median was already in place. Now I collected 5-year crash data from before any countermeasures were implemented at the site. The other 5 sites were thereby added to the sample. As a result the ultimate number of sites was 108. These sites are near 63 signalized intersections. A pin map of these 108 sites distributed in Wake County is shown in Appendix A. Table 3 shows the lane configuration of these 108 sites at their driveways.

Table 3 – Sites lane configuration at the driveway

<table>
<thead>
<tr>
<th>Configuration</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of Configuration</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>Number of Sites</td>
<td>23</td>
<td>10</td>
<td>32</td>
<td>12</td>
<td>20</td>
<td>11</td>
</tr>
</tbody>
</table>

Selecting 108 sites from the initial 674 sites reveals that this research is based on almost 16% of the total original sample. Of course, many sites discarded did not even have driveways. The result I develop will likely apply directly to between one-quarter and one-half of all driveways near to and upstream of signalized intersections in Wake County.
3.2 Data collection

A complete set of data, including geometric, traffic, and crash data, were needed to statistically model the collisions at the access points.

3.2.1 Geometric data

I collected data on all geometric factors that I thought might affect the crash rate of interest at the sites in the sample. These factors included:

- Corner clearance (CC), i.e., the distance between closest driveway and the intersection
- Lane configuration at the driveway on the major road (LCD)
- Lane configuration at the intersection near the access point on the major road (LCI)
- Major road median type at the access point (MM)
- Speed limit on the major road (SL)
- Grade on the major road (GR)
- Driveway angle, i.e., the angle between driveway centerline and the edge of travel way (DA)
- Driveway one-way or two-way operations
- Exclusive right turn into driveway
- Driveway median status, i.e., divided or undivided driveway (DM)
- Driveway width (DW)
- Transition between major road and driveway pavements (TR)
- Driveway radius (RA)
- Existence of another second driveway associated with the same parcel within 150 ft from the main access point (AD)

The researcher had access to the signal timing plans of all the intersections in Wake County. These timing plan sheets provided signal timing information and some relevant geometry data. Other sources of information to complete the geometry data set were aerial images online and site visits.
3.2.2 Traffic data

The most important explanatory variable in a statistical collision prediction model is usually traffic volume. Traffic volume represents the amount of exposure to the risk of collision. The traffic volume used for the purpose of this research consisted of the major road traffic volume and the access point traffic volume. The queue backed up at the intersection near the access point was also an estimated data using Synchro software. The methods to collect and estimate these data are described below.

3.2.2.1 Major road traffic volume

For major road traffic volumes, the Traffic Survey Group of the NCDOT provides comprehensive AADT data for the major road segments in North Carolina. The AADT database is presented in the form of map sheets on which road AADTs are spotted manually. Figure 8 illustrates a snapshot of the AADT map adjacent to NC State University.

![Figure 8 – A snapshot of the AADT map (39)](image)

The best condition would be using an average of the AADTs for the period of 5 years between 2005 and 2009. Since these data are typically provided every few years, an average was computed from the available data from 2005 to 2009.
Google Earth Pro, Version 5.2.1, added a new feature called “US Daily Traffic Counts”. This feature shows the AADTs for some of the major roads in the US as illustrated in Appendix B. I obtained some of the missing AADTs from Google Earth Pro.

Since the NCDOT information was included in the Google Earth Pro, I validated accuracy of the Google Earth Pro data by comparing them with the NCDOT AADT Maps for over 20 cases. The collected AADT values were almost the same.

In the few cases that the AADT map and Google Earth Pro did not provide the needed AADT, traffic data for the nearest point was assumed to be the same as for the study segment. For a few other cases, where there was no traffic volume available on or near the road segment, the missing data were interpolated by two AADTs located on either side of the road segment. If none of these techniques was helpful to find the AADT, the AADT was estimated using traffic counts obtained from the city engineer or collected manually. The traffic counts were converted to an AADT using Equation 7:

\[
AADT = \frac{\text{Traffic volume both direction}}{K_{\text{Sum}}} \times \text{S. F.}
\]  

(Equation 7)

Where \(K_{\text{Sum}}\) is the proportion of AADT occurring in the hours counted and S. F. is the seasonal factor. These values were provided by Traffic Survey Group of the NCDOT (40). Table 4 shows the \(K\) factors based on the type of route and the hours actually counted. Since most of the counts were collected for more than one hour, sums of the relevant \(K\) values were used. If the manual counts were for the fractions of the hours, an interpolation was done between the corresponding \(K\) factors.
The seasonal factor is used to adjust a daily volume count to AADT estimate. The NCDOT provides the seasonal factor for different types of roads. Table 5 illustrates these values for the types of roads in our research per month and day of the week based on the date the data were collected. If the manual count were collected on two different days, I used an average weekday factor for the month counted.

### Table 4 – K factors (40)

<table>
<thead>
<tr>
<th>Hour</th>
<th>Interstate</th>
<th>US</th>
<th>NC</th>
<th>SR</th>
<th>Local</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00AM</td>
<td>0.0110</td>
<td>0.0064</td>
<td>0.0063</td>
<td>0.0068</td>
<td>0.0068</td>
</tr>
<tr>
<td>1:00AM</td>
<td>0.0081</td>
<td>0.0040</td>
<td>0.0039</td>
<td>0.0037</td>
<td>0.0037</td>
</tr>
<tr>
<td>2:00AM</td>
<td>0.0071</td>
<td>0.0033</td>
<td>0.0032</td>
<td>0.0029</td>
<td>0.0029</td>
</tr>
<tr>
<td>3:00AM</td>
<td>0.0073</td>
<td>0.0036</td>
<td>0.0035</td>
<td>0.0027</td>
<td>0.0027</td>
</tr>
<tr>
<td>4:00AM</td>
<td>0.0098</td>
<td>0.0064</td>
<td>0.0071</td>
<td>0.0047</td>
<td>0.0047</td>
</tr>
<tr>
<td>5:00AM</td>
<td>0.0192</td>
<td>0.0183</td>
<td>0.0189</td>
<td>0.0149</td>
<td>0.0149</td>
</tr>
<tr>
<td>6:00AM</td>
<td>0.0439</td>
<td>0.0441</td>
<td>0.0447</td>
<td>0.0396</td>
<td>0.0396</td>
</tr>
<tr>
<td>7:00AM</td>
<td>0.0664</td>
<td>0.0695</td>
<td>0.0721</td>
<td>0.0780</td>
<td>0.0780</td>
</tr>
<tr>
<td>8:00AM</td>
<td>0.0595</td>
<td>0.0589</td>
<td>0.0595</td>
<td>0.0581</td>
<td>0.0581</td>
</tr>
<tr>
<td>9:00AM</td>
<td>0.0530</td>
<td>0.0535</td>
<td>0.0528</td>
<td>0.0474</td>
<td>0.0474</td>
</tr>
<tr>
<td>10:00AM</td>
<td>0.0534</td>
<td>0.0547</td>
<td>0.0534</td>
<td>0.0472</td>
<td>0.0472</td>
</tr>
<tr>
<td>11:00AM</td>
<td>0.0553</td>
<td>0.0578</td>
<td>0.0589</td>
<td>0.0545</td>
<td>0.0545</td>
</tr>
<tr>
<td>12:00PM</td>
<td>0.0569</td>
<td>0.0607</td>
<td>0.0631</td>
<td>0.0612</td>
<td>0.0612</td>
</tr>
<tr>
<td>1:00PM</td>
<td>0.0593</td>
<td>0.0619</td>
<td>0.0612</td>
<td>0.0606</td>
<td>0.0606</td>
</tr>
<tr>
<td>2:00PM</td>
<td>0.0638</td>
<td>0.0668</td>
<td>0.0681</td>
<td>0.0668</td>
<td>0.0668</td>
</tr>
<tr>
<td>3:00PM</td>
<td>0.0701</td>
<td>0.0751</td>
<td>0.0757</td>
<td>0.0747</td>
<td>0.0747</td>
</tr>
<tr>
<td>4:00PM</td>
<td>0.0743</td>
<td>0.0809</td>
<td>0.0809</td>
<td>0.0801</td>
<td>0.0801</td>
</tr>
<tr>
<td>5:00PM</td>
<td>0.0749</td>
<td>0.0820</td>
<td>0.0807</td>
<td>0.0832</td>
<td>0.0832</td>
</tr>
<tr>
<td>6:00PM</td>
<td>0.0577</td>
<td>0.0600</td>
<td>0.0580</td>
<td>0.0630</td>
<td>0.0630</td>
</tr>
<tr>
<td>7:00PM</td>
<td>0.0432</td>
<td>0.0419</td>
<td>0.0409</td>
<td>0.0480</td>
<td>0.0480</td>
</tr>
<tr>
<td>8:00PM</td>
<td>0.0348</td>
<td>0.0328</td>
<td>0.0319</td>
<td>0.0382</td>
<td>0.0382</td>
</tr>
<tr>
<td>9:00PM</td>
<td>0.0296</td>
<td>0.0263</td>
<td>0.0254</td>
<td>0.0300</td>
<td>0.0300</td>
</tr>
<tr>
<td>10:00PM</td>
<td>0.0234</td>
<td>0.0187</td>
<td>0.0181</td>
<td>0.0200</td>
<td>0.0200</td>
</tr>
<tr>
<td>11:00PM</td>
<td>0.0175</td>
<td>0.0125</td>
<td>0.0120</td>
<td>0.0138</td>
<td>0.0138</td>
</tr>
<tr>
<td>24-Hour</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

The seasonal factor is used to adjust a daily volume count to AADT estimate. The NCDOT provides the seasonal factor for different types of roads. Table 5 illustrates these values for the types of roads in our research per month and day of the week based on the date the data were collected. If the manual count were collected on two different days, I used an average weekday factor for the month counted.

### Table 5 – Seasonal factors (40)

<table>
<thead>
<tr>
<th>Month</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thr</th>
<th>Fri</th>
<th>Avg</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.11</td>
<td>1.14</td>
<td>1.10</td>
<td>1.09</td>
<td>0.99</td>
<td>1.09</td>
</tr>
<tr>
<td>2</td>
<td>1.07</td>
<td>1.07</td>
<td>1.06</td>
<td>1.02</td>
<td>0.92</td>
<td>1.03</td>
</tr>
<tr>
<td>3</td>
<td>1.04</td>
<td>1.05</td>
<td>1.05</td>
<td>0.99</td>
<td>0.94</td>
<td>1.01</td>
</tr>
<tr>
<td>4</td>
<td>1.01</td>
<td>0.99</td>
<td>0.98</td>
<td>0.95</td>
<td>0.87</td>
<td>0.96</td>
</tr>
<tr>
<td>5</td>
<td>1.00</td>
<td>0.97</td>
<td>0.95</td>
<td>0.93</td>
<td>0.82</td>
<td>0.93</td>
</tr>
<tr>
<td>6</td>
<td>0.96</td>
<td>0.96</td>
<td>0.97</td>
<td>0.93</td>
<td>0.83</td>
<td>0.93</td>
</tr>
<tr>
<td>7</td>
<td>0.95</td>
<td>0.96</td>
<td>0.93</td>
<td>0.96</td>
<td>0.85</td>
<td>0.93</td>
</tr>
<tr>
<td>8</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.93</td>
<td>0.83</td>
<td>0.93</td>
</tr>
<tr>
<td>9</td>
<td>1.00</td>
<td>0.98</td>
<td>0.99</td>
<td>0.94</td>
<td>0.83</td>
<td>0.95</td>
</tr>
<tr>
<td>10</td>
<td>0.99</td>
<td>0.99</td>
<td>0.96</td>
<td>0.93</td>
<td>0.81</td>
<td>0.94</td>
</tr>
<tr>
<td>11</td>
<td>1.02</td>
<td>1.00</td>
<td>0.97</td>
<td>1.03</td>
<td>0.89</td>
<td>0.98</td>
</tr>
<tr>
<td>12</td>
<td>1.03</td>
<td>1.05</td>
<td>1.06</td>
<td>1.02</td>
<td>0.91</td>
<td>1.01</td>
</tr>
</tbody>
</table>
3.2.2.2 Access point traffic volume

The AM and PM peak hour traffic volumes entering and exiting the access point are one of the key independent variables. I used the following techniques to collect or estimate these volumes.

Some of the access points are unsignalized intersections of the major road and a minor road. Peak hour turning movement counts for these intersections were obtained from the city engineer or collected manually on site or in the office. A camera was used to record the events as shown in Appendix C.

If the access point was associated with a residential subdivision, or a small parcel such as a gas station, convenience store, or fast food restaurant, the square footage of the buildings were derived from the Wake County GIS website or the Google Earth Pro tool (US Parcel Data). Then the Trip Generation Manual (41) was used to estimate the peak hour traffic volumes entering and exiting the access point. The higher peak hour volume was chosen among the AM and PM peak hours.

Multi-use development is defined by ITE as a single real estate project that consists of two or more ITE land-use classifications between which trips can be made without using the off-site road system. For the multi-use land developments, the trip generation process could be more complicated because the Trip Generation Manual would likely over-estimate the generated trips due to the internal circulating traffic and activities between on-site land uses.

The ITE Trip Generation Handbook (42) recommended a procedure to estimate these internal trips but it captures the internal trips only during the PM peak hour and not the AM peak hour. It also limits the data to only retail shopping centers, general office buildings, and residential. Recently Bochner et al. (43) conducted a study to improve the methodology used by the ITE Trip Generation Handbook. They examined and provided internal trip information for AM and PM peak hours for the six primary land uses of office, retail, restaurant, residential, cinema, and hotel. Dispensing with minor differences, I presumed all the studied land uses were among these six primary land uses. Then I deducted these internal capture estimates from the original trip generation estimations.
While reviewing the aerial historical images and visiting the study site areas if I noticed an empty parking lot during regular business hours, that particular land use was assumed to be vacant. Exceptions were made for some business with time-limited activities such as churches.

For subdivisions with multiple access points the typical methodology used for traffic impact analysis (TIA) was implemented for trip distribution. The geographic position of each access point and adjacent road led the analyst to come up with a reasonable assumption of the percentage of produced and attracted traffic that would use each access point. If such an assumption was not reasonable, I assumed that each access point was used proportionally to the adjacent road’s AADT.

**3.2.2.3 Queue at the intersection near the access point**

Queue length was estimated using the Synchro software because it was more convenient than actual site data collection and queue observation and because it has been widely used by analysts in traffic engineering for the last few years. The available turning movement and signal timing information were used as inputs to Synchro to obtain the estimated 95% queue, in feet at the intersection near the access point. The 95% queue in Synchro is the maximum back of queue with 95 percentile traffic volume (44).

Figure 9 illustrates an example of the queue formed on Oneal Road near the signalized intersection of Leesville Road and Oneal Road in Raleigh, NC. One can see that the queue blocked the access point from entering or exiting.
3.2.2.3.1 Queue validation

The queue length estimated values using the Synchro software were validated in this section. I randomly selected 14 sites from the 108 study sample size and observed the peak hour maximum number of vehicles at the intersection near the access point. Since I planned to compare the observed queue to the Synchro estimation that was in feet, I used a 25 feet average length of vehicles bumper to bumper to convert these queue lengths to feet. Table 6 shows the observed number of cars, maximum length of the queue and Synchro estimated queue values for 14 sites.
Table 6 – Observed and Synchro estimated maximum queue values

<table>
<thead>
<tr>
<th>Observed queue (#cars)</th>
<th>Observed queue (ft)</th>
<th>Estimated queue (ft)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>75</td>
<td>149</td>
</tr>
<tr>
<td>11</td>
<td>275</td>
<td>394</td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td>87</td>
</tr>
<tr>
<td>14</td>
<td>350</td>
<td>523</td>
</tr>
<tr>
<td>8</td>
<td>200</td>
<td>224</td>
</tr>
<tr>
<td>3</td>
<td>75</td>
<td>123</td>
</tr>
<tr>
<td>7</td>
<td>175</td>
<td>121</td>
</tr>
<tr>
<td>11</td>
<td>275</td>
<td>325</td>
</tr>
<tr>
<td>9</td>
<td>225</td>
<td>377</td>
</tr>
<tr>
<td>19</td>
<td>475</td>
<td>638</td>
</tr>
<tr>
<td>5</td>
<td>125</td>
<td>219</td>
</tr>
<tr>
<td>6</td>
<td>150</td>
<td>89</td>
</tr>
<tr>
<td>5</td>
<td>125</td>
<td>157</td>
</tr>
<tr>
<td>14</td>
<td>350</td>
<td>541</td>
</tr>
<tr>
<td>Mean</td>
<td>207</td>
<td>283</td>
</tr>
</tbody>
</table>

These values are plotted in Figure 10 to compare the Synchro estimated and observed queues.

![Figure 10 – Synchro estimated queue versus observed queue](image)

A regression analysis shows a linear relation between queues with a slope of 1.37 and standard error of 0.144. This means if an analyst collects the maximum queue at the intersection, this value can be multiplied by 1.37 to be usable by this research prediction model. In the other
words, on average the Synchro estimated queue lengths are 37% higher than the observed maximum queues. This different could be because:

- The definition of 95% queue in Synchro is the maximum back of queue with 95th percentile traffic volume while the maximum queue observed was derived from the maximum of the vehicles waiting at the intersection during the peak hour traffic volume.
- I considered an average of 25 feet between bumper to bumper of the vehicles but this value could be different depending on the number of trucks and other vehicles.
- Synchro estimation was using the available signal timing plans’ information. These could have been changed when the observed maximum number of vehicles was collected.

### 3.2.3 Collision data

Collisions reported on any roadway in North Carolina can be obtained from the collision database Traffic Engineering Accident Analysis System (TEAAS). In case a road has more than one road name, reported collisions can refer to different road names, so it was important to obtain all roadway names.

To develop a statistical model I collected 5 years of collision reported data, from January 2005 to December 2009. Law enforcement agencies are not required to submit non-reportable crashes and as a result these collisions are not included in our research. The reportable crashes must meet at least one of these criteria: a) results in a fatality, b) results in a non-fatal personal injury, c) results in property damage of $1,000 or greater, or d) results in property damage of any amount to a vehicle seized. A seized vehicle is a vehicle involving alcohol or other drugs in sufficient amount to constitute a DWI, a stolen, repossessed, or with other such history. The reportable crashes must also occur on public roadways. A copy of the North Carolina Police report is shown in Appendix D. Collisions were collected on both side of the major road on a full movement access point, but for the RIRO design only the collisions on the side of the driveway location were collected.

In this research, collision data refer to all the crashes reported within 150 ft of the access point. If the intersection was located within 150 ft of the access point, the collected collisions were
limited to those outside the physical area of the intersection. The physical area was shown in Figure 1.

3.2.3.1 Obtaining individual collision IDs and collision reports

To obtain the actual police reports from NC DMV website, each individual collision ID was needed. TEAAS was used to find the collision IDs near the intersections. The following steps were followed to make sure all the collisions were reviewed and no useful collision data were missed.

1- Coinciding names: As mentioned earlier, some of the roads have one or more coinciding names or local city names. Depending on the police officer notes, TEAAS provides crash IDs to different collisions for each one of the names. For example, the segment of Capital Boulevard between I-440 and New Hope Church/ Buffalo Road is also known as US-1, US-401, and SR-2179. Using different methods such as a feature in TEAAS called "Feature Names, Secondary Routes and Exceptions", Google Earth Pro images, and site visits I identified what I hope are all the possible route names for the sample segments.

2- 1000 ft from the intersection: As mentioned earlier, access points within 800 ft upstream of the intersection are in the feasible study area of this research. On the other hand collisions within 150 ft away from the access point were considered for modeling. The researcher used a feature of TEAAS called "Intersection Analysis Report" for the period of 5 years to extract the reported crashes. To avoid missing any useful data, all the collision data up to 1000 ft away from the intersection were collected.

3- Review the police reports: The collision data obtained up to this point were reviewed individually and filtered to find the qualifying crashes. This was done by scanning the police reports one by one to find the collisions within the desired segment near the access point. Figure 11 shows a snapshot of a sample police report.
A complete list of the collected collisions IDs is presented in Appendix E.

### 3.2.3.2 Filtering the collision reports

Since access – related collisions are of the interest of this study, first all the crashes within 150 ft from the access point were considered and then I filtered them down to the ones related to the access point within the said distance. To collect the access – related collisions it was essential to derive a logical set of criteria to distinguish the crashes that occurred because of the existence of the access point. One of the indications listed in the police report (number 49) states the vehicle maneuver/ action. This field states if the vehicles are making right turn, left turn, U turn, slowing, parking, etc. and helped the researcher in finding the access point – related crashes. A collision needed to satisfy at least one of the criteria below to be included in the access – related collision analysis.

- The narrative or diagram of the collision indicated that at least one of the vehicles involving the collision is clearly headed to or from the driveway, or
- Any word like "driveway" or "access point" is in the narrative section of the police report, or
- Field number 49 of the police report (shown in Appendix D) stated that one of the vehicles was making a turn to or from the driveway.
The criteria that were followed to delete collisions from the analysis were:

✓ Collisions involving pedestrian or animal
✓ Collisions involving vehicles travelling on different major intersection approaches, or
✓ Collisions that happened in the physical area of the intersection (as shown in Figure 1).

3.3 Data analysis

This section describes the collected geometry data qualitatively and quantitatively.

3.3.1 Corner clearance

The distance between the closest driveway and the intersection is defined as corner clearance and is measured from the curb’s intersecting tangent lines at the intersection to the curb’s tangent lines at the access point. These data were collected for all eligible intersections and are illustrated in Figure 12.

![Figure 12 – Corner clearance of the 108 sample sites](image-url)
Figure 12 shows that the distribution of corner clearances covers a wide range of distance. The database provides lots of small and large corner clearances that should help calibrate a reasonable model for this variable.

The North Carolina Driveway Manual (6) defines the corner clearance as "at an intersecting street or highway, the distance measured from the edge of the pavement curb line or the intersection of the right of-way lines to the beginning of outside driveway radius". This definition is different from what was used in this research. Figure 13 illustrates these two corner clearance measurements with “CC” representing the corner clearance.

On average, corner clearance measured according to the NCDOT Driveway Manual is 65 ft less than the measurement in this research.

3.3.2 Lane configuration at the driveway

The lane configurations on the major road at the driveway were shown in Table 3. As discussed earlier, I decided to collect the data for 6 major categories of the lane configurations at the driveways.

3.3.3 Lane configuration at the intersection

The lane configurations at the intersection near the access point are shown in Table 7. The table also shows the frequency of each collected category.
Table 7 – Lane configuration at the intersection

<table>
<thead>
<tr>
<th>LCI*</th>
<th>LCI* shape</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L,T,TR</td>
<td>42</td>
</tr>
<tr>
<td>2</td>
<td>L,TR</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>L,T,R</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>LT,R</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>L,T,T,R</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>L,L,T,R</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>L,L,TR</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>108</td>
</tr>
</tbody>
</table>

* LCI: Lane configuration at the intersection

In the table "L", "T", and "R" are left, through, and right lanes, respectively. "LT" is a left and through shared lane while "TR" refers to a through and right shared lane.

3.3.4 Major road median type

As mentioned earlier, the only left-over site was eliminated from the sample. Table 8 shows the number of full movement and RIRO sites.

Table 8 – Full movements and RIRO sites

<table>
<thead>
<tr>
<th>Median type</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Movement</td>
<td>77</td>
</tr>
<tr>
<td>RIRO</td>
<td>31</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
</tr>
</tbody>
</table>

3.3.5 Speed limit on the major road

Major road speed limit was considered as one of the variables in the collision model. Out of 108 eligible sites, one site is operating with no posted speed limit which implies 35 miles per hour (mph). Table 9 shows the posted limit in mph for the sample. Forty-eight percent of the sample operate with a speed limit of 35 mph or less and 52% operate with 45 mph or more.
Table 9 – Posted speed limit on the major roads

<table>
<thead>
<tr>
<th>Speed limit (mph)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>35</td>
<td>42</td>
</tr>
<tr>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>53</td>
</tr>
<tr>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>55</td>
<td>3</td>
</tr>
<tr>
<td>No Posted Speed Limit</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
</tr>
</tbody>
</table>

3.3.6 Grade on the major road

Steep roads have negative impacts on operations and safety. Usually these issues occur at intersection approach grades 4% or higher. Since most of the roads in Wake County are in flat terrain I decided to study the roads with grades of less than 4% and delete the steeper sites. There was only one site with a grade of 0.5. To ease the analysis this site assumed to be flat. Table 10 shows grades of all sample sites at the access point.

Table 10 – Major roads grade at the access point

<table>
<thead>
<tr>
<th>Major road grade (percent)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
<td>12</td>
</tr>
<tr>
<td>-2</td>
<td>18</td>
</tr>
<tr>
<td>-1</td>
<td>15</td>
</tr>
<tr>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>2</td>
<td>26</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
</tr>
</tbody>
</table>

3.3.7 Driveway angle

The angle between driveway centerline and the edge of major road travel way is defined as driveway angle. A higher driveway angle allows drivers to turn into the driveway at higher speed. If this angle is lower, drivers need to slow down or possibly stop before entering the driveway. Figure 14 shows how the driveway angle is measured.
Table 11 lists the driveway angles for the sample sites. Almost 68% of the sites have a right angle.

<table>
<thead>
<tr>
<th>Driveway angle (degrees)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-59</td>
<td>1</td>
</tr>
<tr>
<td>60-69</td>
<td>1</td>
</tr>
<tr>
<td>70-79</td>
<td>8</td>
</tr>
<tr>
<td>80-89</td>
<td>2</td>
</tr>
<tr>
<td>90-99</td>
<td>73</td>
</tr>
<tr>
<td>100-109</td>
<td>7</td>
</tr>
<tr>
<td>110-119</td>
<td>9</td>
</tr>
<tr>
<td>120-129</td>
<td>6</td>
</tr>
<tr>
<td>130-139</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
</tr>
</tbody>
</table>

According to the NC Driveway Manual (6), a driveway stem is defined as "the portion of a driveway between the public roadway and the internal roadway network or area where parking maneuvers occur". The State of NC requires a minimum of 100 feet of driveway stem for any new development with an internal roadway network. The direction of the driveway stem also shows the direction of traffic on driveway. In reality, there are parking lots, gas stations, and residences with no or very small driveway stems like the case shown in Figure 15.
For these cases, it was assumed that the vehicles' moving direction on the driveway would be perpendicular to the major road and as a result a 90-degree angle was recorded.

### 3.3.8 Driveway one-way or two-way operations

Initially I expected that some of the driveways may be restricted to one-way only. After examining the sites in the sample it appeared that all driveways are operating in two directions. As a result this variable was not used in model calibration.

### 3.3.9 Exclusive right turn into driveway

I was expecting some of the access points to have exclusive right turn lanes on the major road entering into the driveway. However, none of the studied sites provided an exclusive right turn lane. Therefore, this variable was also deleted from the list.

### 3.3.10 Driveway median status

I grouped the driveways into divided and undivided driveways. Table 12 shows the frequencies for this category. Over 86% of the studied sample sites had no median.
Table 12 – Driveway median type

<table>
<thead>
<tr>
<th>Driveway median type</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Undivided</td>
<td>93</td>
</tr>
<tr>
<td>Divided</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
</tr>
</tbody>
</table>

The 93 undivided sites included 70 full movement and 23 RIRO driveways. The 15 divided sites included 8 with full movements and 7 RIRO. For the purpose of this research driveway median length and width are not considered to be significant factors.

