2009

Level-of-service And Traffic Safety Relationship: An Exploratory Analysis Of Signalized Intersections And Multiland High-speed Arterial Corridors

Ana Almonte-Valdivia
University of Central Florida

LEVEL-OF-SERVICE AND TRAFFIC SAFETY RELATIONSHIP: AN EXPLORATORY ANALYSIS OF SIGNALIZED INTERSECTIONS AND MULTILANE HIGH-SPEED ARTERIAL CORRIDORS

by

ANA MARIA ALMONTE VALDIVIA
B.S. Michigan State University, 2007

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Civil, Environmental and Construction Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Spring Term
2009

ABSTRACT

Since its inception in 1965, the Level-of-Service (LOS) has proved to be an important and practical "quality of service" indicator for transportation facilities around the world, widely used in the transportation and planning fields. The LOS rates these facilities' traffic operating conditions through the following delay-based indicators (ordered from best to worst conditions): A, B, C, D, E and F. This LOS rating has its foundation on quantifiable measures of effectiveness (MOEs) and on road users' perceptions; altogether, these measures define a LOS based on acceptable traffic operating conditions for the road user, implying that traffic safety is inherent to this definition. However, since 1994 safety has been excluded from the LOS definition since it cannot be quantified nor explicitly defined. The latter has been the motivation for research based on the LOS-Safety relationship, conducted at the University of Central Florida (UCF). Using data from two of the most studied transportation facility types within the field of traffic safety, signalized intersections and multilane high-speed arterial corridors, the research conducted has the following main objectives: to incorporate the LOS as a parameter in several traffic safety models, to extend the methodology adopted in previous studies to the subject matter, and to provide a platform for future transportation-related research on the LOS-Safety relationship.

A meticulous data collection and preparation process was performed for the two LOS-Safety studies comprising this research. Apart from signalized intersections' and multilane-high speed arterial corridors' data, the other required types of information corresponded to crashes and road features, both obtained from FDOT's respective databases. In addition, the Highway Capacity Software (HCS) and the ArcGIS software package were extensively used for the data

preparation. The result was a representative and robust dataset for each LOS-Safety study, to be later tested and analyzed with appropriate statistical methods.

Regarding the LOS-Safety study for signalized intersections, two statistical techniques were used. The Generalized Estimating Equations (GEEs), the first technique, was used for the analyses considering all periods of a regular weekday (i.e. Monday through Friday): Early Morning, A.M. Peak, Midday, P.M. Peak and Late Evening; the second technique considered was the Negative Binomial, which was used for performing an individual analysis per period of the day. On the other hand, the LOS-Safety study for multilane high-speed arterial corridors made exclusive use of the Negative Binomial technique. An appropriate variable selection process was required for the respective model building and calibration procedures; the resulting models were built upon the six following response variables: total crashes, severe crashes, as well as rear-end, sideswipe, head-on and angle plus left-turn crashes.

The final results proved to be meaningful for the understanding of traffic congestion effects on road safety, and on how they could be useful within the transportation planning scope. Overall, it was found that the risk for crash occurrence at signalized intersections and multilane high-speed arterial corridors is quite high between stable and unacceptable operating conditions; it was also found that this risk increases as it becomes later in the day. Among the significant factors within the signalized intersection-related models were LOS for the intersection as a whole, cycle length, lighting conditions, land use, traffic volume (major and minor roads), left-turn traffic volume (major road only), posted speed limit (major and minor roads), total number of through lanes (major and minor roads), overall total and total number of left-turn lanes (major road only), as well as county and period of the day (dummy variables). For multilane-high speed arterial corridors, the final models included LOS for the road section, average daily traffic

(ADT), total number of through lanes in a single direction, total length of the road section, pavement surface type, as well as median and inside shoulder widths. A summary of the overall results per study, model implications and each LOS indicator is presented. Some of the final recommendations are to develop models for other crash types, to perform a LOS-Safety analysis at the approach-level for signalized intersections, as well as one that incorporates intersections within the arterial corridors' framework.

To God, the Holy Family and the Saints, for without their blessings and guidance

I could not have become the person I am today;

to my parents, Limber and Nilda, and my brother, Rafael, for their unconditional love and support, especially throughout this exciting journey in the U.S.A.;

and to everyone I have met, since each of you has made a beautiful difference in my life by just being yourself.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Mohamed Abdel-Aty, for his valuable advice and guidance throughout my graduate studies at UCF, especially for providing me the opportunity of attending such beautiful university –these have truly been two of the best years of my life, in every sense–. Likewise, I would like to thank the rest of my committee members, Dr. Essam Radwan and Dr. Xuesong Wang; same as my advisor, I admire the contributions you have made to our profession, and your dedication will continue inspiring me in the years ahead.

Also, a very special acknowledgment goes to my research team: Abhishek Das, Kirolos Haleem, Hany Hassan, Mohamed Ahmed, Ali Darwiche, Prem Chand Devarasetty, Choudhury Siddiqui (Sid) and Al-Ahad Ekram (Turen), as well as Alexis Nevarez, Dr. Anurag Pande, Fraya Cassier, Dr. Helai Huang, Ming Ma, Dr. Yuren Chen, and also to the rest of my colleagues within our Transportation Systems Engineering Program; all of you are the ones with whom I have shared most of my moments while at UCF, and I am very grateful for your friendship and for all what we have accomplished together.

Finally, thanks to UCF's Center for Advanced Transportation Systems Simulation (CATSS), the Florida Department of Transportation (FDOT), as well as the different local transportation authorities and entities which made this investigation a reality through all the resources provided.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS/ABBREVIATIONS

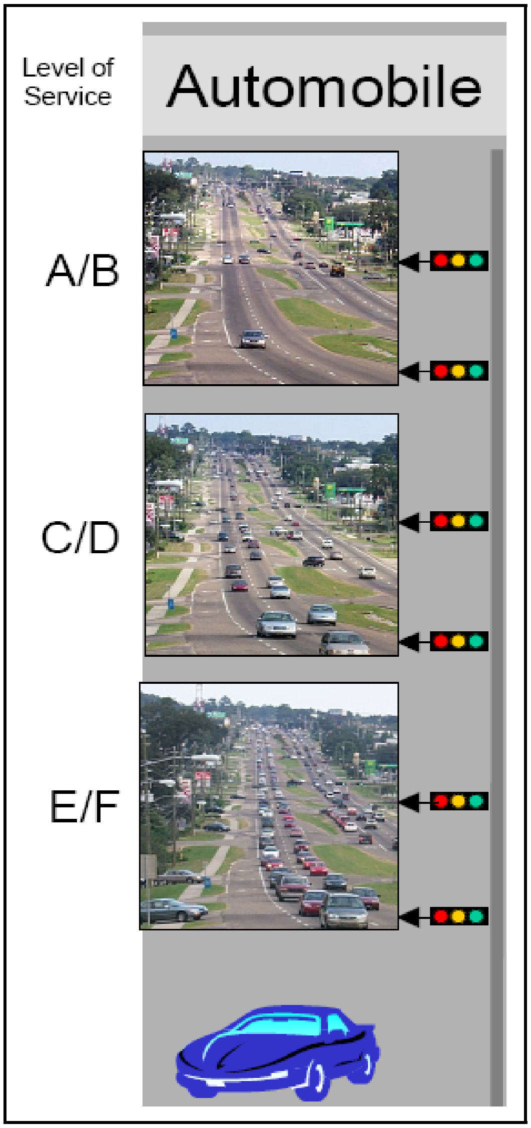
ADT	Average Daily Traffic
CATSS	Center for Advanced Transportation Systems Simulation
DHSMV	Department of Highway Safety and Motor Vehicles
FDOT	Florida Department of Transportation
GEE	Generalized Estimating Equations
GIS	Geographic Information Systems
GLM	Generalized Linear Model
HCM	Highway Capacity Manual
HSM	Highway Safety Manual
MPO	Metropolitan Planning Organization
NCHRP	National Cooperative Highway Research Program
LOS	Level-of-Service
LRTP	Long Range Transportation Plan
PDO	Property Damage Only
SAFETEA-LU	Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users
TAZ	Traffic Analysis Zone
TRB	Transportation Research Board

CHAPTER 1: GENERAL INTRODUCTION

1.1 Background

Since its inception in 1965 by the Transportation Research Board's (TRB) Highway Capacity Manual (HCM), the Level-of-Service (LOS) has proved to be a practical "quality of service" indicator for transportation systems around the world. Categorical in nature, the LOS rates the functionality of a transportation facility through six levels –from best to worst– based on delay as follows: A (excellent), B (very good), C (good), D (satisfactory), E (acceptable) and F (unacceptable) (AASHTO, 2005). This LOS rating has its foundation on quantifiable measures of effectiveness (MOE) that are unique to each facility's traffic flow conditions; some MOE examples include delay for signalized intersections, travel speed and density for freeways, walking speed for pedestrians, etc. In addition, the LOS includes user perception-related measures such as comfort and convenience when using a certain transportation facility, freedom to maneuver, traffic interruptions and travel time (HCM, 2000). As it can be seen, all these measures define a LOS based on acceptable operational conditions for the road user, implying that safety is an inherent part of this definition. Records indicate that all 1965, 1985 and 1992 editions of the HCM included the term safety in the LOS definition; however, starting with the 1994 edition of the HCM the term safety has been excluded since it cannot be quantified nor explicitly defined. The latter has been the motivation for some researchers in the field, who have recommended further studies on the relationship between the LOS and safety of transportation systems and on how this traffic safety could be made explicit or quantified through LOS-related research. Though very few studies have been undertaken on the topic, their contributions have propelled a series of initiatives for making safety a priority in the evaluation and planning of transportation projects; for example, the coming 2009 release of the interim document of the Highway Safety Manual (HSM), also created by the TRB, reflects the importance, need and potential for more LOS-Safety studies.

Regarding safety assessments for transportation facilities, intersections have constantly been a recurrent subject. Formed by the junction of two or more roads, both signalized and unsignalized intersections carry a very high crash risk at or near their influence area due to the formation of conflict points; as a result, subjects ranging from drivers to pedestrians are quite exposed to traffic accidents at these locations and/or have a high probability of being involved. In the State of Florida alone, a total of 256,206 traffic crashes were reported in 2007 by the State's Department of Highway Safety and Motor Vehicles (DHSMV) (FDHSMV, 2007). In particular, it is the group of signalized intersections that constantly receive the attention of transportation planners and road safety authorities: though their design apparently makes them more controlled than their unsignalized counterparts, signalized intersections have more crashes associated with them; for example, a study from Bhesania (1991) showed that signalized intersections had an average number of 9.6 crashes per year whereas intersections with stop or yield signs had an average number of 2 crashes per year. In summary, statistics from past and present demonstrate that there are many factors to be considered with regards to the safety of signalized intersections. Whether it is due to their large size in terms of geometry and/or traffic volume, or due to their control type and operational features, meaningful insights on signalized intersections' safety conditions could be obtained by analyzing more in depth some of their performance indicators such as LOS. Figure 1-1 below illustrates the LOS concept as applied to automobiles, which are the focus of this thesis.



(Source: Quality/Level of Service Handbook)

Figure 1-1: Example of LOS for Automobile Mode for Urban Roadways

At a larger scale, multilane high-speed arterial corridors have also been the main subject of several studies focusing on the evaluation of safety and operational conditions throughout the transportation network. In contrast with uninterrupted flow facilities like expressways and freeways, which handle vehicle travel at higher speeds, arterials account for a larger number of severe and fatal crashes. The National Highway Traffic Safety Administration (NHTSA) reported in 2006 that 57% of Florida's fatal crashes occur at arterial corridors (NHTSA, 2006); worth of attention, this statistic suggests an imperative need for safety improvements along this type of facilities. Since arterials are composed by road segments connected by a series of signalized and unsignalized intersections, this need can be translated into safety studies that treat LOS at the macroscopic level.

As it can be seen, there is much that can be discovered from investigating LOS-related parameters and road safety conditions. Findings like these not only benefit researchers in the field but can also have the potential to make Metropolitan Planning Organizations (MPOs) and related state agencies fulfill the goals in their Long Range Transportation Plans (LRTP) in new, well-defined and practical ways (Kramer, 2005). For example, Handy (2008) states that it is imperative for current practitioners to count with innovative performance measures applicable to the transportation planning process that would explicitly describe safety. For example, planners and traffic engineers would finally be able to explicitly judge the trade-offs between efficient transport mobility and safety (Ha and Berg, 1995). Related studies could help incorporate safety along with efficiency indicators (e.g. delay, etc.) at the planning stage of transportation projects, as required by the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (SAFETEA-LU) (2005).

Recently, researchers at the University of Central Florida (UCF) in Orlando, FL, have been conducting similar studies towards enhancing the safety of signalized intersections and multilane high-speed arterial corridors. The first study motivating this thesis' topic is based on research conducted by Wang et al. (2009) that investigated the relationship between the LOS and traffic safety focusing on signalized intersections; the corresponding paper, “Incorporating Traffic Operation Measures in Safety Analysis at Signalized Intersections”, was recently accepted by TRB’s committee of Safety Data, Analysis and Evaluation, fact that reflects the topic’s potential in innovative research. The other study considered for this thesis is based on the project “Reducing Fatalities and Severe Injuries on Florida’s High-Speed Multi-Lane Arterial Corridors”, based on a project sponsored by the Florida Department of Transportation (FDOT); similarly, publications coming from this project have been accepted by committees from TRB.

In general terms, this thesis is an extension of the first study from Wang et al. (2009) by incorporating two additional time periods, Early Morning and Late Evening, in the overall study so that insights of the LOS-Safety relationship at signalized intersections can be obtained for the whole day. In addition, the author incorporated the LOS-Safety analysis just described in multilane high-speed arterial corridors; these results would contribute with insights on the LOS-Safety relationship at a different facility type. Based on this, this thesis consists of two main LOS-Safety studies: one focusing on signalized intersections and the other focused on multilane high-speed arterial corridors. Though exploratory in nature, this research aims to contribute to the understanding of traffic congestion effects on road safety.

1.2 Research Objectives

Based on its two studies, the objectives of this research are as follows:

1. To critically review previous studies on how LOS may relate to traffic safety.

2. To demonstrate the possibility of incorporating the LOS as a parameter in traffic safety models.
3. To extend the methodology adopted in previous studies to the subject matter.
4. To create several models for the safety analysis of signalized intersections and multilane high-speed arterial corridors by incorporating LOS as the predominant indicator.
5. To provide a platform for future transportation-related research on the LOS-Safety relationship.

1.3 Thesis' Structure

The current thesis has been developed based on a LOS and safety study conducted for the transportation settings introduced in this chapter's first section: signalized intersections and multilane high-speed arterial corridors. Chapter 2 contains a detailed review of past and recent literature published on the topic being discussed. Chapter 3 describes the foundation of the study's framework, which details the research design as well as the methods for data preparation, parameter calculation and data assembly. Chapter 4 focuses on the data's preliminary trends, including the respective analyses. Chapter 5 details the statistical techniques used as well as the resulting models for a more in-depth analysis of the topic; it also reveals the results obtained through the modeling approach as well as details on the models' assessment. Finally, Chapter 6 summarizes the research conducted, including the most important findings, limitations and contributions, as well as recommendations for future studies.

CHAPTER 2: LITERATURE REVIEW

2.1 Overview

A thorough literature review of past work related to the topic of interest was performed since the earlier stages of this research study. This approach aided the author in formulating related questions, identifying any problems that could be solved, refining the hypothesis and in defining the significant contributions the overall research can provide. More specifically, this review would support the use of the LOS indicator in the evaluation of traffic safety performance of signalized intersections and multilane high-speed arterial corridors.

This chapter contains the key points and/or framework of the studies considered as most relevant to the research being discussed; overall, it has been divided based on ideas and topics relevant to this research. The documentation that was found, albeit somewhat narrow in scope, proved to be meaningful as it aided in meeting the aforementioned objectives.

2.2 Operational Conditions within the Context of Safety Performance

2.2.1 Incorporating LOS in Traffic Safety Studies

Worth of attention, the area of traffic safety is continuously evolving towards finding new ways of characterizing road safety performance. Based on this premise, one of the earliest attempts to relate LOS-related measures in traffic safety studies is reflected by the work of Frantzeskakis and Iordanis (1987). Their study studied the potential relationship between the volume-to-capacity (v/c) ratio to traffic accident occurrence on interurban four-lane highways in Greece; the use of the v/c ratio in the study enabled them to translate their results into Level-of-Service (LOS) terms. Using some of the standard equations provided by the 1985 edition of the Highway Capacity Manual (HCM), they studied different crash categories by incorporating factors such as type of vehicle, weather, light and pavement conditions, as well as time of day (i.e. day and night). Their findings suggested that accident rates on non-hazardous locations are almost constant for v/c ratios up to 0.65 (i.e. for LOS A, LOS B and LOS C), whereas a considerable increase of these rates was related to v/c ratios greater than 1.0.

Similarly, Persaud and Dzbik (1993) performed a capacity evaluation for freeways in Ontario, Canada, that accounts for traffic safety. Through regression prediction models that focused on A.M. and P.M. peak periods, they found that crash risk increases as the LOS worsens –for freeways in general–, as well as during P.M. peak periods –for expressways only–; they also found that collectors have a higher accident risk than expressways overall.

As a new way to express the LOS in traffic safety studies, Ha and Berg (1995) developed a safety-based LOS indicator in order to evaluate safety conditions at isolated signalized intersections. This indicator was defined by the number of hazards per crossing vehicle and number of conflicts characterizing the intersection(s). Overall, this study is considered to be exploratory in nature.

Similarly, Lee and Berg (1998) developed a safety-based LOS criterion for 2-way stop-controlled intersections. By using the total number of conflicts/crashes per year per crossing vehicle at the intersection, as well as the total hazards per year per crossing vehicle, these totals were divided in order to produce 6 LOS indicators, analogous to the ones in the existing scale; the 1994 edition of the HCM was considered. Through regression models they incorporated parameters such as geometry, vehicular volume and composition, available sight distance and pavement condition; then, simulation models of the crossing maneuvers at the intersection estimated the frequency of conflicts influenced by sight distance restrictions. The intersections with the highest number of conflicts and/or hazards per crossing vehicle were assigned with the most deteriorated LOS.

Another relevant study is that of Persaud and Nguyen (1998, a), who analyzed 107 four-legged signalized intersections located within rural areas of Ontario, Canada. Their main goal was to explore the relationship between the LOS and safety performance at these facilities. Also a temporal analysis, it considered the A.M. and P.M. peak periods of the day under the assumption that these are similar in terms of severity model parameters (1998, b); in addition, a 6-year crash data (1988-1993) was considered, consisting in a total of 970 crashes linked to the intersections of interest. Using the Highway Capacity Software (HCS), under the guidelines of the 1994 edition of the HCM, the main parameters were obtained for the analysis: peak hour average stopped delay, capacity and LOS. As a result of the modeling process, they found that having separate calibrated parameters for a signalized intersection's major and minor traffic flows improved the models' performance; they also found that, considering only the peak periods of the day, both LOS B and LOS C are associated with the highest crash frequencies, whereas LOS D and LOS E are associated with the lowest ones.

Kononov and Allery (2003) introduced their concept of Level of Service of Safety (LOSS), an indicator that qualitatively assigns a degree (i.e. magnitude) of safety or unsafety to roadway facilities. The LOSS concept was built within the framework of Safety Performance Functions (SPFs) (i.e. crash prediction models relating safety with traffic exposure) and problem diagnostics (i.e. the issue of diagnosing the cause of any safety problem). This LOSS indicator was designed with 4 levels: LOSS I (low potential for accident reduction), LOSS II (better than expected safety performance), LOSS III (less than expected safety performance) and LOSS IV (high potential for accident reduction). Through the LOSS concept, their main objective was to provide a frame of reference with regards to the safety performance of roadway facilities (e.g. expected crash frequencies and severity norms); this could be applicable to safety- as well as non-safety-motivated projects (e.g. corridor segment improvements as well as resurfacing or reconstruction projects, respectively). They used crash data (1989-2001) from segments belonging to two-lane rural roads, as well as from rural and urban freeways in Colorado. Results from this study suggested that the LOSS can be useful only for describing the magnitude of a traffic safety problem; the nature of the problem has to be determined through direct diagnostics and pattern recognition methods.

Zhang and Prevedouros (2003) also conducted a study that used a signalized intersection's LOS as a parameter that would account for traffic safety risk. Adhering to the 2000 version of the HCM, they developed a methodology that quantified both vehicle-to-vehicle as well as vehicle-to-pedestrian conflicts related to left-turns (i.e. consideration of left-turning and opposing through vehicles only); this method would model in an explicit way the tradeoff between traffic safety (i.e. related to the number of conflicts) and efficiency (i.e. related to the delay and LOS). The result was a Delay and Safety (DS) Index, a comprehensive LOS indicator resulting from a model that attempted to combine both delay and safety. They conducted a case study based on only 2 signalized intersections. Their findings suggested that, provided that the potential for traffic conflicts is not considered, an acceptable LOS (i.e. less delay) is generally obtained when the intersection has a permitted left-turn phasing; on the other hand, provided that the potential for traffic conflicts is considered, an acceptable LOS is generally obtained when the signalized intersection has a protected left-turn phasing.

Within the context of unsignalized intersections, Lu et al. (2008) developed a procedure for evaluating safety performance of highway unsignalized intersections in China. For this purpose, they developed the Level-of-Safety-Service: a six-level scale, from A through F, that would evaluate the provision of safety service by this type of intersections to the traveling public (e.g. motorized and non-motorized vehicles, as well as pedestrians). The models developed were based on these intersections' traffic conflict points and the characteristics of these points; minor factors were used only for adjusting the models that were initially based on traffic conflict points alone. The authors noted the effectiveness of this method after its successful application throughout many places in China, corroborated by highway engineers and transportation authorities.

A recent study by Wang et al. (2009) incorporated traffic operation measures, including the LOS, as a promising way to analyze crash occurrence at signalized intersections. A representative sample of 164 four-legged signalized intersections located in Central Florida, of both coordinated and isolated types, were analyzed through Generalized Estimating Equation (GEE) models with Negative Binomial link function; being an extension of the Generalized Linear Models (GLMs), GEE models are capable of handling balanced and continuous longitudinal data. Considering the 2000 edition of the HCM, the 2005 version of HCS was used for calculating the LOS of each signalized intersection in the study sample. As a temporal analysis, this investigation studied the correlation between the LOS and crash occurrence throughout the main periods of a regular weekday: A.M. Peak, Midday and P.M. Peak. Assuming the GEE unstructured correlation structure, crash frequency models were developed for total, severe, rear-end and sideswipe, as well as angle and left-turn crashes; the LOS was used as the main parameter in all models. Results indicated that the six-level LOS indicator was better than the delay factor when predicting crashes at signalized intersections; however, the LOS has to be accompanied by other types of factors (e.g. intersections' geometric and design features) in order to model safety conditions more efficiently. It was also found that LOS D, the fourth level of the LOS scale, was associated with the lowest probability for total, rear-end and sideswipe, as well as right-angle and left-turn crash occurrence, which makes it into a desirable level to attain. In addition, the authors oriented their efforts towards making a positive impact in both traffic safety analysis and transportation planning with these results.

2.2.2 Assessing Crash Risk through Other Operational Measures

Hall and Polanco de Hurtado (1992) conducted a study that incorporated different variables of the traffic stream as well as operational measures. With the purpose of investigating how congestion affects accident rates, they used both traffic volume and crash data corresponding to 260 urban signalized intersections located in Albuquerque, New Mexico. The traffic volume data (1989-1990) was for one-hour peak periods (A.M. Peak beginning either at 7:15 or 7:30 a.m., and P.M. Peak beginning at 4:30 p.m.)., weekdays only; these volumes were used for calculating the respective LOS and v/c ratios based on the methods from the 1985 edition of the HCM. With regards to the crash data (1987-1989), crash occurrences were converted to crash rates. The results showed a small but positive correlation between accident rates and the traffic volume entering the signalized intersection; however, it was found that the v/c ratio-based equations that were developed for the study showed high standard errors. The use of crash data in the form of rates, as well as using ordinary linear regression, could have caused those errors in such safety analysis (Hauer, 2002).

A similar study by Zhou and Sisiopiku (1997) used the v/c ratio as a predictor for collision rates. A 16-mile segment from I-94 in Detroit, Michigan, was considered for the study. The resulting models gave insights on the correlation between v/c ratios and accident rates by day of the week, and different crash types, etc; this correlation, when plotted, displayed a consistent U-shaped pattern.

Garber and Subramanyan (2001) conducted a case study of freeways in the State of Virginia. Their goal was to incorporate crash rates within the development of congestion-mitigation strategies; for this purpose, models were developed using real-time data (e.g. speed and occupancy) in addition to information on traffic flow. The results of this study suggested that peak traffic flows are not related to peak accident rates, and that real-time data can be useful for the modeling of crashes and traffic conditions.

Regarding other facility types, Lord, Manar and Vizioli (2005) developed models for predicting crash occurrence on rural and urban freeway segments. The respective v/c ratios and traffic densities were used for describing single and multivehicle crashes; they concluded that these parameters are necessary for characterizing crashes on this type of road segments.

2.3 Relevant Statistical Tools

2.3.1 Analysis through Generalized Estimating Equations (GEEs)

Liang and Zeger (1986) were the first ones to introduce GEE models, an extension of the GLM approach, for studying the correlation among longitudinal data; particularly, their study analyzed repeated observations over time. Since then, their approach to GEEs has been reviewed by several researchers. For example, Ziegler, Kastner and Blettner (1998) prepared an annotated bibliography on the advantages and several applications of GEE models. Similarly, Zorn (2001) also reviewed applications of the GEE technique for modeling correlated data; he emphasized its usefulness for most scientific disciplines, its ability to estimate models having event counts as the outcome variable (e.g. crash frequencies), as well as its flexibility in choosing correlation structures –depending on the subject(s) being studied–. Overall, and based on the aforementioned work samples, the GEE technique makes it possible to understand the properties of correlated data's empirical dependencies, by having an identical estimate interpretation to that for models with uncorrelated data (e.g. logit and probit); furthermore, this technique is accessible to most researchers thanks for the range of software packages available.

Regarding transportation-related applications, Wang (2006) made use of GEEs for the temporal, spatial and site correlation analyses of crash occurrence at signalized intersections; the analyses were done at both the intersection- and approach-level. Results from this investigation indicated that the autoregression structure is best for both temporal and spatial analyses, whereas the unstructured correlation structure is more appropriate for right angle crash analysis at both roadway and approach-level as well as for left-turn crashes at the approach-level.

Recalling the study of Wang et al. (2009), they also made use of the GEE technique when analyzing crash data (2000-2005) corresponding to signalized intersections in Central Florida. More specifically, they used the unstructured correlation structure in order to evaluate the correlation between data from the A.M. Peak, Midday and P.M. Peak periods; this would give insights on how operational conditions (e.g. LOS) and safety of signalized intersections perform throughout these three periods of the day.

2.3.2 Using the Negative Binomial

Abdel-Aty and Radwan (2000) conducted a study of accident occurrence and involvement by analyzing data from a principal arterial in Central Florida. By using 3-year crash data (1992-1994), they fitted a negative binomial model which had the frequency of accident occurrence as the outcome variable; the AADT, degree of horizontal curvature, land use, road section's length, as well as lane, shoulder and median widths were used as independent variables. In addition, another negative binomial model was fitted for predicting crash occurrence by considering demographic characteristics of the driver population (e.g. age and gender). The negative binomial proved to be a valuable tool for identifying contributing factors within the context of accident occurrence and involvement.

Considering a more theoretical approach, the study from Lord, Washington and Ivan (2005) provided a defensible guidance on how to model crash data more appropriately. Their work recognizes the wide range of statistical models commonly used in crash-related analyses (e.g. binomial, Poisson, negative binomial, zero-inflated Poisson and negative binomial models, as well as multinomial probability models). With regards to the negative binomial, also known as

Poisson-gamma, they stated that this method approximates the crash process in good statistical terms and that it is efficient for the modeling of random events.

2.4 Insights from Research and Current Practices

2.4.1 Signalized Intersections

Under the notion that heavily congested intersections have a higher accident risk, Ogden et al. (1994) conducted a comprehensive study of the factors affecting crash occurrence at signalized intersections. High, normal and low accident frequency data (1987-1991), as well as traffic volumes of signalized intersections located in Melbourne, Australia, were used. They found that accident variation was explained by factors other than traffic volumes; however, crash occurrence was not attributed to a single factor. A list of safety guidelines was identified for reducing accident risk at intersections.

Lord and Persaud (2000) studied a 6-year period of data corresponding to 868 four-legged signalized intersections in Toronto, Canada. Using only entering traffic flows as explanatory factors, their study confirmed that traffic flow is the most important factor for modeling crash occurrence at signalized intersections.

In terms of severe crashes, Abdel-Aty and Keller (2004) performed an exploratory analysis of the overall and specific injury severity levels of crashes at signalized intersections. The depth of the investigation also gave insights on the completeness and quality of crash data usually available for this type of studies. Data from 832 signalized intersections in four counties within Central Florida were used. A total of 33,592 crashes corresponding to the intersection sample were considered; in addition to the injury-related ones, these also comprised the minor/no injury crashes. The individual severity levels and the factors associated with each of them were explored by using the ordered-probit modeling and tree-based regression techniques, respectively. Results showed that the highest prediction rates of injury level could be obtained when using intersection characteristics' data combined with the specific information for each crash; for example, minor roadways that are divided or that have a high speed limit are associated with a reduced injury level. Also, the incorporation of minor/no injury crashes within the modeling process depicted significant differences when modeling these events, thus improving the overall results.

Souleyrette et al. (2004) studied the safety effect of all-red clearance intervals in reducing crash occurrence at signalized intersections. Cross-section analysis as well as regression modeling were used for studying 4 years of data corresponding to low-speed urban 4-way signalized intersections in Minneapolis, Minnesota. An additional before-and-after study was conducted using 11 years of data for evaluating both short- and long-term safety effects. The overall results did not indicate any defined safety benefit; short-term crash reductions were the only ones observed. If a safety measure based on all-red clearance intervals were considered, they recommended that extended clearance intervals should be implemented specifically for off-peak hours; this would improve traffic safety (e.g. reduction of signal violations) while not interfering with peak hour traffic flow.

Abdel-Aty et al. (2006) proposed a simple and practical approach for identifying the expected number of crash events, by severity and type, at signalized intersections. After analyzing a large dataset of 1,335 signalized intersections from five counties in the State of Florida, along with 26,603 crashes (1999-2001) associated with these, an intersection's size was found to be the most significant factor. Crash events can be identified based on that factor: the larger the intersection (i.e. more traffic, longer cycle length, more number of phases), the higher the expected crash frequency.

Temporal and spatial studies of crash occurrence at signalized intersections were conducted by Wang et al. (2006) and Abdel-Aty and Wang (2006), respectively. With regards to the entering traffic volume factor, these studies revealed that a signalized intersection's traffic intensity (defined as total entering traffic volume divided by the total number of lanes) is the most significant form of this factor.

Regarding the differences between vehicle crashes, it is generally considered that there are different contributing factors (e.g. geometric, traffic, environmental conditions, etc.) for each crash type. Thus, it is recommended to model crashes by its different types. As an example, Wang and Abdel-Aty (2006) thoroughly investigated the occurrence of rear-end crashes. In addition, Hauer et al. (1988) took into account the different crash mechanisms throughout a regular day by fitting different models for A.M. and P.M., as well as off-peak periods; overall models were also obtained. Lord, Manar and Vizioli (2005) also emphasized the importance of creating separate models by crash type (e.g. single- vs. multi-vehicle). For instance, to fit a signalized intersection's number of crashes by crash type can reveal the different effects of its related factors or injury level(s) associated with it.

2.4.2 Multilane High-Speed Arterial Corridors

In terms of corridor safety improvement efforts, and considering today's common practices, Jernigan (1999) conducted a comparative and evaluation study of Corridor Safety Improvement Programs (CSIPs) that have been implemented by several State DOTs across the

U.S.A. The study's main purpose was to identify the factors related to the effectiveness of these programs; in addition, Jernigan developed model guidelines to support these CSIPs.

Similarly, Green and Agent (2002) conducted a study of high traffic crash corridors using data from the State of Kentucky. Their main objectives were to determine a method for identifying corridors with a high crash risk, and to develop a data analysis procedure that will lead to the formulation of safety countermeasures. A ranking methodology was the method of choice, through which corridor routes –traveling through more than one county– would be selected, based on their corresponding attributes (e.g. crash frequency, length, traffic volume) and the relative value of each attribute. In the end, their analysis was divided into a corridor analysis and a high crash analysis; a detailed explanation of the ranking-based methodology, specific to each analysis type, was presented.

Along the same line, Plazak and Souleyrette (2003) conducted an access management study of corridors near large urban areas in Iowa; specifically, they focused on four-lane expressways and two-lane arterials, facilities more likely to serve extensive commuter traffic. By using GIS spatial and statistical data corresponding to these corridors (e.g. crash records, demographics, land use, orthophotography, satellite imagery, roadway configuration and traffic composition) they developed a ranking method for identifying the routes that need the most attention with regards to access management. The analysis revealed that frequency and loss are highly rank-correlated (reason for which they were not used together in the final composite priority rankings). Other findings revealed that two-lane rural cross-sections were the highest ranked ones, followed by few four-lane expressways having at-grade private driveways and public road intersections, and that the regions with the fastest population, employment and commuting activity growth were linked to the highly ranked corridors in terms of safety deficiencies. Overall, the results led to the implementation of this corridor management assessment method at some of the highest-ranked corridors identified through this research.

Rees (2003) also conducted a study on effective corridor management practices. His study focused on U.S. 20, one of eastern Iowa's corridors characterized by high traffic volumes as well as a growing population and development; these characteristics made it a candidate for being treated with the measures proposed in this study. These measures were based on identifying the access-related problems along the corridor in order to develop the corresponding strategies; in addition, Rees emphasized the important role that planning authorities have throughout the process. GIS spatial information and crash data (1997-2000), as well as driveway and signal inventories, were combined in order to develop treatments for the selected corridors with the highest crash rates; also, the economic costs assigned by the DOT were associated with each crash according to its severity level. With the implementation of these measures, it was demonstrated that crash-related costs (e.g. minimum property damage, reduced injuries and saved lives) had the potential of being significantly reduced.

The importance of assessing safety performance of corridors is also reflected in many of today's transportation Long Range Transportation Plans (LRTPs). Such an example is that of the U.S. 31 Kokomo Corridor Project (INDOT, 2003), having Howard County, Indiana, as the study area. The motivation of this project is to provide overall transportation improvements for this "Statewide Mobility Corridor" (i.e. a corridor that provides safe, free-flowing and fast connections among States in the U.S.A.). One of the main goals of this project was to decrease crash rates through a reduction in delay and traffic congestion; for example, crash rates could be reduced by making changes to the facility type (i.e. corridor) and vehicle distribution, as well as by improving operational conditions (i.e. achieving a minimally acceptable LOS D or better).

Fontaine and Read (2006) documented the results from a study of three Highway Safety Corridor (HSC) programs adopted and applied in Virginia since their implementation in 2003. By definition, HSC programs deal with safety issues by implementing countermeasures based on the “3 Es”: enforcement, education and engineering. Based on 2004 and 2005 preliminary data, they reported that the program produced positive and negative results in terms of speed reduction and safety benefits; most of these variations are due to a lack of enough site data and allocated resources for a systematic implementation of countermeasures.

Focusing on more specific measures, Green and Blower (2007) investigated the effect of signal timing optimization on corridors’ safety conditions. They used intersection-related data from the southeast Michigan area, as well as crash data (2001-2005) corresponding to conditions of both before and after signal timing optimization. After mapping 130 intersections through GIS spatial analysis software, the respective crash statistics were obtained for injury severity, crash type, as well as time of day and day of the week; for example, rear-end crashes were found to be the most frequent at these locations. A before-and-after statistical model indicated that there was not a single effect from signal timing optimization; instead, the results varied from overall crash reduction, no change at all, and crash increase after implementing signal optimization. In addition, the effects of signal optimization on severe and some crash types (e.g. single-vehicle, head-on, angle, rear-end, sideswipe, unknown) were investigated; it was found that higher percentages of angle crashes and lower percentages of same-direction crashes were linked to intersections that had their crash occurrence reduced after signal timing optimization, contrary to those that did not show any change in terms of crashes.