3.3.11 Driveway width

The narrowest width of driveway measured parallel with the edge of traveled way is defined as the driveway width. This is measured from edge to edge of the pavement for a driveway with no median. In the case of a divided driveway, this value is the sum of two directions' lanes from edge to edge of the pavement. Table 13 lists the driveway widths in feet with a minimum of 8 ft, a maximum of 58 ft, and average of 27.5 ft.

Table 13 – Driveway width

<table>
<thead>
<tr>
<th>Driveway width (ft)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-9</td>
<td>1</td>
</tr>
<tr>
<td>10-19</td>
<td>8</td>
</tr>
<tr>
<td>20-29</td>
<td>58</td>
</tr>
<tr>
<td>30-39</td>
<td>31</td>
</tr>
<tr>
<td>40-49</td>
<td>8</td>
</tr>
<tr>
<td>&gt; 50</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
</tr>
</tbody>
</table>

About 54% of the sample driveway width were 20 ft to 29 ft, 29% were 30 ft to 39 ft, and only 8% were less than 19 ft. The NC Driveway Manual (6) suggests a minimum width of 20 ft and a maximum width of 36 ft for two-way operations for the new driveway to be permitted.

3.3.12 Transition

This variable expresses the transition between major road and driveway pavements. A driveway may connect to the major road very smoothly or elevated by a small hitch. As Figure 16 shows, this variable could impact a driver's behavior on how to enter and exit the access point.
Table 14 shows the frequencies of the transitions.

Table 14 – Transition between major road and driveway pavements

<table>
<thead>
<tr>
<th>Driveway transition</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elevated</td>
<td>17</td>
</tr>
<tr>
<td>Smooth</td>
<td>91</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
</tr>
</tbody>
</table>

3.3.13 Driveway radius

The radius between major road and driveway is categorized as large or small based on a rational judgment. As an example in Figure 17, the left photo shows a driveway with a large radius and the right photo shows a driveway with a small radius.

Table 15 shows that almost 54% of the sites had a large radius and the rest had a small radius.
### Table 15 – Driveway radius

<table>
<thead>
<tr>
<th>Driveway radius</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large</td>
<td>58</td>
</tr>
<tr>
<td>Small</td>
<td>50</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
</tr>
</tbody>
</table>

#### 3.3.14 Second driveway

It was mentioned earlier that if a different parcel has a driveway within 150 ft from the main access point, this site was eliminated from the candidate site list. On the other hand if the same parcel had other driveways within the said distance, a binary variable was defined to distinguish between these sites and the sites with a single access point inside the said distance.

Table 16 shows that there were 94 sites with a single driveway in the study area and 14 sites with more than one driveway within 150 ft from the main access point for the same parcel.

### Table 16 – Existence of second driveway

<table>
<thead>
<tr>
<th>Second driveway</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>15</td>
</tr>
<tr>
<td>No</td>
<td>93</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
</tr>
</tbody>
</table>

#### 3.4 Summary of data

Appendix F shows a summary of the entire collected data and the collision reports from the 108 study sites. The collisions and other data were observed within a small section on the major road and on the driveway. This section contains the driveway width plus 150 ft from each side of the driveway limited to the intersection stop bar. The minimum corner clearance in this study is 28 ft, so the segment length on the major road ranges between 186 ft (150 + 8 + 28) and 358 ft (150 + 58 + 150). The LCI, Q-L, Q-T, and Q-R may or may not be located within the said distance depending on if the corner clearance was less or more than 150 ft.
4. Analysis Methodology

In recent years numerous collision prediction models have been developed to find the effects of different variables on road collisions. Crash frequency (number of crashes per year) and crash rate (number of crashes per million vehicle miles) have been used in recent research. Since typically crashes are not linearly correlated to the road traffic volume and other variables, researchers prefer to use crash frequency rather than crash rate (45).

Collisions could involve many possible contributing factors such as vehicle condition, road condition, climate condition, and human behavior. Some of these factors, like human behavior, are very difficult to measure and as a result it requires very extensive experiments to reveal the appropriate link between these factors and roadway collisions. Usually researchers choose a limited and reduced number of variables to be included in a model.

One of the objectives of this research was to find the statistical relationship between driveway related collisions and casual factors. A model predicts the chance of a collision based on the
known factors using previously collected historical data. It’s based on the assumption that the influence of the factors is reflected in the data. Previously it was revealed that the most commonly implemented technique in safety modeling is multivariable analysis. I decided to implement this method in this collision prediction analysis because of its long and successful use in collision analyses. This chapter describes the access point – related collision statistical modeling in detail. As shown in

Figure 18, first the method to choose the appropriate model is described showing the data traits. Then, the detail of the collision modeling process including the candidate forms, ID method to recognize the candidate functional forms, coefficients estimation, and model selection are explained. Finally, a method to test the quality of the fitted model is developed.

4.1 Modeling method

Modeling tools optimize the coefficients of the chosen model frame. It’s very important to adopt a proper model frame, since the outcome of modeling effort could result in different models even with the same data set. Also without an appropriate modeling tool, even with a good set of data and proper model selection, a reasonable model wouldn’t be possible. This section discusses choosing a proper modeling method and selecting the model frames for the statistical collision modeling.

4.1.1 Data traits

Characteristics of the data dictate how to choose the appropriate model fitting the data. For example if the relationship between two variables is not linear, a linear regression model cannot fit the data sufficiently. It’s also critical to know the probability distribution of the data. This helps the researcher to pick a better modeling frame fitting the data. For instance, a modeling method can be different for a normally-distributed dataset and a Poisson-distributed dataset. The data used in this research consist of number of reported access point – related collisions per year as the dependent variable and the following 15 independent variables:

1. Major road AADT in vehicles per day for both directions.
2. Driveway traffic volume in vehicles per hour.
a. Driveway entrance traffic volume in vehicles per hour.
b. Driveway exit traffic volume in vehicles per hour.

3. Corner clearance, i.e., the distance between closest driveway and the intersection, in feet.
4. Synchro queue at the signalized intersection on access point leg, in feet.
5. Driveway width in feet.
6. Lane configuration on the major road at the intersection near the access point. Table 7 shows that our research consists of 7 types of lane configurations.
7. Lane configuration on the major road at the access point (types I, II, III, IV, V, and VI). Table 3 shows these configurations.
8. Grade on the major road.
9. Speed limit on the major road in mile per hour.
10. Major road median type at the access point: For our research this variable is only categorized as full movement and RIRO.
11. Driveway median status, i.e., divided or undivided driveways.
12. Driveway angle, i.e., the angle between driveway centerline and the edge of travel way.
13. Driveway radius: If this radius is large or small.
14. Transition between major road and driveway pavements.
15. Existence of second driveway (for the same activity) within 150 feet each side of the study driveway.

As mentioned in Chapter 2, the only dependent variable in this research, number of collisions, follows a negative binomial distribution. During the 5-year monitoring period on 108 study sites there were a total of 56.6 reported access point – related collisions per year overall, ranging from 0 to 6.4 collisions per year per site. Since fatalities and injuries due to access point – related collisions are too infrequent to be analyzed alone, total access point related collisions were considered for analysis and prediction.

The major road AADT ranged from 1,230 to 54,000 vehicles per day with an average of 15,600 vehicles per day over the 108 study sites. This variable is expressed as the AADT/10,000 to reduce the high number of zeroes and simplify model development.
Even if there was some dependence between the 108 study sites, I believe the effects to be negligible. Figure 19 illustrates the relationship between the dependent variable, annual access–related collisions, and the primary independent variable, AADT/10,000. Visually it’s very hard to discover the relationship between these two variables since there are no associated data between two clusters of collisions (over 2.5 collisions per year and under 1.5 collisions per year).

![Figure 19 – Correlation between access–related collisions per year and AADT/10,000](image)

Relationships between the access–related collisions and the other independent variables are also very hard to find visually. It was difficult to see if the relationships were linear or nonlinear, or if they followed a particular functional form. More on this will be discussed in the next sections.

### 4.2 Modeling Process

In this research there were fifteen independent variables, including binary, categorical, and continuous variables. The only dependent variable was the number of collisions which was a discrete, nonnegative variable and assumed to follow a negative binomial distribution. Figure 20 illustrates the modeling methodology that will be discussed in following chapters.
Figure 20 – Modeling methodology overview (46)
4.2.1 Candidate functional forms

The ultimate form of the model may have multiplicative or additive components. The additive parts represent the characteristics of the hazardous points such as narrow bridges. On the other hand the multiple components represent the influence of the variables with a continuous role along the road regardless of the road section length (35). The data collection section described that access – related collisions and other variables in this study are for a small section of the road ranging between 186 ft and 358 ft. Within this section all the variables are continuous meaning that their effect is multiplied to the overall crashes and not added. For example even though the median type on the major road (full movement or RIRO) looks to be a point hazard and not continuous, drivers notice the median type probably even before entering the study segment and as a result they decide how to react (turning left, right, or going straight) based on the median condition long before the actual median location in the database. Therefore, it is considered a continuous variable. The same logic guides the treatment of the other variables such as LCI and LCD. Therefore, in this study the functional form of the expected collision prediction model is

\[
y = f\left(\frac{AADT}{10,000}\right) \times g(DW) \times h(Q) \times ...
\]  

(Equation 8)

Where \(y\) is the expected number of access point related collisions per year, “DW” is driveway width in feet, “Q” is the Synchro queue length in feet, and \(g()\) and \(h()\) are the functions to be formed and estimated. Traffic data on the major road was the dominant independent variable in developing the collision model and was considered as the principal (initial) variable because it is the exposure to the risk of getting involved in a collision. Then the other independent variables are introduced one after another. These variables could be continuous, categorical (class), or binary like major road traffic volume, lane configuration at the intersection, or speed limit, respectively. Modeling the continuous functions using the Empirical Integral Function (EIF) method will be described in next chapter.

I decided to consider a one-way ANOVA model for the categorical (discrete) variables. To do so, a categorical function \(f(x)\) written as:

\[
g(x) = \beta_1(x = x_1) + \beta_2(x = x_2) + \cdots + \beta_i(x = x_i)
\]  

(Equation 9)
In which $\beta_1, \beta_2, ..., \beta_i$ are the parameters to be estimated using the NLMIXED procedure. NLMIXED and the parameter estimation process will be discussed in upcoming chapters. The values in parenthesis, $(x = x_i)$ are denoted only 0 or 1 and are considered 0 if $x \neq x_i$ and it is 1 if $x = x_i$. As mentioned earlier, since this function is multiplied by the collision prediction model, if, for example, $x = x_1$, the former function is multiplied by $\beta_1$.

The same methodology was used for binary variables. Indeed a binary variable is a categorical variable with only two possible values. Thus it can be written as:

$$g(x) = \beta_1 (x = x_1) + \beta_2 (x = x_2)$$  \hspace{1cm} (Equation 10)

Previous research (46) modeled the binary variables as $g(x) = 1 + \beta x$ assuming the variable $x$ is 0 or 1 and thus $g(x)$ is estimated as 1 or $(1 + \beta)$, respectively. Since this model is multiplied to the initial model, $f(AADT/10,000)$, the problem is that it’s assumed one of the values of the binary variables does not affect collisions while the other one has some effect. With the methodology used in this study I presume that with either scenario of the binary variable, there is some effect on collisions.

For example consider the variable “major road median type”. As mentioned earlier, in this study the median type on the major road at the access point is categorized as full movement or RIRO and therefore it’s a binary variable. Each of the scenarios may have an effect on overall collisions, with an expectation of the latter having fewer collisions. A ‘no effect’ scenario could be expected for a no–driveway condition. The estimation results in a collision prediction model of:

$$y = f_1(AADT/10,000)[0.50(\text{Median} = \text{RIRO}) + 1.06(\text{Median} = \text{FULL})]$$ \hspace{1cm} (Equation 11)

This model reveals a significant 50% decrease for a RIRO median and a 6% increase for a full movement median relative to the collision prediction model using only $AADT/10,000$.  

50
4.2.2 Integrate – differentiate method

The integrate – differentiate (ID) method was first developed by Hauer et al. in 1997 (47) to recognize candidate functional forms between a dependent variable and a continuous independent explanatory variable. The ID method can be used when only the model form is a multiplication of the independent variable functions. The model forms are examined one by one to find the best fit to the integral function. Thus this method cannot be an automated procedure and requires consideration of each variable function individually.

In some cases there were simpler (perhaps linear) forms to be considered. I reviewed all the possibilities even though it required extra time and work because 1) this was consistent with the other functions, 2) this would cover all possible functions, 3) safety research professionals in traffic and safety have previously done this, and 4) usually the outcome turned out to be the same anyway, meaning the result will be linear.

The purpose of this section was to show how the ID method was used to develop the model function form and find candidate relationships between number of collisions and independent variables such as the AADT/10,000. Scatterplots of AADT/10,000 with a total of 108 points were shown in Figure 19. The majority of the sites had no collisions and only a few sites had more than 2 collisions. The correlation between AADT and the number of collisions is not strong in Figure 19.

By drawing the EIF this relationship could be better understood. First data are sorted by AADT/10,000 and then bin widths and bin heights for each site are determined to calculate the bin area. To determine the bin width, the left boundary is half-way to the nearest lower AADT/10,000 and the right boundary is half-way to the nearest higher AADT/10,000. This is shown in Figure 21. As a result the bin width is half of the difference between the nearest higher and nearest lower AADT/10,000. The bin height is the number of collisions at the site. If more than one site has the same AADT/10,000, an average of their collisions was used for bin height. Each value of the EIF at the right boundary is the cumulative bin area at that boundary. This is the sum of all bin areas from the lowest AADT/10,000 up to that point.
The EIF of the AADT data is shown in Figure 22. Some pattern is seen in this figure.
The concept of the ID method is that if functional form of the dataset can be determined by the EIF even if the scatterplots don’t show any discernible pattern. The shape in Figure 22 could be fitted by different curves having similar cumulative functions. Assume one of the curves to be fitted is \( y = f(x) \) that relates number of collisions, \( y \), to the AADT/10,000 (x). Then \( Y = F(X) \) is the integral function of \( y = f(x) \) between \( x = 0 \) and \( x = X \). On the other hand, a summation of the bin areas between these two points is the EIF and is denoted by \( F_E(X) \) which is an estimation of the \( F(X) \). As a result the functional form of the \( f(x) \) can be found by taking a derivative of the \( F_E(X) \). Now the question is how to find an appropriate fit for the \( F_E(X) \).

For the number of collisions and the traffic volume data, two example candidate functions are shown in Figure 23 fitting the cumulative function, \( F(X) \) with high \( R^2 \) values. The candidates are \( F_1(X) = 0.1455X^2 + 0.0808X \) and \( F_1(X) = 0.2437X^{1.6771} \). As a result the corresponding functions will be \( f_1(x) = 0.2910x + 0.0808 \) and \( f_1(x) = 0.4087x^{0.6771} \), respectively. Soon it will be shown that these are the candidate function forms to be fitted to the data.

![Figure 23 – Example of two candidate functions fitting the cumulative function](image)

Some candidate \( f(x) \) functions that could link the number of collisions and daily traffic and their corresponding \( F(X) \) are illustrated below. According to Baek (46) and Hauer (47), these candidate functions are relatively well known among traffic engineering researchers.
Table 17 – Candidate functions to fit the EIF and their corresponding functions

<table>
<thead>
<tr>
<th>#</th>
<th>F(X)</th>
<th>f(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\beta_0 X + \beta_1 X^2$</td>
<td>$\beta_0 + 2\beta_1 x$</td>
</tr>
<tr>
<td>2</td>
<td>$\beta_0 X^{\beta_1 + 1}$</td>
<td>$\beta_0 (\beta_1 + 1)x^{\beta_1}$</td>
</tr>
<tr>
<td>3</td>
<td>$\beta_0 X^2 + \beta_1 X^3$</td>
<td>$2\beta_0 x + 3\beta_1 x^2$</td>
</tr>
<tr>
<td>4</td>
<td>$\beta_0 X^3 + \beta_1 X^4$</td>
<td>$3\beta_0 x^2 + 4\beta_1 x^3$</td>
</tr>
<tr>
<td>5</td>
<td>$(\beta_0/\beta_1^2) e^{\beta_1 X} (\beta_1 X - 1)$</td>
<td>$\beta_0 x e^{\beta_1 x}$</td>
</tr>
<tr>
<td>6</td>
<td>$\beta_0 X^2 + e^{\beta_1 x}$</td>
<td>$2\beta_0 x + \beta_1 e^{\beta_1 x}$</td>
</tr>
<tr>
<td>7</td>
<td>$-x^{\beta_0 + 1} e^{-\beta_1 X}$</td>
<td>$x^{\beta_0 e^{\beta_1 x}}$</td>
</tr>
</tbody>
</table>

Table 17 includes simple linear, polynomial (quadratic and cubic), power, Hoerl’s, and other functions. The unknown coefficients in models are estimated using NLMIXED in SAS®. Fitting criteria will also be compared to select an appropriate model among the candidates. These methods are discussed in following sections.

### 4.2.3 Estimation of the model coefficients

In the previous section I described how to choose a functional form. Now to develop the statistical models it’s essential to estimate the model coefficients using NLMIXED in SAS®, by maximizing an approximation to the log likelihood integrated over the random effects (49). The reason NLMIXED was chosen is the ability to fit nonlinear models, as the predicted models in this study may have nonlinear forms for each independent variable. On the other hand, the number of collisions follows a negative binomial distribution that is one of the response probability distributions supported by NLMIXED as discussed below.

#### 4.2.3.1 Negative binomial log likelihood model

SAS® software using NLMIXED performs the maximum log likelihood estimation method. NLMIXED searches parameter values that make the log likelihood largest using some optimizing techniques. The NLMIXED optimizes parameter values in a model so as to maximize the logarithm of the probability of the observed collision counts. The log likelihood function to be maximized for the negative binomial distribution can be expressed as (46):

\[
L(y, \mu, k) = \sum_i l_i
\]

\[
l_i = \log(f(y_i, \mu_i, k)) = y_i \log(k\mu_i) - (y_i + 1/k) \log(1 + k\mu_i) + \log \left( \frac{\Gamma(y_i + 1/k)}{\Gamma(y_i + 1)\Gamma(1/k)} \right)
\]  

(Equation 12)

(Equation 13)
Where

\[ L = \text{Log likelihood function} \]
\[ l_i = \text{Individual contribution to the log likelihood} \]
\[ y_i = \text{Response} \]
\[ \mu_i = \text{Estimated mean} \]
\[ k = \text{Dispersion parameter (shown in Section 2.4)} \]

This analysis is based on the assumption that the introduced variables are independent among the 108 study sample sites.

### 4.2.3.2 Optimization techniques in NLMIXED

A variety of alternative optimization techniques such as trust region optimization and dual quasi-Newton method are available in NLMIXED to maximize the log likelihood estimation, but the default is a dual quasi-Newton algorithm (49). The trust region method uses the gradient vector (first order derivatives) and the Hessian matrix (second order derivatives) for optimization and it requires that the objective function have continuous first and second order derivatives inside the feasible region. The trust region method performs well for small- to medium- sized problems. On the other hand the dual quasi-Newton method uses the gradient vector and doesn’t need the second order derivatives for optimization. It works well for medium to moderately large optimization problems where the objective function and the gradient are much faster to compute than the Hessian. These two methods were tried in the modeling process and since there was no discernible difference in the result, the default method was adopted for the modeling in this study.

The result of the NLMIXED was validated for the negative binomial log likelihood function by previous research (46). I verified this validation with our study data, applying the same data in NLMIXED and in GENMOD, another procedure widely used to fit Poisson and negative binomial regression models for multinomial data. GENMOD can perform a negative binomial regression analysis with a log link function that is also called log-linear model (Equation 2). I fit a model with 2 independent variables using three coefficients \((b_0, b_1, b_2)\) related to the number of annual collisions. The codes used in SAS® for the GENMOD procedure and the output are shown in Figure 24.
```sas
data ARCollisions;
input Config C C AR_C RAD1_10k Dri_Vol Q_L Q_I Q_R HTOD LQI MT SL Grade Ang Dri_IW Dri_Div Dri_Ir Dri_Rad;
data lines:
6 112 1.0 2.30 186 6 322 62 0 5 0 1 52 90 88 0 0 1
1 148 0.8 0.42 66 51 36 36 0 2 1 35 -3 39 35 0 0 1
....
5 206 0.0 1.60 75 18 195 195 0 1 1 45 -3 30 20 0 0 1
5 460 0.2 1.44 140 376 35 35 0 1 1 35 0 0 30 0 0 0
;
run;

proc genmod data = ARCollisions;
model AR_C = Dri_IW AADT_10k/
dist = nb
link = log;
run;
```

---

**The SAS System**

**The GENMOD Procedure**

### Criteria For Assessing Goodness Of Fit

<table>
<thead>
<tr>
<th>Criterion</th>
<th>DF</th>
<th>Value</th>
<th>Value/DF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deviance</td>
<td>105</td>
<td>73.8124</td>
<td>0.7000</td>
</tr>
<tr>
<td>Scaled Deviance</td>
<td>105</td>
<td>73.8124</td>
<td>0.7000</td>
</tr>
<tr>
<td>Pearson Chi-Square</td>
<td>105</td>
<td>105.1701</td>
<td>1.0016</td>
</tr>
<tr>
<td>Scaled Pearson X2</td>
<td>105</td>
<td>105.1701</td>
<td>1.0016</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td></td>
<td>-97.3103</td>
<td></td>
</tr>
<tr>
<td>Full Log Likelihood</td>
<td></td>
<td>-202.6205</td>
<td></td>
</tr>
<tr>
<td>AIC (smaller is better)</td>
<td></td>
<td>202.6205</td>
<td></td>
</tr>
<tr>
<td>AICC (smaller is better)</td>
<td></td>
<td>203.0089</td>
<td></td>
</tr>
<tr>
<td>BIC (smaller is better)</td>
<td></td>
<td>213.3430</td>
<td></td>
</tr>
</tbody>
</table>

### Analysis Of Maximum Likelihood Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Wald 95% Confidence Limits</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>-3.0617</td>
<td>0.6607</td>
<td>-4.3566</td>
<td>-1.7660</td>
<td>0.0001</td>
</tr>
<tr>
<td>Dri_IW</td>
<td>1</td>
<td>0.0593</td>
<td>0.0175</td>
<td>0.0250</td>
<td>0.0936</td>
<td>0.0007</td>
</tr>
<tr>
<td>AADT_10k</td>
<td>1</td>
<td>0.3914</td>
<td>0.1557</td>
<td>0.0863</td>
<td>0.8385</td>
<td>0.0119</td>
</tr>
<tr>
<td>Dispersion</td>
<td>1</td>
<td>0.4917</td>
<td>0.2425</td>
<td>-0.0496</td>
<td>0.9970</td>
<td></td>
</tr>
</tbody>
</table>

---

Figure 24 – Codes and output for GENMOD procedures in SAS®
The codes used in SAS® for the NLMIXED procedure and the output are shown in Figure 25.

```
data ARCollisions;
  input Config CC AR_C AADT_10k Dri_Vol Q_L Q_T Q_R WTOPD LCI MT SL Grade Ang Dri_LW Dri_Div Dri_Ts Dri_Rad;
  datalines;
  6 112 1.0 2.30 186 4 322 62 0 5 0 45 2 90 50 0 0 1
  1 198 0.8 0.42 66 54 36 36 0 2 1 35 -3 90 33 0 1 0
  ......
  5 208 0.0 1.60 75 18 185 185 0 1 1 45 -3 90 20 0 0 1
  5 460 0.2 1.44 160 376 35 35 0 1 1 35 0 80 30 0 0 0
; run;
proc NLMIXED data = ARCollisions maxfunc = 10000 maxiter = 1000;
  para b0=1 b1=1 b2=1;
  mu = exp(b0+b1*Dri_LW+b2*AADT_10k);
  loglike = AR_C*log(k*mu)-(AR_C+(1/k)))*log(1+k*mu)+lgamma(AR_C+(1/k))-lgamma(AR_C+1)-lgamma(1/k);
  model AR_C ~ general(loglike);
run;
```

**The SAS System**

**The NLMIXED Procedure**

**Fit Statistics**

-2 Log Likelihood: 194.6
AIC (smaller is better): 202.6
AICC (smaller is better): 203.6
BIC (smaller is better): 213.3

**Parameter Estimates**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Error</th>
<th>DF</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>Alpha</th>
<th>Lower</th>
<th>Upper</th>
<th>Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>b0</td>
<td>-3.0617</td>
<td>0.6607</td>
<td>108</td>
<td>-4.63</td>
<td>&lt;.0001</td>
<td>0.05</td>
<td>-4.3713</td>
<td>-1.7521</td>
<td>-0.00002</td>
</tr>
<tr>
<td>b1</td>
<td>0.05933</td>
<td>0.01750</td>
<td>108</td>
<td>3.39</td>
<td>0.0010</td>
<td>0.05</td>
<td>0.02485</td>
<td>0.09402</td>
<td>0.00193</td>
</tr>
<tr>
<td>b2</td>
<td>0.3914</td>
<td>0.1557</td>
<td>108</td>
<td>2.51</td>
<td>0.0134</td>
<td>0.05</td>
<td>0.08280</td>
<td>0.7000</td>
<td>-0.00041</td>
</tr>
<tr>
<td>k</td>
<td>0.4317</td>
<td>0.2425</td>
<td>108</td>
<td>1.70</td>
<td>0.079</td>
<td>0.05</td>
<td>-0.04302</td>
<td>0.3124</td>
<td>0.000094</td>
</tr>
</tbody>
</table>

Figure 25 – Codes and output for NLMIXED procedures in SAS®
The GENMOD procedure model output and NLMIXED procedure output are also shown in Figure 24 and Figure 25. One can see that the model estimation results of both procedures, including dispersion parameters, are identical (or very close). This validates the negative binomial log likelihood function and the nonlinear optimization algorithm specified in NLMIXED. For this study the NLMIXED procedure was chosen because it can fit nonlinear models while GENMOD can only fit generalized linear functions.

4.2.3.3 Initial values for each parameter

The NLMIXED procedure estimates the model coefficients based on the initial values but it’s not guaranteed that estimated optimum solutions are global (rather than locally-optimized) in general nonlinear models. Indeed the possibility of terminating at a local optimal solution based on the initial values is very undesirable. Figure 26 shows an example depicting local and global optimal solution concept. One can see that if the initial value is chosen to be “A”, the optimization algorithm will end up at a local solution “C” and if it starts from the initial value of “D”, the result will be the global solution of “G”. The broken lines are the objective function level curves.

![Figure 26 – Optimum local solution and global solution (48)](image)

The default initial value in NLMIXED is 1 but it is recommended (48) to choose a reasonable initial value to start the NLMIXED procedure. It’s very important to know the graphical
behavior of the function to choose an appropriate initial value. As an example Figure 27 shows
different graphical shapes of a quadratic function per different values for “a” and “b” parameters
for a quadratic function in the form of $y = ax^2 + bx$.