Finally, a study conducted by Das et al. (2008) provided insights on the safety assessment of urban arterial corridors in Florida; specifically, the study focused on the effect that intersections along these corridors exert on crashes that occurred at or even beyond their physical area. Through the simultaneous estimation of crash injury severity and crash location (intersections vs. road segments), they were able to account for common factors affecting the severity and location characterizing a crash. Some of the results of this simultaneous estimation suggested that crashes along corridors tend to be less severe if they occurred on blacktop surfaces and/or during afternoon peak traffic conditions; on the other hand, more severe crashes are likely to happen at higher speed limits, wider pavement surfaces, and at a lower than median AADT. Furthermore, it was found that dry pavements (i.e. pavement condition) are significant when differentiating intersection- vs. segment-related crashes at low influence distance thresholds ($\leq$ 50 ft), whereas blacktop surfaces (i.e. pavement type) are significant at higher thresholds ($\geq$ 150 ft).

2.5 Insights on the Development of Traffic Safety

2.5.1 Relating Safety with Transportation Planning

Recent literature on transportation planning practices denotes the need to develop new and innovative performance measures applicable to the transportation planning process. Based on that premise, the work of Kononov and Allery (2004) fits in this category. With the aim to fulfill part of the objectives within the Transportation Equity Act for the 21st Century (TEA-21) of 1998, their proposed methodology would facilitate the explicit consideration of safety within the transportation planning process; they implemented the following safety-based performance measures and/or standards into two case studies: Safety Performance Functions (SPFs), Level of

Service of Safety (LOSS), and Diagnostic Analysis. Through these studies, they demonstrated that these safety-based measures would provide planners with an anticipated level of safety when developing transportation-related projects.

In a similar way, Ladrón de Guevara, Washington and Oh (2004) also noticed how critical it is to account for safety in the transportation planning process. Recalling the aforementioned TEA-21, they recognized that road safety is an important but also commonly neglected area within transportation planning, and that there is a need for reliable tools that would forecast safety at the regional planning scale. In order to cover these objectives, they developed a simultaneous negative binomial model that would use demographic data as well as information on crashes (1998 and 1999) corresponding to Tucson, Arizona; their model would predict crashes at the planning level (i.e. at the Traffic Analysis Zone, TAZ, level or higher). Overall, their study demonstrated that planning-level safety models are feasible and should be inherent to the planning activities of tomorrow; however, they also recommend not relying solely on this type of models for countermeasure selection or policy decision-making, since these models are developed for making long-range approximate forecasts solely. Their effort would serve as an incentive for the promoting the implementation of safety improvement programs.

Handy (2008) conducted a study that assessed today's regional transportation planning practices in the U.S.A. More specifically, her goal was to examine ways in which the transportation planning process can evolve with regards to its technical aspects and/or methodologies as well as the policies governing these. She based their study on the regional transportation plans from four MPOs in the U.S.A.: 2 from California, 1 from Washington State and 1 from Minnesota. After studying the latest goals, performance measures and strategies of these entities, she concluded that traffic congestion continues driving the planning process because of their well-established travel demand models; so, if the emerging and new goals are important, then new tools also need to be developed and entrenched in the planning field. In addition, she recommended having performance measures more qualitative in nature (i.e. to use simpler counts or rating scales, etc.).

2.5.2 On the Road toward the Highway Safety Manual (HSM)

As it is known in the transportation field, continuous efforts have been recently done towards the release of a new HSM –the equivalent, in safety terms, of the HCM–. An example of such efforts is represented by the work of Fitzpatrick, Schneider IV and Carvell (2006), who presented the results from their application of the Draft Prototype Chapter (DPC) for rural two-lane highways to be included in the HSM. Through two case studies for the Texas Department of Transportation (TxDOT), they explored the applicability and effectiveness of the DPC's methodology, as well as the different approaches for the calculation of calibration factors; this would constitute a good reference for improving future DPC editions along with the rest of HSM chapters. Based on their results, they conclude that more guidance is needed for road segment identification (e.g. definition of segment length thresholds, etc.), for prediction combinations (e.g. individual road segments vs. the roadway as a whole, etc.), on recommending assumptions, and on how transportation authorities and professionals could make the most of the predictions resulting from the proposed DPC's methodology.

Sun et al. (2006) also developed safety prediction models having a transferable methodology (i.e. applicable to different locations/regions and transportation facilities – due to the differences in crash data recording and driving behaviors–), optimum degree of accuracy, efficient yet simple calibration process, as well as good reliability, that have been proposed for inclusion in Part III of the soon to be released HSM. They based their models on road segment information as well as crash data (1999-2001) corresponding to rural two-lane highways in Louisiana; these models included predictions for severe (e.g. fatal, injury and PDO) as well as different single- and multi-vehicle crash types. Results suggested that the crash frequency models performed reasonably well; minor differences were observed between the observed and predicted crash frequencies.

CHAPTER 3: METHODOLOGY

3.1 Overview

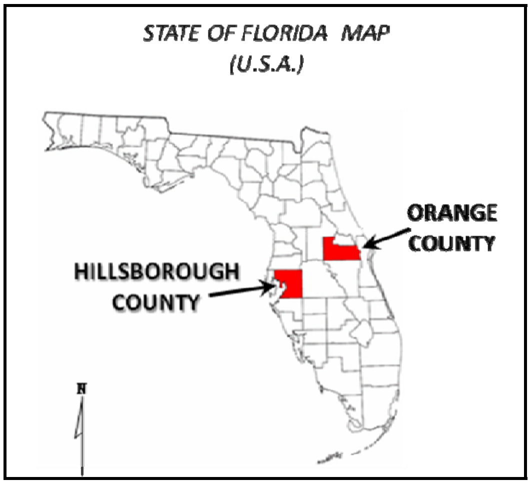
In research, some of the most important requirements towards producing meaningful contributions are to count with trusted resources and a good plan to follow. Based on this premise, this research incorporated data from varied sources in order to obtain representative samples to work with. In addition, after obtaining insights on the procedures from the studies referred to in the previous pages, the methodology proposed here proves to have become a good foundation for the subsequent chapters of this thesis.

This chapter provides a thorough view of the methods used for each of the two LOS-Safety studies' research design. Ranging from the minimum details to the most general ones, the overall description of this methodology will guide the reader towards understanding the true nature and value of the data being used.

3.2 LOS-Safety Study for Signalized Intersections

3.2.1 Study Area

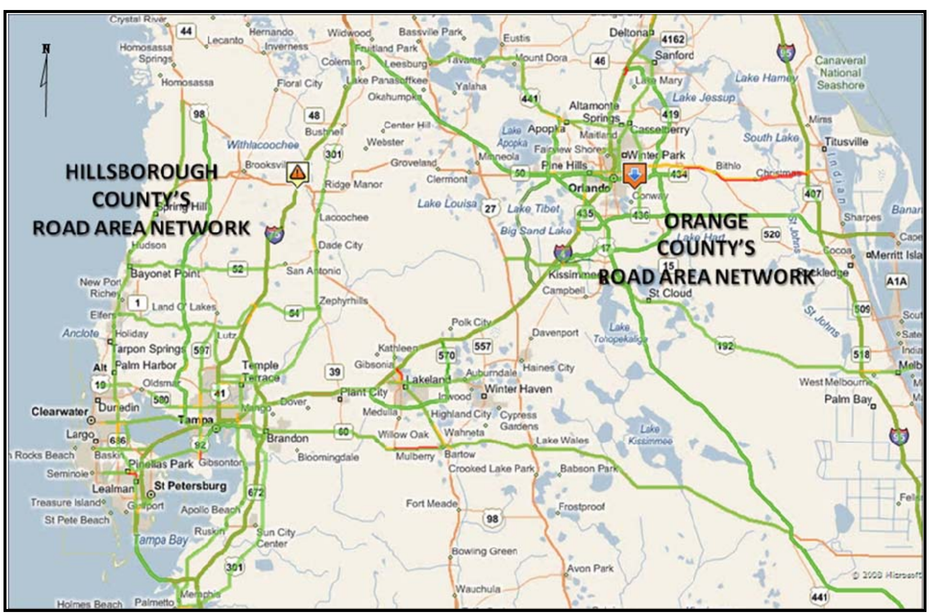
This first study used signalized intersection data that correspond to two major counties in the State of Florida: Orange and Hillsborough. Located in the central and west central Florida areas, respectively, Orange and Hillsborough counties produce and attract a big majority of the trips that take place in the State (see Figure 3-1).



(Source: http://www.floridacountiesmap.com)

Figure 3-1: Map of the State of Florida Displaying Orange and Hillsborough Counties

The fast-paced growth of these two counties is mostly influenced by tourism, appealing weather conditions, location and overall economic activity, which in turn have demanded these counties to count with a connected and well-integrated transportation network (see Figure 3-2).



(Source: http://maps.live.com)

Figure 3-2: Road Network Maps for Orange and Hillsborough Counties

Currently, the roads of Orange and Hillsborough counties are characterized by a high degree of traffic congestion that has been increasing at an almost continuous pace. This deteriorating condition has caught the attention of local authorities and transportation officials for improving traffic flow, a phenomenon that is highly affected by the operations of signalized intersections along these roads. Overall, this has raised awareness for road safety and on how this traffic congestion may have been contributing to crash occurrence at these locations. For these reasons, this first study has focused on how LOS, as a measure of traffic congestion, can be used in a safety evaluation framework specific for signalized intersections. Figure 3-3 provides a better representation of the sample of signalized intersections considered for the study.

(Source: http://maps.live.com)

Figure 3-3: Map Displaying the Major Concentration of Signalized Intersections in the State of Florida, Including Orange and Hillsborough Counties

3.2.2 Data Preparation

3.2.2.1 Intersection Data

Overall, the intersection data used was a subset of information and raw data coming from previous research presented by Wang (2006). For the purpose of this study, the data that was used consisted of 52 intersections from Orange County and 97 intersections belonging to Hillsborough County, being each intersection formed by at least one state road. The result was a set of 149 signalized intersections; being all 4-legged, this constituted a total of 596 intersection approaches. Also, it is to be noted that the sample included coordinated intersections (i.e. having a pre-timed phasing and/or signal timing pattern) as well as some of the isolated type (i.e. having a "free" signal timing pattern actuated through a loop detector). See Appendix A for the complete list of intersections composing the study sample.

Indeed, the signalized intersections' sample that was used was a very representative one; apart from the variation with regards to geometric configuration and design, as well as to signal operations, the sample's observed land use composition (37% urban, 55% suburban and 8% rural) is also a factor that confirms this fact. Overall, three main types of data were necessary for the LOS calculation to be performed afterwards; following are the details on these data types.

3.2.2.1.1 Geometric Configuration and Design Features

Information on the geometric configuration and design features for each of the signalized intersections in the sample was organized in a very large and detailed dataset. Most of the information on these features had been obtained back in 2004 with the aid of blue prints and CAD files provided by the respective counties. Google Earth (2008) was accessed to refine and to determine additional parameters. Google Earth is one of the most practical and efficient aerial imagery tools available; this software application permitted an easy access to each signalized intersection's actual location, information that could be retrieved by the respective "ID" assigned to the intersection (e.g. OC 93, HC 1162, etc.). Figure 3-4 depicts the display in Google Earth of one of the intersections in the study sample; features such as number of lanes, orientation of both major and minor roads, as well as the presence of signals, markings, medians and such, can be retrieved through the display.

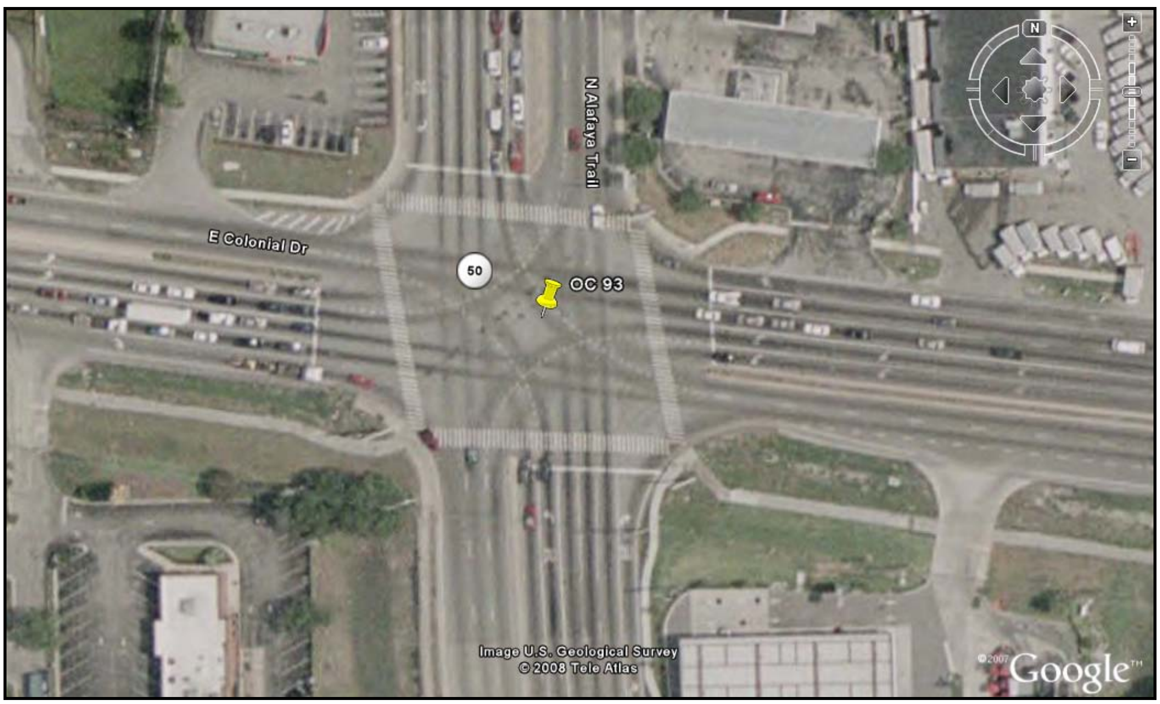


Figure 3-4: Google Earth View of Signalized Intersection "OC 93" (Colonial Dr. (SR 50) with Alafaya Trail (SR 434))

Due to the large amount of features within the aforementioned dataset, an educated selection of the variables that would be most relevant to the study was made. Next are the details on the content of the final dataset:

- *Intersection ID*: A unique name given to each signalized intersection (e.g. OC 93, HC 1162, etc.); this ID was created for the purposes of this study only.
- *County Node*: County node number corresponding to each intersection (e.g. 3, 1162, etc.); this is a unique number provided by the respective county offices and/or authorities which matches their databases.
- *SR Node*: State road number corresponding to each intersection (e.g. 5199, 8075, etc.); this is a unique number provided by the respective county offices and/or authorities which matches their databases.
- *Mile Point*: Number indicating the exact location of the signalized intersection along a specific road's length (in miles) from the study area (e.g. 11.736, 4.818, etc.); this is a unique value provided by the respective county offices and/or authorities which matches their databases.
- *Corridor*: Indicates the name of the main corridor forming the signalized intersection (e.g. SR 50 (East), SR 597, etc.)
- *Direction*: Indicates the orientation/direction of the signalized intersection's major road (e.g. West-East, North-South, etc.)
- *Major and Minor Roads*: Indicates the names of the major and minor roads forming the signalized intersection (e.g. Colonial Dr. and Alafaya Tr., Fletcher Ave. and Dale Mabry Hwy., etc.).
    - In this study, the major road is considered to be the one with the highest total number of through lanes and/or highest Annual Average Daily Traffic (AADT).

- *Number of Left-Turn Lanes*: Indicates the observed total number of left-turn lanes along the major and minor roads forming the signalized intersection.
- *Number of Through Lanes*: Indicates the observed total number of through lanes along the major and minor roads forming the signalized intersection.
- *Number of Right-Turn Lanes*: Indicates the observed total number of right-turn lanes along the major and minor roads forming the signalized intersection.
- *Length of Left-Turn Lanes*: Indicates the observed maximum length (in feet) for the group of left-turn lanes along the major and minor roads forming the signalized intersection.
- *Length of Right-Turn Lanes*: Indicates the observed maximum length (in feet) for the group of right-turn lanes along the major and minor roads forming the signalized intersection.
- *Median Type*: Indicates the observed type of median along the major and minor roads forming the signalized intersection (e.g. No median, Narrow median and Wide median).
- *Speed Limit*: Indicates the observed speed limit (in miles per hour) along the major and minor roads forming the signalized intersection.
- *Land Use*: Indicates the land use corresponding to each signalized intersection (e.g. Rural, Suburban and Urban).

3.2.2.1.2 Signal Timing

Information pertaining to the signal timing of each intersection was provided by the respective Traffic Engineering offices from Orange and Hillsborough counties. This was available for each intersection in the sample and was contained in the form of individual signal timing plans, all in electronic format (.xls and .pdf).

Albeit each county showed a unique way to organize their signal timing plans, the information had common and basic elements that would still facilitate the LOS calculation for the intersection sample to be analyzed (see Appendix B). Following is a relation of the most relevant information provided by these plans:

- *Intersection Information*
  - *Location*: Indicates the where the signalized intersection is located by providing the names of its corresponding major and minor roads (e.g. Colonial Dr. and Alafaya Tr., Fletcher Ave. and Dale Mabry Hwy., etc.)
  - *Node number*: County node number corresponding to each intersection (e.g. 3, 1162, etc.); this is a unique number provided by the respective county offices and/or authorities which matches their databases.
  - *Date*: Indicates the date at which the signal timing plan was approved by the respective Professional Engineer (P.E.) for implementation.
- *Phases and Basic Timing*
  - *Phase (Ø)*: Denotes the 8 main phases (labeled 1 through 8) for a given cycle; there is a phase per approach (Eastbound, Westbound, Northbound and Southbound) and per left-turning movement (Eastbound Left, Westbound Left, Northbound Left and Southbound Left).
    - ➢ The phasing sequence of each turning movement varies per signalized intersection.