![Figure 27 – Different graphical behavior of quadratic function of $f(x) = ax^2 + bx$](image)

Knowing the possible shape of the function helps the researcher find the proper relationship.
When $a = 0.4$ and $b = 0.3$ (solid curve), the graph is very similar to that estimated using the
NLMIXED and shown in Figure 23 fitting the cumulative bin areas by quadratic function.

### 4.2.4 Model selection

In this chapter I first discussed model selection for the first variable, AADT/10,000. Then it is
shown how other independent variables are selected and introduced to expand the model.

#### 4.2.4.1 Model selection for one variable

When the candidate functions are estimated using NLMIXED, one of them is to be picked as the
optimized solution. To do this, as Figure 25 shows, NLMIXED provides a table called “fit
statistics” to compare different nonlinear mixed models. This table lists the final maximized
value of the log-likelihood function and information criterion including $-2$ times the log
likelihood, the Akaike information criterion (AIC), a small sample bias corrected version of AIC
(AICC), and the Bayesian information criterion (BIC). These criteria are defined as shown in Equation 14 (49).

\[
\begin{align*}
\text{AIC} &= -2L(\hat{\theta}) + 2p \\
\text{AICC} &= -2L(\hat{\theta}) + 2pn/(n - p - 1) \\
\text{BIC} &= -2L(\hat{\theta}) + p \log(s)
\end{align*}
\]

(Equation 14)

Where \(L(\hat{\theta})\) is the marginal log likelihood function, \(\hat{\theta}\) is the vector of parameter estimates, \(p\) is the number of parameters, \(n\) is the number of observations (108, in this study), and \(s\) is the number of subjects which in this study is the same as the number of observations.

The value \(-2\log\text{likelihood}\) is used to compare different models and the AIC, AICC, and BIC criteria are implemented to add a penalty for using too many parameters. The AIC, AICC, and BIC criteria are in the “smaller is better” form. The BIC was developed in 1978 by Schwarz (50) and penalizes the number of parameters stronger than AIC and AICC. As an example, for a model with 3 parameters and 108 observations, let’s assume the value of \(-2\log\text{likelihood}\) is 199.7. Then:

\[
\begin{align*}
\text{AIC} &= -2L(\hat{\theta}) + 6.0 = 205.7 \\
\text{AICC} &= -2L(\hat{\theta}) + 6.3 = 206.0 \\
\text{BIC} &= -2L(\hat{\theta}) + 14.1 = 213.8
\end{align*}
\]

(Equation 15)

This calculation shows that the BIC charges a penalty of 14.1 while AIC and AAIC charge 6.0 and 6.3, respectively. So the BIC was chosen for this study to emphasize that the result is practically useful.

**4.2.4.2 Model selection for more than one variable**

When a new variable was added into the current model, the new model selection process between the current model and the new model (including the new variable) was performed. Indeed two models are called nested if both contain the same terms but one has at least one additional term. For example model \(y_1 = x_1^2 + 2x_1\) is nested within \(y_2 = x_1^2 + 2x_1 + x_2^2\) (with a new variable). So in this case \(y_1\) is the nested model and \(y_2\) is an alternative model with a new
variable. The nested or “null” model is a special case of the new model. In statistics, a likelihood ratio test (LRT) is used to compare the fit of null and alternative models. Equations 16 and 17 show the computation of the LRT.

$$\text{LRT} = -2 \log \left( \frac{\text{Likelihood of null model}}{\text{Likelihood of alternative model}} \right)$$  \hspace{1cm} (Equation 16)

$$\text{LRT} = \left[ -2 \log(\text{Likelihood of null model}) \right] - \left[ -2 \log(\text{Likelihood of alternative model}) \right]$$  \hspace{1cm} (Equation 17)

This value will be compared with the chi-square value in which $n$ is the degree of freedom or the difference between numbers of parameters of two models. The null hypothesis is that all the parameters of the newly added variables are zero in which two models are identical. As an example, look back to the example of $y_1$ and $y_2$ above. Assuming that the $-2 \log$ likelihood of $y_1$ and $y_2$ are 123.6 and 118.7, respectively, the LRT shows a reduction of 4.9 ($123.6 - 118.7$). On the other hand the chi-square critical value with one degree of freedom (the difference between the number of parameters of $y_1$ and $y_2$) is 3.84 at level 0.05. Since the LRT statistic value is bigger than the chi-square, model $y_2$ is statistically better than the model $y_1$ and parameter $x_2$ should be introduced in modeling.

**4.2.4.3 Model expansion**

It was mentioned earlier that the desired model was formed by multiplying the functions by each other. After fitting the principal variable, AADT/10,000 by $f_1()$, the collision model should be expanded by adding more variables. As shown in Figure 20, I must decide the order in which the variables will be introduced and which variables will be included in modeling. Assume that the second variable to be introduced after AADT/10,000 is $x_2$ and its corresponding function is $f_2()$. As a result, the expected number of annual collisions is estimated as:

$$y = f_1(\text{AADT}/10,000) \times f_2(x_2)$$  \hspace{1cm} (Equation 18)

The preceding section described how to use the LRT to compare the models. The LRT was used to find the contribution of the new variable to the improvement of the log likelihood, if there is
any. To do this, the functions \( f() \) and \( f_2(x_2) \) were determined first. Then the newly defined function can be rephrased as:

\[
f_2(x_2) = \frac{y}{f_1(AADT/10,000)} \quad \text{(Equation 19)}
\]

Where \( f_2(x_2) \) is the ratio of the recorded annual number of collisions to the annual number of collisions predicted by the principal variable, AADT/10,000, as shown in Equation 20:

\[
f_2(x_2) = \frac{\text{Recorded annual number of collisions}}{\text{Predicted annual number of collisions}(x_1)} \quad \text{(Equation 20)}
\]

If this ratio is greater than 1, the initial model predicted too few collisions, and if it is below 1, the prediction was more than the actual number of collisions. Ideally this value needs to be 1 so that the recorded and predicted numbers of collisions are identical.

When the functional forms of the variables were chosen, parameters of the expanded model, \( f_1(AADT/10,000) \times f_2(x_2) \) were re-estimated using NLMIXED. Then the LRT method was used to compare the new model with the initial model to choose the variables to be introduced into the initial model. The high the LRT, the sooner the particular variable will be introduced.

The same procedure is repeated for the other variables as:

\[
y = f_1(AADT/10,000) \times f_2(x_2) \times ... \times f_{i-1}(x_{i-1}) \times f_i(x_i) \quad \text{(Equation 21)}
\]

\[
f_i(x_i) = \frac{y}{[f_1(AADT/10,000) \times f_2(x_2) \times ... \times f_{i-1}(x_{i-1})]} \text{, or} \quad \text{(Equation 22)}
\]

\[
f_i(x_i) = \frac{\text{Recorded annual number of collisions}}{\text{Predicted annual number of collisions}(x_1,x_2,...,x_{i-1})} \quad \text{(Equation 23)}
\]

### 4.3 Goodness of fit

A basic test of model quality is the residual test. A residual is defined as the difference between the annual number of recorded collisions and the annual number of predicted collisions. The residual graph shows how well the model was chosen and fits the data. In road safety applications this method doesn’t reveal much information to judge a model. Assuming a prediction collision model of \( f(x) = x^{1.7249} e^{-0.7486x} \), where \( x \) is the principal independent
variable, AADT/10,000, and \( f(x) \) is the annual access – related collisions. Figure 28 shows the residual plot using this model and the collected data information.

![Figure 28 – Residual plot between annual collected and predicted collisions](image)

In 1997 Hauer suggested that the cumulative residual (CURE) plot methodology reveals a better view of the model fit. To produce this plot, the residuals are determined for each variable value. Then they are sorted in the increasing order of the variable and residuals are accumulated. If the predicted model fits the data well in all ranges of the variable, the CURE plot oscillates around the horizontal axis, \( y = 0 \), and if it is all above or below \( y = 0 \), the model doesn’t fit the data properly. Depending on the parameters estimated, the CURE plot should also end close to \( y = 0 \). Figure 29 shows the CURE plot for the previous example model of \( f(x) = x^{1.7249} e^{-0.7486x} \).
Statistically the CURE plot method can be considered as a “random walk”. It means that limits are defined that the CURE plot should cross very rarely. If \( N \) is the total number of the data points and \( n \) is an integer between 1 and \( N \), assuming \( S(n) \) to be the cumulative residuals from 1 to \( n \), then \( \sigma^2(n) \) is the variance of the \( S(n) \). Hauer concluded that the probability density of the \( S(n) \) is normal with a mean of zero and variance of \( \sigma^* = \sigma^2(n) \left[ 1 - \frac{\sigma^2(n)}{\sigma^2(N)} \right] \). He also showed that for a well-fit model, the CURE plot will be contained within \( \pm 2\sigma^* \). Figure 29 shows that our example model’s CURE plot is fit well within the \( \pm 2\sigma^* \).
5. MODELING PROCESS

I used the methodology described in Chapter 4 to conduct a prediction model using the data that was collected in Chapter 3. As shown in Figure 30, this chapter described the adopted modeling methodology on developing access – related collision model starting from the initial model to the final model.

Initial Model
Section 5.1

Second Model
Section 5.2

Third Model
Section 5.3

Fourth Model
Section 5.4

Figure 30 – Modeling process overview

5.1 Initial model

As discussed before, the principal independent variable was AADT/10,000. The initial model was the relationship between annual number of access – related collisions and AADT/10,000. The scatterplots of this variable were shown previously in Figure 19 and the candidate functions that could fit it with the collisions were shown in Table 17. As discussed in the preceding chapter, the NLMIXED procedure in SAS® was used to estimate the parameters of the candidate functions. Table 18 shows a summary of some of these functions and their estimated parameters and the corresponding BIC criteria values. In the table $-2LL$ represents the $-2$ log likelihood value.
Table 18 – Initial model BIC values and estimated parameters

<table>
<thead>
<tr>
<th>#</th>
<th>Parameters</th>
<th>BIC</th>
<th>-2LL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$f_2(x_2) = \beta_{1-1} x_1 + \beta_{1-2} x_1^2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$\beta_{1-1} + \beta_{1-2} x_1$</td>
<td>0.081</td>
<td>0.291</td>
</tr>
<tr>
<td></td>
<td>$\beta_{1-1} + \beta_{1-2} x_1$</td>
<td>0.409</td>
<td>0.677</td>
</tr>
<tr>
<td>2</td>
<td>$\beta_{1-1} x_1^2 + \beta_{1-2} x_1^3$</td>
<td>0.498</td>
<td>-0.074</td>
</tr>
<tr>
<td></td>
<td>$\beta_{1-1} x_1^2 + \beta_{1-2} x_1^3$</td>
<td>0.319</td>
<td>-0.056</td>
</tr>
<tr>
<td>3</td>
<td>$\beta_{1-1} x_1^2 + \beta_{1-2} x_1^3$</td>
<td>0.606</td>
<td>-0.294</td>
</tr>
<tr>
<td>4</td>
<td>$\beta_{1-1} x_1 e^{\beta_{1-2} x_1}$</td>
<td>-4.891</td>
<td>-2.125</td>
</tr>
<tr>
<td>5</td>
<td>$\beta_{1-1} x_1 e^{\beta_{1-2} x_1}$</td>
<td>0.349</td>
<td>-125.87</td>
</tr>
<tr>
<td>6</td>
<td>$x_1 e^{\beta_{1-1} x_1} - 0.519$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>$x_1 e^{\beta_{1-1} x_1} - 0.519$</td>
<td>0.297</td>
<td>0.332</td>
</tr>
<tr>
<td>8*</td>
<td>$x_1 e^{\beta_{1-1} x_1}$</td>
<td>1.725</td>
<td>-0.749</td>
</tr>
<tr>
<td>9</td>
<td>$x_1 e^{\beta_{1-1} x_1}$</td>
<td>-0.061</td>
<td>0.565</td>
</tr>
<tr>
<td>10</td>
<td>$x_1 e^{\beta_{1-1} x_1}$</td>
<td>-0.519</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>$x_1 e^{\beta_{1-1} x_1}$</td>
<td>0.297</td>
<td>0.332</td>
</tr>
</tbody>
</table>

* Selected model

This result shows that function number 8 was the best fit to predict annual access – related collisions based on AADT/10,000 as shown in Equation 24:

$$y = f_1(x_1) = (\text{AADT}/10,000)^{1.725} e^{-0.749(\text{AADT}/10,000)} \quad \text{(Equation 24)}$$

This was because its BIC value was the lowest among the candidates. The model dispersion parameter, $k$, was estimated as 0.495. It’s important to mention that I tried fitting a few other functional forms that could fulfill the EIF form such as $x_1^{\beta_1} + e^{\beta_1 x_1}$, $\beta_0 x_1^3 + \beta_1 x_1^4$, and other forms. However, none of them provided any better BIC value. Some of them were not even close to those in Table 18 in terms of BIC value.

The difference between two model’s BIC values and their corresponding $-2 \text{ LL}$ values were identical for all the above models. This was because they all had 3 parameters including the dispersion parameter estimated in the NLMIXED procedure as shown in Figure 25. This revealed that the $-2 \text{ LL}$ can also be used to pick the best model among the other models.

Next I tried fitting the selected model’s CURE plot within $\pm 2\sigma^*$ as discussed in previous chapter. The CURE plot of the above selected model and the $\pm 2\sigma^*$ were shown previously in Figure 29. The plot proved that this model fitted the data well. To compare how this model
predicted annual collisions, Figure 31 displayed and compared the annual collisions derived from the first model and the collected annual collisions.

The predicted collision scatterplot shows that the annual access – related collisions increased when major road traffic volume increased up to about 25,000 vehicles per day. Then, annual collisions decreased above 25,000 vehicles per day. This decrease could be because drivers pay more attention and slow down more on busier roads.

5.2 Second model

The second model was the relationship between the dependent variable, annual access – related collisions, and two independent variables – AADT/10,000 and a secondary variable. As discussed earlier, the second variable was multiplied by the principal variable. The annual access – related collisions were predicted in this stage as:

\[
y = f_1(\text{AADT}/10,000) \times f_2(x_2) \quad \text{(Equation 25)}
\]

Where,

\[
f_1(\text{AADT}/10,000) = (\text{AADT}/10,000)^{\beta_1} e^{-\beta_1(AADT/10,000)} \quad \text{(Equation 26)}
\]
\( x_2 \) is the second selected variable, and \( f_2() \) is the corresponding function that had parameters estimated using the NLMIXED Procedure per Equation 14. Table 19 shows the selected models for each variable and the estimated parameters. Even though the parameters \( \beta_{1-1} \) and \( \beta_{1-2} \) were already estimated, they were re-estimated every time a new variable was introduced.

Table 19 – Second model estimated parameters

<table>
<thead>
<tr>
<th>( x^* )</th>
<th>( \beta_{1-1} )</th>
<th>( \beta_{1-2} )</th>
<th>( \beta_{2-1} )</th>
<th>( \beta_{2-2} )</th>
<th>( \beta_{2-3} )</th>
<th>( \beta_{2-4} )</th>
<th>( \beta_{2-5} )</th>
<th>( \beta_{2-6} )</th>
<th>( \beta_{2-7} )</th>
<th>(-2\text{LL})</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV</td>
<td>2.304</td>
<td>-1.173</td>
<td>0.008</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>194.3</td>
</tr>
<tr>
<td>CC</td>
<td>1.606</td>
<td>-0.610</td>
<td>-0.001</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>199.0</td>
</tr>
<tr>
<td>Q-T</td>
<td>2.009</td>
<td>-1.145</td>
<td>0.002</td>
<td>195.3</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>195.3</td>
</tr>
<tr>
<td>DW</td>
<td>1.436</td>
<td>-0.474</td>
<td>0.002</td>
<td>1.722</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>189.6</td>
</tr>
<tr>
<td>LCI</td>
<td>1.303</td>
<td>-0.417</td>
<td>0.561</td>
<td>0.431</td>
<td>1.292</td>
<td>0.231</td>
<td>0.713</td>
<td>1.465</td>
<td>0.538</td>
<td>191.5</td>
</tr>
<tr>
<td>LC</td>
<td>1.481</td>
<td>-0.578</td>
<td>0.762</td>
<td>0.523</td>
<td>0.870</td>
<td>1.320</td>
<td>0.468</td>
<td>0.973</td>
<td>---</td>
<td>195.6</td>
</tr>
<tr>
<td>GR</td>
<td>1.443</td>
<td>-0.579</td>
<td>0.467</td>
<td>0.908</td>
<td>0.379</td>
<td>0.717</td>
<td>0.974</td>
<td>1.239</td>
<td>0.366</td>
<td>192.4</td>
</tr>
<tr>
<td>SL</td>
<td>1.865</td>
<td>-0.674</td>
<td>1.148</td>
<td>0.614</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>196.2</td>
</tr>
<tr>
<td>MM</td>
<td>1.743</td>
<td>-0.685</td>
<td>0.507</td>
<td>1.064</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>195.3</td>
</tr>
<tr>
<td>DM</td>
<td>1.609</td>
<td>-0.657</td>
<td>0.891</td>
<td>0.860</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>199.7</td>
</tr>
<tr>
<td>DA</td>
<td>1.615</td>
<td>-0.680</td>
<td>0.327</td>
<td>0.989</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>196.9</td>
</tr>
<tr>
<td>RA</td>
<td>1.603</td>
<td>-0.670</td>
<td>0.815</td>
<td>0.984</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>199.3</td>
</tr>
<tr>
<td>TR</td>
<td>1.580</td>
<td>-0.636</td>
<td>0.896</td>
<td>0.702</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>199.4</td>
</tr>
<tr>
<td>AD</td>
<td>1.691</td>
<td>-0.714</td>
<td>0.874</td>
<td>1.463</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>198.1</td>
</tr>
</tbody>
</table>

The process of model selection for each variable was described in the following sections. Also in these sections, all the variables were tested to determine which one was selected as the secondary variable. First the continuous variables, such as driveway volume, and then categorical (discrete) variables, like lane configuration at the intersection, were introduced. Lastly binary variables like speed limit were tested for possible inclusion in the model.

5.2.1 Driveway volume

In this section the driveway traffic volume was considered as the second variable. Figure 32 shows the scatterplots of \( f_2(x_2) \) and driveway traffic volume.
The scatterplots were not really helpful in finding a fit among data, so the EIF method was used as shown in Figure 33.

Table 20 presents a summary of the candidate functions considered to be a good fit for the driveway traffic volume function. The BIC criteria value of the second function in Table 20, \( f_2(x_2) = 1 + \beta_2 x_2 \), has the lowest BIC of the candidate functions and was chosen as the best
functional form. Table 19 showed the parameter’s estimation and the $-2\text{LL}$ result for this function.

Table 20 – $f_2(x_2)$ BIC values for the driveway traffic volume

<table>
<thead>
<tr>
<th>#</th>
<th>$f_2(x_2)$</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1 + \beta_{2-1}\log(x_2)$</td>
<td>216.8</td>
</tr>
<tr>
<td>2*</td>
<td>$1 + \beta_{2-1}x_2$</td>
<td>213.1</td>
</tr>
<tr>
<td>3</td>
<td>$\beta_{2-1} + \beta_{2-3}x_2$</td>
<td>216.5</td>
</tr>
<tr>
<td>4</td>
<td>$x_2\beta_{2-1}$</td>
<td>216.0</td>
</tr>
<tr>
<td>5</td>
<td>$\beta_{2-1}x_2\beta_{2-2}$</td>
<td>216.4</td>
</tr>
<tr>
<td>6</td>
<td>$e^{\beta_{2-1}x_2}$</td>
<td>213.4</td>
</tr>
<tr>
<td>7</td>
<td>$\beta_{2-1}e^{\beta_{2-2}x_2}$</td>
<td>217.4</td>
</tr>
<tr>
<td>8</td>
<td>$x_2\beta_{2-1}\beta_{2-2}x_2$</td>
<td>218.0</td>
</tr>
</tbody>
</table>

* Selected model

The LRT statistic for this model, which is the $-2\text{LL}$ value for AADT model alone minus this $-2\text{LL}$ value, was 5.4 ($199.7 - 194.3$). The chi-square critical value with one degree of freedom was 3.84 with a p value of 0.02 at level 0.05. Since the LRT was larger than the chi-square, this model was preferred over the initial model.

The traffic volume entering and exiting the access point were also available in this study. As discussed before, some of these volumes were collected on site and some of them were estimated using the Trip Generation Manual. I used the same method shown above to test the behavior of these two variables. Table 21 shows that using these variables instead of the total driveway traffic volume reduced the $-2\text{LL}$ value of the chosen model very slightly.

Table 21 – BIC comparison between driveway entering, exiting, and total traffic

<table>
<thead>
<tr>
<th>#</th>
<th>$f_2(x_2)$</th>
<th>Total Driveway</th>
<th>Driveway Enter</th>
<th>Driveway Exit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1 + \beta_{2-1}\log(x_2)$</td>
<td>216.8</td>
<td>215.8</td>
<td>215.8</td>
</tr>
<tr>
<td>2*</td>
<td>$1 + \beta_{2-1}x_2$</td>
<td><strong>213.1</strong></td>
<td><strong>212.7</strong></td>
<td><strong>213.6</strong></td>
</tr>
<tr>
<td>3</td>
<td>$\beta_{2-1} + \beta_{2-3}x_2$</td>
<td>216.5</td>
<td>216.1</td>
<td>217.3</td>
</tr>
<tr>
<td>4</td>
<td>$x_2\beta_{2-1}$</td>
<td>216.0</td>
<td>214.9</td>
<td>215.0</td>
</tr>
<tr>
<td>5</td>
<td>$\beta_{2-1}x_2\beta_{2-2}$</td>
<td>216.4</td>
<td>215.8</td>
<td>216.4</td>
</tr>
<tr>
<td>6</td>
<td>$e^{\beta_{2-1}x_2}$</td>
<td>213.4</td>
<td>213.2</td>
<td>214.4</td>
</tr>
<tr>
<td>7</td>
<td>$\beta_{2-1}e^{\beta_{2-2}x_2}$</td>
<td>217.4</td>
<td>217.2</td>
<td>218.7</td>
</tr>
<tr>
<td>8</td>
<td>$x_2\beta_{2-1}\beta_{2-2}x_2$</td>
<td>218.0</td>
<td>217.9</td>
<td>219.0</td>
</tr>
</tbody>
</table>

* Selected model

As a result I decided to keep the total traffic volume on driveway as an independent variable and ignore the entering and exiting traffic volumes.
5.2.2 Corner clearance

Assuming the selected second variable as the corner clearance, \( f_2(x_2) \) was estimated using Equation 20. Figure 34 shows scatterplots of the \( f_2(x_2) \) and the corner clearance.

![Figure 34](image)

Figure 34 – \( f_2(x_2) \) when \( x_2 \) is the corner clearance

Since it was not possible to find a good model fitting these data points, I used the EIF method, shown in Figure 35.

![Figure 35](image)

Figure 35 – EIF of \( f_2(x_2) \) when \( x_2 \) is the corner clearance
The functions shown in Table 22 were considered to be relatively good fits for the integral function.

Table 22 – $f_2(x_2)$ BIC values for the corner clearance

<table>
<thead>
<tr>
<th>#</th>
<th>$f_2(x_2)$</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1 + \beta_{2-1}\log(x_2)$</td>
<td>218.4</td>
</tr>
<tr>
<td>2</td>
<td>$1 + \beta_{2-1}x_2$</td>
<td><strong>217.8</strong></td>
</tr>
<tr>
<td>3</td>
<td>$x_2^\beta_{2-1}$</td>
<td>218.4</td>
</tr>
<tr>
<td>4</td>
<td>$\beta_{2-1}x_2^\beta_{2-2}$</td>
<td>223.1</td>
</tr>
<tr>
<td>5</td>
<td>$e^{\beta_{2-1}x_2}$</td>
<td>218.0</td>
</tr>
<tr>
<td>6</td>
<td>$1 + \beta_{2-1}x_2e^{\beta_{2-2}x_2}$</td>
<td>223.1</td>
</tr>
<tr>
<td>7</td>
<td>$\beta_{2-1}x_2 + e^{\beta_{2-2}x_2}$</td>
<td>222.4</td>
</tr>
<tr>
<td>8</td>
<td>$x_2^\beta_{2-1}e^{\beta_{2-2}x_2}$</td>
<td>222.4</td>
</tr>
</tbody>
</table>

* Selected model

Since the BIC criteria value of the function $f_2(x_2) = 1 + \beta_{2-1}x_2$ was the lowest of the candidate functions, it was chosen as the best function form. Table 19 displayed the parameters estimation values of this function. The LRT statistic for this model was 0.7 (199.7 – 199.0). The chi-square critical value with one degree of freedom was 3.84 at level 0.05, and the p value was 0.40. Since LRT was much smaller than chi-square, this model was not statistically better than the initial model, but since the $-2\, LL$ was improved, the corner clearance was retained for possible introduction into the model.

Since the corner clearance was one of the variables of interest in this research, I tried a different methodology to test it. Even though the corner clearance is a continuous variable, it was also tested as a categorical variable as shown in Figure 12, ranging from 0-49 feet, 50-99 feet, ..., 750-800 feet, but the result did not provide a significant LRT statistic. I also tested the interaction of this variable with a few statistically significant variables but the results were not satisfactory.

5.2.3 Synchro queue length

Figure 36, Figure 37, and Figure 38 show the scatterplots of the estimated $f_2(x_2)$ for Synchro queue lengths in the left, through, and right lanes, respectively, using Equation 20. These were 95th percentile queues during the peak hour estimated using Synchro.
Figure 36 – $f_2(x_2)$ when $x_2$ is the left turn lane Synchro queue

Figure 37 – $f_2(x_2)$ when $x_2$ is the through lane Synchro queue
Figure 38 – $f_2(x_2)$ when $x_2$ is the right turn lane Synchro queue

The scatterplots were not helpful in finding a function with a good fit. Therefore, the EIF method was used as shown in Figure 39, Figure 40, and Figure 41 for left, through, and right lanes, respectively.

Figure 39 – EIF of $f_2(x_2)$ when $x_2$ is the Synchro queue for left movement
As a result the only function forms that could plausibly be chosen are shown in Table 23. This table also shows the corresponding BIC and $-2LL$ values for each candidate model.
Table 23 – Candidate models of the left, through, and right turns’ Synchro queues

<table>
<thead>
<tr>
<th>#</th>
<th>(f_2(x_2))</th>
<th>Q_{Left}</th>
<th>Q_{Through}</th>
<th>Q_{Right}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>BIC</td>
<td>-2LL</td>
<td>BIC</td>
</tr>
<tr>
<td>1*</td>
<td>(1 + \beta_{2-1}x_2)</td>
<td>217.4</td>
<td>198.7</td>
<td>214.0</td>
</tr>
<tr>
<td>2</td>
<td>(e^{\beta_{2-1}x_2})</td>
<td>217.4</td>
<td>198.7</td>
<td>214.0</td>
</tr>
<tr>
<td>3</td>
<td>(\beta_{2-1} + \beta_{2-2}x_2)</td>
<td>221.8</td>
<td>198.4</td>
<td>216.5</td>
</tr>
<tr>
<td>4</td>
<td>(1 + \beta_{2-1}x_2e^{\beta_{2-2}x_2})</td>
<td>218.6</td>
<td>195.2</td>
<td>227.2</td>
</tr>
<tr>
<td>5</td>
<td>(\beta_{2}x_2 + e^{\beta_{3}x_2})</td>
<td>222.1</td>
<td>198.7</td>
<td>217.4</td>
</tr>
</tbody>
</table>

* Selected model

Table 23 shows that the models 1 and 2 have the lowest BIC values for left, through, and right movements. I decided to choose the linear model, number 1, for its simplicity. Recall that the initial model \(-2\) LL was 199.7 and the chi-square was 3.84 with one degree of freedom at level 0.05. So the LRT statistics for left, through, and right movements are 1.0, 4.4, and 0.0, respectively. Thus if the Synchro queue was considered as the second variable, the only model that was statistically better than the initial model was the model including the intersection through movement Synchro queue. Table 19 showed the estimated parameters of this variable using the NLMIXED procedure.