- *Split*: Denotes the duration (in seconds) of a individual phase; there is a split per approach (Eastbound, Westbound, Northbound and Southbound) and per left-turning movement (Eastbound Left, Westbound Left, Northbound Left and Southbound Left).
- *Direction*: Denotes the basic movements (8 in total) that are allowed at the intersection; there is a direction per approach (Eastbound, Westbound, Northbound and Southbound) and per left-turning movement (Eastbound Left, Westbound Left, Northbound Left and Southbound Left).
- *Lead/Lag*: Denotes whether a left-turning phase is leading or lagging during a given cycle.
- *Minimum Green*: Denotes the minimum duration of green times (in seconds) for the signalized intersection; there is a minimum green time per approach (Eastbound, Westbound, Northbound and Southbound) and per left-turning movement (Eastbound Left, Westbound Left, Northbound Left and Southbound Left).
- *Maximum Green*: Denotes the maximum duration of green times (in seconds) for the signalized intersection; there is a maximum green time per approach (Eastbound, Westbound, Northbound and Southbound) and per left-turning movement (Eastbound Left, Westbound Left, Northbound Left and Southbound Left).
    - There are usually two sets of maximum green times as a minimum; each is considered for a specific signal pattern (e.g. coordinated and/or isolated).

- *Clearance*: Denotes the duration of yellow times (in seconds) for the signalized intersection; there is a clearance time per approach (Eastbound, Westbound, Northbound and Southbound) and per left-turning movement (Eastbound Left, Westbound Left, Northbound Left and Southbound Left).
- *Red*: Denotes the duration of all-red times (in seconds) for the signalized intersection; there is a red time per approach (Eastbound, Westbound, Northbound and Southbound) and per left-turning movement (Eastbound Left, Westbound Left, Northbound Left and Southbound Left).
- *Vehicle Extension*: Denotes the available vehicle gap (in seconds) at the signalized intersection; there is a vehicle extension per approach (Eastbound, Westbound, Northbound and Southbound) and per left-turning movement (Eastbound Left, Westbound Left, Northbound Left and Southbound Left).
- *Recall/Memory*: Denotes an actuated signal pattern.
- *Flash*: Denotes whether a Flashing pattern takes place at the signalized intersection.
    - ➢ It is to be noted that the Flash pattern was not considered in this study, even though it could have fit within the analysis for the Early Morning period (1-3 A.M.). This was decided for the following reasons: 1) not all intersections had a flashing pattern, and 2) this created several limitations for the respective LOS calculation with HCS.

- *Coordination Patterns*
    - *Cycle Number*: Denotes the assigned name of a specific cycle and/or coordination pattern (e.g. 1/1/1, 2/1/1 and 3/1/1; Cycle 1, Cycle 2 and Cycle 3; etc.)

- *Cycle Length*: Denotes the minimum and maximum duration (in seconds) allowed for a specific cycle and/or coordination pattern, resulting from the summation of the respective phases' time splits.
- *Time*: Denotes the time at which a specific time split starts at the signalized intersection (e.g. 9:00, 15:00, 23:00, etc.)
- *Day*: Denotes the day(s) on which a specific cycle and/or coordination pattern takes place at the signalized intersection (e.g. 1=Sunday, 2=Monday, 3=Tuesday, 4=Wednesday, 5=Thursday, 6=Friday and 7=Saturday).

3.2.2.1.2.1 Phase Diagrams

In order to have a better representation of the data contained in the signal timing plans, separate phase diagrams were made for each of the coordinated and isolated intersections in the study sample. By definition, a phase diagram is a graphical representation of how a complete signal timing cycle behaves for a particular intersection; it displays the sequence and overlaps (if any) of all phases as well as the corresponding types of turning movements allowed in each phase.

A phase diagram was made for each period of the day considered in the study: Early Morning (1-3 A.M.), A.M. Peak (7-9 A.M.), Midday (12-2 P.M.), P.M. Peak (4-6 P.M.) and Late Evening (8-10 P.M.) In addition, signal timing data only corresponding to regular weekdays (i.e. Monday through Friday) were considered; this was done since traffic safety conditions differ considerably over the weekend. The result was a good graphic representation of the sequence (e.g. leading, lagging, etc.), type of protection (e.g. protected, permissive and combined) of the main turning movements for each of the intersection's approaches. As an example, Figure 3-5 shows one of the phase diagrams made for signalized intersection “OC 93” (Colonial Dr. (SR 50) with Alafaya Trail (SR 434)); it was labeled by Orange County’s Traffic Engineering Office as “Coordination Plan 3/1/1”, which run from Monday (Day 2) through Friday (Day 6), from 3 P.M. to 7 P.M.

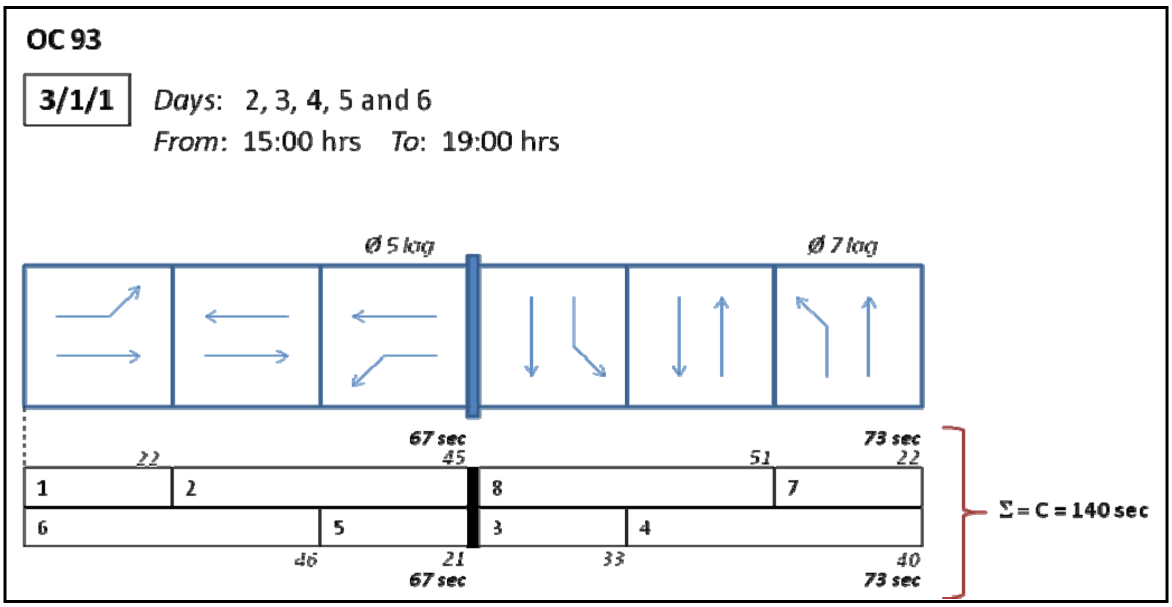


Figure 3-5: Sample Phase Diagram for Signalized Intersection “OC 93” (Colonial Dr. (SR 50) with Alafaya Trail (SR 434)) Corresponding to the P.M. Peak Period

Overall, it has to be emphasized that engineering judgment played an important role when interpreting the signal timing plans for each intersection; the task not only consisted in using the data straightforward, but also in making sure that the resulting turning movement patterns per phase and cycle lengths are realistic with regards to the study area.

3.2.2.1.3 Turning Movement Counts and Approach Volume

Accurate data on the turning movement counts were also provided by the respective offices from Orange and Hillsborough counties. For each intersection, turning movement counts were provided for one year only, one within the 6-year period of interest (2000-2005); also, since the investigation focused its observations on the average traffic volume, most of these counts corresponded to the middle year (2003). These data represent the approach traffic volume for each intersection in the sample and were also stored in electronic format (.xls and .pdf). Appendix C contains a sample sheet containing the respective turning movement counts, as well as other related parameters, provided by Orange County.

These volumes were recorded for 15-minute intervals and covered the periods of the day to be analyzed in this study. This type of data complements the two other intersection data types (i.e. geometric configuration/design features and signal timing) in order to perform the respective LOS calculations for different periods of the day. Following is a relation of the information contained in the traffic count files:

- *Approach*: Indicates the signalized intersection's approach for which particular traffic counts have been collected (e.g. Eastbound, Westbound, Northbound and Southbound).
- *Traffic counts*: Indicates the total traffic counts collected for the signalized intersection; there is traffic count data collected per approach (Eastbound, Westbound, Northbound and Southbound) and per turning movement (Left, Through and Right).
- *Time*: Indicates the time at which a particular 15-minute traffic count started (e.g. 7:00, 7:15, 7:30, 7:45, 8:00, etc.)
- *Peak-Hour Factor (PHF)*: Indicates the PHFs corresponding to the signalized intersection; there is a PHF per approach (Eastbound, Westbound, Northbound and Southbound) and per turning movement (Left, Through and Right).

- All PHFs were calculated for each period of the day considered in the study; this was done by using data from the hour (i.e. four 15-minute intervals) having the highest total volume corresponding to the period of interest.

At some instances, when the traffic counts themselves were not available for some of the intersections, particularly for the Early Morning and Late Evening periods of the day, these had to be estimated using an adequate engineering judgment and considering the overall characteristics (i.e. size and location, etc.) of the corresponding intersection. Also, when other volume-related data were not provided, these had to be calculated following the procedures recommended by current practice. As an example, Equation 1 below represents the procedure for calculating the PHF, as specified in the HCM (HCM, 2000):

$$PHF = \frac{V}{4 \times V_{15}} \qquad (1)$$

where PHF is the calculated Peak-Hour Factor, V is the traffic volume per hour (in vehicles/hour) and V_{15} is the volume corresponding to the peak 15 minutes within the respective peak hour (in vehicles/15 minutes).

3.2.2.2 Crash Data

A representative crash data set corresponding to the study area was used. This set of information was downloaded from the Crash Analysis Reporting (CAR) system, a database maintained by the Florida Department of Transportation's (FDOT) Safety Office, and which contains the most details belonging to crash reports from around the State; for this reason, the CAR database is frequently accessed by researchers in the field. The crash data considered corresponded to the years 2000 through 2005, a 6-year span compatible with the year(s) and/or time frame of the other types of data (i.e. intersections', roadway features', etc.) for this study.

The resulting crash data set reported a total crash frequency of 5,532 events; these crashes were linked to the 149 signalized intersections of interest and correspond specifically to crashes occurring on weekdays and within the time periods of interest (refer to Section 3.2.2.1.2.1). An additional condition for these data was to be catalogued as "at intersection" or "intersection-related" only. Details corresponding to the crash events' severity levels, types, as well as other related information, were included in this dataset.

With regards to severity, this is an important indicator of the injury level sustained by the driver, any pedestrian(s) involved, or by both, at any crash event. Florida's DHSMV, the agency that regulates highway safety matters within the State, has established a 5-level injury severity scale, denoted by numbers, which police officers have to consider when completing a crash report (FDHSMV, 2007); consequently, these levels also characterize the crash data used for this study. Following is the description of these injury severity levels:

- *No Injury (1)*: Level indicating that none of the individuals involved in the crash were physically harmed. This is the equivalent to Property Damage Only (PDO) crashes.
- *Possible Injury (2)*: Level indicating that no visible signs of injury were observed after the crash event; however, complaints of temporary unconsciousness and/or pain were reported.
- *Non-Incapacitating Evident Injury (3)*: Level indicating that visible signs of injury were observed after the crash event (e.g. abrasions, bruises, limping, etc.)

- *Incapacitating Injury (4)*: Level indicating that in addition to having visible signs of injury after the crash event (e.g. abrasions, bruises, limping, etc.) the individuals involved had to be carried away from the scene.
- *Fatal Injury (5)*: Level indicating that the injury or injuries sustained after the crash event resulted in the death of the individual(s) involved, this happening within 30 days after the crash.

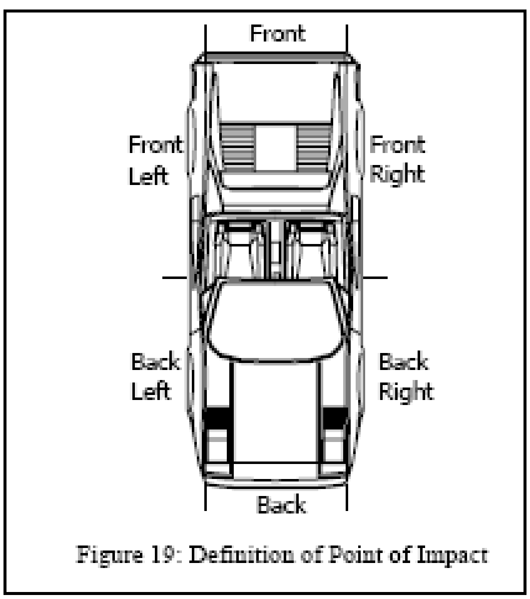
For the purpose of this study, only two severity levels were used: level 4 (incapacitating injury) and level 5 (fatal injury) were considered since these two levels fall within the severe crash category.

In addition, from all the crash types reported in the CAR database, only five were considered for the study; following is their description:

- *Rear-End*: Crash that occurs when a vehicle hits another vehicle from behind; the two were traveling along the same direction. This is the most common type of crash in the U.S.
- *Sideswipe*: Crash that occurs when a vehicle, when trying to pass another vehicle, strikes the latter along the side; the two were traveling along the same direction. This type of crash can also occur when a vehicle is traveling through a divided highway in the wrong direction.
- *Head-On*: Crash that occurs when two vehicles' front ends hit each other in a frontal or angular manner; the two were traveling in opposite directions. This is the crash type associated with most fatal crashes due to the large magnitude of the impact's force involved.

- *Right Angle*: Crash that occurs when two vehicles collide as a result of one of them omitting the Stop or Yield sign, runs a red light, or has not cleared the intersection while the conflicting approach's green signal has already been activated; the two were traveling in non-opposing directions
- *Left Turn*: Crash that occurs when a vehicle attempts to make a left-turn in front of an opposing vehicle already in motion.

The crash types just listed are frequent at signalized intersections. Studies like the one conducted by Wang et al. (2009) reflect this fact; for this reason, and for consistency, it was considered appropriate to use these crash types in order to make further contributions with the study presented here. Figure 3-6 depicts a regular passenger car with its possible points of impact during a crash event; from these, the different crash types are derived.



(Source: Wang, 2006)

Figure 3-6: A Passenger Vehicle's Possible Points of Impact

3.2.2.3 Additional Road Features Data

When necessary, the Roadway Characteristics Inventory (RCI), database also maintained by the FDOT, was accessed for retrieving road features information that had not been initially provided within the set of intersection data detailed in Section 3.2.2.1. For this reason, data from the years applicable to the study were downloaded. Considered to be a complementary dataset, this was to be merged with the overall signalized intersections' data to be used in the study.

3.2.3 LOS Data Preparation

3.2.3.1 Preliminary Steps

Having already prepared a final dataset per main type of data, as detailed in Section 3.2.2, the next key step in the methodology was to use the information in these datasets as inputs for obtaining the LOS of each signalized intersection in the study sample. Based on this, the most appropriate tool for calculating the LOS was the Highway Capacity Software (HCS), one of the most widely used and trusted software packages, opinion shared by several transportation capacity analysts and experts worldwide (McTrans, 2008); following is a brief overview of this valuable tool.

3.2.3.1.1 Highway Capacity Software (HCS)

Developed by the Center for Microcomputers in Transportation (McTrans), an entity established by the Federal Highway Administration (FHWA) in 1986, HCS is considered to be "the most widely used transportation software package in the world" (McTrans, 2008). This software has served since 1987 as a companion of the HCM by firmly adhering to the procedures defined and prescribed in the manual. Its latest version has been released in 2008, resulting in a total of 6 versions of the software up to this day. It is to be noted that the study presented here used the 2005 release (i.e. 5th version of the software), based on availability, also labeled as HCS+ (see Figure 3-7); this release adheres to the guidelines of the 2000 edition of the HCM.

(Source: http://mctrans.ce.ufl.edu/hcs/downloads/)

Figure 3-7: Welcome Screen Display of the Highway Capacity Software (HCS+)

For this study's purposes, HCS was mainly used to retrieve the LOS at the intersection level. Still, the complete output from the software included additional information on LOS-related parameters (e.g. lane group capacity, flow rate, volume to capacity ratio (v/c), effective green time ratio (g/C) and delay). HCS' calculation accuracy, data processing efficiency, as well as its good output storage capability, all facilitated the traffic capacity analysis of the 149 signalized intersections in the study sample.

The following section provides a thorough description of a typical LOS calculation process, as it was performed for the study. It is to be noted that such a calculation was made for each period of the day, and for each intersection, considered in the study.

3.2.3.2 LOS Calculation

The first step when using HCS was to create a new file, per period of the day to be studied, indicating the type of operational analysis to be performed; specifically, the study discussed here required to choose the "Signalized Intersections Operational Analysis" option. Second, the option "Single Period" was selected as the analysis type; this option indicates that the analysis being performed corresponds to a single period of the day only. In the end, a total of 5 files were created per signalized intersection.

As an additional note, it was made sure that all assumptions and engineering judgment-based decisions complied with the corresponding procedures and acceptable default values for LOS calculation and highway capacity analysis already contained within HCS, set by the HCM (2000), and the recently released NCHRP Report 599: Default Values for Highway Capacity and Level of Service Analyses (Zegeer et al., 2008).

3.2.3.2.1 Geometry and Volume

For this first stage, the type of data to be entered was the number of outflowing and receiving lanes –all assumed to have a width of 12 ft–, indicating whether these are exclusive and/or shared, etc; regarding the available queue storage lengths for each approach, these were based on the length of their respective exclusive left turn lane(s). Other input data were the respective traffic volumes and PHF values. For the analysis, right turns on red (RTOR) were not computed since no exact turning movement data for these were available. All these considerations were applied for each of the intersection's approaches (i.e. Eastbound, Westbound, Northbound and Southbound).

3.2.3.2.2 Operating Parameters

This second stage consisted in entering basic operational parameters corresponding to the cycle being analyzed with the software. For example, the corresponding arrival type was specified (i.e. type 3 for uncoordinated and type 4 coordinated intersections, respectively), unit extension, start-up lost time, extension of effective green, etc.

3.2.3.2.3 Phasing Design

The first part of this stage consisted in presetting the respective phasing. Minimum and maximum green times, as well as clearance and red times were the first input values; at the same time, the total number of phases was set within the software's interface. For this process, the respective phase diagram was used (refer to Section 3.2.2.1.2.1) as it facilitated the input of turning movement types allowed per approach, their sequence (e.g. lead and/or lag, etc.), type of protection (e.g. protected, permissive and combined), etc. Another important component of this stage was the selection of the corresponding actuation type; whereas coordinated intersections required the pre-timed type, the isolated ones required the actuated type. Also, since isolated intersections do not count with a specific cycle length to consider (i.e. their cycle is not coordinated), the time estimation option had to be used for these since it allows the software to estimate adequate signal timing values in order to get the most realistic cycle length possible; as shown in Figure 3-8, this option allows the input of more information related to the parameters affecting the actuation of a traffic signal (refer to Section 3.2.2.1.2).