5.2.4 Driveway width

The driveway width was measured in feet and was shown in Table 13. The \(f_2(x_2)\) value from Equation 20 was estimated and shown in Figure 42.
It was difficult to determine what model can be fit from Figure 42, so the EIF was used as shown in Figure 43.

Many function forms were candidates to fit the integral function. Some of these functions are shown in Table 24.
Table 24 – \( f_2(x_2) \) BIC values for the driveway width

<table>
<thead>
<tr>
<th>#</th>
<th>( f_2(x_2) )</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( 1 + \beta_{2-1} \log(x_2) )</td>
<td>218.1</td>
</tr>
<tr>
<td>2</td>
<td>( 1 + \beta_{2-1}x_2 )</td>
<td>215.5</td>
</tr>
<tr>
<td>3</td>
<td>( x_2^\beta_{2-1} )</td>
<td>217.9</td>
</tr>
<tr>
<td>4*</td>
<td>( \beta_{2-1}x_2^{\beta_{2-2}} )</td>
<td>213.0</td>
</tr>
<tr>
<td>5</td>
<td>( \beta_{2-1}x_2 )</td>
<td>214.1</td>
</tr>
<tr>
<td>6</td>
<td>( x_2^{\beta_{2-1}}e^{\beta_{2-2}x_2} )</td>
<td>213.9</td>
</tr>
<tr>
<td>7</td>
<td>( \beta_{2-1}x_2 + \beta_{2-2}x_2 )</td>
<td>214.8</td>
</tr>
<tr>
<td>8</td>
<td>( 1 + \beta_{2-1}x_2 + \beta_{2-2}x_2^2 )</td>
<td>213.5</td>
</tr>
<tr>
<td>9</td>
<td>( 1 + \beta_{2-1}x_2^2 + \beta_{2-2}x_2^3 )</td>
<td>215.4</td>
</tr>
</tbody>
</table>

* Selected model

The BIC value was the lowest for model number 4 with a function form \( f_2(x_2) = \beta_{2-1}x_2^{\beta_{2-2}} \). The model estimated parameters were shown in Table 19. The LRT statistic of this model was 10.1 (199.7 – 189.6). Since this value was higher than the chi-square critical value with two degrees of freedom (5.99) at level 0.05, this model was statistically preferred over the initial model.

5.2.5 Lane configuration at the intersection

As was shown in Table 7, lane configuration at the intersection near the access point was a categorical (class) variable with 7 possible configurations. Section 4.2.1 described how to model these type of variables. Thus the function of the lane configuration at the intersection can be expressed as:

\[ f_2(x_2) = \beta_{2-1}(x_2 = 1) + \beta_{2-2}(x_2 = 2) + \cdots + \beta_{2-7}(x_2 = 7) \]  

(Equation 27)

Table 19 showed the parameter’s estimated result for this model. The LRT statistic was 8.2 (= 199.7 – 191.5) that was less than chi-square with seven degrees of freedom (14.07) at level 0.05. Thus this variable was not statistically preferred over the initial model but since it contributed to the reduction of the \(-2\text{LL}\), it was retained for the possible consideration in modeling.
5.2.6 Lane configuration at the driveway

Lane configuration on the major road at the driveway was shown in Table 3 with 6 possible configurations. As discussed before, the function of this categorical variable was expressed as:

\[ f_2(x_2) = \beta_{2-1}(x_2 = 1) + \beta_{2-2}(x_2 = 2) + \cdots + \beta_{2-6}(x_2 = 6) \]  
(Equation 28)

Table 19 presented the parameter’s estimated result. The LRT statistic for the lane configuration at the driveway was 4.1 (= 199.7 – 195.6) which was less than the chi-square for six degrees of freedom (12.59) at level 0.05. Therefore, statistically the function with new variable of LCD was not better than the initial model, but since it reduced the −2LL, it was retained for the further analysis and for possible later introduction.

5.2.7 Grade on the major road

Grade on the major road was shown in Table 10. One can see that this was also a categorical variable with possible values of −3, −2, ..., +2. Thus its functional form was:

\[ f_2(x_2) = \beta_{2-1}(x_2 = -3) + \beta_{2-2}(x_2 = -2) + \cdots + \beta_{2-7}(x_2 = +3) \]  
(Equation 29)

Using the NLMIXED procedure the parameters were estimated as shown in Table 19. The LRT statistic was 7.3 (199.7 – 192.4) and the chi-square for seven degrees of freedom was 14.07 at level 0.05. Since the LRT statistic was less than the chi-square, the function did not statistically improve the initial model, but because of the −2LL reduction, it was retained for the possible later introduction.

5.2.8 Speed limit

The speed limit values in this research were shown in Table 9. About 48% of the 108 studied sites operated with a speed limit of 35 mph or less and 52% of the sites operate with 45 mph or more. The function of this variable was shown as:

\[ f_2(x_2) = \beta_{2-1}(x_2 \leq 35) + \beta_{2-2}(x_2 \geq 45) \]  
(Equation 30)
The estimated parameters were shown in Table 19. It was noticeable that since $\beta_{2-1} > 1$ and $\beta_{2-2} < 1$, if all other variables were the same, sites with higher than 45 mph speed limit create fewer collisions than sites with below 35 mph speed limit. This is different from the public expectation that higher speed limits are less safe than lower speed limit sites. Indeed this could be the result of implementing proper speed limits by state DOT engineers, meaning they set (or increase) a higher speed limit for safer roads. The calculated value of the LRT statistic was 3.5 (199.7 – 196.2) which was close to the chi-square critical value with two degrees of freedom (5.99) at level 0.05 so this model could be considered better than the initial model. The p value was 0.06.

### 5.2.9 Major road median

Median type on the major road was the next variable to be considered. As mentioned earlier, the only median types studied in this research were full movement and RIRO. The function of this variable was defined as:

$$f_2(x_2) = \beta_{2-1}(x_2 = \text{RIRO}) + \beta_{2-2}(x_2 = \text{Full})$$

(Equation 31)

Table 19 presented the parameter estimation of the median type function. One can see that assuming all else is the same, a RIRO median reduces the collisions almost by 50% compared to the full movement median. The LRT statistic was 4.4 (199.7 – 195.3) which was less than the chi-square critical value with two degrees of freedom (5.99) at level 0.05. The p value was 0.1. Thus this model was not statistically better than the initial model, but because the $-2LL$ value was reduced, median type variable was retained for possible introduction into the model.

### 5.2.10 Driveway median

The next variable expressed whether the driveway was divided or not. Since this was also a binary variable, its function was expressed as:

$$f_2(x_2) = \beta_{2-1}(x_2 = \text{Undivided}) + \beta_{2-2}(x_2 = \text{Divided})$$

(Equation 32)
Table 19 showed the result of the model estimation. The calculated $-2 \text{LL}$ for this model was not reduced in comparison with that of the initial model. This means even if there was any influence of the driveway median status, it was not noticeable using this model. As a result this variable was eliminated from the candidate variables list.

### 5.2.11 Driveway angle

The driveway angle was shown in Figure 14 and Table 11. After consideration of the driveway angle and the possible statistical modeling of this variable I decided to group the 108 studied sites as shown in Table 25.

<table>
<thead>
<tr>
<th>#</th>
<th>Driveway Angle</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DA &lt; 90</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>DA $\geq$ 90</td>
<td>96</td>
</tr>
</tbody>
</table>

Therefore, this binary variable’s function was expressed as:

$$f_2(x_2) = \beta_{2-1}(x_2 < 90) + \beta_{2-2}(x_2 \geq 90) \quad \text{(Equation 33)}$$

The result of parameter estimation of this grouped variable using NLMIXED was shown in Table 19. This model’s LRT was $2.8 \ (199.7 - 196.9)$. This value was less than the chi-square critical value with two degrees of freedom (5.99) at level 0.05 and the model was not statistically better than the initial model, but it was still retained because it made some contribution to the reduction of the $-2 \text{LL}$.

### 5.2.12 Driveway radius

Driveway radius was categorized as large or small between the major road and driveway and was shown in Figure 17. Thus the function form was expressed as:

$$f_2(x_2) = \beta_{2-1}(x_2 = \text{Small}) + \beta_{2-2}(x_2 = \text{Large}) \quad \text{(Equation 34)}$$

Table 19 showed the result of the model estimation. The LRT statistic for this model was $0.4 \ (199.7 - 199.3)$ which was less than the chi-square critical value with two degrees of
freedom (5.99) at level 0.05. The p value was 0.58. Therefore, this model was not statistically better than the initial model. Although the effect of this variable seemed very small, there was still some contribution from it to reducing the $- 2 \text{ LL}$, so it was not eliminated from the modeling yet.

5.2.13 Transition

This binary variable is the transition between the major road and the driveway and was shown in Table 14. The functional form of this variable was expressed as:

$$f_2(x_2) = \beta_{2-1}(x_2 = \text{Not elevated}) + \beta_{2-2}(x_2 = \text{Elevated})$$  \hspace{1cm} (Equation 35)

Table 19 showed the result of the model estimation. The LRT statistic for this model was 0.3 (199.7 – 199.4) which was less than the chi-square critical value with two degrees of freedom (5.99) at level 0.05. The p value was 0.58. Therefore, this model was not statistically better than the initial model but since there was still some contribution of this variable to reducing the $- 2 \text{ LL}$, it was retained for possible introduction into the model.

5.2.14 Second driveway

This binary variable, which is related to the existence of more than one driveway within 150 ft from the main access point, was shown in Table 16. Within the said distance if the same parcel has more than one access point, this variable was denoted a “Yes”, with a “No” value for when there was not a secondary driveway. So the function of this variable was:

$$f_2(x_2) = \beta_{2-1}(x_2 = \text{No}) + \beta_{2-2}(x_2 = \text{Yes})$$  \hspace{1cm} (Equation 36)

Table 19 presented the parameter estimation result and showed that a site with more than one driveway creates higher collisions than a driveway with only one driveway. The LRT statistic for this model was 1.6 (199.7 – 198.1) but that was less than the chi-square critical value with two degrees of freedom (5.99) at level 0.05. The p value was 0.22. This model was not statistically better than the initial model. Since there was some contribution of this variable to reducing the $- 2 \text{ LL}$, it was retained for possible introduction into the model.
5.2.15 Result

Now that all the variables were considered one by one, Table 26 summarizes the order in which the variables will be introduced to the model. In this table “MV” is the major road traffic volume, AADT/10,000, in vehicles per day.

Table 26 – Summary of the second variable result

<table>
<thead>
<tr>
<th>Order</th>
<th>Variable</th>
<th>−2LL</th>
<th>LRT</th>
<th>Chi-square</th>
<th>Select</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MV</td>
<td>199.7</td>
<td>---</td>
<td>---</td>
<td>Yes</td>
</tr>
<tr>
<td>2*</td>
<td>DW</td>
<td>189.6</td>
<td>10.1</td>
<td>5.99</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>LCI</td>
<td>191.5</td>
<td>8.2</td>
<td>14.07</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>GR</td>
<td>192.4</td>
<td>7.3</td>
<td>14.07</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>DV</td>
<td>194.3</td>
<td>5.4</td>
<td>3.84</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>Q-T</td>
<td>195.3</td>
<td>4.4</td>
<td>3.84</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>MM</td>
<td>195.3</td>
<td>4.4</td>
<td>5.99</td>
<td>Yes</td>
</tr>
<tr>
<td>8</td>
<td>LCD</td>
<td>195.6</td>
<td>4.1</td>
<td>12.59</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>SL</td>
<td>196.2</td>
<td>3.5</td>
<td>5.99</td>
<td>Yes</td>
</tr>
<tr>
<td>10</td>
<td>DA</td>
<td>196.9</td>
<td>2.8</td>
<td>5.99</td>
<td>Yes</td>
</tr>
<tr>
<td>11</td>
<td>AD</td>
<td>198.1</td>
<td>1.6</td>
<td>5.99</td>
<td>Yes</td>
</tr>
<tr>
<td>12</td>
<td>CC</td>
<td>199.0</td>
<td>0.7</td>
<td>3.84</td>
<td>Yes</td>
</tr>
<tr>
<td>13</td>
<td>RA</td>
<td>199.3</td>
<td>0.4</td>
<td>5.99</td>
<td>Yes</td>
</tr>
<tr>
<td>14</td>
<td>TR</td>
<td>199.4</td>
<td>0.3</td>
<td>5.99</td>
<td>Yes</td>
</tr>
<tr>
<td>15</td>
<td>DM</td>
<td>199.7</td>
<td>0.0</td>
<td>5.99</td>
<td>No</td>
</tr>
</tbody>
</table>

* Selected second variable

Table 26 shows that all the variables except the driveway median status were retained for possible consideration in modeling. Higher LRT values in the table suggested that a variable should be introduced in modeling sooner. The driveway width had the highest LRT statistic and thus was selected as the second variable to be introduced. The final form of the model with AADT/10,000 and the driveway width was expressed as:

\[ y = x_1^{\beta_1-1} e^{\beta_1 x_1} (\beta_2-1 x_2^{\beta_2-2}) = x_1^{1.436} e^{-0.474 x_1} (0.002 x_2^{1.722}) \]  

(Equation 37)

Where

\[ y = \text{Annual access – related collisions} \]
\[ x_1 = \text{AADT/10,000 in vehicles per day} \]
\[ x_2 = \text{Driveway width in feet} \]
The model dispersion parameter was estimated as 0.346. The functional form of the second variable, driveway width, has a polynomial form, presented as \( f_2(x_2) = 0.002 \cdot x_2^{1.722} \), and was chosen from Table 24. Driveway width ranges between 8 ft and 58 ft in this study data set. Within this range \( f_2(x_2) \) function would increase if the driveway width increases.

From a traffic engineering point of view, I expected that the access – related number of collisions to be higher for small driveway widths, then to decrease when driveway width increases and at some point to bounce back and increase. This is considering the fact that drivers have difficulties getting in and out of the driveway when the driveway is small but the conditions would be improved for moderate driveway widths and then would get bad again for larger driveways because drivers may think it is safe to drive faster. This expectation was not fulfilled by the predicted model. It is hard to say whether the traffic engineering expectation or the provided model is accurate because in reality there are other factors that could affect drivers’ behavior.

The CURE plot of this variable and the \( \pm 2\sigma^* \) shown in Figure 44 illustrates the accuracy of this variable and the estimated parameters. The function started and ended near \( y = 0 \) and oscillated within \( \pm 2\sigma^* \). This confirmed that the suggested model was preferred over the first model.

![CURE plot](image)

Figure 44 – CURE plot fitted in \( +2\sigma^* \) and \( -2\sigma^* \) for second model

Figure 45 shows the annual access – related collisions vs. AADT/10,000 scatterplots of the collected data and the projected data using the second variable. One can see that the predicted
values were getting closer to matching the actual observations and definitely better than the first model scatterplots shown in Figure 31.

![Figure 45 – Comparison between annual collected and the second model predicted collisions](image)

One can see in Figure 45 that, as for the initial model, the annual access – related collisions increased when major road traffic volume increased up to about 25,000 vehicles per day. Then, annual collisions decreased above 25,000 vehicles per day. This decrease could be because drivers pay more attention and slow down more on busier roads, or could be because of some factor at the intersection I did not consider.

### 5.3 Third model

The function form with the third variable was set to be:

\[
y = f_1(AADT/10,000) \times f_2(DW) \times f_3(x_3)
\]  

(Equation 38)

In this section different efforts to find an appropriate third variable are described. The order in which variables are introduced as the third variable was shown in Table 26. Table 27 presents the selected models for each variable and the estimated parameters. The parameters \( \beta_{1-1}, \beta_{1-2}, \beta_{2-1}, \) and \( \beta_{2-2} \) were re-estimated when introducing the new variables.
Table 27 – Third model estimated parameters

<table>
<thead>
<tr>
<th></th>
<th>$\beta_{1-1}$</th>
<th>$\beta_{1-2}$</th>
<th>$\beta_{2-1}$</th>
<th>$\beta_{2-2}$</th>
<th>$\beta_{3-1}$</th>
<th>$\beta_{3-2}$</th>
<th>$\beta_{3-3}$</th>
<th>$\beta_{3-4}$</th>
<th>$\beta_{3-5}$</th>
<th>$\beta_{3-6}$</th>
<th>$\beta_{3-7}$</th>
<th>$-2\text{LL}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCI</td>
<td>1.111</td>
<td>-0.269</td>
<td>0.005</td>
<td>1.416</td>
<td>0.841</td>
<td>0.666</td>
<td>1.423</td>
<td>0.317</td>
<td>0.898</td>
<td>1.345</td>
<td>0.835</td>
<td>185.2</td>
</tr>
<tr>
<td>GR</td>
<td>1.183</td>
<td>-0.322</td>
<td>0.003</td>
<td>1.589</td>
<td>0.654</td>
<td>1.053</td>
<td>0.485</td>
<td>0.746</td>
<td>1.292</td>
<td>1.416</td>
<td>0.501</td>
<td>183.5</td>
</tr>
<tr>
<td>DV</td>
<td>1.549</td>
<td>-0.555</td>
<td>0.002</td>
<td>1.433</td>
<td>0.305</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>186.2</td>
</tr>
<tr>
<td>Q-T</td>
<td>0.933</td>
<td>-0.375</td>
<td>0.0004</td>
<td>1.841</td>
<td>0.009</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>181.3</td>
</tr>
</tbody>
</table>

In the following sections the variables in the order shown in Table 26 will be tested to decide which one will be selected as the third variable.

5.3.1 Lane configuration at the intersection

In third model as shown in Table 27, lane configuration at the intersection was introduced. As before when a new variable was introduced, all the parameters were re-estimated. As discussed earlier, the lane configuration at the intersection near the access point was a categorical variable with 7 possible forms as shown in Table 7. Since lane configuration at the intersection is a categorical variable, as shown in Table 7, the function was formed as:

$$f_3(x_3) = \beta_{3-1}(x_3 = 1) + \beta_{3-2}(x_3 = 2) + \cdots + \beta_{3-7}(x_3 = 7)$$  \hspace{1cm} (Equation 39)

Table 27 showed the estimation result of this variable and re-estimation of the previous variables using NLMIXED procedure. Since the previous $-2\text{LL}$ value for the DW was 189.6, the LRT statistic for this variable was estimated as 4.4 ($189.6 - 185.2$). The chi-square value for seven degrees of freedom was 14.07 at level 0.05. The LRT was much smaller than the chi-square which means that the new model was not statistically better than the previous model, therefore it was eliminated from introduction in modeling. A comparison of the scatterplots of the annual recorded and predicted collision data also confirmed that this model did not improve the prediction of the overall collision data.

5.3.2 Grade on the major road

The next candidate as the third variable was grade on the major road. As discussed before, this variable was shown in Table 10, with seven possible values. The function form was expressed as:

$$f_3(x_3) = \beta_{3-1}(x_3 = -3) + \beta_{3-2}(x_3 = -2) + \cdots + \beta_{3-7}(x_3 = +3)$$  \hspace{1cm} (Equation 40)
The parameter’s estimation was illustrated in Table 27. The LRT statistic was 6.1 (189.6 – 183.5) that was less that the chi-square of 14.07, for seven degrees of freedom at level 0.05. Thus this variable was also eliminated from modeling because it was not statistically better than the model with two variables of AADT/10,000 and the driveway width.

5.3.3 Driveway volume

The next variable in order was the driveway traffic volume. The scatterplots in Figure 46 show the estimated \( f_3(x_3) \) from Equation 23.

![Figure 46 – \( f_3(x_3) \) when \( x_3 \) is the driveway traffic volume](image)

As before, it was not possible to predict what model would fit well among these data points. The EIF was used as shown in Figure 47.
Figure 47 – EIF of $f_3(x_3)$ when $x_3$ is the driveway traffic volume

The function forms in Table 28 were the possible candidates to fit the integral function.

### Table 28 – $f_3(x_3)$ BIC values for the driveway volume

<table>
<thead>
<tr>
<th>#</th>
<th>$f_3(x_3)$</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1 + \beta_{3.1}x_3$</td>
<td>214.8</td>
</tr>
<tr>
<td>2*</td>
<td>$x_3^{\beta_{3.1}}$</td>
<td><strong>214.3</strong></td>
</tr>
<tr>
<td>3</td>
<td>$x_3^{\beta_{3.1}}x_2^{\beta_{3.2}}$</td>
<td>216.3</td>
</tr>
<tr>
<td>4</td>
<td>$x_3^{\beta_{3.1}}x_3^{\beta_{3.2}}$</td>
<td>218.9</td>
</tr>
<tr>
<td>5</td>
<td>$e^{\beta_{3.1}x_3}$</td>
<td>215.3</td>
</tr>
<tr>
<td>6</td>
<td>$\beta_{3.1}e^{\beta_{3.2}x_3}$</td>
<td>220.0</td>
</tr>
<tr>
<td>7</td>
<td>$\beta_{3.1}x_3^{\beta_{3.2}}$</td>
<td>219.0</td>
</tr>
</tbody>
</table>

* Selected model

One can see that the model number 2 had the minimum BIC value among above models. Previously Table 27 showed the result of the NLMIXED estimation for model parameters. Even though this model has reduces the $-2\text{LL}$, since the LRT statistic, 3.4 (189.6 - 186.2), was less than the chi-square of 3.84 for one degree of freedom at level 0.05, this model was not statistically better than the model with two variables of the AADT/10,000 and driveway lane width. Therefore, this variable was not introduced into the model.
5.3.4 Through movement Synchro queue

Table 26 showed that the variables through lane Synchro queue and major road median type had the same LRT values. Thus either one can be considered as the next variable to be introduced at this point. First I considered the through movement Synchro queue. Figure 48 shows the scatterplots of the estimated $f_3(x_3)$ vs. through movement Synchro queue using Equation 23.

![Figure 48 – $f_3(x_3)$ when $x_3$ is the through movement Synchro queue](image)

The scatterplot was again not helpful in finding a good fit among these data. Therefore, the EIF method was used as shown in Figure 49.
Figure 49 – EIF of $f_3(x_3)$ when $x_3$ is the trough movement Synchro queue

The EIF method reveals that the only candidate function fitting the integral function is a linear function, $f_3(x_3) = 1 + \beta_{3,1}x_3$. Table 27 presented the model parameter estimation result. The LRT statistic of this model was 8.3 (189.6 – 181.3) which was higher than the chi-square of 3.84 for one degree of freedom at level 0.05. This model was statistically preferred over the model with two variables and therefore the Synchro queue length at the intersection for the through movement was selected to be the third variable in this analysis.

The LRT values for major road median type and through movement Synchro queue are the same as shown in Table 26. If the major road median type had been considered first as the third variable instead of the through movement Synchro queue, the NLMIXED procedure estimation showed a $-2LL$ value of 186.9. With a LRT of 2.7 (189.6 – 186.9). This variable was not statistically significant and therefore it was eliminated as the third variable. Later this variable will be tried and introduced as a fourth variable.
5.3.5 Result

The final form of the third model therefore was expressed as:

\[ y = 0.000396 x_1^{0.9326} e^{-0.3748 x_1 x_2^{1.841}} (1 + 0.008532 x_3) \]  
(Equation 41)

Where,

- \( y \) = Annual access–related collisions at the access point near a signalized intersection
- \( x_1 \) = AADT/10,000 in vehicles per day,
- \( x_2 \) = Driveway width in feet, and
- \( x_3 \) = Synchro peak hour 95% queue at the intersection for through movement in feet.

The model dispersion parameter, \( K \), was estimated as 0.268. The functional form of the third variable, intersection through movement Synchro peak hour 95% queue, has a linear form of \( f_3(x_3) = 1 + 0.008532 x_3 \) that was estimated in previous section. Synchro through movement 95% queue ranges between zero and 1,120 feet in this study data set. Within this range the \( f_3(x_3) \) function would increase as the driveway width increases.

From a traffic engineering point of view, I expected the access–related number of collisions to increase when the intersection through movement queue got larger. This could be because larger queues block the access point, making the operation more complicated, and as a result access–related collisions go up. This expectation was fulfilled by the predicted model.

The CURE plot of this variable and the ±2\( \sigma^* \) shown in Figure 50 confirmed that this model was a good fit and the estimated parameters were accurate. The function starts and ends near \( y = 0 \) and oscillates within ±2\( \sigma^* \). Thus the suggested model was preferred over the second model.
Figure 50 – CURE plot fitted in +2σ* and −2σ* for third model

Figure 51 illustrates the scatterplots of the annual access-related collisions vs. AADT/10,000 for both collected and projected data using the third variable. This shows improvement over the second model.

One can see in Figure 45 that, as for the initial and second models, the annual access-related collisions increased when major road traffic volume increased up to about 25,000 vehicles per day. Then, annual collisions decreased above 25,000 vehicles per day. This decrease could be
because drivers pay more attention and slow down more on busier roads, or could be because of some factor at the intersection I did not consider.

### 5.4 Fourth model

I kept looking for possible variable introduction per the order in Table 26. Assuming that there will be a fourth variable, the overall annual access – related collision model will look like:

\[
y = f_1(\text{AADT}/10,000) \times f_2(\text{DW}) \times f_3(\text{Q_L}) \times f_4(x_4) \quad \text{(Equation 42)}
\]

This section discusses the detail of this effort. Table 29 summarizes the selected models for each variable and the estimated parameters. The parameters \(\beta_{1-1}, \beta_{1-2}, \beta_{2-1}, \beta_{2-2}, \text{ and } \beta_{3-1}\) were re-estimated when introducing the new variables.