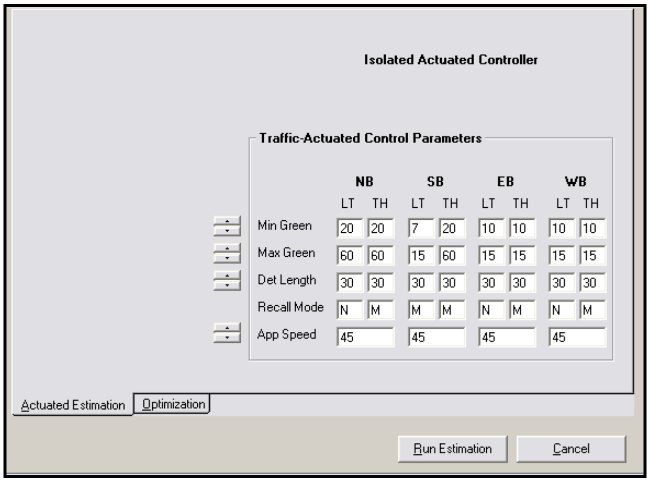


Figure 3-8: View of the Actuated Estimation's Window in HCS+

The second and last part of this stage was based on saturation flow rate adjustments, a very important component for determining the final LOS. For example, this consisted in specifying the saturation flow rate, parameter that was kept at its default value of 1,900 pcphgpl at all instances; also, the percent for heavy vehicles was set at 2%. Apart from that, the analysis did not consider parking maneuvers, bus stops, neither bikes nor pedestrian flows. All these assumptions were equally applied for each of the intersection's approaches.

3.2.3.2.4 Results

Finally, all operational measures corresponding to the signalized intersection were provided by HCS. In particular, this resulted in a more complete source of LOS-related information which now contained information on lane group capacity, flow rate, volume to capacity ratio (v/c), effective green time ratio (g/C), delay and LOS.

3.2.4 Final Data Assembly

Having completed the required data preparation processes, the last step for the study's research design was to assemble the respective data in a way that would facilitate the statistical analysis to be performed afterwards.

First, all final data and related outputs had to be saved in MS Excel spreadsheets, sorted mainly by the respective intersection ID and period of the day. Provided data corresponding to the variables roadway ID, mile point, node number and county node were all included in the spreadsheets, these were then imported into the Statistical Analysis Software (SAS). Next, SAS accurately merged the data contained in the spreadsheets; then, it sorted these by signalized intersection and by period of the day. The final result was an organized and robust dataset consisting of 237 columns (237 variables) and 745 rows (149 intersections x 5 periods of the day = 745 rows of data). Overall, this way of assembling the data was considered to be the most appropriate for the modeling approach considered, as will be detailed in Chapter 5, since it looks for the existing correlation among all periods of the day.

Furthermore, SAS was also used in the creation of variables not initially present in the original datasets. For example, new variables were derived from some of the geometry configuration and design features as well as from traffic volume data; these new variables were classified by the roads forming the intersection (i.e. major and minor). In addition, crash-related frequencies were computed for total crashes, as well as for severe crashes and the crash types being considered (see Section 3.2.2.2). Deemed as important, the new variables would also be used in this study's final analysis.

3.3 LOS-Safety Study for Multilane High-Speed Arterial Corridors

3.3.1 Study Area

In contrast to the study referred in Section 3.2, the data used for this second study correspond to road segments from multilane high-speed arterial corridors located in Hillsborough County, part of the west central Florida region (see Figure 3-9). As was previously mentioned, the transportation network across this county carries a large portion of the trips that take place in the State and is characterized by a continuous growth attributed to several factors (refer to Section 3.2.1).

(Source: http://www.floridacountiesmap.com)

Figure 3-9: Map of the State of Florida Displaying Hillsborough County

As it will be shown in the forthcoming analyses, most corridors in Hillsborough County are characterized not only by a deteriorated LOS but also by a high number of severe crashes; this also raises the question on whether traffic congestion may be contributing to crash occurrence along these roads. Consequently, the study presented in this section constitutes a second investigation of the LOS-Safety relationship but this time having corridors from multilane high-speed arterials as the context of interest.

3.3.2 Data Collection through GIS

3.3.2.1 Overview

The corridor data used came in the form of a Geographic Information Systems (GIS) file, which mapped a very large set of road sections belonging to the study area for which the respective LOS had been obtained. These data were provided by Hillsborough County's Metropolitan Planning Organization (MPO), through its City County Planning Commission (see Figure 3-10).

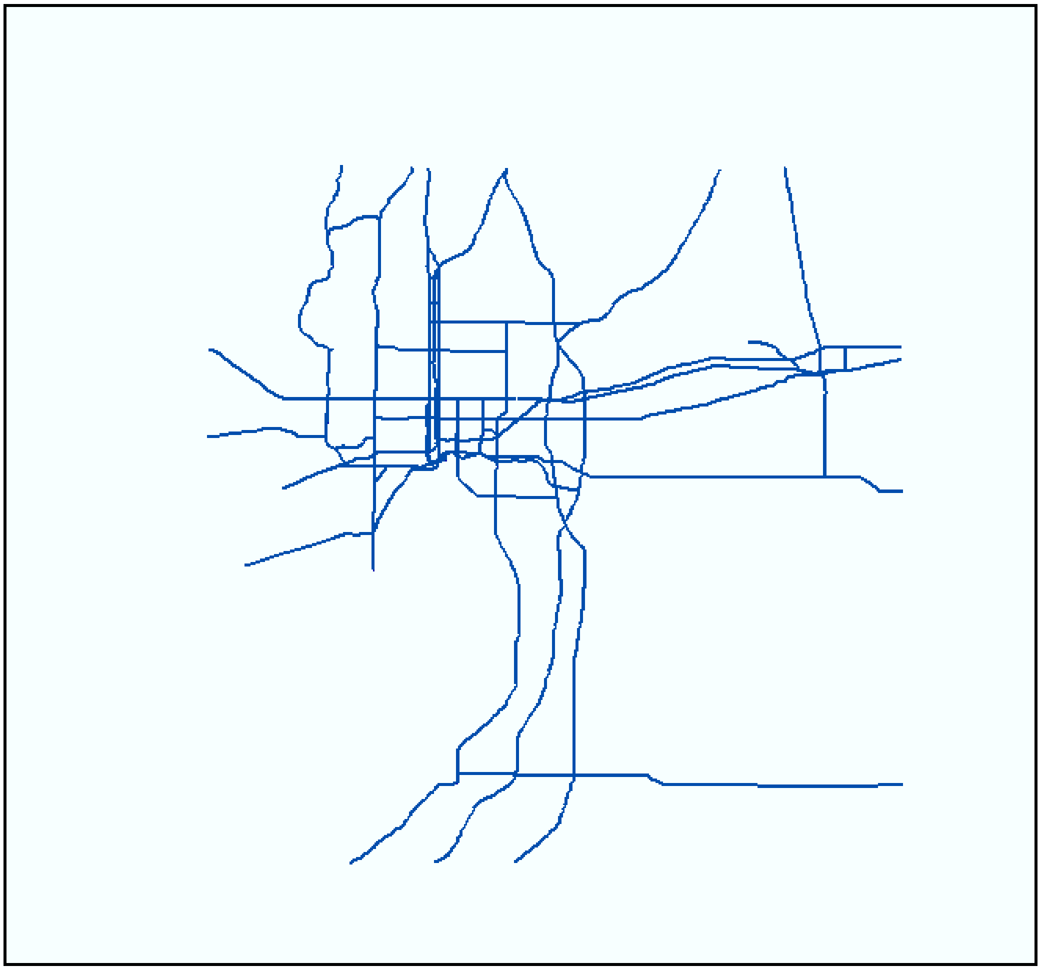
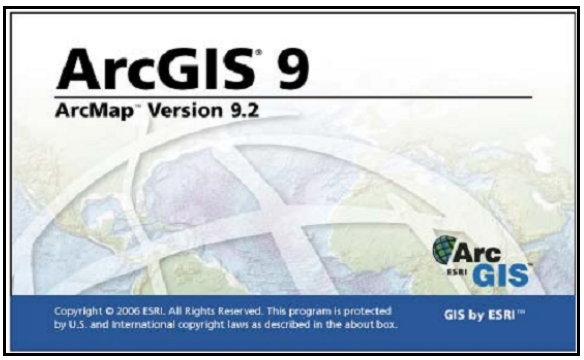


Figure 3-10: ArcMap View of Original GIS Layer File (with 1999 LOS and Road Data) Provided by Hillsborough County

Despite the fact that the road segments' LOS, parameter of utmost importance for the research conducted, was already included within the GIS file just mentioned, the absence of data corresponding to the variables roadway ID and mile point was found to be a very significant drawback from the start; these variables are critical for proceeding with the study's analysis since without them no merging with the data from the CAR (i.e. crashes) and RCI (i.e. road features) databases can take place. Fortunately, this could be overcome by using the ArcMap software package; following is a brief description of how this valuable tool served this study's purposes.

3.3.2.1.1 ArcMap

Originally developed by the Environmental Systems Research Institute (ESRI), a former consulting firm specializing in land use-related projects (ESRI, 2008), ArcMap is one of the components from the ArcGIS family of software packages; it is widely used by analysts in the planning and transportation fields around the world. This study used the 9.2 version of the software, corresponding to ArcGIS 9 (see Figure 3-11).



(Source: http://mctrans.ce.ufl.edu/hcs/downloads/)

Figure 3-11: Welcome Screen Display of ArcMap Version 9.2

For this study's purposes, the GIS features from ArcMap were used for obtaining the roadway IDs and beginning and end mile points for each of the road sections in the study sample. Following is a thorough description of such process.

3.3.2.2 <u>Data Collection</u>

3.3.2.2.1 GIS Layer Matching

As the name suggests, this first step of missing data collection consisted in matching layers of GIS files. Counting already with one GIS file in hand, the other one corresponded to a complete 2006 GIS road network map corresponding to the same study area (i.e. Hillsborough County), provided by the FDOT (see Figure 3-12).

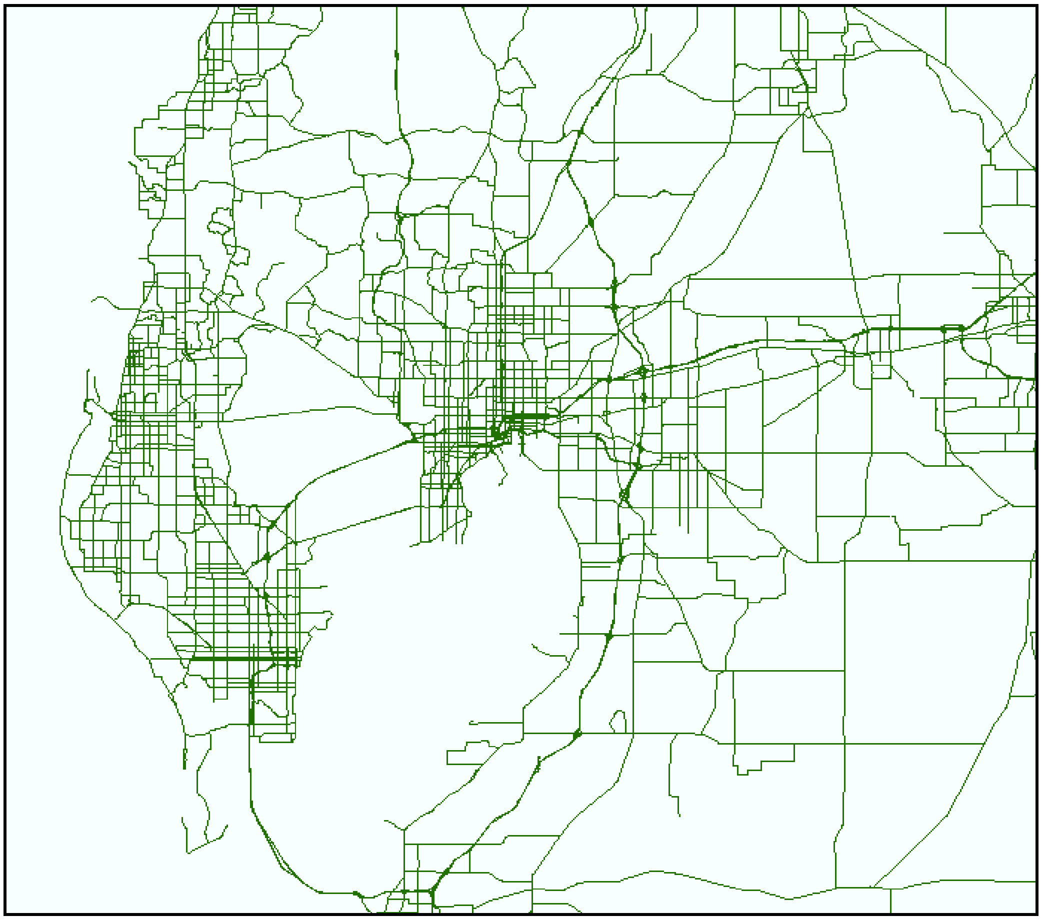


Figure 3-12: ArcMap View of Original GIS Layer File (with 2006 Road Data) Provided by FDOT

An advantage of the file just shown was that the data for all of Hillsborough County's roads were built on a series of individual road segments (i.e. consecutive road segments makes a road section). In addition, all the roadway IDs and beginning and end mile points for all road segments in the study area were contained in this GIS file. Despite the fact that both files corresponded to different years, 1999 vs. 2006, this did not interfere with the study's methodology. The matching sets of mapping coordinates allowed these two road network maps from Hillsborough County could be easily displayed on the software's interface one above the other (see Figure 3-13).

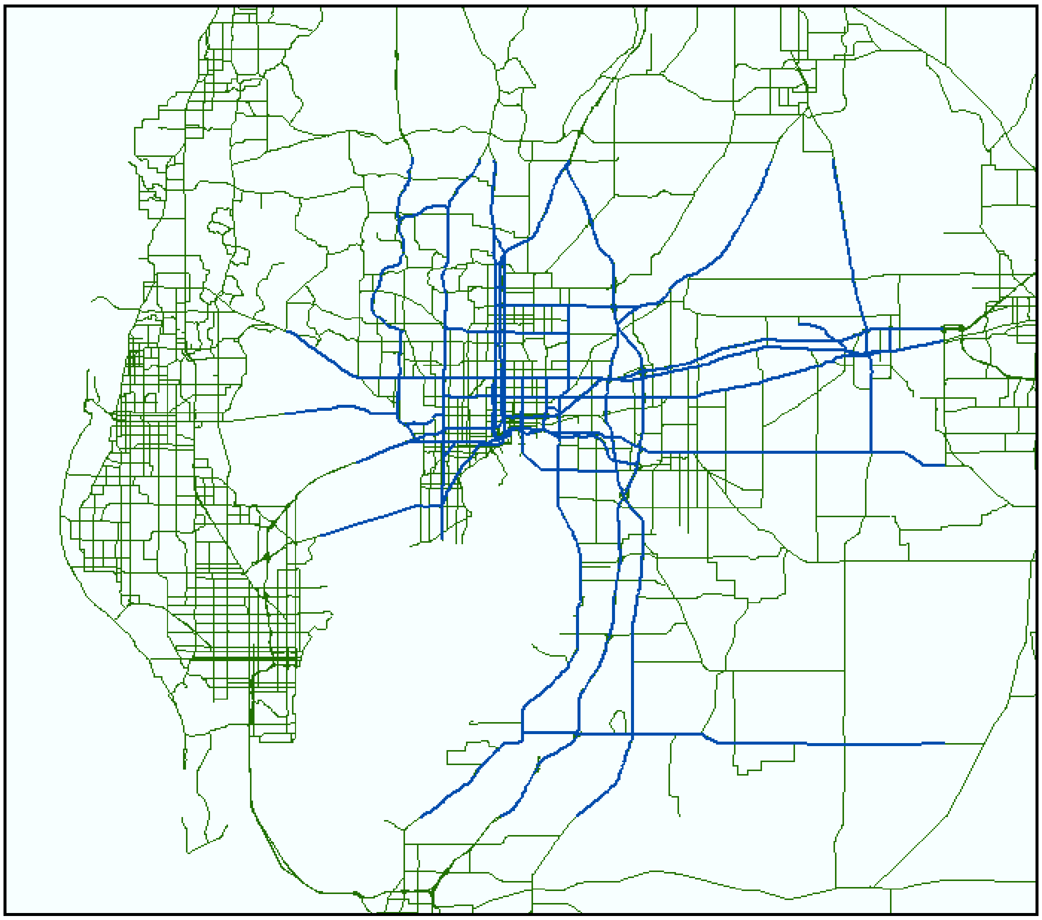


Figure 3-13: ArcMap View of the 2 Original GIS Layer Files (Hillsborough County's and FDOT's) Matching with Each Other

3.3.2.2.2 Data Retrieval

Having the road network layers from both GIS files matching with one another, the data that was missing could be retrieved. This second phase demanded a large number of hours for completion due to the large set of LOS data available (i.e. for 399 road sections); however, the process was not complicated to perform.

The retrieval of missing data started by having both GIS layers displayed on the computer screen. Next, a segment, for which a recorded LOS was available, had to be selected (i.e. highlighted) while having the 2006 road network GIS layer still displayed on the screen; then, the attributes table of the respective 2006 road segment was prompted on the screen, which listed a series of data including the roadway ID and beginning and end mile points for that segment; finally, these data were stored in a MS Excel spreadsheet (i.e. .xls). This procedure was done for each of the remaining road sections; appropriate engineering judgment was also applied during the process.

3.3.3 Data Preparation

3.3.3.1 Road Segments and LOS Data

After the successful data collection process performed with GIS applications, a data preparation procedure was needed in order to enhance and organize the rest of data available, particularly LOS data that was initially provided for each road section. These LOS have been obtained based on the guidelines from FDOT's Quality/Level-of-Service Handbook, taken from the ARTPLAN methodology (i.e. for arterials), which is also based on HCM 2000 standards (FDOT, 2002).