<table>
<thead>
<tr>
<th>(x_4)</th>
<th>(\beta_{1-1})</th>
<th>(\beta_{1-2})</th>
<th>(\beta_{2-1})</th>
<th>(\beta_{2-2})</th>
<th>(\beta_{3-1})</th>
<th>(\beta_{4-1})</th>
<th>(\beta_{4-2})</th>
<th>(\beta_{4-3})</th>
<th>(\beta_{4-4})</th>
<th>(\beta_{4-5})</th>
<th>(\beta_{4-6})</th>
<th>(-2LL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MM</td>
<td>1.077</td>
<td>-0.419</td>
<td>0.009</td>
<td>1.741</td>
<td>0.008</td>
<td>0.044</td>
<td>0.075</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>179.0</td>
</tr>
<tr>
<td>LCD</td>
<td>0.890</td>
<td>-0.367</td>
<td>0.001</td>
<td>1.805</td>
<td>0.009</td>
<td>0.795</td>
<td>0.990</td>
<td>1.028</td>
<td>0.910</td>
<td>0.501</td>
<td>0.000</td>
<td>178.8</td>
</tr>
<tr>
<td>SL</td>
<td>1.003</td>
<td>-0.350</td>
<td>0.004</td>
<td>1.643</td>
<td>0.011</td>
<td>0.183</td>
<td>0.118</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>179.5</td>
</tr>
<tr>
<td>DA</td>
<td>0.938</td>
<td>-0.388</td>
<td>0.005</td>
<td>1.784</td>
<td>0.008</td>
<td>0.041</td>
<td>0.106</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>179.2</td>
</tr>
<tr>
<td>AD</td>
<td>0.998</td>
<td>-0.418</td>
<td>0.014</td>
<td>1.773</td>
<td>0.008</td>
<td>0.036</td>
<td>0.050</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>180.6</td>
</tr>
<tr>
<td>CC</td>
<td>0.973</td>
<td>-0.383</td>
<td>0.0005</td>
<td>1.800</td>
<td>0.008</td>
<td>-0.0005</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>181.1</td>
</tr>
<tr>
<td>RA</td>
<td>0.927</td>
<td>-0.375</td>
<td>0.022</td>
<td>1.896</td>
<td>0.008</td>
<td>0.014</td>
<td>0.018</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>180.8</td>
</tr>
<tr>
<td>TR</td>
<td>0.886</td>
<td>-0.333</td>
<td>0.009</td>
<td>1.898</td>
<td>0.008</td>
<td>0.036</td>
<td>0.025</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>180.6</td>
</tr>
</tbody>
</table>

In this section the variables in the order presented in Table 26 were tested to select the fourth variable.

### 5.4.1 Major road median

Since the only type of medians introduced in this research were full movement and RIRO, the function form for this variable was \(f_4(x_4) = \beta_{4-1}(x_4 = \text{RIRO}) + \beta_{4-2}(x_4 = \text{Full})\). The result of the parameter estimation was shown in Table 29. The LRT statistic is calculated to be 2.3 (181.3 – 179.0). The LRT statistic value was less than the chi-square that was 5.99 for two degrees of freedom at level 0.05. Thus this variable was not statistically significant and should not be included in modeling.
5.4.2 Lane configuration at the driveway

As shown before, the lane configuration at the driveway shown in Table 3 was a categorical variable and so the function was expressed as:

\[ f_4(x_4) = \beta_{4-1}(x_4 = 1) + \beta_{4-2}(x_4 = 2) + \cdots + \beta_{4-6}(x_4 = 6) \]  

(Equation 43)

The estimated parameters were shown in Table 29. The LRT statistic was 2.5 (181.3 − 178.8) which was less than the chi-square for six degrees of freedom (12.59) at level 0.05. Therefore, this variable contributed no value to the overall -2LL of the model and was eliminated from modeling.

5.4.3 Speed limit

As shown before, the binary variable of speed limit function was expressed as:

\[ f_4(x_4) = \beta_{4-1}(x_4 \leq 35) + \beta_{4-2}(x_4 \geq 45) \]  

(Equation 44)

The parameter estimation results were shown in Table 29. The LRT statistic was calculated to be 1.8 (181.3 − 179.5). This value was less than the chi-square of 5.99 for two degrees of freedom at level 0.05. Thus speed limit was not statistically important enough to be included in modeling and was eliminated.

5.4.4 Driveway angle

The function of this binary variable was:

\[ f_4(x_4) = \beta_{4-1}(x_4 < 90) + \beta_{4-2}(x_4 \geq 90) \]  

(Equation 45)

The parameter estimation was summarized in Table 29. LRT statistic value was 2.1 (181.3 − 179.2) which was less than the chi-square of 5.99 for two degrees of freedom at level 0.05. Therefore, the angle between driveway and the major road was not statistically important enough to be included in modeling and was eliminated.
5.4.5 Second driveway

As discussed before, this variable was classified as “Yes” or “No” when there is or there is not second driveway, respectively. The function of this variable was:

\[ f_4(x_4) = \beta_{4-1}(x_4 = \text{No}) + \beta_{4-2}(x_4 = \text{Yes}) \]  \hspace{1cm} (Equation 46)

The parameters were estimated using the NLMIXED and the results were shown in Table 29. The LRT statistic was calculated to be 0.7 (181.3 – 80.6) that was less than the chi-square of 5.99 for two degrees of freedom at level 0.05. Since this model was not statistically better than the previous model the variable was eliminated from inclusion in modeling.

5.4.6 Corner clearance

The study sites’ corner clearance distribution was illustrated in Figure 12. Since this is a continuous variable, the \( f_4(x_4) \) needs to be estimated using Equation 23. Figure 52 shows the scatterplots of the \( f_4(x_4) \) and the corner clearance.

![Corner Clearance Scatterplot](image)

Figure 52 – \( f_4(x_4) \) when \( x_4 \) is the corner clearance

It was not easy if possible to find a good model fitting these data points. Thus the EIF method, shown in Figure 53, was used.
The possible candidate functions fitting the curve above are shown in Table 30. The model’s parameter estimation result was shown in Table 29.

Table 30 – \( f_4(x_4) \) BIC values for the corner clearance

<table>
<thead>
<tr>
<th>#</th>
<th>( f_4(x_4) )</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( 1 + \beta_{4,1} \log(x_4) )</td>
<td>213.9</td>
</tr>
<tr>
<td>2*</td>
<td>( 1 + \beta_{4,1} x_4 )</td>
<td>213.9</td>
</tr>
<tr>
<td>3</td>
<td>( \beta_{4,1} + \beta_{4,2} x_4 )</td>
<td>218.5</td>
</tr>
<tr>
<td>4</td>
<td>( x_4^{\beta_{4,1}} )</td>
<td>214.0</td>
</tr>
<tr>
<td>5</td>
<td>( \beta_{4,1} x_4^{\beta_{4,2}} )</td>
<td>218.6</td>
</tr>
</tbody>
</table>

* Selected model

Models 1 and 2 have the minimum BIC value but model 2 was chosen because of its simplicity. The LRT statistic was 0.2 (181.3 – 181.1). This is much less than the chi-square (3.84 for one degree of freedom at level 0.05) and so this variable was not worth adding to the model.

5.4.7 Driveway radius

As shown before, the DR model was expressed as:

\[
\begin{align*}
f_4(x_4) &= \beta_{4\text{-}1}(x_4 = \text{Small}) + \beta_{4\text{-}2}(x_4 = \text{Large})
\end{align*}
\]  

(Equation 47)
The parameter estimation was shown in Table 29. The LRT statistic was calculated to be 0.5 (181.3 – 180.8). The LRT statistic value was less than the chi-square, which was 5.99 for two degrees of freedom at level 0.05. Thus driveway radius was not statistically significant and was eliminated from modeling.

5.4.8 Driveway transition

The function of this binary variable was shown before as:

\[ f_4(x_4) = \beta_{4-1}(x_4 = \text{Not elevated}) + \beta_{4-2}(x_4 = \text{Elevated}) \]  \hspace{1cm} \text{(Equation 48)}

The result of the parameter estimation was shown in Table 29. The LRT statistic was 0.7 (181.3 – 180.6). This value was less than the chi-square of 5.99 for two degrees of freedom at level 0.05. Thus the model with this variable was not statistically better than the model without it. Therefore, the driveway transition was also eliminated from the modeling process. In the end, no fourth variable was worthy of induction into the model.

5.5 Modeling result

A summary of the SAS® output for the first, second, and final models is presented in Appendix G. As a result of the modeling effort, the final annual access – related collision prediction model functional form was estimated as:

\[ y = 0.000396 x_1^{0.9326} e^{-0.3748 x_1} x_2^{1.841} (1 + 0.008532 x_3) \] \hspace{1cm} \text{(Equation 49)}

Where,
\[ y = \text{Annual access – related collisions at the access point near a signalized intersection} \]
\[ x_1 = \text{AADT/10,000 in vehicles per day,} \]
\[ x_2 = \text{Driveway width in feet, and} \]
\[ x_3 = \text{Synchro peak hour 95\% queue at the intersection for through movement in feet.} \]
The standard errors for the parameters, 0.000396, 0.9326, -0.3748, 1.841, 0.008532 were 0.001, 0.786, 0.463, 0.608, 0.011, respectively. The model dispersion parameter, K, was estimated as 0.268.

The model was developed using a dataset including the major road AADT ranging between 1,230 and 54,000 vehicles per day, driveway width ranging between 8 ft and 58 ft, and Synchro through movement 95% queue ranging between zero and 1,120 feet.

To use the model, consider a site with a driveway near a signalized intersection that meets all the requirements discussed in the site selection section. The major road AADT is 19,000 vehicles per day, the driveway width is 20 feet, and the intersection through movement Synchro peak hour 95% queue length is 704 feet. This would result in an estimated 0.6 access – related collisions per year.

To understand the behavior of the predicted collisions using the model, Figure 54 shows scatterplots of the residual (observed minus predicted) of annual access – related collisions. This graph shows that small mean values correspond to small variances and larger mean values corresponds to larger variances.

![Figure 54](image)

Figure 54 – Prediction model residuals versus predicted annual access – related collisions

Figure 55 shows the predicted versus observed annual access – related collisions for the studied 108 sites. A linear regression analysis using SAS® revealed a linear form of $y = 0.879 x + 0.06$
in which \( y \) is the observed annual access – related collision and \( x \) is the predicted annual access – related collision. The standard errors for slope and intercept are 0.186 and 0.131, respectively. As expected, the linear function slope is not statistically significantly different from 1 and the intercept is not different from 0.

![Figure 55 – Predicted versus observed annual access – related collisions](image)

The prediction model shows that the longer the through movements Synchro queue at the intersection, the higher the number of access – related crashes. This is because the queue at the intersection could block or interfere with movements at the access point. Drivers entering or exiting the driveway upstream of the intersection face difficulties maneuvering and therefore the collision frequency goes up. Statistically our model shows that every 100 ft increase in length of the 95% peak hour through movement Synchro queue at the intersection results in 0.15 access – related crashes increase per year on average when the major road traffic volume is 15,600 vehicles per day and the driveway width is 30 ft.

The final model also revealed that narrower driveways are associated with fewer access – related crashes. This could be because drivers pay more attention entering and exiting the driveway when it is narrower. The driveway width result may also relate to the driveway traffic volume, driveway median design, and other factors. The model form reveals that increasing the driveway width from 20 ft to 30 ft relates to almost 0.33 access – related collision raise per year when the
major road traffic volume is 15,600 vehicles per day and the intersection 95% through movement peak hour Synchro queue is 300 ft.

Annual collisions related to the access point were also a function of the major road traffic volume. Assuming a driveway width of 55 ft and the intersection 95% through movement Synchro queue of 300 ft, the model shows that if the major road traffic volume increases from 7,000 to 17,000 vehicles per day, access – related collisions would increase 0.71 per year. It also reveals that if the major road traffic volume is increased from 35,000 to 45,000 vehicles per day, the access – related collisions decrease 0.33 per year.

As shown in Figure 51, the scatterplot of the final prediction model revealed that the annual access – related collisions increased when major road traffic volume increased up to about 25,000 vehicles per day. Then, annual collisions decreased above 25,000 vehicles per day. This decrease could be because drivers pay more attention and slow down more on busier roads, or could be because of some factor at the intersection I did not consider.

Table 31 summarizes the result of finding proper variables to be included in the model.

<table>
<thead>
<tr>
<th>Variables</th>
<th>–2LL</th>
<th>LRT</th>
<th>x2</th>
<th>Select</th>
</tr>
</thead>
<tbody>
<tr>
<td>AADT (initial model)</td>
<td>199.7</td>
<td>---</td>
<td>----</td>
<td>Yes</td>
</tr>
<tr>
<td>AADT, DW (second model)</td>
<td>189.6</td>
<td>10.1</td>
<td>5.99</td>
<td>Yes</td>
</tr>
<tr>
<td>AADT, DW, LCI</td>
<td>185.2</td>
<td>4.4</td>
<td>14.07</td>
<td>No</td>
</tr>
<tr>
<td>AADT, DW, GR</td>
<td>183.5</td>
<td>6.1</td>
<td>14.07</td>
<td>No</td>
</tr>
<tr>
<td>AADT, DW, DV</td>
<td>186.2</td>
<td>3.4</td>
<td>3.84</td>
<td>No</td>
</tr>
<tr>
<td>AADT, DW, Q-T (final model)</td>
<td>181.3</td>
<td>8.3</td>
<td>3.84</td>
<td>Yes</td>
</tr>
<tr>
<td>AADT, DW, Q-T, MM</td>
<td>179.0</td>
<td>2.3</td>
<td>5.99</td>
<td>No</td>
</tr>
<tr>
<td>AADT, DW, Q-T, LCD</td>
<td>178.8</td>
<td>2.5</td>
<td>12.59</td>
<td>No</td>
</tr>
<tr>
<td>AADT, DW, Q-T, SL</td>
<td>179.5</td>
<td>1.8</td>
<td>5.99</td>
<td>No</td>
</tr>
<tr>
<td>AADT, DW, Q-T, DA</td>
<td>179.2</td>
<td>2.1</td>
<td>5.99</td>
<td>No</td>
</tr>
<tr>
<td>AADT, DW, Q-T, AD</td>
<td>180.6</td>
<td>0.7</td>
<td>5.99</td>
<td>No</td>
</tr>
<tr>
<td>AADT, DW, Q-T, CC</td>
<td>181.1</td>
<td>0.2</td>
<td>3.84</td>
<td>No</td>
</tr>
<tr>
<td>AADT, DW, Q-T, RA</td>
<td>180.8</td>
<td>0.5</td>
<td>5.99</td>
<td>No</td>
</tr>
<tr>
<td>AADT, DW, Q-T, TR</td>
<td>180.6</td>
<td>0.7</td>
<td>5.99</td>
<td>No</td>
</tr>
<tr>
<td>AADT, DW, Q-T, DM</td>
<td>---</td>
<td>---</td>
<td>3.84</td>
<td>N/A</td>
</tr>
</tbody>
</table>
As mentioned previously, the CURE plot was used in this study to measure the goodness-of-fit of the models. For the highlighted models in Table 31, the CURE plot started, ended, and oscillated around $y = 0$ axis. This confirmed the accuracy of these models.

Other variables such as the distance of the driveway from the signalized intersection were not included in the model but could still be practically important. The reason they were eliminated from modeling might be that the engineers and the policy makers at NCDOT are properly specifying corner clearance to reduce collisions.

The proposed collision model could be used by state DOTs and municipal traffic engineers to address access management issues. The model is very easy to use because it is a function of only three variables (AADT, driveway width, and Synchro 95% through movement queue at the intersection near the access point). These variables are relatively easy to measure or estimate. The AADT can be found from AADT map sheets or Google Earth Pro. The driveway width can be obtained from aerial images or site visits. The intersection through movement 95% queue length should be estimated using traffic software, Synchro, using intersection turning movements and signal timing plans. If this value is observed by site visits, an adjustment factor of 1.37 should be multiplied to the observed peak hour maximum length of the queue.

5.6 Model validation

To validate the model shown in previous section, I randomly chose 53 sites from the original 739 signalized intersections within Wake County. These sites then were filtered down according to required scope and limitation described in Section 1.3. At the end, I had 27 sites that were not used previously available for model calibration. The prediction model had three independent variables of AADT, driveway width, and through movement Synchro queue length plus annual number of access – related collisions. These independent variables and the annual access – related collisions were collected using the same techniques described in Section 3.2 and are shown in Table 32.
Table 32 – Independent and dependent variables for model validation

<table>
<thead>
<tr>
<th>Site #</th>
<th>DW*</th>
<th>MV*</th>
<th>Q-T*</th>
<th>Observed**</th>
<th>Predicted***</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23</td>
<td>3.9</td>
<td>1201</td>
<td>0.60</td>
<td>1.18</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>0.8</td>
<td>491</td>
<td>0.80</td>
<td>0.90</td>
</tr>
<tr>
<td>3</td>
<td>32</td>
<td>1.3</td>
<td>967</td>
<td>0.80</td>
<td>1.70</td>
</tr>
<tr>
<td>4</td>
<td>28</td>
<td>0.4</td>
<td>792</td>
<td>0.00</td>
<td>0.53</td>
</tr>
<tr>
<td>5</td>
<td>38</td>
<td>1.9</td>
<td>1191</td>
<td>1.60</td>
<td>3.19</td>
</tr>
<tr>
<td>6</td>
<td>23</td>
<td>0.5</td>
<td>693</td>
<td>0.00</td>
<td>0.36</td>
</tr>
<tr>
<td>7</td>
<td>18</td>
<td>1.3</td>
<td>386</td>
<td>0.00</td>
<td>0.27</td>
</tr>
<tr>
<td>8</td>
<td>31</td>
<td>1.4</td>
<td>976</td>
<td>1.00</td>
<td>1.66</td>
</tr>
<tr>
<td>9</td>
<td>24</td>
<td>0.9</td>
<td>269</td>
<td>0.00</td>
<td>0.29</td>
</tr>
<tr>
<td>10</td>
<td>19</td>
<td>0.9</td>
<td>597</td>
<td>0.40</td>
<td>0.35</td>
</tr>
<tr>
<td>11</td>
<td>14</td>
<td>1.5</td>
<td>1183</td>
<td>0.20</td>
<td>0.47</td>
</tr>
<tr>
<td>12</td>
<td>39</td>
<td>0.5</td>
<td>667</td>
<td>1.00</td>
<td>0.91</td>
</tr>
<tr>
<td>13</td>
<td>25</td>
<td>1.9</td>
<td>76</td>
<td>0.00</td>
<td>0.22</td>
</tr>
<tr>
<td>14</td>
<td>35</td>
<td>0.1</td>
<td>40</td>
<td>0.00</td>
<td>0.04</td>
</tr>
<tr>
<td>15</td>
<td>37</td>
<td>1.9</td>
<td>73</td>
<td>0.00</td>
<td>0.44</td>
</tr>
<tr>
<td>16</td>
<td>30</td>
<td>1.4</td>
<td>257</td>
<td>0.00</td>
<td>0.54</td>
</tr>
<tr>
<td>17</td>
<td>16</td>
<td>0.8</td>
<td>255</td>
<td>0.00</td>
<td>0.12</td>
</tr>
<tr>
<td>18</td>
<td>27</td>
<td>1.6</td>
<td>185</td>
<td>0.00</td>
<td>0.37</td>
</tr>
<tr>
<td>19</td>
<td>29</td>
<td>0.7</td>
<td>68</td>
<td>0.00</td>
<td>0.16</td>
</tr>
<tr>
<td>20</td>
<td>34</td>
<td>1.0</td>
<td>138</td>
<td>0.20</td>
<td>0.38</td>
</tr>
<tr>
<td>21</td>
<td>21</td>
<td>0.6</td>
<td>168</td>
<td>0.00</td>
<td>0.12</td>
</tr>
<tr>
<td>22</td>
<td>34</td>
<td>0.4</td>
<td>66</td>
<td>0.00</td>
<td>0.14</td>
</tr>
<tr>
<td>23</td>
<td>29</td>
<td>1.9</td>
<td>198</td>
<td>0.00</td>
<td>0.47</td>
</tr>
<tr>
<td>24</td>
<td>37</td>
<td>0.5</td>
<td>98</td>
<td>0.20</td>
<td>0.24</td>
</tr>
<tr>
<td>25</td>
<td>36</td>
<td>1.3</td>
<td>110</td>
<td>1.60</td>
<td>0.44</td>
</tr>
<tr>
<td>26</td>
<td>30</td>
<td>1.6</td>
<td>1023</td>
<td>0.00</td>
<td>1.72</td>
</tr>
<tr>
<td>27</td>
<td>31</td>
<td>1.0</td>
<td>107</td>
<td>0.80</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Average: 0.34 | 0.65

* DW: Driveway volume, in feet
  MV: Major road volume (AADT/10,000), in vehicles per day
  Q-T: Synchro estimated 95% peak hour through movement queue, in feet
** Annual observed access - related collisions
*** Annual predicted access - related collisions

The independent variables were plugged in the model to find the predicted annual access – related collisions, as shown in Table 32. These values then were compared with the observed annual access – related collisions as shown in Figure 56. One can see in Table 32 that annual predicted collisions are continuous while the annual observed collisions have an increment of
0.2. This is due to the fact that I collected the annual – related collisions for a period of 5 years and then divided it by 5 to obtain average annual collisions over these 5 years.

![Figure 56 – Model validation (predicted versus observed annual access – related collisions)](image)

A linear regression analysis using SAS® revealed a relation of $y = 0.447x + 0.051$ in which $y$ is the predicted annual access – related collision and $x$ is the observed annual access – related collision. The standard errors for slope and intercept are $0.11$ and $0.10$, respectively. This reveals that the prediction model is over-predicting the number of access – related collisions by a factor of about 2. Although this is not desired, it was not surprising because of the multiple sources of error in data collection, especially the number of access – related collisions. This also could be because I didn’t consider other factors that might be important such as sight distance, driveway grade, etc. In the next chapter I provide a different point of view to try to understand what was special about the higher collisions sites.
6. Higher collision site identification

Collision prediction models cannot necessarily describe collision–related behavior in all circumstances. This could be because it is hard to model all the factors causing collisions such as human behavior. In Chapter 5 I modeled the access–related collisions using data from 108 sites in Wake County, but the 7 outlier sites shown in Figure 51 were covered by the prediction model less effectively. These 7 sites had 2.8 to 6.4 access–related collisions per year (or 14 to 32 collisions during the 5-year study) and will be referred as “higher collision sites”. The remaining 101 sites have less than 1.4 annual access–related collisions each and are referred to as “lower collision sites”.

In this chapter, as shown in Figure 57, I described the methodology that was used to collect more data and information on the higher collision sites. Studies of the higher collision sites and statistical analyses were performed to compare them to the rest of the sites. At the end recommendations are provided to traffic engineers, planners, and decision-makers to avoid permitting and building higher collision sites in the future.
6.1 Methodology

The data collection approach for the study sites was discussed in Chapter 3. To further analyze and understand the characteristics of the higher collision sites, this section reviews the methodology implemented particularly for the 7 higher collision sites.

I traveled to the 7 higher collision sites and collected characteristics such as sight distance, major road and driveway grades, access point channelization, traffic signal operation, and lane
configurations on driveway and intersection. Some of this information was collected using Google Earth aerial images. An informal interview was also done with business owners, passing pedestrians, and drivers. I was cautious to notice and keep record of any unusual road geometry, driver behavior, and other factors that possibly could affect the normal circumstances and cause unusually high numbers of collisions. Other information such as the driveway peak hour turning volumes was collected manually.

To understand the characteristics of the higher collision sites, a comparison was made between the 7 higher collision sites and the 101 lower collision sites. In some cases the comparison was done between these 7 sites, which all had full movement, and the rest of 70 full movement sites. If the assessment of a variable of interest was not possible among the 70 full movement lower collision sites because the variable was too difficult to collect, a sample of 15 randomly selected lower collision full movement sites was picked to be evaluated against the higher collision sites.

6.2 Case studies

The higher collision sites were named as site 1 to site 7 with the site 1 having the highest number of access – related collisions and the site 7 having the least number of access – related collisions among the 7 higher collision sites. This section reviews the higher collision sites individually. Site 1 is located in Raleigh, site 2 is in North Raleigh, site 3 is in Cary, and site 4 is in Fuquay-Varina. Sites 5, 6, and 7 are 3 legs of one intersection in Fuquay-Varina, near site 4. For each site, information such as the detail of their location, major road AADT, driveway peak hour traffic volume and the land use type are presented.

Using the access – related collisions, collision diagrams were prepared and presented for each site. These diagrams were used to help identify similar collision patterns by providing information such as time of the day and direction of vehicles involved in collision. In this research the diagrams also show if the collision occurred in dark or daylight. A discussion on each collision diagram were also provided.
6.2.1 Site 1

Site 1 is located south central side of Raleigh on Lake Wheeler Road (AADT 15,000 vehicles per day) near Tryon Road (AADT 20,000 vehicles per day) and provides access to the Raleigh Oaks Shopping Center. The site is only 4.6 miles away (8 minutes) from the State Capitol. This shopping center was built in 1989 with 60,480 square feet of gross leasable area. An aerial image of the site in Figure 58 shows that it has three access points, one on Tryon Road and two on Lake Wheeler Road.

As of 2011 Raleigh Oaks Shopping Center’s big box stores were Food Lion and Dollar General. The other stores are Braids and Beauty Salon, Veterinarian, Barber Shop, Cricket, Nationwide Insurance Agency, China Wok, Members Credit Union, Family Dentistry, Checker’s Pizza, Happy Nails, Raleigh Oak Laundry, and Big Will’s Barber Shop. There is also a Han-Dee Hugo’s 12-pump gas station and convenience store and a Cash Point ATM in the shopping center.

![Site 1 aerial image](image)

Figure 58 – Site 1 aerial image

The collision diagram in Figure 59 shows that this site had 32 access – related collisions during the 5-year data collection period between January 2005 and December 2009.
The collision diagram shows that over 90% of the total access–related collisions involve a vehicle traveling southbound on Lake Wheeler Road. Almost all of the crashes occurred between 3 PM and 7 PM, and two-thirds of the crashes occurred during the afternoon peak hours of 5 PM to 7 PM. I noticed that the PM directional split on Lake Wheeler Road is almost 65% on southbound, the same side as the driveway location. So this could be one of the significant factors causing all the high access–related collisions.

The other interesting point is that 14 out of the total 32 collisions involved a vehicle leaving the driveway and turning left on Lake Wheeler Road. Since there is another driveway north of the major access point, by converting this full movement access point to a RIRO, drivers would use the nearby access point and this could reduce the overall collisions at this point significantly. Also 15 crashes are for the left turn traffic from Lake Wheeler into the access point. A RIRO solution may also eliminate these collisions.
6.2.2 Site 2

This access point is located on New Hope Church Road (AADT 21,000 vehicles per day) very near Atlantic Avenue (AADT 27,000 vehicles per day) in north east Raleigh. The site has 8,000 SF of retail including convenience store and a 4-pump gas station. An interview with the business owners revealed that during the study period of January 2005- December 2009 the retail store in this location was a professional cleaning company called “Cox Textile Services” providing laundry maid service, dry cleaning, and laundry business services. This retail store is not there now and currently the office location is vacant. An aerial image of this site in Figure 60 shows that the back of this access point is indirectly connected to Brentwood Animal Hospital.

![Figure 60 – Site 2 aerial image](image)

Figure 61 illustrates the collision diagram of this access point on New Hope Church Road. A total of 21 access – related collisions were recorded during the period of January 2005 to December 2009. The collision diagram doesn’t provide a strong temporal pattern among collisions but it reveals that 85% of the total access – related collisions involve a vehicle traveling westbound on New Hope Church Road. After a site visit I discovered that the lack of an
appropriate sight distance for drivers leaving the access point could be a reason for these collisions.

Figure 61 – Site 2 collision diagram

The profile for New Hope Church Road from Google Earth Pro in Figure 62 shows a very steep hill upstream of the access point. The grade at the access point is 2% so the site qualified for this study but its grade is much higher (up to 7%) upstream of the access point. Because of this issue drivers leaving the driveway do not have a full view to the left and that could result in collisions.