As mentioned in Section 3, the main corridor data used came in the form of a very large dataset, to be read through software compatible with GIS; the dataset's content was composed by records taken for road segments belonging to the study area. These data were provided by

Hillsborough County's Metropolitan Planning Organization (MPO), through its City County Planning Commission, and corresponded to the year 1999. The latter fact was considered a limitation; having started the study in 2008, it was somewhat difficult to obtain other traffic safety-related data for 1999, especially since the databases to be accessed (e.g. CAR and RCI) are constantly being updated. In order to overcome this limitation, the author considered years close enough to 1999 for the other traffic safety-related data that had to be used, also depending on availability; good engineering judgment was applied when making these decisions.

After a preliminary evaluation of the data in hand, the original dataset was reduced to 399 road sections, each composed by one or more state road (SR) segments, corresponding to multilane high-speed arterial corridors in Hillsborough County (see Figure 3-14). In addition, these road sections met the following classification(s) (in order to comply with the respective project proposal): minor and principal arterials (both rural and urban types), total number of lanes of 4 or greater, posted speed limit of 40 mph or greater, and could or not have an intersection along its path; again, it has to be emphasized that although these sections may be influenced by intersections, this thesis has considered a LOS-Safety analysis for the road section in general, since no data on the intersections in the area were available. Appendix D contains the complete list of road sections composing the final study sample.

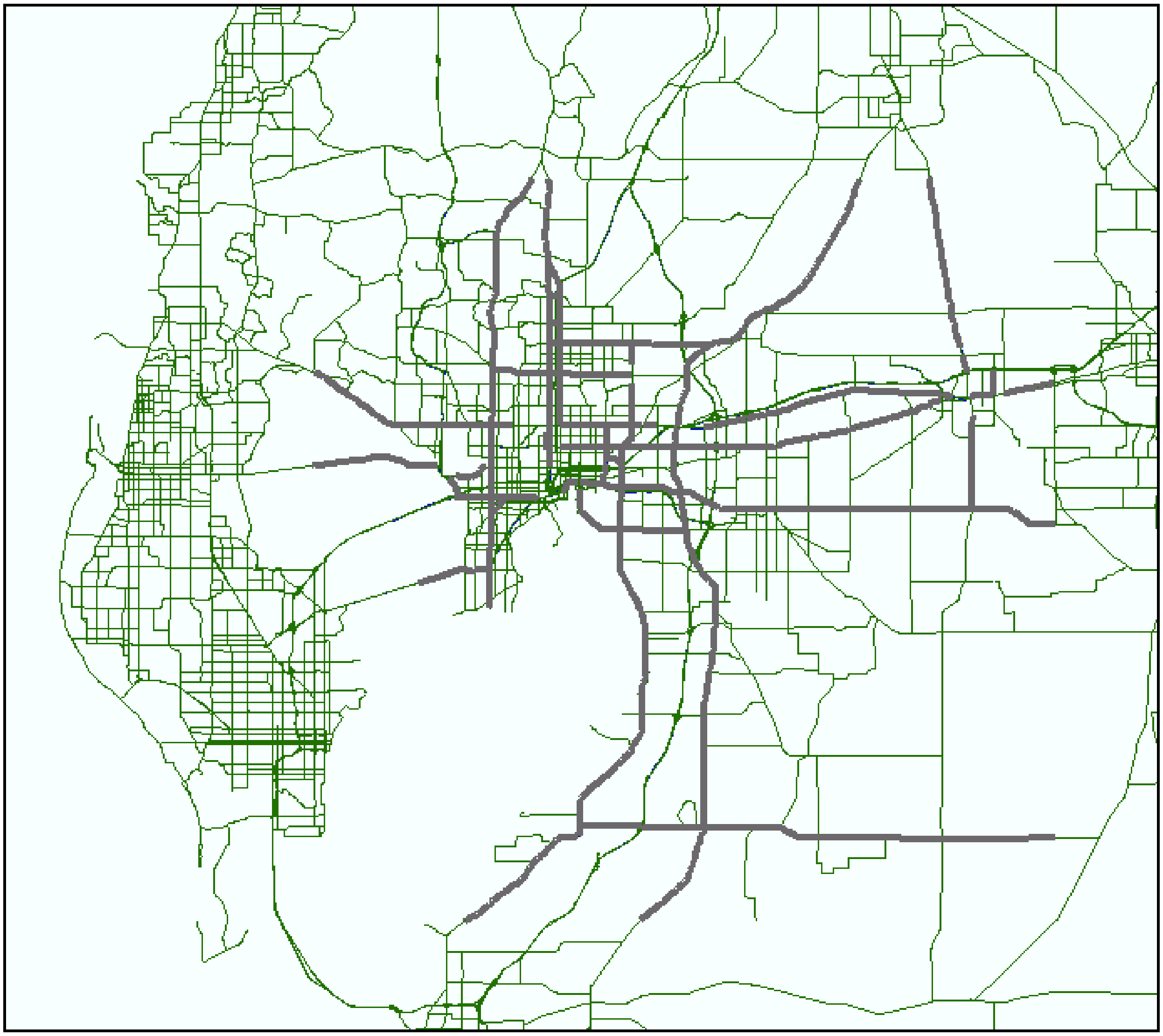


Figure 3-14: ArcMap View Highlighting the 399 Road Sections Considered for the Study

Having defined the road sections to be studied, GIS would provide a better view of the spatial LOS and crash distributions for the study. Following are Figures 3-15 through 3-20, which show the respective LOS distribution. In addition, Figure 3-21 will show the distribution of severe crashes (incapacitating and fatal), which are of utmost importance for the FDOT project motivating this research. Overall, these spatial distributions denote the importance of analyzing crash occurrence at multilane high-speed arterials in general, how to alleviate their operating conditions and to understand how the LOS and road safety conditions interact with each other.

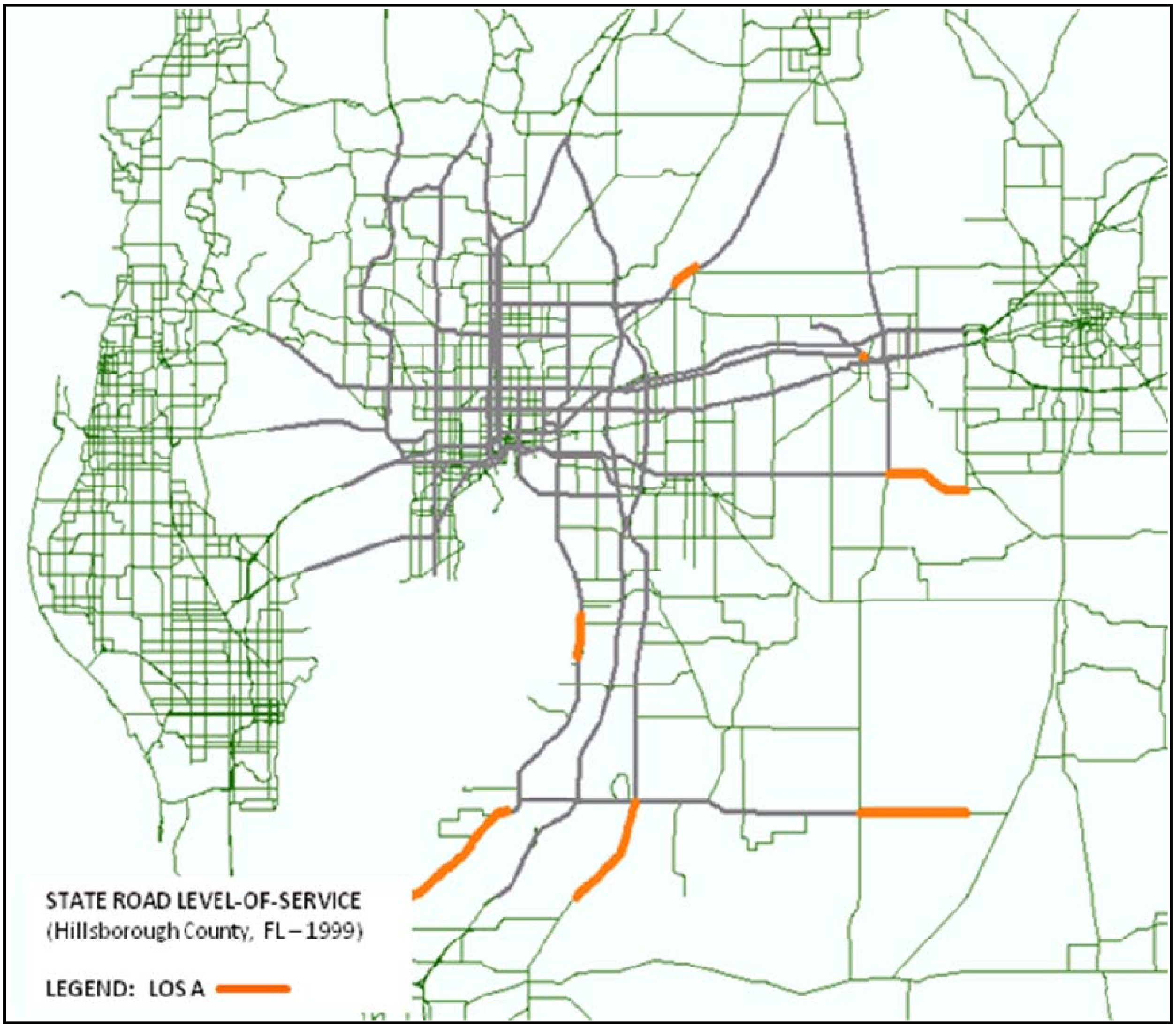


Figure 3-15: Distribution of LOS A (Free Flow) for the Study of Multilane High-Speed Arterial Corridors (Based on 1999 Data from Hillsborough County)

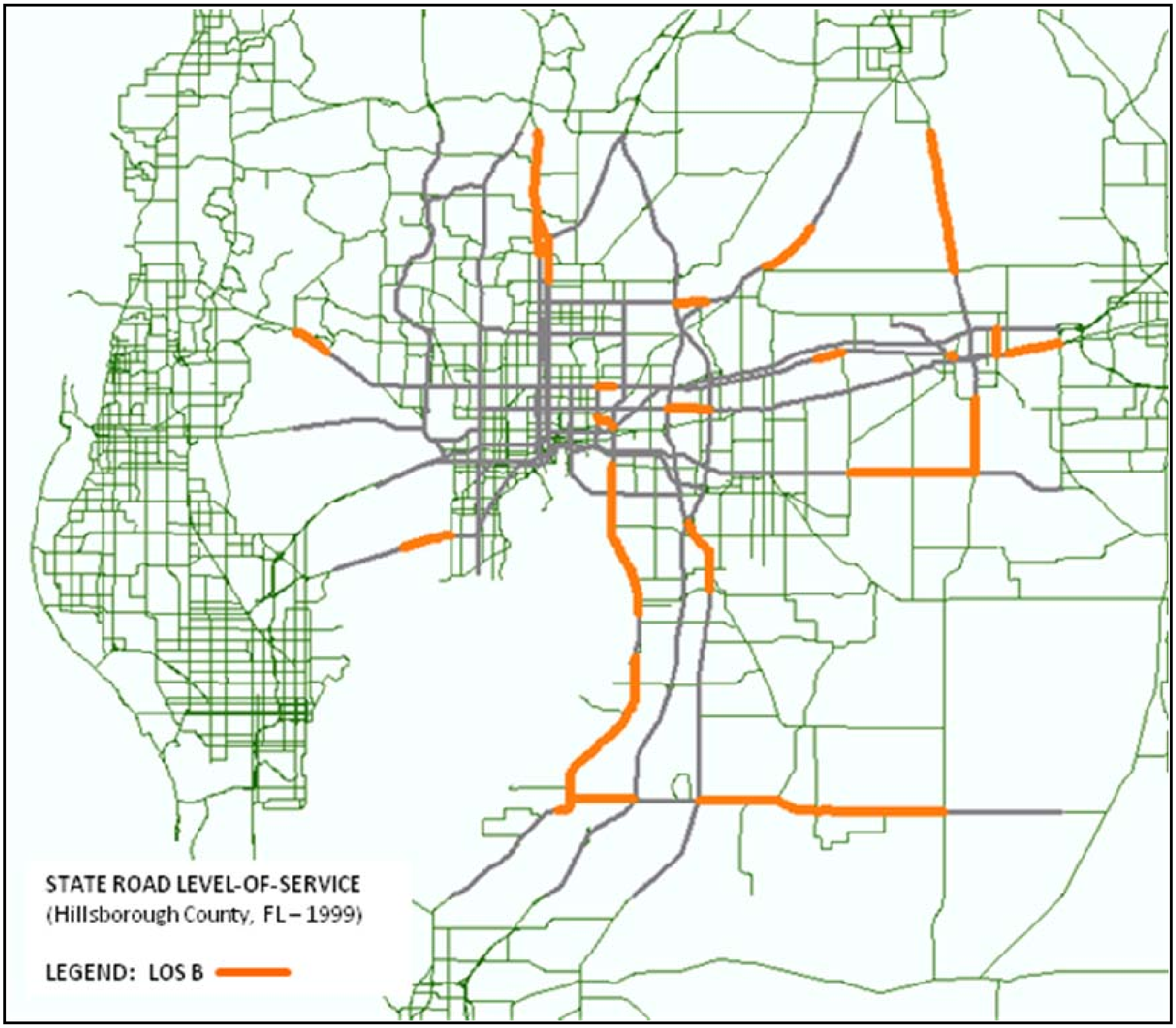


Figure 3-16: Distribution of LOS B (Reasonably Free Flow) for the Study of Multilane High-Speed Arterial Corridors (Based on 1999 Data from Hillsborough County)

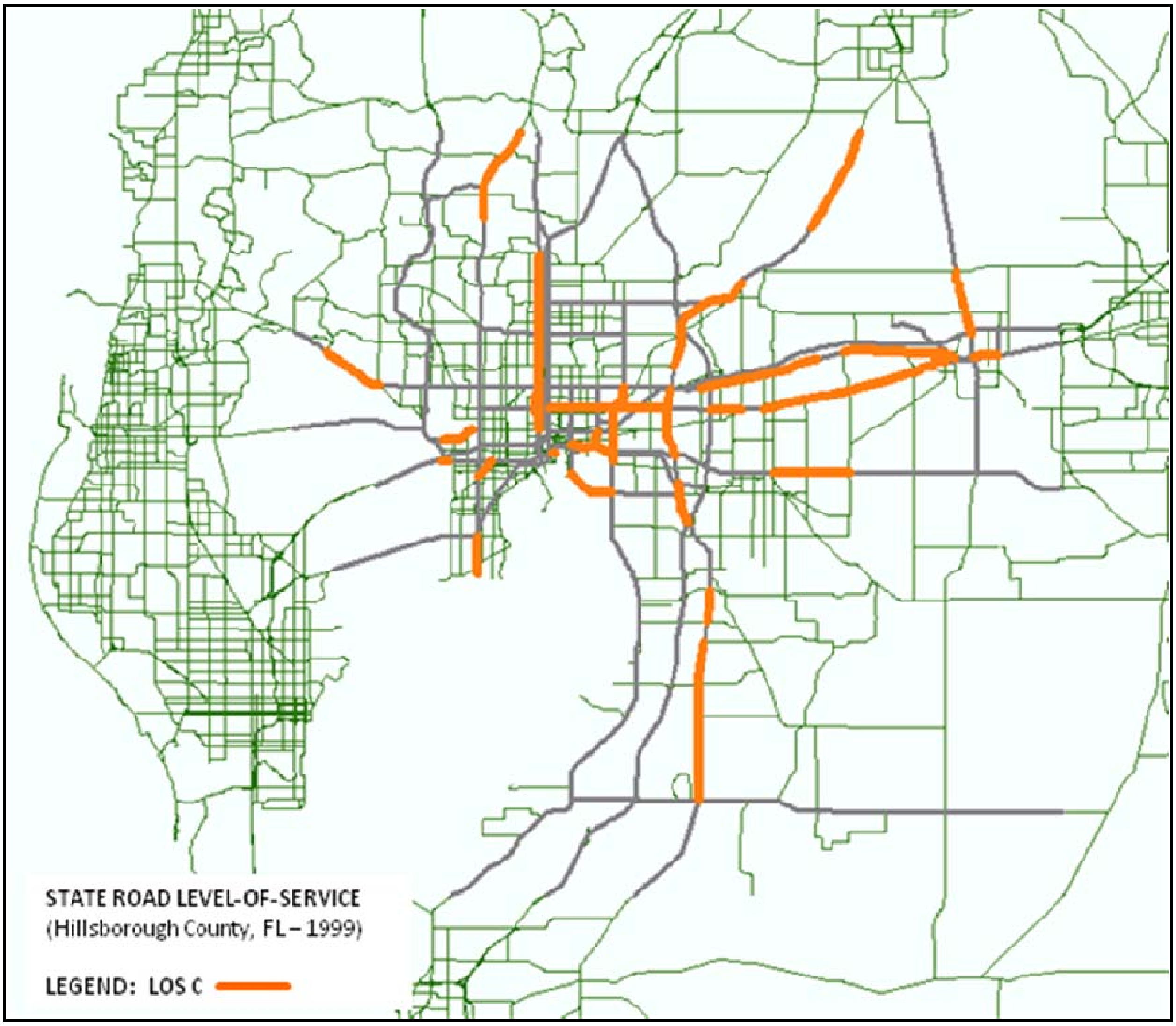


Figure 3-17: Distribution of LOS C (Stable Flow) for the Study of Multilane High-Speed Arterial Corridors (Based on 1999 Data from Hillsborough County)