Figure 62 – Steep hill upstream of the access point at site 2
A detailed review of the collisions at site 2 showed that almost half of the access – related collisions at this site involved a pick-up truck, or a van. This could be because the professional cleaning store attracted larger vehicles than passenger cars. The aerial image of Figure 60 confirms this assumption.

Further review of the collision diagram reveals that over half of these collisions involved a vehicle leaving the driveway and turning left. A left-over design would eliminate all left turn collisions out of the driveway at this access point. Turning left turn from New Hope Church Road to the driveway doesn’t seem to be an issue because there is only one such collision occurred. If the New Hope Church Road is not wide enough for a left-over, a RIRO median can be installed.

### 6.2.3 Site 3

Site 3 is located on Maynard Road (AADT 15,000 vehicles per day), north of High House Road (AADT 20,000 vehicles per day) in Cary, NC. Maynard Road is a loop road around downtown Cary and High House Road at this intersection is located between north-west and south-west Maynard Road. This driveway is one of the five access points to the Maynard Road Shopping Center. With a total of over 122,000 square feet of gross leasable area, Maynard Road Shopping Center’s big box store is Kroger. As of 2011, other stores in the shopping center were Brigs Restaurant, Oxford Learning Center, Orient Garden, Theatrical Dance Arts Academy, Hallmark, Prime Mobile, State Farm Insurance, Great Clips, Rapid Refill Ink, The UPS Store, Jet's Pizza, Michelangelo's Italian Grill, INTREX Computers, GNC, KJ Natural Stone, US Nails, Planet Beach, The Corner Tavern and Grill, KOMO KOMO, Time Warner Cable, Pearle Vision, Wolf Camera, Doctors Express, Salon Jon Clay, Bellsouth Tower, and Wendy's. The immediate access at the driveway in question is to the Wendy's fast food restaurant and the shopping center parking lot. Figure 63 shows an aerial image of this full movement site.
As mentioned in Section 3.1 on site selection, a few sites were chosen based on a recent median change on the access point. The reason was that I presumed the improvements on the median took place because of complaints received by residents or investigations done by traffic engineers as a result of higher collisions at those sites. Site 3 is one of those sites that used to operate full movement but a left-over design was implemented in 2005. So the crash data on this site were collected before the improvement (between October 2000 and November 2004). During this time 18 access – related collisions were recorded as shown in the collision diagram in Figure 64.
Two-thirds of all access-related collisions (12 out of 18 collisions) involved a vehicle traveling southbound on Maynard Road, but there is no time pattern among the collisions. Nine collisions involved a vehicle leaving the access point and turning left, with 2 others that could be through or left turn movements. So with at least half of the collisions involving a left turn from access point, it seems that the left-over design that was implemented was a good solution to restrict this movement.

6.2.4 Site 4

Site 4 is located in the Town of Fuquay-Varina, almost 16 miles south-west of downtown Raleigh. The access point is called Market View Lane and it intersects with Purfoy Road, south of the intersection of Sunset Lane/ Purfoy Road (AADT 13,000 vehicles per day) and Main Street (AADT 22,500 vehicles per day) as shown in Figure 65. The access point provides direct access to Bojangles’, Arby's, and Milano Pizza fast food restaurants and also indirect access to Dairy Queen and Waffle House restaurants. Market View Lane ultimately connects to a few other retail stores including Harris Teeter, Fast-Med Urgent Care, Bellini Fine Italian Cuisine, and Tacos Mexican Restaurant.
Main Street (US-401) is the major arterial between Raleigh and the Town of Fuquay-Varina. Currently the site is RIRO but as discussed in Section 3.1, this is one of the sites chosen because the median improvement was implemented in 2005. Before this channelization the access point was operating full movement. Collisions and other data were collected between February 2000 and January 2004. Seventeen access-related collisions, as shown in Figure 66, were observed during this period. Almost 60% of these collisions involved a vehicle traveling southbound on Purfoy Road. The PM directional split was over 65% for Purfoy Road southbound. No significant temporal pattern was discovered among the crashes.

It is noticeable that almost 70% of the overall access - related collisions (11 out of 17) involved a vehicle exiting the access point, crossing Purfoy Road, and traveling toward the retail stores across the road. The big box retail stores across the access point are Roses Discount Store and Food Lion and the small businesses include Fidelity Bank, First Citizen Bank, Subway, and other stores. Attraction to this shopping center from the study site was the main reason for high crossing traffic volumes leading to numerous collisions.
All the access – related collisions, except one, would be eliminated by having a RIRO at this access point. The NCDOT traffic engineers recognized this properly and currently the site is operating with a RIRO median. A recent site visit revealed that sight distance for drivers leaving the access point is very limited. Also, the stop bar is not visible and needs repainting. As Figure 67 shows, the stop sign location is not matched well with the stop bar location.
After reviewing the issues at this access point and the possible causes of collisions, it became obvious that NCDOT engineers were aware of the problems and made an appropriate decision in installing a RIRO median to restrict the access point through and left turn movements.

6.2.5 Site 5

This site is almost half a mile east of site 4 on Main Street (AADT 22,500 vehicles per day), near the intersection of Judd Parkway (AADT 15,500 vehicles per day) in the Town of Fuquay-Varina. Figure 68 shows that the site provides immediate access to Steve Ashworth Performance (auto parts and repair center), Glam-O-Rama (dry-cleaning and laundry services), and KFC fast food restaurant. It also has access to Quality Inn, Fuquay Eye Care, and First Federal Bank.

The collision data were collected between January 2005 and December 2009 and classified to find the 16 access-related collisions shown in Figure 69. The directional split on Main Street is almost half-half on both directions during the day and the collisions are scattered throughout the day. A RIRO median could have prevented 12 collisions of the 16 collisions. Two collisions were rear-end collisions including vehicles crossing Main Street and probably happened because of traffic in front of the lead vehicle. These may also be eliminated by a RIRO design.
Apparently 12 collisions (75% of all collisions) involved a vehicle leaving the access point. A site visit, as shown in Figure 70, revealed that the driveway at this location has a steep grade.

Figure 71 with data from Google Earth Pro shows that the grade at this driveway is over 17% and as a result drivers have a hard time making sure no car is coming from the left on Main Street.
6.2.6 Site 6

At the same intersection as site 5, site 6 is on Main Street (AADT 22,500 vehicles per day), east of the intersection of Judd Parkway (AADT 15,500 vehicles per day) in the Town of Fuquay-Varina. This driveway provides access to Burger King and Pizza Hut fast food restaurants. It indirectly connects to Quizno’s Subs, Computers Store, and a Curves fitness center. There are also a few vacant buildings in the shopping area. An aerial view of the site is shown in Figure 72.

The collision data at this location were collected between January 2005 and December 2009. One can see in Figure 73 that 15 access-related collision were recorded in which 7 collisions
involved a vehicle turning left and leaving the driveway, 2 collisions involved a vehicle turning left turn into the driveway, and 4 collisions happened in the opposite direction of the driveway location and were most likely the result of a following vehicle that was trying to turn left into the driveway.

![Collision Diagram](image)

**Figure 73 – Site 6 collision diagram**

Main Street has a two way left turn lane (TWLTL), but close to the intersection, and particularly at the access point, there is no TWLTL. Even though the access point channelization meant to restrict the turning left to and from the access point, all the collisions except two happened because drivers did not follow the rule. No-left turn signs and better channelization could fix the issue. More visible and curved right turn markings from driveway into Main Street could also be helpful.

If possible, a RIRO median or at least a left-over design would eliminate most of these collisions. The restricted left turns could use the other access point further away from the intersection on Main Street or the driveway on Judd Parkway.
6.2.7 Site 7

Site 7 is very close to site 5 and site 6. It is located on Judd Parkway (AADT 15,500 vehicles per day), south of the intersection with Main Street (AADT 22,500 vehicles per day) in the Town of Fuquay-Varina. This driveway provides access to Sunset Plaza as shown in Figure 74.

![Site 7 aerial image](image)

Sunset Plaza was built in 1987 with over 522,000 square feet of gross leasable area. As of 2011, the big box stores were Roses Discount Store and Food Lion. Other stores are Fidelity Bank, First Citizen Bank, Johnson Optometric Associates, H&R Block, Taekwon Do, #1 Nail, Todays Hair, Beauty Supply, Little Caesars, Postnet, Golden China, Subway, and Tobacco Store. The studied driveway is one of the five access points to this area that also has indirect access to other stores like Wendy’s, Rite Aid Pharmacy, and Tractor Supply Company.

Currently this access point is operating with a RIRO median but it is one of the sites chosen because the median improvement was implemented in 2005 as a result of very high collisions (discussed in Section 3.1). The collision data were recorded over a 5-year period between October 1999 and September 2004, when the site was a full movement access point. There were 14 access – related collisions as shown in Figure 75. Almost all the collisions involved a vehicle turning left (or through) from the driveway. Thus traffic engineers and decisions makers were
well-informed of the situation and the RIRO implemented solution likely eliminated almost all the access – related collisions.

![Site 7 collision diagram](image)

**Figure 75 – Site 7 collision diagram**

### 6.3 Analysis

As shown in Table 8, out of the 108 sites, 77 sites were operating as full movement sites while the rest of the 31 sites had a RIRO design. All the 7 higher collision sites were among the full movement sites during the study period.

In this section a comparison was made between the 7 higher collision sites and the 101 lower collision sites. In some cases the comparison was done between these 7 sites and the rest of 70 full movement sites. If the assessment of a variable of interest was not possible among the 70 lower collision sites because the variable was too difficult to collect, a sample of 15 randomly selected sites was picked to be evaluated against the higher collision sites. At the end of this section demographic and socio economic information for the 7 higher collision sites was compared to averages in Wake County.

Table 33 lists the 7 higher collision sites showing all of the parameters collected for the modeling effort.
Using the information provided in Section 6.2 and Table 33, statistical analyses of the quantitative and binary variables were done to find relationships or differences between the higher collision and lower collision sites.

### 6.3.1 Quantitative variables

Table 34 provides a list of quantitative variables and corresponding averages for the 7 higher collision sites and 101 lower collision sites. It also presents t test results revealing if the mean values for the higher collision sites were statistically different than for the lower collision sites. Considering a 95% confidence level, driveway peak hour traffic volume and intersection peak hour through movement Synchro 95% queue were the variables significantly different for the higher collision sites.
Table 34 – t test for quantitative variables

<table>
<thead>
<tr>
<th>Quantitative variables</th>
<th>7 high collision sites’ mean</th>
<th>101 lower collision sites’ mean</th>
<th>P-value t test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Driveway volume (vehicle/ hour)</td>
<td>165</td>
<td>86</td>
<td>0.008</td>
</tr>
<tr>
<td>Through movement Synchro 95% queue (feet)</td>
<td>535</td>
<td>295</td>
<td>0.036</td>
</tr>
<tr>
<td>Driveway width (feet)</td>
<td>33</td>
<td>27</td>
<td>0.108</td>
</tr>
<tr>
<td>Major Road Grade (percent)</td>
<td>1</td>
<td>0</td>
<td>0.138</td>
</tr>
<tr>
<td>Speed limit (mile/ hour)</td>
<td>41</td>
<td>39</td>
<td>0.574</td>
</tr>
<tr>
<td>Corner clearance (feet)</td>
<td>264</td>
<td>258</td>
<td>0.887</td>
</tr>
</tbody>
</table>

This means that, generally, the higher driveway traffic volume, the greater number of access – related collisions. Even though this fact was not shown in the modeling process in the previous chapter probably due to the small sample of higher collision sites, this relationship was expected.

The t test also shows the importance of the intersection through movement Synchro 95 percentile queue. This variable was statistically significant in the modeling process. This shows that the larger length of the intersection through movement Synchro queue, the higher number of access – related collisions.

6.3.2 Binary variables

For the binary variables, I chose the chi-square contingency table to test for relationships. Table 35 shows our binary variables and the way they were categorized in this research. It also provides frequencies of the higher collision and lower collision sites and the result of the chi-square test. If the p-values are less than 0.10, it means we have evidence that these two sets are statistically different. These p-values for the existence of a secondary driveway and the major road median type as shown in Table 35, are less than 0.10. This means that these variables were statistically different for the higher and lower collisions sites’ samples and as a result they contributed significantly to the high number of collisions.
Table 35 – Chi-square test for binary variables

<table>
<thead>
<tr>
<th>Binary variable</th>
<th>Category</th>
<th>High collision</th>
<th>Low collision</th>
<th>Chi-square test p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second driveway</td>
<td>Yes</td>
<td>3</td>
<td>12</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>4</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Major road median</td>
<td>Full</td>
<td>7</td>
<td>70</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td>RIRO</td>
<td>0</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Driveway angle</td>
<td>&lt; 90</td>
<td>0</td>
<td>12</td>
<td>0.333</td>
</tr>
<tr>
<td></td>
<td>≥ 90</td>
<td>7</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Driveway radius</td>
<td>Large</td>
<td>4</td>
<td>53</td>
<td>0.811</td>
</tr>
<tr>
<td></td>
<td>Small</td>
<td>3</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>Driveway transition</td>
<td>Smooth</td>
<td>6</td>
<td>85</td>
<td>0.913</td>
</tr>
<tr>
<td></td>
<td>Elevated</td>
<td>1</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Driveway median</td>
<td>Divided</td>
<td>1</td>
<td>14</td>
<td>0.975</td>
</tr>
<tr>
<td></td>
<td>Undivided</td>
<td>6</td>
<td>87</td>
<td></td>
</tr>
</tbody>
</table>

Although the chi-square test in Table 35 shows that the variable for the existence of a secondary driveway may be important, I was not able to support the statistical result with logical backup. Attempts were made using statistical tests to find a correlation between this variable and variables such as driveway volume and AADT, but none of them was successful. This could be because the secondary driveway variable had small sample size in each category for higher collision sites. Other than a correlation with an important variable, I could not think of a good logical reason that the presence of a second driveway would influence a site to have more collisions. Thus, due to the lack of theoretical support, the variable for the existence of a secondary driveway was deleted from the list of factors that could contribute to the checklist.

As discussed earlier, all of the higher collision sites were operating with a full movement at the access point during the study period so none of the 31 RIRO median sites were among the higher collision sites. This shows that the chance of a RIRO median to be one of the higher collision sites is slim. The test result in Table 35 confirms this observation as the chi-square value is below 0.10. We can conclude that higher collision sites will almost certainly be full movement driveways.

6.3.3 Driveway left turn proportion

More than two-thirds of all access – related collisions involve a left-turning vehicles (51). Although the effect of the driveway volume was reflected in this research and was shown to be
statistically significant, I noticed that the proportion of left turn traffic volume leaving the access point to the major road AADT could be a contributing factor. For example this proportion for site 1 with 42 vehicles per hour left turn vehicles leaving the driveway and the major road AADT of 15,500 vehicles per day is $100 \times (42/15,500) = 0.271$. I collected left, through, and right turn counts leaving the access point for 6 higher collision sites and a random sample of 15 full movement sites among lower collision sites. As stated Section 6.2, at some point in the last few years at 3 higher collision sites (Sites 3, 4, and 7) a RIRO median was installed and so their study period was when they were operating as full movement driveways. I found the old traffic counts for sites 4 and 7 from previous traffic analyses. Manual collection of the turning counts for all lower collision sites would be time consuming, so I decided to pick a random sample of 15 full median sites among the lower collision sites. It was assumed that distribution of these 15 sites effectively represents the entire sample size of the full median lower collision sites.

A comparison was made between the 6 higher collision sites and the 15 lower collision sites. For these sites, proportions of left traffic volumes from the access point were calculated. Their mean values were 0.28 and 0.10 for the higher collision and lower collision sites, respectively. A t test p-value of 0.002 proved that these two data sets were dissimilar at a 99% confidence level. Higher collision sites likely have higher left turn proportions than lower collision sites.

6.3.4 Categorical variables

The categorical variables of interest, lane configuration at the driveway and intersection lane configuration, are shown in Table 36. For lane configuration at the driveway the highest frequency of both higher and lower collision sites are for configuration 3 and the rest of sites are distributed proportionally very closely to other configurations. Lane configurations at the driveway were shown in Table 3. Table 36 also shows that the lane configurations at the intersection are proportionally equally distributed between the higher and lower collision sites. The lane configuration shapes at the driveway were shown in Table 7. It appears that neither of these variables had much effect on the sites being among the higher or lower collision sets.
### Table 36 – Categorical variables

<table>
<thead>
<tr>
<th>Configuration type</th>
<th>Higher collision sites</th>
<th>Lower collision sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCD*</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>LCI**</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

* Lane configuration at the driveway
** Lane configuration at the intersection

### 6.3.5 Hourly distribution of collisions

The other factor considered was the distribution of collisions over the 24 hours of the day. Figure 76 shows that almost 40% of the total 133 collisions at the higher collision sites occurred between 5 PM – 7 PM while 12 PM – 1 PM had 12% and the rest of the hours were all below 10% of the overall collisions.
The hourly distribution of 7 higher collision sites (with a total of 133 access – related collisions) was compared with the rest of 70 full movement sites (with a total of 93 access – related collisions). Figure 77 shows the histograms of these two data sets.
A Pearson chi-square test is used to compare the two hourly distributions. The test p-value was 0.154, which is greater than 0.05, which confirms that these samples are not statically different. Despite the strong temporal pattern at higher collision site 1, it appears that the hourly distribution of collisions is overall not a particularly useful piece of information to distinguish the higher collision sites from lower collision sites.

6.3.6 Socio economics

The other aspects of the higher collision sites considered were the demographic and socio economic factors. This was not a part of the modeling process but could a factor helping distinguish higher and lower collision sites. Figure 78, Figure 79, and Figure 80 show the household income, education level, and average car ownership, respectively for sites 1, 2, and 3. Since sites 4, 5, 6, and 7 are very close to each other, one data point is shown for them in the graph. All these data then are compared to the average for Wake County.

![Image](image.png)

**Figure 78 – Household income**

128
Figure 79 – Education level

Figure 80 – Car ownership
One can see that the income level in sites 1 and 2 was lower than the Wake County average while site 3 had a higher income level and site 4, 5, 6, and 7 are close to the Wake County average. The same situation existed for the education level. The car ownership graph shows very steady car ownership among all higher collision sites near the average for Wake County. Overall it appears that the demographic and socio economic factors are not particularly useful to distinguish the higher collision sites from lower collision sites.

6.4 Conclusion

Quantitative, binary, and categorical variables that could possibly distinguish the higher collision sites were investigated. As a result I concluded that a full movement access point, a high driveway peak hour volume, a long intersection through movement Synchro queue, and a high left turn proportion from the driveway are the factors related to higher access – related collisions. Other factors such as driveway width, major road grade, corner clearance, speed limit, driveway median type, angle, radius, transition, lane configurations at the driveway and access point were shown generally not to help distinguish between higher collision and lower collision sites. The variable for the existence of a secondary driveway was not pursued as a contributing factor due to the lack of theoretical support. Collisions hourly distribution and socio economic factors were also tested but were not able to classify the higher collision sites from lower collision sites.

During the site visits I noticed some other factors that could have changed the research path if they were considered as variables during the modeling procedure. These factors were: a) lack of adequate sight distance at the access point, as observed in higher collision sites 2 and 5, b) driveway grade, which was certainly a factor for the higher collision site 5, c) the driveway at higher collision site 1 that was located on the same side as the PM peak directional traffic which caused over 90% of the access – related collisions, and d) driveway channelization form that was noticed to be a contributing factor at higher collision site 6. Small numbers of higher collision sites and consequently few samples with the described causing factors do not allow us to expand our conclusion and discover patterns with these variables; however.
7. **CONTRIBUTING FACTORS**

Site master plans are prepared by architects and site planners and usually it is too late to make changes by the time transportation and safety professionals are informed of the plan because the site is under construction or open and any change would require extra capital cost. Sadly, decision-makers often do not approve any site plan change before they see evidence of actual collisions.

It would be useful if there was checklist available to be used by decision-makers before finalizing the site plans to avoid possible collisions. In this section, as shown in Figure 81, I tested the access-related collision prediction model from Chapter 5 and the higher collision site results from Chapter 6 to try to provide site plan checklist related to driveways near signalized intersections for planners and decision-makers. At the end a conclusion that summarizes the policy checklist is presented.

![Contributing Factors](image)

**Figure 81 – Contributing factors overview**

### 7.1 Methodology

The classic hit and miss statistical method from information theory was used to specify policy checklist. In this method there are four regions including true positive (hit), false positive (type I error), false negative (type II error), and true negative (miss) (52). As shown in Table 37, in this research these regions show the relationships between actual and identified higher and lower collision sites.
Table 37 – Hit and miss statistical method

<table>
<thead>
<tr>
<th>Sites</th>
<th>Identified higher collision</th>
<th>Identified lower collision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual higher collision</td>
<td>True positive (hit)</td>
<td>False negative (type II error)</td>
</tr>
<tr>
<td>Actual lower collision</td>
<td>False positive (type I error)</td>
<td>True negative (miss)</td>
</tr>
</tbody>
</table>

We call a “hit” a site that is identified as higher collision and indeed will be a higher collision site. From safety point of view, we prefer to have a lower number in this cell but identifying the higher collision sites is an important safety matter because they can be fixed and the possible access – related collisions can be prevented in advance if predicted accurately.

Type I error sites are the sites identified as higher collision that will not have any safety issues in reality. It is possible that unnecessary countermeasures will be recommended and implemented at these locations, and the improvements would increase the project cost even though they are not particularly helpful. This cell is expected to have a zero value for an accurate prediction method.

Type II error sites are the ones that will have high collisions but for some reason were not identified as higher collision sites. This is the group that traffic engineers and safety analysts are very concerned about because the identifying method was not able to pick them up as higher collision sites and so lives will be at risk at these locations. The ultimate goal is having a zero value for this cell.

The last cell in the table is “miss” that refers to the sites that neither will be nor were identified as higher collision sites. In the other words there won’t be any safety issues at these locations and the prediction method can successfully identify them as lower collision sites.

7.2 Prediction model

It was stated earlier that all the 7 higher collision sites were operating with a full movement median. To find the altering factors these 7 higher collision sites were best compared to the remaining 70 full movement sites that have lower collisions. By substituting the required variables of major road AADT, driveway width, and intersection through movement Synchro 95 percentile queue from the 77 full movement sites into the Equation 49, the corresponding annual predicted access – related collisions’ mean and standard deviation were estimated 0.510 and
0.443, respectively. The 85 percentile is defined as a base to differentiate the higher collision sites (53). Mean plus one standard deviation represents the 85 percentile distribution, which is 0.95 annual access – related collisions in this research to identify higher collision sites.

Table 38 shows the actual higher and lower collision sites and those predicted using the model and a standard of mean plus one standard deviation. Although 5 hits are reasonable, there were still 2 type II error sites. These were the higher collision sites that were identified as lower collision sites using the model. There were also 8 false positive sites. These were not really higher collision sites but were identified as such. The 62 sites were true negative meaning they were neither actually nor were identified as high collisions sites.

<table>
<thead>
<tr>
<th>Sites</th>
<th>Identified higher collision</th>
<th>Identified lower collision</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual higher collision</td>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Actual lower collision</td>
<td>8</td>
<td>62</td>
<td>70</td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>64</td>
<td>77</td>
</tr>
</tbody>
</table>

At the 8 false positive sites improvements and countermeasures would likely be recommended but would not be necessary. It would be costly, but no safety risks would be involved at these locations. Safety analysts would be more concerned about the two false negative sites that the model would not be able to identify but would have caused safety issues.

**7.3 Higher collision site identification**

In this section higher collision sites are identified by applying the contributing factors discussed in Chapter 6 and using the statistical hypothesis test method described previously. Based on the information provided in Table 34, it was determined that quantitative variables of driveway volume and the intersection through movement Synchro 95 percentile queue were the primary factors contributing to the seven higher collision sites in this database. Among the 7 higher collision sites, the minimum value for driveway peak hour traffic volume was 120 vehicles per hour (or 2 vehicles per minute) and the minimum value for the intersection through movement Synchro 95% queue length was 230 feet. Assuming these minimum values as thresholds, among all the 77 full movement sites in the database the hit cell had 7 sites and there was no site identified as false negative. However there were 6 false positives while 64 sites were in the miss
cell. Although this result is better than the identification produced by applying the model, the unnecessary countermeasures at the 6 sites identified as higher collision sites that would not actually be higher collision sites would be costly.

At this stage I decided to consider the driveway left turn movement proportion as an additional discriminatory factor, since it proved to be a significant as shown in Table 35. As discussed earlier, driveway left turn counts were collected for a sample of 15 sites among lower collision full movement sites. Since these 15 sites were selected randomly, I expect that the driveway left turn proportion distribution for these 15 sites is the same as for all 70 lower collision full movement sites.

The minimum value of the left turn proportion at the 7 higher collision sites was 0.2. Assuming this value as threshold, all the 7 higher collision sites were identified as higher collision sites. Reviewing the 15 lower collision sites revealed that all of them were also identified as lower collision sites using this factor. Table 39 summarizes this analysis and shows that no false negative values would cause any safety hazardous issues in this set of 22 sites and subsequently no unnecessary improvements would be required because the false positive were estimated to be zero. To the extent that the left turn proportion behaves the same way over the full set of lower collision sites it appears to be a good discriminatory factor.

Table 39 – Higher collision sites hit and miss

<table>
<thead>
<tr>
<th>Sites</th>
<th>Identified higher collision</th>
<th>Identified lower collision</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual higher collision</td>
<td>7</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Actual lower collision</td>
<td>0</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
<td>15</td>
<td>22</td>
</tr>
</tbody>
</table>

7.4 Conclusion

In this section I recommend the use of factors that appear to contribute to higher collision driveways near the signalized intersections. This could be used by site planners and decision-makers to ensure that no hazardous driveway is constructed. Different methods were tested to conduct an appropriate recommendation.
Based on the results from previous sections, major road median type, driveway traffic volume, intersection through movement Synchro 95% queue, and driveway left turn proportion were the contributing factors to be considered. The driveway traffic volume is the Trip Generation Manual (41) estimation result usually done as a part of traffic impact analysis (TIA) as discussed in Section 3.2.2. The intersection peak hour through movement 95% queue can be estimated using traffic software, Synchro, using intersection turning movements and signal timing plans information. If the peak hour maximum length of the queue length is observed on site, it should be multiplied by a factor of 1.37 before plugging into the model.

The left turn proportion is the driveway left turn traffic volume divided by the major road AADT. The driveway left turn is also a part of information taken from TIA and the AADT is found using the Google Earth Pro or the NCDOT map sheets as discussed in Section 3.2.2.

If site planners and decision-makers consider these factors before construction, the chance of a higher number of access–related collisions at the access points near a signalized intersection should diminish greatly. Almost certainly a site within the scope of this research would turn out to be a higher collision site if all the criteria below are met:

1. The access point is planned to operate full movement on the major road, and
2. Driveway peak hour traffic volume is predicted to be over 120 vehicles per hour, and
3. Signalized intersection peak through movement Synchro 95% queue exceeds 230 feet, and
4. The driveway left turn proportion (100 × left turns from driveway per peak hour /AADT) is predicted to be over 0.2.

A site to be included in the list of higher collision sites should meet all the 4 criteria. Dropping even one criterion would increase the probability of error types I and II.

A site with these characteristics would likely lead to high collisions if not considered for further investigation and if safety countermeasures were not implemented. A site meeting all these contributing factors should be considered as a red flag, when site developers and decision-makers should inform or hire traffic or safety engineers. Further safety investigations and
implementing countermeasures would likely be needed to avoid future access – related collisions.

It was shown in the case studies section that some other factors such as sight distance, driveway and major road steep grade, and driveway channelization could have contributed to high collisions. Even though they were not included in this research checklist, they’re strongly recommended to be avoided.

It is important to know that driveways not meeting all thresholds may still benefit from good safety countermeasures. In other words, this research does not imply that all access points not meeting the thresholds are off the hook.

The potential solutions at an identified site could be limiting the access point to RIRO or providing another access point far from the intersection. This could reduce the traffic load on the access point of interest and consequently make it safer. Other countermeasures, including increasing the access point corner clearance, re-arranging the permissible land use type to control the driveway traffic volume more efficiently, and closing the access point, could also be considered.
8. CONCLUSIONS AND RECOMMENDATIONS

8.1 Research summary

According to NHTSA (2), every year over five million collisions occur in the US in which over two million people are injured. In 2009 alone 33,000 people died due to highway collisions. Over half of the total crashes are intersection and access point – related. Most collision reporting systems do not provide the necessary level of information to identify access – related collisions but collision data, where available, indicate a high incidence of access – related collisions. The objectives of this research were therefore developing a valid statistical model to estimate the number of access point – related collisions occurring at access points near signalized intersections and providing checklist for site planners and decision-makers to distinguish higher collision sites from lower collision sites and avoid constructing higher collision sites.

Over 200 intersections in Wake County, North Carolina were chosen randomly. Each intersection could consist of 3, 4, or 5 sites depending on the number of the legs. Rigid site selection criteria were implemented to find 108 appropriate sites on near 63 signalized intersections for this research. Geometric, traffic, and access – related collision data over a 5-year period were collected for the 108 sites. Since fatalities and injuries due to access point – related collisions are too infrequent to be analyzed alone, total access – related collisions were considered for analysis and prediction.

In the modeling process the first variable was the major road AADT and then 14 independent variables were investigated, including driveway traffic volume, corner clearance, the intersection thorough movement peak hour Synchro 95% queue, driveway width, lane configuration at the intersection, lane configuration on the major road at the driveway, major road grade, speed limit, major road median, driveway median status, driveway angle, driveway radius, transition between major road and driveway, and existence of another access point. These variables were introduced into the model one by one in a multiplicative form. This is based on a methodology using the negative multinomial likelihood function suggested by Hauer (36). The empirical integral function method was applied for the continuous variables and the one-way ANOVA was
considered for categorical (class) and binary variables to find the best functional forms. Then the
NLMIXED, a nonlinear optimization method, in SAS® was adopted for estimating parameters
using a negative binomial log likelihood function. Models were selected based on the -2 log
likelihood and BIC statistics. Finally the cumulative residuals (CURE) plot technique was used
to check the goodness of fit of the models.

To provide checklist for site planners to distinguish the higher collision sites from lower collision
sites, the data that were previously collected and some new information such as demographic and
socio economic were used. The higher collision sites were investigated one by one. Quantitative,
binary, categorical, demographic, and socio economic variables were analyzed and compared
between the higher and lower collision sites. Statistical tests were used to find the contributing
factors and provide checklist to certify no access points will be constructed before the safety
issues are considered.

8.2 Findings and conclusions

As a result of the modeling process, the access – related collision prediction model functional
form was estimated as:

\[ y = 0.000396 x_1^{0.9326} e^{-0.3748 x_1 x_2^{1.844}} (1 + 0.008532 x_3) \]  
(Equation 50)

Where \( y \) = annual access – related collisions at the access point near the signalized intersection,
\( x_1 = \) AADT/10,000 in vehicles per day, \( x_2 = \) driveway width in feet, and \( x_3 = 95\% \) peak hour
Synchro queue at the intersection for the through movement in feet. The standard errors for the
parameters, 0.000396, 0.9326, -0.3748, 1.841, 0.008532 were 0.001, 0.786, 0.463, 0.608, 0.011,
respectively. The model dispersion parameter, \( K \), was estimated as 0.268.

The model was developed using a dataset including the major road AADT ranging between
1,230 and 54,000 vehicles per day, driveway width ranging between 8 ft and 58 ft, and Synchro
through movement 95\% queue ranging between zero and 1,120 feet.

As an example, a site with a driveway near a signalized intersection, major road AADT of
19,000 vehicles per day, the driveway width of 20 feet, and the intersection through movement
peak hour 95% Synchro queue length of 704 feet would result in an estimated 0.6 access – related collisions per year.

Other variables such as the distance of the driveway from the signalized intersection were not included in the model but could still be practically important. The reason they were eliminated from modeling might be that the engineers and the policy makers at NCDOT are properly specifying corner clearance to reduce collisions.

Based on the detailed analysis of the seven higher collision sites conducted in Chapter 6, the research proposed policy checklist to find potential higher collision sites. The policy checklist proposed stated that an access point of concern was full movement, had a driveway peak hour volume of over 120 vehicles per hour, had an intersection peak hour through movement 95% Synchro queue of over 230 feet, and had a driveway left turn proportion (100× left turns from driveway per hour /AADT) of over 0.2.

If site planners and decision-makers consider these factors before construction, the chance of a higher number of access – related collision at the access point near a signalized intersection should diminish greatly. Almost certainly a site would turn out to be a higher collision site if all of the criteria are met.

It was shown in case studies section that some other factors such as sight distance, driveway and major road steep grade, and driveway channelization could have contributed to high collisions. Even though they were not included in the proposed checklist, they could be altering factors and so they are recommended to be checked. It is important to know that driveways not meeting all thresholds may still benefit from good safety countermeasures. In the other words this research does not imply that all access points not meeting the thresholds are off the hook.
8.3 Future research recommendations

This research made several contributions to the fields of safety and traffic engineering. However, due to the scope of the research, there were some restrictions. In this section I focused on the limitations and recommended future research paths.

This research focused only on the crashes that occurred at the access point. In reality when the left turn at the access point is restricted, drivers would make a U turn at the next intersection or median opening. This is shown in Figure 82. Future research is recommended to cover this issue and look at the access point issue as a whole and not isolated from the rest of the system.

![Figure 82 – Restricted left turn makes a U-turn at the intersection](image)

I did not consider pedestrian – related collisions. As a result the presented model would not be useful in very urbanized areas such as downtown or school campus. Future research could consider these effects on the collisions.

Researchers usually consider a limited number of variables in modeling. In this research I was able to examine the independent variable of access – related collisions and 15 dependent variables. Even though this variable set was more extensive than was considered by most previous research, I feel that other variables such as the land use type, sight distance, number of intersection legs, etc. could be introduced in future modeling.

In this research 108 sites were chosen from the available 674 sites in Wake County. This was done because the rest of the sites were not eligible to be included in our sample size. The reason behind this filtering was discussed in Section 3.1. One can notice that selecting 108 sites from
the initial 674 sites means that this research is based on almost 16% of the total original sample. Of course, many sites discarded did not even have driveways. The result I developed would likely apply directly to between ¼ and ½ of all driveways near to and upstream of signalized intersections in Wake County. More investigation on the remaining ½ to ¾ of the sites is needed to see if there are any ways that the methodology in this research can be expanded to cover the entire sample.

The location at which the collision occurred is an important piece of information and in this research the police collision report was the only resource to provide that. Not knowing the exact location of the collision could cause significant errors in developing the collision models. Unfortunately the police collision reports are not very accurate in providing a good estimation on the location for the purpose of this particular research. Considering that almost all police vehicles are now equipped with a laptop or navigator device, I hope that accurate geographic information such as latitude and longitude be included in the future police collision reports in NC and other states to help researchers.

One other source of limitation and possible error that could be fixed in future research is the methodology for filtering collisions to find the access – related collisions. For the purpose of this research I decided to specify a few criteria such as: 1) the narrative or diagram of the collision indicated that at least one of the vehicles involving the collision was clearly headed to or from the driveway, 2) any word like "driveway" or "access point" was in the narrative section of the police report, or 3) an indication in the police report that states one of the vehicles was making a turn to or from the driveway. The problem was that even with all that effort, the access – related collisions may not be accurate. For example, if a vehicle was entering the driveway, the following vehicles may stop or slow down and as a result collide each other. Usually police reports state that the vehicles slowed down or stopped because of traffic ahead but don’t mention the reason. Indeed, traffic ahead could be the queue backed up at the intersection or a vehicle maneuvering at the driveway. I didn’t have any way of recognizing these types of collisions as access – related collisions and do not know how to fix this issue. Further investigation is required to find a solution for this possible source of error.
During the site visits I noticed some other factors that could have changed the research path if they were considered as variables during the modeling procedure. These factors were lack of adequate sight distance at the access point, driveway grade, and driveway channelization. Another potential factor could be because the driveway was located on the same side as the peak directional traffic. In this research small numbers of higher collision sites and consequently few cases of the described factors did not allow me to expand the conclusion and discover patterns towards these variables, but they could be considered in the future research.

The access–related collision model provided in this research can be used to develop a collision modification factor for the existence of an access point near signalized intersection. The Highway Safety Manual (54) provides a model to predict the number of collisions occurring at intersections. Using a modification factor, one can bring access points into the equation and compare the safety of access points in different circumstances.
References


40. Provided by the NCDOT, June 2010.


Appendices
Appendix A – Distribution of the study sites in Wake County
Appendix B – Google Earth Pro daily traffic counts
Appendix C – Traffic counts data collection method
Appendix D – Crash Report Form DMV-349
### Appendix F – Summary of study sites and collected data in Wake County

<table>
<thead>
<tr>
<th>Intersection #</th>
<th>LCD</th>
<th>CC</th>
<th>ARC</th>
<th>MV</th>
<th>Q_T</th>
<th>AD</th>
<th>LCI</th>
<th>MM</th>
<th>SL</th>
<th>GR</th>
<th>DA</th>
<th>DW</th>
<th>DM</th>
<th>TR</th>
<th>RA</th>
<th>Near Intersection of</th>
</tr>
</thead>
<tbody>
<tr>
<td>05-0015</td>
<td>6</td>
<td>112</td>
<td>1.0</td>
<td>2.30</td>
<td>186</td>
<td>322</td>
<td>No</td>
<td>5</td>
<td>RIRO 45</td>
<td>2</td>
<td>90</td>
<td>58</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>SR 1793 (Glenwood Avenue) at Oberlin Road</td>
</tr>
<tr>
<td>05-0025</td>
<td>1</td>
<td>148</td>
<td>0.8</td>
<td>0.42</td>
<td>66</td>
<td>36</td>
<td>No</td>
<td>2</td>
<td>Full 35</td>
<td>-3</td>
<td>90</td>
<td>33</td>
<td>Undivided</td>
<td>Elevated</td>
<td>Small</td>
<td>SR 1676 (Dixie Trail) at SR 3007 (Hillsborough Street)</td>
</tr>
<tr>
<td>05-0025</td>
<td>1</td>
<td>28</td>
<td>0.2</td>
<td>0.42</td>
<td>32</td>
<td>89</td>
<td>No</td>
<td>2</td>
<td>Full 25</td>
<td>-3</td>
<td>90</td>
<td>20</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>SR 1676 (Dixie Trail) at SR 3007 (Hillsborough Street)</td>
</tr>
<tr>
<td>05-0116</td>
<td>1</td>
<td>50</td>
<td>0.0</td>
<td>0.30</td>
<td>4</td>
<td>45</td>
<td>No</td>
<td>2</td>
<td>Full 35</td>
<td>1</td>
<td>55</td>
<td>30</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>SR 1564 (Wilmington Street) at Layden Street and Pecan Road</td>
</tr>
<tr>
<td>05-0116</td>
<td>5</td>
<td>110</td>
<td>0.0</td>
<td>0.87</td>
<td>4</td>
<td>55</td>
<td>No</td>
<td>1</td>
<td>RIRO 45</td>
<td>-1</td>
<td>90</td>
<td>27</td>
<td>Undivided</td>
<td>Elevated</td>
<td>Small</td>
<td>SR 1564 (Wilmington Street) at Layden Street and Pecan Road</td>
</tr>
<tr>
<td>05-0139</td>
<td>6</td>
<td>76</td>
<td>0.4</td>
<td>2.90</td>
<td>26</td>
<td>264</td>
<td>No</td>
<td>5</td>
<td>RIRO 45</td>
<td>2</td>
<td>90</td>
<td>26</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>US 401 at SR 2036-2108 (New Hope Road)</td>
</tr>
<tr>
<td>05-0162</td>
<td>1</td>
<td>93</td>
<td>0.0</td>
<td>0.42</td>
<td>7</td>
<td>149</td>
<td>Yes</td>
<td>2</td>
<td>Full 25</td>
<td>1</td>
<td>90</td>
<td>8</td>
<td>Undivided</td>
<td>Elevated</td>
<td>Small</td>
<td>US 401 - NC 42 (Main Street) / US 401 at SR 1107/ NC 42 (Academy Street)</td>
</tr>
<tr>
<td>05-0162</td>
<td>1</td>
<td>141</td>
<td>0.4</td>
<td>1.00</td>
<td>191</td>
<td>131</td>
<td>Yes</td>
<td>2</td>
<td>Full 25</td>
<td>1</td>
<td>90</td>
<td>19</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
<td>US 401 - NC 42 (Main Street) / US 401 at SR 1107/ NC 42 (Academy Street)</td>
</tr>
<tr>
<td>05-0162</td>
<td>4</td>
<td>83</td>
<td>0.6</td>
<td>1.55</td>
<td>88</td>
<td>197</td>
<td>No</td>
<td>3</td>
<td>Full 25</td>
<td>-2</td>
<td>90</td>
<td>50</td>
<td>Undivided</td>
<td>Elevated</td>
<td>Large</td>
<td>US 401 - NC 42 (Main Street) / US 401 at SR 1107/ NC 42 (Academy Street)</td>
</tr>
<tr>
<td>05-0201</td>
<td>2</td>
<td>42</td>
<td>0.2</td>
<td>0.30</td>
<td>100</td>
<td>46</td>
<td>No</td>
<td>4</td>
<td>Full 35</td>
<td>1</td>
<td>90</td>
<td>29</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>US 70 (Glenwood Avenue) at SR 1741 (Holly Ridge Road)</td>
</tr>
<tr>
<td>05-0212</td>
<td>4</td>
<td>353</td>
<td>6.4</td>
<td>1.55</td>
<td>138</td>
<td>238</td>
<td>Yes</td>
<td>3</td>
<td>Full 35</td>
<td>1</td>
<td>90</td>
<td>44</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>SR 1009-1370 (Tryon Road) at SR 1371 (Lake Wheeler Road)</td>
</tr>
<tr>
<td>05-0225</td>
<td>3</td>
<td>103</td>
<td>0.0</td>
<td>1.50</td>
<td>40</td>
<td>258</td>
<td>No</td>
<td>1</td>
<td>RIRO 45</td>
<td>1</td>
<td>90</td>
<td>23</td>
<td>Undivided</td>
<td>Elevated</td>
<td>Small</td>
<td>SR 2911 (New Bern Avenue) at SR 2921 (Raleigh Boulevard)</td>
</tr>
<tr>
<td>05-0258</td>
<td>5</td>
<td>208</td>
<td>0.8</td>
<td>2.45</td>
<td>68</td>
<td>394</td>
<td>Yes</td>
<td>1</td>
<td>Full 35</td>
<td>-1</td>
<td>70</td>
<td>27</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
<td>SR 3007 (Hillsborough Street) at Faircloth Street / Gorman Street</td>
</tr>
<tr>
<td>05-0262</td>
<td>1</td>
<td>185</td>
<td>0.0</td>
<td>0.39</td>
<td>25</td>
<td>87</td>
<td>No</td>
<td>2</td>
<td>Full 35</td>
<td>2</td>
<td>115</td>
<td>20</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>SR 2921 (Raleigh Boulevard) at Oakwood Avenue</td>
</tr>
<tr>
<td>05-0269</td>
<td>4</td>
<td>296</td>
<td>0.6</td>
<td>1.50</td>
<td>77</td>
<td>523</td>
<td>No</td>
<td>3</td>
<td>Full 35</td>
<td>2</td>
<td>90</td>
<td>23</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>SR 1011 (East Chatham Street) at SR 1415 (NE. / SE. Maynard Road)</td>
</tr>
<tr>
<td>05-0387</td>
<td>1</td>
<td>208</td>
<td>0.0</td>
<td>1.05</td>
<td>304</td>
<td>165</td>
<td>Yes</td>
<td>2</td>
<td>Full 25</td>
<td>-3</td>
<td>90</td>
<td>34</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>SR 1728 (Wade Avenue) at Faircloth Street / Commercial Drive</td>
</tr>
<tr>
<td>05-0417</td>
<td>2</td>
<td>278</td>
<td>0.2</td>
<td>2.80</td>
<td>68</td>
<td>35</td>
<td>Yes</td>
<td>4</td>
<td>Full 35</td>
<td>0</td>
<td>90</td>
<td>32</td>
<td>Undivided</td>
<td>Elevated</td>
<td>Small</td>
<td>SR 2000 (Old Wake Forest Road) at Creekside Drive</td>
</tr>
<tr>
<td>05-0417</td>
<td>5</td>
<td>275</td>
<td>0.4</td>
<td>2.90</td>
<td>281</td>
<td>137</td>
<td>No</td>
<td>1</td>
<td>Full 35</td>
<td>0</td>
<td>90</td>
<td>37</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>SR 2000 (Old Wake Forest Road) at Creekside Drive</td>
</tr>
<tr>
<td>05-0444</td>
<td>6</td>
<td>137</td>
<td>0.2</td>
<td>1.85</td>
<td>46</td>
<td>165</td>
<td>No</td>
<td>5</td>
<td>RIRO 45</td>
<td>3</td>
<td>120</td>
<td>25</td>
<td>Undivided</td>
<td>Elevated</td>
<td>Large</td>
<td>NC 50 (Creedmoor Road) at SR 1822 (Leesville Road/Millbrook Road)</td>
</tr>
<tr>
<td>05-0444</td>
<td>6</td>
<td>131</td>
<td>0.8</td>
<td>1.90</td>
<td>255</td>
<td>176</td>
<td>Yes</td>
<td>5</td>
<td>RIRO 45</td>
<td>-1</td>
<td>90</td>
<td>25</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
<td>NC 50 (Creedmoor Road) at SR 1822 (Leesville Road/Millbrook Road)</td>
</tr>
<tr>
<td>05-0460</td>
<td>1</td>
<td>155</td>
<td>0.6</td>
<td>0.37</td>
<td>115</td>
<td>130</td>
<td>No</td>
<td>2</td>
<td>Full 25</td>
<td>2</td>
<td>125</td>
<td>31</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
<td>SR 2034 (New Hope Church Road) at Brentwood Road</td>
</tr>
<tr>
<td>05-0460</td>
<td>5</td>
<td>684</td>
<td>1.4</td>
<td>2.15</td>
<td>154</td>
<td>278</td>
<td>No</td>
<td>1</td>
<td>Full 35</td>
<td>2</td>
<td>90</td>
<td>37</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
<td>SR 2034 (New Hope Church Road) at Brentwood Road</td>
</tr>
<tr>
<td>05-0606</td>
<td>6</td>
<td>254</td>
<td>1.0</td>
<td>5.40</td>
<td>190</td>
<td>927</td>
<td>No</td>
<td>5</td>
<td>RIRO 55</td>
<td>-2</td>
<td>90</td>
<td>21</td>
<td>Divided</td>
<td>Not elevated</td>
<td>Large</td>
<td>US 1 (Capital Boulevard) at SR 2006 (Durant Road) and Perry Creek Road</td>
</tr>
<tr>
<td>05-0608</td>
<td>2</td>
<td>323</td>
<td>0.6</td>
<td>0.88</td>
<td>76</td>
<td>288</td>
<td>No</td>
<td>2</td>
<td>Full 45</td>
<td>1</td>
<td>90</td>
<td>33</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>NC 55 (Williams Street) at SR 1160 (Olive Chapel Road) and Hunter Street</td>
</tr>
<tr>
<td>05-0608</td>
<td>4</td>
<td>102</td>
<td>0.0</td>
<td>0.98</td>
<td>68</td>
<td>224</td>
<td>No</td>
<td>3</td>
<td>RIRO 35</td>
<td>2</td>
<td>90</td>
<td>26</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>NC 55 (Williams Street) at SR 1160 (Olive Chapel Road) and Hunter Street</td>
</tr>
<tr>
<td>05-0608</td>
<td>1</td>
<td>414</td>
<td>0.6</td>
<td>2.00</td>
<td>28</td>
<td>432</td>
<td>No</td>
<td>2</td>
<td>Full 45</td>
<td>1</td>
<td>90</td>
<td>20</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>NC 55 (Williams Street) at SR 1160 (Olive Chapel Road) and Hunter Street</td>
</tr>
<tr>
<td>05-0638</td>
<td>5</td>
<td>260</td>
<td>0.0</td>
<td>0.46</td>
<td>205</td>
<td>0</td>
<td>No</td>
<td>4</td>
<td>Full 35</td>
<td>-1</td>
<td>90</td>
<td>37</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
<td>SR 2030 (Old Wake Forest Road) at Navaho Drive</td>
</tr>
<tr>
<td>Intersection #</td>
<td>LCD</td>
<td>CC</td>
<td>ARC</td>
<td>MV</td>
<td>DV</td>
<td>Q_T</td>
<td>AD</td>
<td>LQ</td>
<td>MM</td>
<td>SL</td>
<td>GR</td>
<td>DA</td>
<td>DW</td>
<td>DM</td>
<td>TR</td>
<td>RA</td>
</tr>
<tr>
<td>---------------</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>05-0816</td>
<td>2</td>
<td>155</td>
<td>0.8</td>
<td>0.51</td>
<td>87</td>
<td>210</td>
<td>No</td>
<td>2</td>
<td>Full</td>
<td>35</td>
<td>1</td>
<td>90</td>
<td>18</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-0816</td>
<td>1</td>
<td>286</td>
<td>0.2</td>
<td>0.56</td>
<td>66</td>
<td>167</td>
<td>No</td>
<td>7</td>
<td>Full</td>
<td>35</td>
<td>-1</td>
<td>90</td>
<td>26</td>
<td>Divided</td>
<td>Elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-0840</td>
<td>3</td>
<td>162</td>
<td>4.2</td>
<td>2.15</td>
<td>121</td>
<td>366</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>35</td>
<td>2</td>
<td>90</td>
<td>31</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-0840</td>
<td>6</td>
<td>235</td>
<td>0.8</td>
<td>2.45</td>
<td>8</td>
<td>1116</td>
<td>No</td>
<td>5</td>
<td>Full</td>
<td>35</td>
<td>3</td>
<td>90</td>
<td>25</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-0857</td>
<td>6</td>
<td>86</td>
<td>0.2</td>
<td>1.60</td>
<td>40</td>
<td>401</td>
<td>No</td>
<td>5</td>
<td>RIRO</td>
<td>45</td>
<td>-1</td>
<td>90</td>
<td>19</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-0857</td>
<td>6</td>
<td>117</td>
<td>0.4</td>
<td>1.90</td>
<td>40</td>
<td>406</td>
<td>No</td>
<td>5</td>
<td>RIRO</td>
<td>45</td>
<td>1</td>
<td>90</td>
<td>38</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-0879</td>
<td>3</td>
<td>120</td>
<td>0.8</td>
<td>0.79</td>
<td>117</td>
<td>123</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>35</td>
<td>-2</td>
<td>90</td>
<td>22</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-0935</td>
<td>3</td>
<td>242</td>
<td>3.4</td>
<td>1.30</td>
<td>284</td>
<td>432</td>
<td>No</td>
<td>3</td>
<td>Full</td>
<td>45</td>
<td>-2</td>
<td>90</td>
<td>43</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-0966</td>
<td>2</td>
<td>272</td>
<td>0.4</td>
<td>1.53</td>
<td>12</td>
<td>296</td>
<td>No</td>
<td>4</td>
<td>Full</td>
<td>35</td>
<td>-2</td>
<td>90</td>
<td>23</td>
<td>Undivided</td>
<td>Elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-0983</td>
<td>2</td>
<td>150</td>
<td>0.2</td>
<td>0.41</td>
<td>60</td>
<td>101</td>
<td>No</td>
<td>3</td>
<td>RIRO</td>
<td>35</td>
<td>-1</td>
<td>90</td>
<td>26</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-0983</td>
<td>1</td>
<td>212</td>
<td>0.0</td>
<td>1.02</td>
<td>49</td>
<td>183</td>
<td>No</td>
<td>3</td>
<td>Full</td>
<td>35</td>
<td>3</td>
<td>77</td>
<td>31</td>
<td>Divided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-0983</td>
<td>4</td>
<td>45</td>
<td>0.0</td>
<td>1.15</td>
<td>38</td>
<td>121</td>
<td>No</td>
<td>3</td>
<td>RIRO</td>
<td>35</td>
<td>-3</td>
<td>90</td>
<td>33</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-0999</td>
<td>2</td>
<td>696</td>
<td>0.0</td>
<td>2.80</td>
<td>197</td>
<td>197</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>3</td>
<td>115</td>
<td>26</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1100</td>
<td>1</td>
<td>314</td>
<td>0.0</td>
<td>0.42</td>
<td>11</td>
<td>262</td>
<td>No</td>
<td>2</td>
<td>Full</td>
<td>35</td>
<td>-2</td>
<td>90</td>
<td>33</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1100</td>
<td>5</td>
<td>189</td>
<td>0.2</td>
<td>1.10</td>
<td>16</td>
<td>425</td>
<td>No</td>
<td>2</td>
<td>Full</td>
<td>35</td>
<td>3</td>
<td>90</td>
<td>29</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1100</td>
<td>3</td>
<td>504</td>
<td>0.0</td>
<td>1.35</td>
<td>29</td>
<td>280</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>-2</td>
<td>90</td>
<td>25</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1107</td>
<td>3</td>
<td>219</td>
<td>0.2</td>
<td>0.60</td>
<td>280</td>
<td>101</td>
<td>No</td>
<td>4</td>
<td>Full</td>
<td>35</td>
<td>0</td>
<td>75</td>
<td>42</td>
<td>Divided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1118</td>
<td>1</td>
<td>157</td>
<td>0.0</td>
<td>1.02</td>
<td>55</td>
<td>324</td>
<td>No</td>
<td>2</td>
<td>Full</td>
<td>35</td>
<td>2</td>
<td>65</td>
<td>26</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1120</td>
<td>2</td>
<td>179</td>
<td>0.0</td>
<td>1.10</td>
<td>51</td>
<td>132</td>
<td>No</td>
<td>2</td>
<td>Full</td>
<td>35</td>
<td>1</td>
<td>75</td>
<td>19</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1120</td>
<td>3</td>
<td>152</td>
<td>0.8</td>
<td>2.70</td>
<td>69</td>
<td>813</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>-2</td>
<td>90</td>
<td>20</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1120</td>
<td>3</td>
<td>250</td>
<td>0.6</td>
<td>3.50</td>
<td>129</td>
<td>465</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>2</td>
<td>125</td>
<td>17</td>
<td>Divided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1122</td>
<td>4</td>
<td>184</td>
<td>0.4</td>
<td>1.15</td>
<td>235</td>
<td>325</td>
<td>No</td>
<td>3</td>
<td>Full</td>
<td>45</td>
<td>-3</td>
<td>100</td>
<td>30</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1129</td>
<td>1</td>
<td>85</td>
<td>0.4</td>
<td>0.94</td>
<td>21</td>
<td>377</td>
<td>Yes</td>
<td>2</td>
<td>Full</td>
<td>45</td>
<td>2</td>
<td>90</td>
<td>42</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1157</td>
<td>3</td>
<td>603</td>
<td>0.4</td>
<td>2.58</td>
<td>26</td>
<td>157</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>2</td>
<td>90</td>
<td>23</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1183</td>
<td>1</td>
<td>40</td>
<td>0.0</td>
<td>0.49</td>
<td>118</td>
<td>249</td>
<td>No</td>
<td>2</td>
<td>Full</td>
<td>35</td>
<td>0</td>
<td>70</td>
<td>28</td>
<td>Undivided</td>
<td>Elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1183</td>
<td>3</td>
<td>158</td>
<td>0.8</td>
<td>2.58</td>
<td>30</td>
<td>354</td>
<td>Yes</td>
<td>1</td>
<td>Full</td>
<td>35</td>
<td>-2</td>
<td>90</td>
<td>44</td>
<td>Divided</td>
<td>Elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1183</td>
<td>3</td>
<td>311</td>
<td>0.6</td>
<td>2.79</td>
<td>15</td>
<td>282</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>0</td>
<td>115</td>
<td>30</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>Intersection #</td>
<td>LCD</td>
<td>CC</td>
<td>ARC</td>
<td>MV</td>
<td>DV</td>
<td>Q_T</td>
<td>AD</td>
<td>LCI</td>
<td>MM</td>
<td>SL</td>
<td>GR</td>
<td>DA</td>
<td>DW</td>
<td>DM</td>
<td>TR</td>
<td>RA</td>
</tr>
<tr>
<td>----------------</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>05-1185</td>
<td>5</td>
<td>211</td>
<td>0.2</td>
<td>0.34</td>
<td>246</td>
<td>46</td>
<td>No</td>
<td>4</td>
<td>Full</td>
<td>35</td>
<td>-1</td>
<td>90</td>
<td>25</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1185</td>
<td>3</td>
<td>516</td>
<td>0.0</td>
<td>3.84</td>
<td>192</td>
<td>632</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>-1</td>
<td>90</td>
<td>33</td>
<td>Divided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1185</td>
<td>3</td>
<td>343</td>
<td>0.0</td>
<td>4.03</td>
<td>69</td>
<td>337</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>0</td>
<td>90</td>
<td>24</td>
<td>Undivided</td>
<td>Elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1188</td>
<td>3</td>
<td>241</td>
<td>0.0</td>
<td>2.00</td>
<td>74</td>
<td>283</td>
<td>No</td>
<td>3</td>
<td>Full</td>
<td>45</td>
<td>3</td>
<td>90</td>
<td>21</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1202</td>
<td>1</td>
<td>135</td>
<td>0.0</td>
<td>0.40</td>
<td>6</td>
<td>460</td>
<td>Yes</td>
<td>2</td>
<td>Full</td>
<td>45</td>
<td>-1</td>
<td>90</td>
<td>25</td>
<td>Undivided</td>
<td>Elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1270</td>
<td>4</td>
<td>101</td>
<td>0.2</td>
<td>1.15</td>
<td>53</td>
<td>698</td>
<td>No</td>
<td>3</td>
<td>Full</td>
<td>35</td>
<td>-1</td>
<td>125</td>
<td>41</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1270</td>
<td>5</td>
<td>328</td>
<td>1.0</td>
<td>1.15</td>
<td>159</td>
<td>632</td>
<td>No</td>
<td>6</td>
<td>Full</td>
<td>35</td>
<td>3</td>
<td>120</td>
<td>35</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1297</td>
<td>5</td>
<td>210</td>
<td>0.0</td>
<td>0.12</td>
<td>32</td>
<td>76</td>
<td>No</td>
<td>4</td>
<td>Full</td>
<td>25</td>
<td>-1</td>
<td>115</td>
<td>22</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1357</td>
<td>4</td>
<td>242</td>
<td>0.8</td>
<td>1.10</td>
<td>54</td>
<td>638</td>
<td>No</td>
<td>3</td>
<td>Full</td>
<td>55</td>
<td>2</td>
<td>90</td>
<td>35</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1380</td>
<td>5</td>
<td>220</td>
<td>0.0</td>
<td>0.33</td>
<td>34</td>
<td>0</td>
<td>Yes</td>
<td>2</td>
<td>Full</td>
<td>15</td>
<td>2</td>
<td>100</td>
<td>30</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1402</td>
<td>1</td>
<td>522</td>
<td>0.0</td>
<td>0.24</td>
<td>290</td>
<td>219</td>
<td>No</td>
<td>2</td>
<td>Full</td>
<td>55</td>
<td>-2</td>
<td>90</td>
<td>32</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1529</td>
<td>3</td>
<td>200</td>
<td>0.2</td>
<td>1.60</td>
<td>134</td>
<td>569</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>0</td>
<td>100</td>
<td>22</td>
<td>Divided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1558</td>
<td>1</td>
<td>314</td>
<td>0.6</td>
<td>0.54</td>
<td>153</td>
<td>155</td>
<td>Yes</td>
<td>6</td>
<td>Full</td>
<td>35</td>
<td>-2</td>
<td>90</td>
<td>42</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1599</td>
<td>9</td>
<td>90</td>
<td>0.6</td>
<td>1.35</td>
<td>84</td>
<td>655</td>
<td>No</td>
<td>3</td>
<td>RIRO</td>
<td>35</td>
<td>2</td>
<td>90</td>
<td>40</td>
<td>Undivided</td>
<td>Elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1559</td>
<td>1</td>
<td>185</td>
<td>2.8</td>
<td>1.55</td>
<td>154</td>
<td>735</td>
<td>Yes</td>
<td>3</td>
<td>Full</td>
<td>35</td>
<td>2</td>
<td>90</td>
<td>26</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1559</td>
<td>3</td>
<td>296</td>
<td>3.2</td>
<td>2.25</td>
<td>140</td>
<td>899</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>2</td>
<td>90</td>
<td>25</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1559</td>
<td>6</td>
<td>150</td>
<td>3.0</td>
<td>2.25</td>
<td>135</td>
<td>662</td>
<td>Yes</td>
<td>5</td>
<td>Full</td>
<td>45</td>
<td>0</td>
<td>105</td>
<td>30</td>
<td>Divided</td>
<td>Elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1561</td>
<td>6</td>
<td>250</td>
<td>0.0</td>
<td>3.10</td>
<td>121</td>
<td>212</td>
<td>No</td>
<td>5</td>
<td>RIRO</td>
<td>35</td>
<td>-1</td>
<td>90</td>
<td>20</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1565</td>
<td>3</td>
<td>462</td>
<td>3.6</td>
<td>1.50</td>
<td>186</td>
<td>412</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>2</td>
<td>90</td>
<td>30</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1565</td>
<td>3</td>
<td>716</td>
<td>0.0</td>
<td>2.00</td>
<td>29</td>
<td>604</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>0</td>
<td>90</td>
<td>22</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1565</td>
<td>3</td>
<td>190</td>
<td>0.0</td>
<td>2.00</td>
<td>45</td>
<td>332</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>1</td>
<td>110</td>
<td>21</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1578</td>
<td>5</td>
<td>346</td>
<td>0.2</td>
<td>2.15</td>
<td>19</td>
<td>667</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>-1</td>
<td>90</td>
<td>22</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
<tr>
<td>05-1615</td>
<td>1</td>
<td>390</td>
<td>0.0</td>
<td>0.43</td>
<td>62</td>
<td>113</td>
<td>No</td>
<td>2</td>
<td>Full</td>
<td>35</td>
<td>-2</td>
<td>100</td>
<td>30</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1615</td>
<td>3</td>
<td>396</td>
<td>0.0</td>
<td>1.85</td>
<td>22</td>
<td>175</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>-3</td>
<td>95</td>
<td>36</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Small</td>
</tr>
<tr>
<td>05-1692</td>
<td>5</td>
<td>268</td>
<td>1.0</td>
<td>2.00</td>
<td>118</td>
<td>506</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>-3</td>
<td>90</td>
<td>20</td>
<td>Undivided</td>
<td>Not elevated</td>
<td>Large</td>
</tr>
</tbody>
</table>