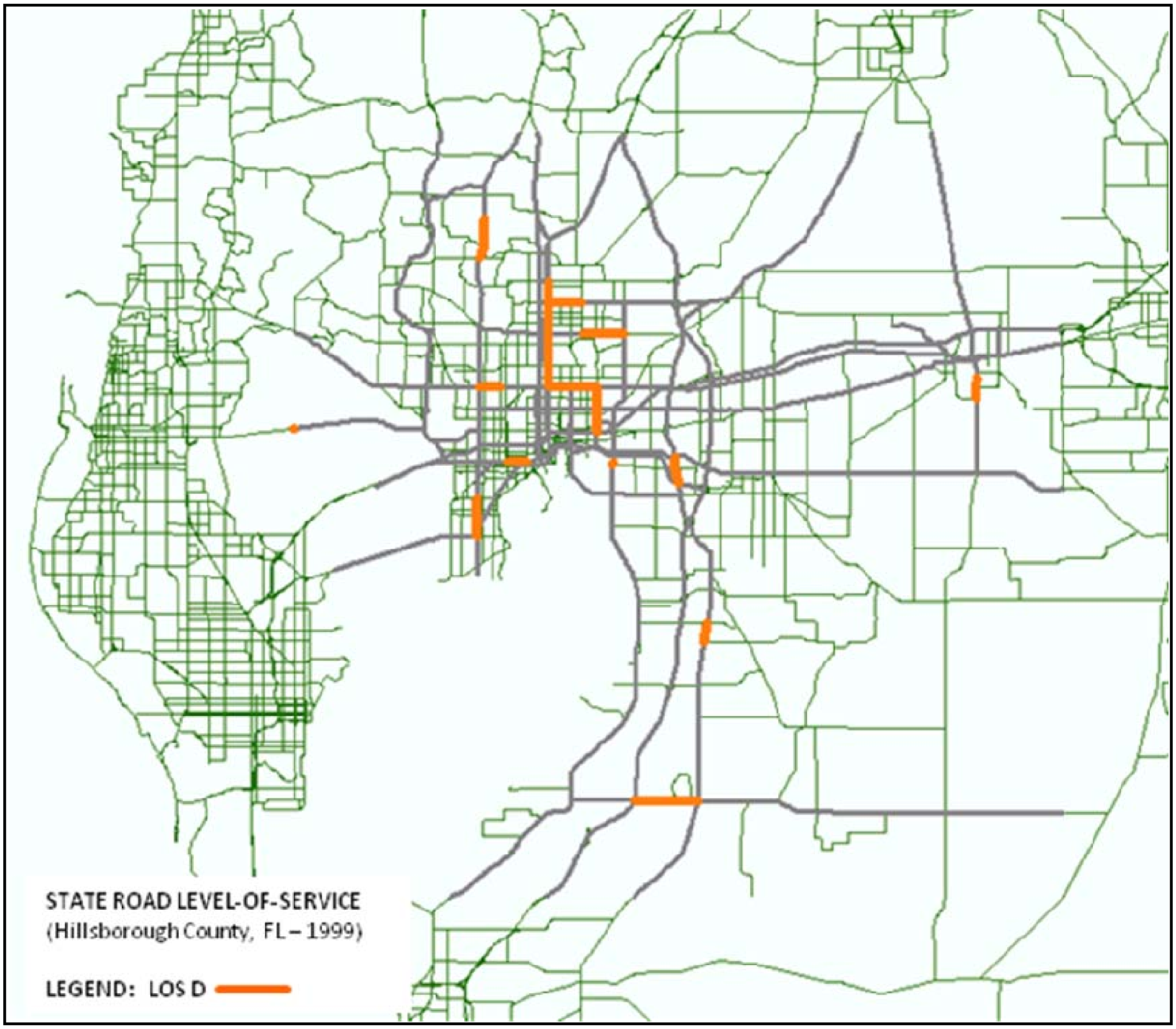


Figure 3-18: Distribution of LOS D (Approaching Unstable Flow) for the Study of Multilane High-Speed Arterial Corridors (Based on 1999 Data from Hillsborough County)

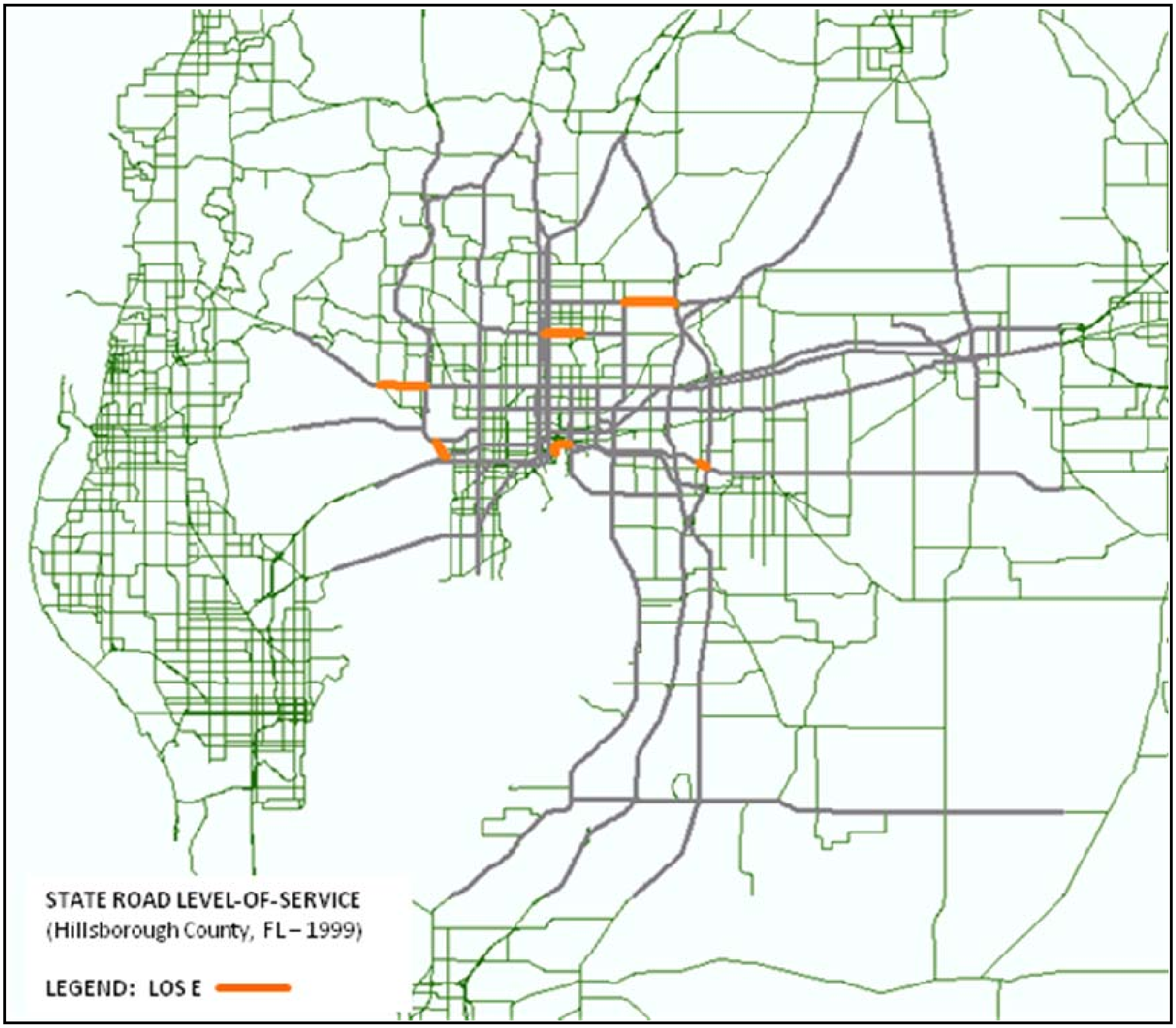


Figure 3-19: Distribution of LOS E (Unstable Flow) for the Study of Multilane High-Speed Arterial Corridors (Based on 1999 Data from Hillsborough County)

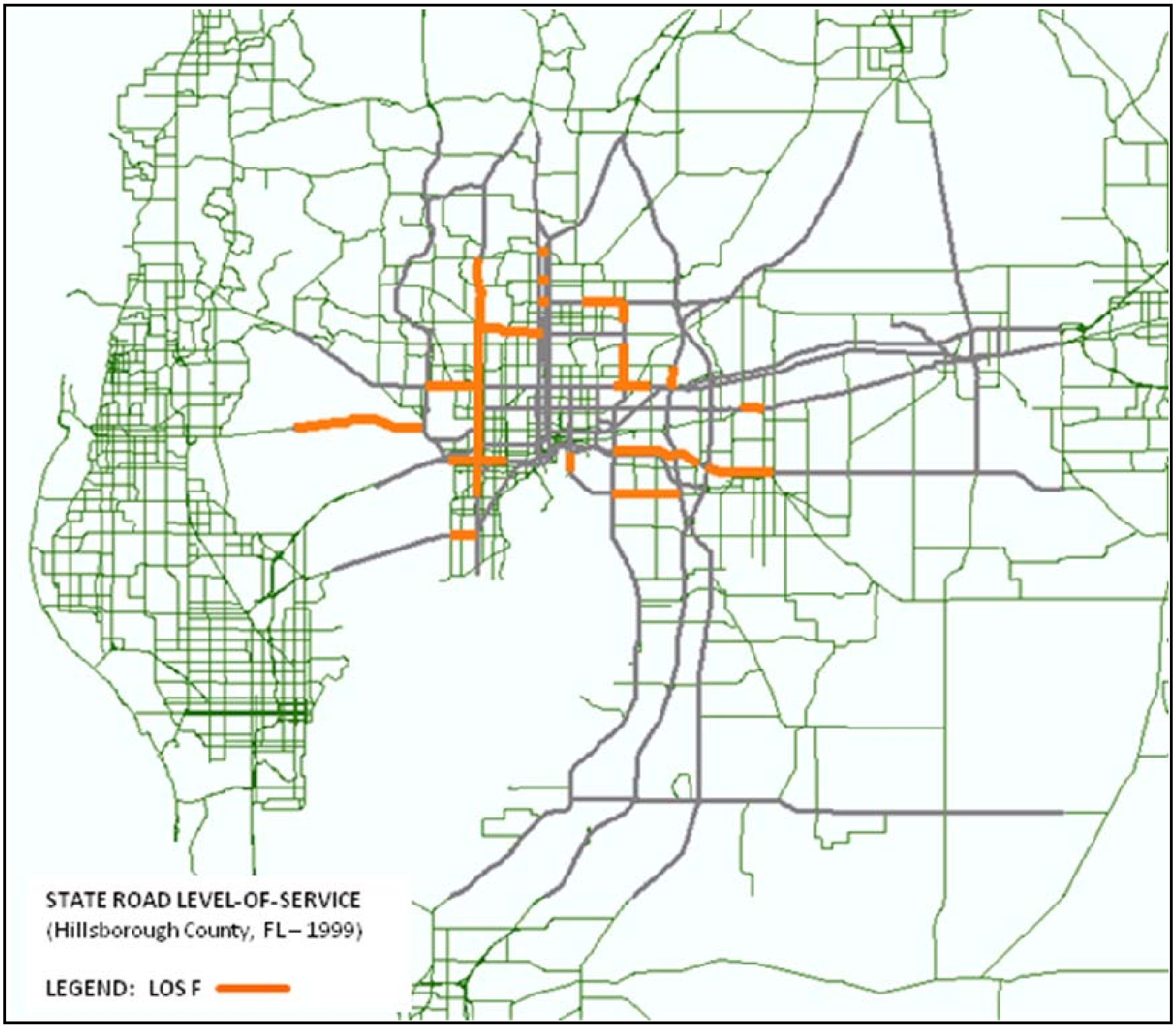


Figure 3-20: Distribution of LOS F (Forced or Breakdown Flow) for the Study of Multilane High-Speed Arterial Corridors (Based on 1999 Data from Hillsborough County)

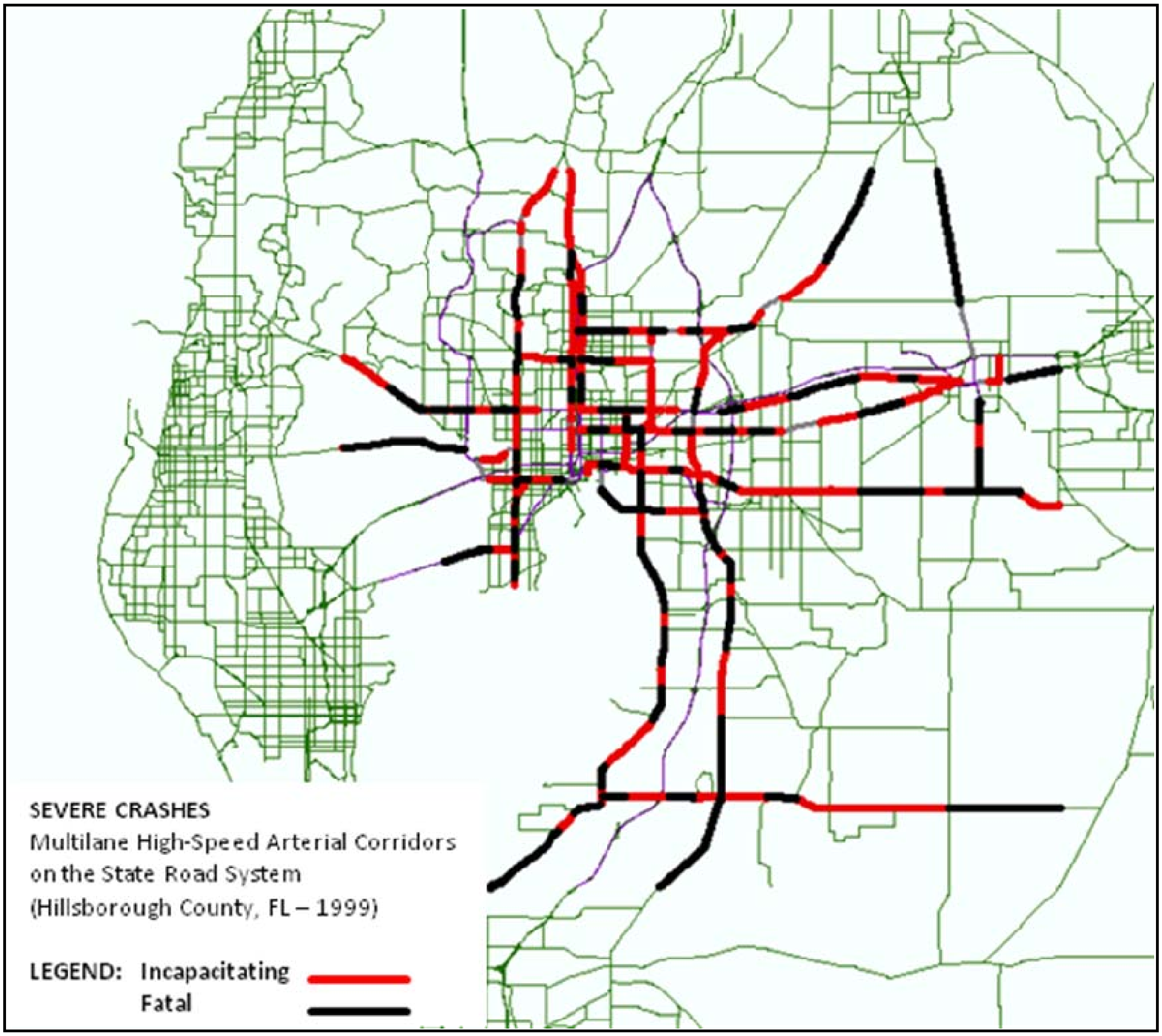


Figure 3-21: Distribution of Severe Crashes for the Study of Multilane High-Speed Arterial Corridors (Based on 1999 Data from Hillsborough County)

3.3.3.2 Crash Data

Similarly, crash records corresponding to the study area were obtained from FDOT's CAR database (refer to Section 3.2.2.2). These data corresponded to the years 2000 and 2001, a

2-year span compatible with the year(s) and/or time frame of the other types of data (i.e. segments' and LOS, road features', etc.) to be used for this study.

The resulting crash data set reported a total crash frequency of 14,339 events; these crashes were linked to the 399 road sections being studied and are applicable to all days of the week as well as the 24 hours in the day. Details corresponding to other related information characterizing these crashes were all included in the dataset.

3.3.3.3 Road Features Data

For this second study, this type of data coming from FDOT's RCI database (refer to Section 3.2.2.3) was very important since it provided the minor details on each road segment's characteristics. In this case, data corresponding to the year 2002 were downloaded since it was the closest year to 1999 with data available in the RCI database. Data for number of lanes, land use, traffic volume, speed limit, as well as for other roadway features were included within the downloaded set of data.

3.3.4 Final Data Assembly

Having completed the required data collection and preparation procedures, the last step within this study's methodology was to assemble the respective data accordingly, so that the respective analyses could be performed afterwards.

First, all final datasets had to be stored in the form of MS Excel spreadsheets, sorted by the respective road segment ID. Provided data corresponding to the variables roadway ID and mile points were already included in the spreadsheets, these were then imported into SAS in order to merge their contents by the two aforementioned variables. The final result was an organized and robust dataset consisting of 67 columns (67 variables) and 399 rows per road section in the sample.

Similar to the study described in Section 3.2, new variables had to be created with the aid of SAS since these had to be considered for this study's modeling approach. After making the respective calculations through programming in SAS, new variables were obtained: vehicle miles traveled (VMT), total section length and others. All new variables were calculated by using some of the road features and traffic volume data available. In addition, the respective frequencies were computed for total, severe, and the crash types being considered for this research (see Section 3.2.2.2).

CHAPTER 4: PRELIMINARY ANALYSIS

4.1 Overview

In order to show the general features and trends from the data just introduced in Chapter 3, a series of charts and plots were prepared for the two studies comprising this investigation. With these, the reader is provided with a general view of the parameters on which this research has been based.

In addition to the aforementioned descriptive statistics, and with the purpose to look beyond this explicit content, this chapter also contains the respective preliminary analyses. Consequently, the following sections denote what was obtained for the two studies being considered.

4.2 LOS-Safety Study for Signalized Intersections

4.2.1 Descriptive Statistics and General Trends

4.2.1.1 Aggregate Analysis (Considering the 5 periods of the Day)

Figure 4-1 shows the frequencies of each LOS (A, B, C, D, E and F) for each period of the day considered in the study.