167
<table>
<thead>
<tr>
<th>Intersection #</th>
<th>LCD</th>
<th>CC</th>
<th>ARC</th>
<th>MV</th>
<th>DV</th>
<th>Q_T</th>
<th>AD</th>
<th>SL</th>
<th>GR</th>
<th>DA</th>
<th>DW</th>
<th>DM</th>
<th>TR</th>
<th>RA</th>
<th>Near Intersection of</th>
</tr>
</thead>
<tbody>
<tr>
<td>05-1692</td>
<td>5</td>
<td>241</td>
<td>0.4</td>
<td>2.35</td>
<td>123</td>
<td>523</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>2</td>
<td>70</td>
<td>32</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1692</td>
<td>6</td>
<td>285</td>
<td>1.2</td>
<td>2.55</td>
<td>224</td>
<td>473</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>2</td>
<td>90</td>
<td>20</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1692</td>
<td>5</td>
<td>233</td>
<td>0.0</td>
<td>2.60</td>
<td>92</td>
<td>485</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>-3</td>
<td>90</td>
<td>20</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1697</td>
<td>4</td>
<td>220</td>
<td>0.0</td>
<td>0.55</td>
<td>98</td>
<td>85</td>
<td>No</td>
<td>3</td>
<td>Full</td>
<td>35</td>
<td>0</td>
<td>90</td>
<td>20</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1715</td>
<td>1</td>
<td>60</td>
<td>0.0</td>
<td>0.14</td>
<td>4</td>
<td>52</td>
<td>No</td>
<td>2</td>
<td>Full</td>
<td>35</td>
<td>-1</td>
<td>70</td>
<td>13</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1748</td>
<td>3</td>
<td>234</td>
<td>0.6</td>
<td>2.00</td>
<td>30</td>
<td>333</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>-3</td>
<td>90</td>
<td>20</td>
<td>Divided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1748</td>
<td>5</td>
<td>380</td>
<td>0.4</td>
<td>2.00</td>
<td>122</td>
<td>344</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>-2</td>
<td>90</td>
<td>28</td>
<td>Divided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1797</td>
<td>4</td>
<td>269</td>
<td>0.0</td>
<td>0.31</td>
<td>106</td>
<td>89</td>
<td>No</td>
<td>3</td>
<td>Full</td>
<td>35</td>
<td>2</td>
<td>100</td>
<td>25</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1814</td>
<td>2</td>
<td>218</td>
<td>0.0</td>
<td>1.25</td>
<td>63</td>
<td>281</td>
<td>No</td>
<td>4</td>
<td>Full</td>
<td>45</td>
<td>-2</td>
<td>90</td>
<td>17</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1814</td>
<td>3</td>
<td>437</td>
<td>0.0</td>
<td>2.35</td>
<td>80</td>
<td>148</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>2</td>
<td>90</td>
<td>25</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1865</td>
<td>5</td>
<td>280</td>
<td>0.0</td>
<td>0.50</td>
<td>45</td>
<td>32</td>
<td>No</td>
<td>2</td>
<td>Full</td>
<td>35</td>
<td>-3</td>
<td>90</td>
<td>30</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1865</td>
<td>5</td>
<td>457</td>
<td>0.0</td>
<td>1.54</td>
<td>28</td>
<td>119</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>-2</td>
<td>110</td>
<td>24</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1865</td>
<td>3</td>
<td>185</td>
<td>0.0</td>
<td>1.66</td>
<td>11</td>
<td>212</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>1</td>
<td>90</td>
<td>30</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1879</td>
<td>4</td>
<td>166</td>
<td>0.2</td>
<td>1.02</td>
<td>96</td>
<td>157</td>
<td>No</td>
<td>3</td>
<td>RIRO</td>
<td>45</td>
<td>1</td>
<td>115</td>
<td>23</td>
<td>Divided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1937</td>
<td>1</td>
<td>580</td>
<td>0.0</td>
<td>1.25</td>
<td>82</td>
<td>979</td>
<td>No</td>
<td>2</td>
<td>Full</td>
<td>45</td>
<td>1</td>
<td>90</td>
<td>22</td>
<td>Undivided</td>
<td>Elevated</td>
</tr>
<tr>
<td>05-1937</td>
<td>5</td>
<td>470</td>
<td>0.0</td>
<td>1.85</td>
<td>40</td>
<td>785</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>1</td>
<td>70</td>
<td>24</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1937</td>
<td>3</td>
<td>215</td>
<td>0.4</td>
<td>2.45</td>
<td>67</td>
<td>227</td>
<td>No</td>
<td>1</td>
<td>RIRO</td>
<td>45</td>
<td>1</td>
<td>85</td>
<td>20</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-1990</td>
<td>5</td>
<td>640</td>
<td>0.2</td>
<td>3.10</td>
<td>251</td>
<td>541</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>2</td>
<td>90</td>
<td>28</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-2055</td>
<td>2</td>
<td>174</td>
<td>0.2</td>
<td>0.41</td>
<td>37</td>
<td>45</td>
<td>No</td>
<td>4</td>
<td>Full</td>
<td>35</td>
<td>-3</td>
<td>120</td>
<td>25</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-2055</td>
<td>3</td>
<td>555</td>
<td>0.0</td>
<td>1.85</td>
<td>24</td>
<td>91</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>2</td>
<td>90</td>
<td>26</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-2055</td>
<td>3</td>
<td>529</td>
<td>0.2</td>
<td>2.00</td>
<td>5</td>
<td>205</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>-2</td>
<td>100</td>
<td>25</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-2080</td>
<td>3</td>
<td>500</td>
<td>0.0</td>
<td>1.36</td>
<td>117</td>
<td>339</td>
<td>No</td>
<td>4</td>
<td>Full</td>
<td>35</td>
<td>2</td>
<td>95</td>
<td>35</td>
<td>Divided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-2080</td>
<td>1</td>
<td>70</td>
<td>0.2</td>
<td>2.00</td>
<td>12</td>
<td>291</td>
<td>No</td>
<td>2</td>
<td>Full</td>
<td>45</td>
<td>-2</td>
<td>115</td>
<td>17</td>
<td>Undivided</td>
<td>Elevated</td>
</tr>
<tr>
<td>05-2080</td>
<td>5</td>
<td>260</td>
<td>0.0</td>
<td>2.00</td>
<td>180</td>
<td>211</td>
<td>No</td>
<td>4</td>
<td>RIRO</td>
<td>45</td>
<td>2</td>
<td>115</td>
<td>22</td>
<td>Divided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-2099</td>
<td>3</td>
<td>198</td>
<td>0.0</td>
<td>1.60</td>
<td>37</td>
<td>132</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>3</td>
<td>90</td>
<td>21</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-2099</td>
<td>3</td>
<td>208</td>
<td>0.0</td>
<td>1.60</td>
<td>75</td>
<td>195</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>45</td>
<td>-3</td>
<td>90</td>
<td>20</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
<tr>
<td>05-2132</td>
<td>3</td>
<td>460</td>
<td>0.2</td>
<td>1.44</td>
<td>140</td>
<td>35</td>
<td>No</td>
<td>1</td>
<td>Full</td>
<td>35</td>
<td>0</td>
<td>80</td>
<td>30</td>
<td>Undivided</td>
<td>Not elevated</td>
</tr>
</tbody>
</table>
Appendix G – SAS® Output for Collision Models

G.1 – Initial Model

data ARCollisions;
input Config CC AR_C AADT_10k Dri_Vol Q_L Q_T Q_R MTOD LCI MT SL Grade Ang Dri_LW Dri_Div Dri_Tr Dri_Rad;
datalines;

5 210 0.0 0.12 32 76 76 33 0 4 1 25 -1 115 22 0 0 0
1 60 0.0 0.14 04 92 52 52 0 2 1 35 -1 70 13 0 0 0
... 3 343 0.0 4.03 69 156 337 337 0 1 1 45 0 90 24 0 1 1
6 254 1.0 5.40 190 73 927 88 0 5 0 55 -2 90 21 1 0 1
;
run;

proc NLMIXED data = ARCollisions maxfunc = 10000 maxiter = 1000;
parms b1_1=1 b1_2=1;
mu = exp(b1_1*log(AADT_10k))*exp(b1_2*AADT_10k);
loglike = AR_C*log(k*mu)-(AR_C+(1/k))*log(1+k*mu)+lgamma(AR_C+(1/k))-
lgamma(AR_C+1)-lgamma(1/k);
model AR_C ~ general(loglike);
run;
The NLMIXED Procedure
Specifications

Data Set                  WORK.ARCOLLISIONS
Dependent Variable       AR_C
Distribution for Dependent Variable General
Optimization Technique   Dual Quasi-Newton
Integration Method       None

Dimensions
Observations Used        108
Observations Not Used    0
Total Observations       108
Parameters               3

Parameters
b1_1   b1_2   k   NegLogLike
1       1       1   236.901232

Iteration History

<table>
<thead>
<tr>
<th>Iter</th>
<th>Calls</th>
<th>NegLogLike</th>
<th>Diff</th>
<th>MaxGrad</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>112.084324</td>
<td>124.8169</td>
<td>28.16235</td>
<td>-1138.08</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>107.142282</td>
<td>4.942043</td>
<td>6.091453</td>
<td>-22.2234</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>99.8605528</td>
<td>2.478E-6</td>
<td>0.006073</td>
<td>-6.38E-6</td>
</tr>
<tr>
<td>11</td>
<td>22</td>
<td>99.8605528</td>
<td>1.391E-7</td>
<td>0.00004</td>
<td>-2.79E-7</td>
</tr>
</tbody>
</table>

NOTE: GCONV convergence criterion satisfied.

Fit Statistics

-2 Log Likelihood       199.7
AIC (smaller is better)   205.7
AICC (smaller is better)  206.0
BIC (smaller is better)   213.8

Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>DF</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>t</th>
<th></th>
<th>Alpha</th>
<th>Lower</th>
<th>Upper</th>
<th>Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>b1_1</td>
<td>1.7249</td>
<td>0.4599</td>
<td>108</td>
<td>3.75</td>
<td>0.0003</td>
<td></td>
<td>0.05</td>
<td>0.8133</td>
<td>2.6365</td>
<td>-0.00003</td>
<td></td>
</tr>
<tr>
<td>b1_2</td>
<td>-0.7486</td>
<td>0.1656</td>
<td>108</td>
<td>-4.52</td>
<td>&lt;.0001</td>
<td></td>
<td>0.05</td>
<td>-1.0769</td>
<td>-0.4203</td>
<td>-0.00004</td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>0.4947</td>
<td>0.2626</td>
<td>108</td>
<td>1.88</td>
<td>0.0622</td>
<td></td>
<td>0.05</td>
<td>-0.02571</td>
<td>1.0151</td>
<td>-8.35E-6</td>
<td></td>
</tr>
</tbody>
</table>
G.2 – Second Model

data ARCollisions;
input Config CC AR_C AADT_10k Dri_Vol Q_L Q_T Q_R MTOD LCI MT SL Grade
Ang Dri_LW Dri_Div Dri_Tr Dri_Rad;
datalines;

5 210 0.0 0.12 32 76 76 33 0 4 1 25 -1 115 22 0 0 0
1 60 0.0 0.14 04 92 52 52 0 2 1 35 -1 70 13 0 0 0
..........................................................
3 343 0.0 4.03 69 156 337 337 0 1 1 45 0 90 24 0 1 1
6 254 1.0 5.40 190 73 927 88 0 5 0 55 -2 90 21 1 0 1
;
run;

proc NLMIXED data = ARCollisions maxfunc = 10000 maxiter = 1000;
parms b1_1=1 b1_2=1 b2_1=1 b2_2=1;
mu = (exp(b1_1*log(AADT_10k))*exp(b1_2*AADT_10k))*b2_1*exp(b2_2*log(Dri_LW));
loglike = AR_C*log(k*mu) -(AR_C+(1/k))*log(1+k*mu)+lgamma(AR_C+(1/k)) -
lgamma(AR_C+1) -lgamma(1/k);
model AR_C ~ general(loglike);
run;
The NLMIXED Procedure
Specifications

Data Set WORK.ARCOLLISIONS
Dependent Variable AR_C
Distribution for Dependent Variable General
Optimization Technique Dual Quasi-Newton
Integration Method None

Dimensions

Observations Used 108
Observations Not Used 0
Total Observations 108
Parameters 5

Parameters

<table>
<thead>
<tr>
<th>b1_1</th>
<th>b1_2</th>
<th>b2_1</th>
<th>b2_2</th>
<th>k</th>
<th>NegLogLike</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>546.45679</td>
</tr>
</tbody>
</table>

Iteration History

<table>
<thead>
<tr>
<th>Iter</th>
<th>Calls</th>
<th>NegLogLike</th>
<th>Diff</th>
<th>MaxGrad</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>136.31106</td>
<td>410.1457</td>
<td>92.96568</td>
<td>-3975.99</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>111.258372</td>
<td>25.05269</td>
<td>17.09568</td>
<td>-136.877</td>
</tr>
<tr>
<td>46</td>
<td>96</td>
<td>94.7788658</td>
<td>5.573E-6</td>
<td>0.568397</td>
<td>-0.00001</td>
</tr>
<tr>
<td>47</td>
<td>98</td>
<td>94.7788656</td>
<td>1.764E-7</td>
<td>0.235612</td>
<td>-2.56E-7</td>
</tr>
</tbody>
</table>

NOTE: GCONV convergence criterion satisfied.

Fit Statistics

-2 Log Likelihood 189.6
AIC (smaller is better) 199.6
AICC (smaller is better) 200.1
BIC (smaller is better) 213.0

Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Error</th>
<th>DF</th>
<th>t Value</th>
<th>Pr &gt;</th>
<th>t</th>
<th></th>
<th>Alpha</th>
<th>Lower</th>
<th>Upper</th>
<th>Gradient</th>
</tr>
</thead>
<tbody>
<tr>
<td>b1_1</td>
<td>1.4360</td>
<td>0.7369</td>
<td>108</td>
<td>1.95</td>
<td>0.0539</td>
<td>0.05</td>
<td>-0.02462</td>
<td>2.8966</td>
<td>-0.00028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b1_2</td>
<td>-0.4742</td>
<td>0.4718</td>
<td>108</td>
<td>-1.01</td>
<td>0.3171</td>
<td>0.05</td>
<td>-1.4093</td>
<td>0.4610</td>
<td>-0.00099</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b2_1</td>
<td>0.002176</td>
<td>0.004448</td>
<td>108</td>
<td>0.49</td>
<td>0.6257</td>
<td>0.05</td>
<td>-0.0664</td>
<td>0.01099</td>
<td>-0.23561</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b2_2</td>
<td>1.7218</td>
<td>0.5506</td>
<td>108</td>
<td>3.13</td>
<td>0.0023</td>
<td>0.05</td>
<td>0.6305</td>
<td>2.8131</td>
<td>-0.00169</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>0.3455</td>
<td>0.2245</td>
<td>108</td>
<td>1.54</td>
<td>0.1266</td>
<td>0.05</td>
<td>-0.09936</td>
<td>0.7904</td>
<td>8.505E-6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
G.3 – Thirds (final) Model

data ARCollisions;
input Config CC AR_C AADT_10k Dri_Vol Q_L Q_T Q_R MTOD LCI MT SL Grade Ang Dri_LW Dri_Div Dri_Tr Dri_Rad;
datalines;
5 210 0.0 0.12 32 76 76 33 0 4 1 25 -1 115 22 0 0 0
1 60 0.0 0.14 04 92 52 52 0 2 1 35 -1 70 13 0 0 0
....................................................................
3 343 0.0 4.03 69 156 337 337 0 1 1 45 0 90 24 0 1 1
6 254 1.0 5.40 190 73 927 88 0 5 0 55 -2 90 21 1 0 1
;
run;

proc NLMIXED data = ARCollisions maxfunc = 10000 maxiter = 1000;
parms b1_1=1 b1_2=1 b2_1=1 b2_2=1 b3_1=1;
mu =
(exp(b1_1*log(AADT_10k))*exp(b1_2*AADT_10k))*(b2_1*exp(b2_2*log(Dri_LW )))) *(1+b3_1*Q_T);
loglike = AR_C*log(k*mu)-(AR_C+(1/k))*log(1+k*mu)+lgamma(AR_C+(1/k))-lgamma(AR_C+1)-lgamma(1/k);
model AR_C ~ general(loglike);
run;
The NLMIXED Procedure
Specifications

Data Set WORK.ARCOLLISIONS
Dependent Variable AR_C
Distribution for Dependent Variable General
Optimization Technique Dual Quasi-Newton
Integration Method None

Dimensions

Observations Used 108
Observations Not Used 0
Total Observations 108
Parameters 6

Parameters

<table>
<thead>
<tr>
<th>b1_1</th>
<th>b1_2</th>
<th>b2_1</th>
<th>b2_2</th>
<th>b3_1</th>
<th>k</th>
<th>NegLogLike</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1122.65018</td>
</tr>
</tbody>
</table>

Iteration History

<table>
<thead>
<tr>
<th>Iter</th>
<th>Calls</th>
<th>NegLogLike</th>
<th>Diff</th>
<th>MaxGrad</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>550.619606</td>
<td>572.0306</td>
<td>210.4699</td>
<td>-11101.5</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>116.583678</td>
<td>434.0359</td>
<td>59.47979</td>
<td>-351.151</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>119</td>
<td>90.6472918</td>
<td>7.304E-7</td>
<td>0.843425</td>
<td>-1.27E-6</td>
</tr>
<tr>
<td>61</td>
<td>121</td>
<td>90.6472918</td>
<td>3.817E-8</td>
<td>0.478096</td>
<td>-5.68E-8</td>
</tr>
</tbody>
</table>

NOTE: GCONV convergence criterion satisfied.

Fit Statistics

-2 Log Likelihood 181.3
AIC (smaller is better) 193.3
AICC (smaller is better) 194.1
BIC (smaller is better) 209.4

Parameter Estimates

| Parameter | Estimate | Error  | DF | t Value | Pr > |t| | Alpha | Lower | Upper | Gradient |
|-----------|----------|--------|----|---------|-------|---| |       |       |       |          |
| b1_1      | 0.9326   | 0.7858 | 108| 1.19    | 0.2379| 0.05| | -0.6250| 2.4903| 0.000181|
| b1_2      | -0.3748  | 0.4632 | 108| -0.81   | 0.4202| 0.05| | -1.2930| 0.5434| 0.00049 |
| b2_1      | 0.000396 | 0.001033| 108| 0.38    | 0.7020| 0.05| | -0.00165| 0.002444| 0.478096|
| b2_2      | 1.8409   | 0.6077 | 108| 3.03    | 0.0031| 0.05| | 0.6363 | 3.0454| 0.000574|
| b3_1      | 0.008532 | 0.01073| 108| 0.80    | 0.4281| 0.05| | -0.01273| 0.02979| 0.022506|
| k         | 0.2680   | 0.1998 | 108| 1.34    | 0.1826| 0.05| | -0.1280| 0.6639| 0.000227|