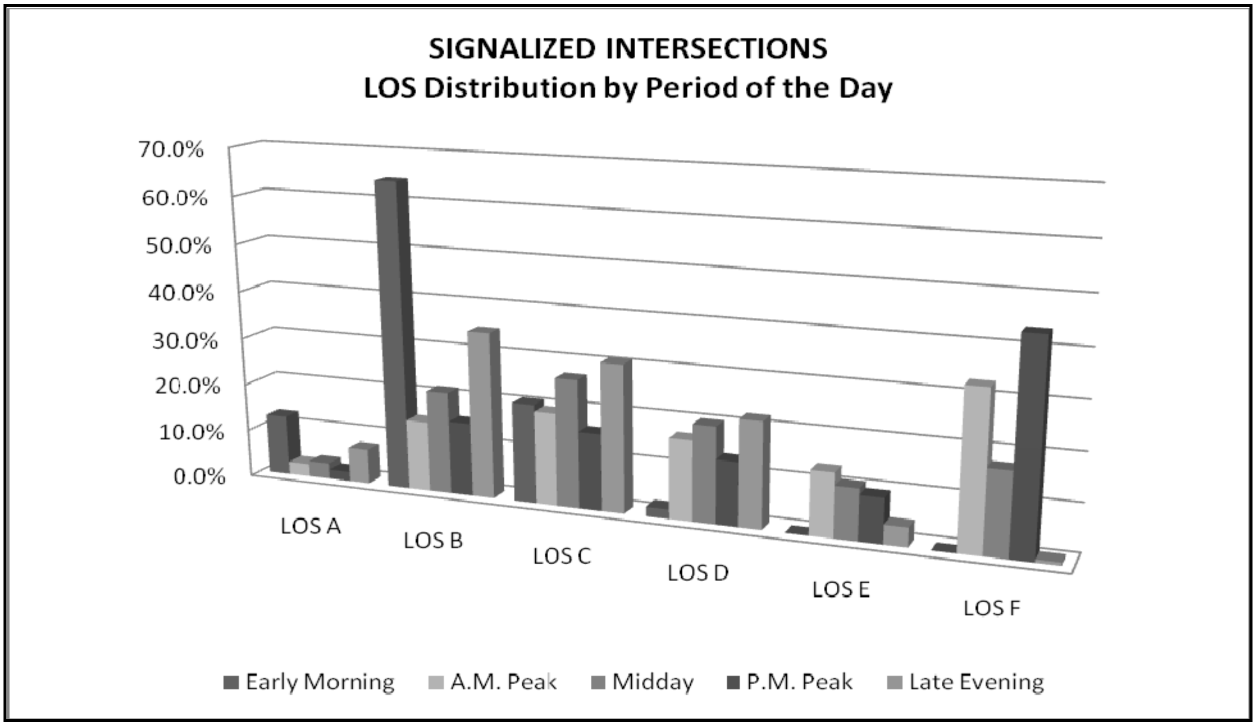


Figure 4-1: Study Sample's LOS Distribution for Signalized Intersections, by Period of the Day

As it can be seen, each LOS has a different frequency per period of the day. Table 4-1 below accompanies the histogram in Figure 4-1 with the corresponding proportions.

Table 4-1: Study Sample's LOS Distribution for Signalized Intersections, by Period of the Day

SIGNALIZED INTERSECTIONS					
LOS	Early Morning (1 - 3 A.M.) Proportion	A.M. Peak (7 - 9 A.M.) Proportion	Midday (12 - 2 P.M.) Proportion	P.M. Peak (4 - 6 P.M.) Proportion	Late Evening (8 - 10 P.M.) Proportion
A	12.8%	2.7%	3.4%	2.0%	7.4%
B	64.4%	14.8%	21.5%	15.4%	34.9%
C	20.8%	19.5%	26.8%	16.1%	30.9%
D	2.0%	16.8%	20.1%	13.4%	22.1%
E	0.0%	13.4%	10.7%	9.4%	4.0%
F	0.0%	32.9%	17.4%	43.6%	0.7%
TOTAL	**100%**	**100%**	**100%**	**100%**	**100%**

4.2.1.1.1 Interpretation and Analysis

With regards to LOS A (insignificant delay), it can be seen that this level is a predominant characteristic of the Early Morning period (1-3 A.M.) of any regular weekday; it also characterizes most Late Evening periods (8-10 P.M.) but not as much as Early Mornings. On the other hand, peak hours of the day are not usually characterized by a LOS A; note that a weekday's P.M. Peak hours (4-6 P.M.) are far from having such minimal delays.

In the case of LOS B (minimal delay), it can be seen that this level is a more predominant characteristic of a regular weekday's Early Morning period (1-3 A.M.) when compared to LOS A; it also characterizes most Late Evening periods (8-10 P.M.) but not as much as Early Mornings. In addition, peak hours of the day are not usually characterized by a LOS B; note that a weekday's peak hours (7-9 A.M. and 4-6 P.M.) are far from having such reasonably insignificant and/or minimal delays on the road.

LOS C (acceptable delay) appears to be a predominant characteristic of a regular weekday's Late Evening (8-10 P.M.) and Midday (12-2 P.M.) periods; to a second degree, it characterizes most Early Morning (1-3 A.M.) and A.M. Peak (7-9 A.M.) periods. In addition, peak hours of the day are not usually characterized by a LOS C; note that a weekday's P.M. Peak hours (4-6 P.M.), for example, are not usually depicted with records of LOS C for signalized intersections.

Going further down the scale, LOS D (tolerable delay) appears to be a predominant characteristic of a regular weekday's Late Evening (8-10 P.M.) and Midday (12-2 P.M.) periods; to a second degree, it characterizes most A.M. Peak periods (7-9 A.M.), followed by P.M. Peak periods (4-6 P.M.). On the other hand, the earliest hours of the day (1-3 A.M.) are not usually characterized by a LOS D.

The fifth level in the scale, LOS E (significant delay), is the only level showing a defined descending/sequential trend in both terms of frequency and time of day; it is also one of the two levels in the scale not characterizing Early Mornings (1-3 A.M.) at all. LOS E appears to be a predominant characteristic for a regular weekday's A.M. Peak period (7-9 A.M.), then for the Midday (12-2 P.M.), P.M. Peak (4-6 P.M.) and Late Evening (8-10 P.M.) periods.

Lastly, LOS F (excessive delay) seems to be a very predominant characteristic of weekdays' P.M. Peak periods (4-6 P.M.). To a lesser degree, it also characterizes A.M. Peak (7-9 A.M.), Midday (12-2 P.M.) and Late Evening (8-10 P.M.) periods. In addition, LOS F does not characterize Early Mornings (1-3 A.M.) at all, same as LOS E.

4.2.1.2 Disaggregate Analysis (Considering one Period of the Day at a time)

4.2.1.2.1 Early Morning Period

As shown in Figure 4-2, the LOS distribution that was obtained for the Early Morning Period is composed by only 4 LOS categories: LOS A, LOS B, LOS C and LOS D. The time frame considered for this period is from 1 A.M. to 3 A.M. This is the only period of the day not having LOS E and LOS F within this distribution; this could be attributed to the time frame for this period, time at which roads have their lowest level of congestion.

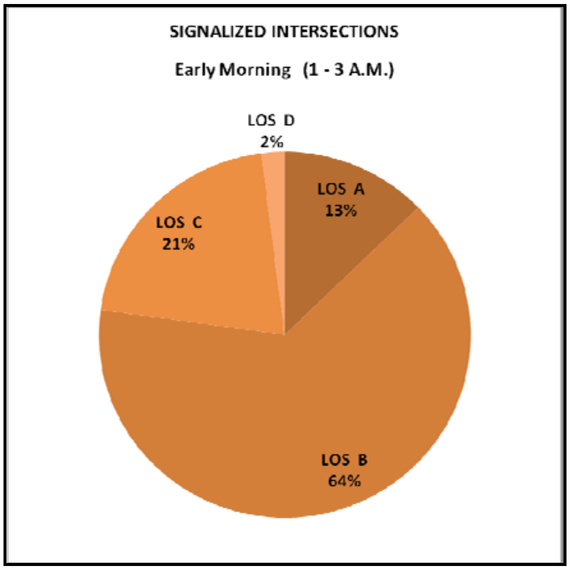


Figure 4-2: Study Sample's LOS Distribution for Signalized Intersections (Early Morning Period)

4.2.1.2.1.1 Interpretation and Analysis

Matching the data in Figure 4-2, and listed in descending order, the following LOS proportions were obtained for the Early Morning Period:

Table 4-2: Study Sample's LOS Distribution for the Early Morning Period (for Signalized Intersections)

SIGNALIZED INTERSECTIONS					
Early Morning					
LOS	Delay (seconds)			Count (Intersections)	Proportion
	HCM 2000 Standards	Observed			
	Range	Min.	Max.		
A	≤ 10.0	3.8	9.9	19	13%
B	10.1 - 20.0	10.1	19.9	96	64%
C	20.1 - 35.0	20.1	34.8	31	21%
D	35.1 - 55.0	35.3	42.5	3	2%
E	55.1 - 80.0	N/A	N/A	0	0%
F	≥ 80.0	N/A	N/A	0	0%
				TOTAL **149**	**100%**

From the content displayed in Table 4-2, it can be noticed that the most frequent level is LOS B (minimal delay), corresponding to traffic delays within range of 10.1-20.0 seconds. These statistics also seem to match the author's expectations since it is commonly known that during the earliest hours of the day (e.g. from midnight to sunrise) vehicular volume is at its lowest, which is related to ideal or free flow conditions and/or the lowest level of traffic congestion.

The second and third most frequent levels recorded were LOS C (acceptable delay), followed by LOS A (insignificant delay). As it can be seen, LOS A did not surpass LOS C in terms of frequency for the Early Morning period; this could be attributed to the fact that there is a lesser probability of having ideal traffic conditions (i.e. delay $\leq$ 10.0 seconds) at such busy transportation network like Florida's, so to have more of LOS C, and LOS B, for these roads was more realistic.

Finally, LOS D (tolerable delay) was observed to be the least frequent level. Only 3 of the 149 intersections in the sample reported a LOS D for the Early Morning period; these outliers were:

- From Orange County
  - OC 93 = Colonial Dr. (SR 50) with Alafaya Tr. (SR 434)
  - OC 97 = Colonial Dr. (SR 50) with Semoran Blvd. (SR 436)
- From Hillsborough County
  - HC 1006 = Hillsborough Ave. (US 92) with Parsons Ave.

From the outliers above, whereas OC 93 and OC 97 are large in terms of number of lanes and traffic volume, HC 1006's major and minor approaches have 2 lanes at the most; these factors can be attributed to the LOS D (i.e. delay 35.1-55.0 seconds) during the Early Morning period recorded for those intersections.

Thus, it can be concluded that the study sample was well-representative of signalized intersections in Florida; furthermore, these realistic observations denote that the methodology considered was close to accurate and appropriate for the study.

4.2.1.2.2 A.M. Peak Period

As shown in Figure 4-3, the LOS distribution that was obtained for the A.M. Peak period is composed by all LOS categories: LOS A, LOS B, LOS C, LOS D, LOS E and LOS F. The time frame considered for this period is from 7 A.M. to 9 A.M.

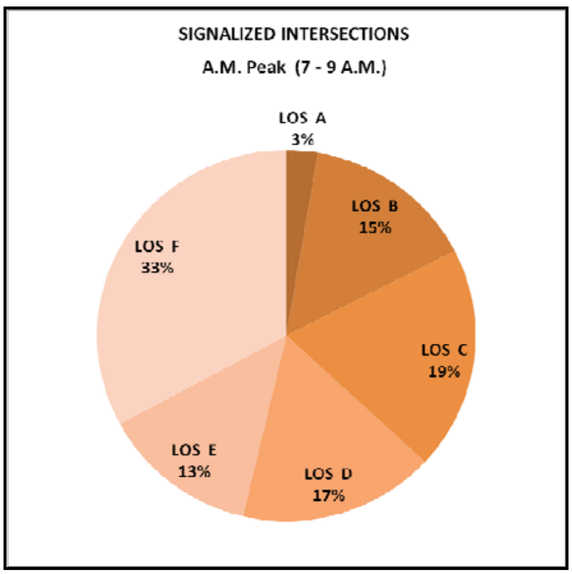


Figure 4-3: Study Sample's LOS Distribution for Signalized Intersections (A.M. Peak Period)

4.2.1.2.2.1 <u>Interpretation and Analysis</u>

Matching the data in Figure 4-3, and listed in descending order, the following LOS proportions were obtained for the A.M. Peak period:

Table 4-3: Study Sample's LOS Distribution for Signalized Intersections (A.M. Peak Period)

SIGNALIZED INTERSECTIONS					
A.M. Peak					
LOS	Delay (seconds)			Count (Intersections)	Proportion (%)
	HCM 2000 Standards	Observed			
	Range	Min.	Max.		
A	< 10.0	4.9	9.1	4	3%
B	10.1 - 20.0	10.6	19.4	22	15%
C	20.1 - 35.0	20.4	34.7	29	19%
D	35.1 - 55.0	35.7	53.3	25	17%
E	55.1 - 80.0	55.3	79.7	20	13%
F	> 80.0	80.1	360.1	49	33%
				TOTAL **149**	**100%**

From the content displayed in Table 4-3, it can be noticed that the most frequent level is LOS F (excessive delay), corresponding to traffic delays > 80.0 seconds. These statistics also seem to meet expectations since it is known that vehicular volume is very high during a regular weekday's A.M. Peak hour (e.g. excessive traffic congestion before getting to school and/or work during the morning rush hour, etc.)

Next in the scale, having frequencies within the 13-19% range, follow LOS C (acceptable delay), LOS D (tolerable delay), LOS B (minimal delay) and LOS E (significant delay).

Finally, LOS A (insignificant delay) was the least frequent level. Only 4 of the 149 intersections in the sample reported a LOS A for this A.M. Peak period; these outliers were:

- From Hillsborough County
  - HC 1027 = Brandon Blvd. (SR 60) with Brandon Crossings Entrance
  - HC 1103 = Mission Hills Dr. with 56th St.
  - HC 1152 = 131st St. with Nebraska Ave.

- HC 1397 = Temple Heights Rd. with 56th St.

The intersections above are medium in size, and seem to be connected to some access points, located in residential/suburban areas; these factors can be attributed to the LOS A (i.e. delay $\leq$ 10.0 seconds) during the A.M. Peak period. This low frequency was also expected since a traffic delay $\leq$ 10.0 seconds is not that typical during the morning rush hour. For this reason, LOS A was considered to be an outlier level for the A.M. Peak period overall.

From these observations, it can be concluded once again that the study sample was well-representative of signalized intersections in Florida.

4.2.1.2.3 Midday Period

As shown in Figure 4-4, the LOS distribution that was obtained for the Midday Period is composed by all LOS categories: LOS A, LOS B, LOS C, LOS D, LOS E and LOS F. The time frame considered for this period is from 12 P.M. to 2 P.M.

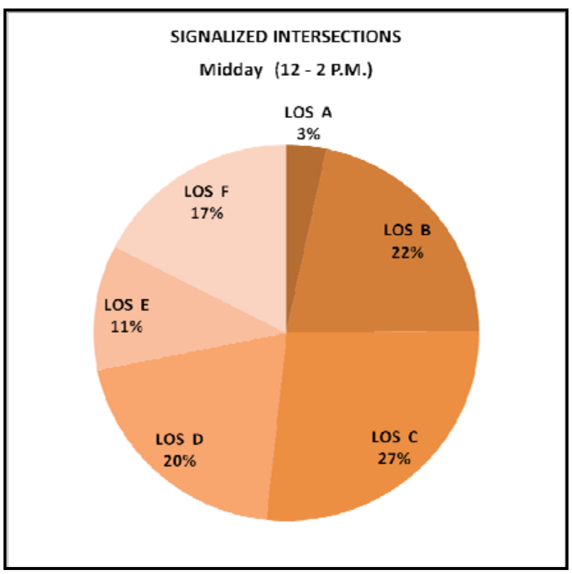


Figure 4-4: Study Sample's LOS Distribution for Signalized Intersections (Midday Period)

4.2.1.2.3.1 Interpretation and Analysis

Matching the data in Figure 4-4, and listed in descending order, the following LOS proportions were obtained for the Midday period:

Table 4-4: Study Sample's LOS Distribution for Signalized Intersections (Midday Period)

SIGNALIZED INTERSECTIONS					
Midday					
LOS	Delay (seconds)			Count (Intersections)	Proportion
	HCM 2000 Standards	Observed			
	Range	Min.	Max.		
A	≤ 10.0	6.5	9.2	5	3%
B	10.1 - 20.0	10.8	19.3	32	22%
C	20.1 - 35.0	20.2	34.9	40	27%
D	35.1 - 55.0	35.1	53.8	30	20%
E	55.1 - 80.0	56.9	78.3	16	11%
F	≥ 80.0	84.6	247.9	26	17%
				TOTAL **149**	**100%**

From the content displayed in Table 4-4, it can be noticed that the most frequent level is LOS C (acceptable delay), corresponding to traffic delays within the range of 20.1-35.0 seconds. These statistics also seem to match the author's expectations since vehicular volume tends to be stable (i.e. not too high neither too low) during Midday hours, without producing an extreme level of congestion.

Next in the scale, having frequencies within the 11-22% range, follow LOS B (minimal delay), LOS D (tolerable delay), LOS F (excessive delay) and LOS E (significant delay). Like the A.M. Peak period (refer to Table 4-3), once again, LOS E is the 2^{nd} least frequent level.

Finally, LOS A (insignificant delay) was the least frequent level. Only 5 of the 149 intersections in the sample reported a LOS A for this Midday period; these outliers were:

- From Hillsborough County
  - HC 1103 = Mission Hills Dr. with 56th St.
  - HC 1106 = Puritan Rd. with 56th St.
  - HC 1121 = Fowler Ave. with Riverhills Dr.
  - HC 1365 = Hillsborough Ave. (US 92) with McIntosh Rd.
  - HC 1397 = Temple Heights Rd. with 56th St.

The intersections above are medium in size, and seem to be connected to some access points, located in residential/suburban areas; these factors can be attributed to the LOS A (i.e. delay $\leq$ 10.0 seconds) during the Midday period. This low frequency was also expected since a traffic delay $\leq$ 10.0 seconds is not that typical during the 12-2 P.M. time frame (e.g. lunch hour, etc.) For these reasons, LOS A was considered to be an outlier level for the Midday period overall.

4.2.1.2.4 P.M. Peak Period

As shown in Figure 4-5, the LOS distribution that was obtained for the P.M. Peak period is composed by all LOS categories: LOS A, LOS B, LOS C, LOS D, LOS E and LOS F. The time frame considered for this period is from 4 P.M. to 6 P.M.

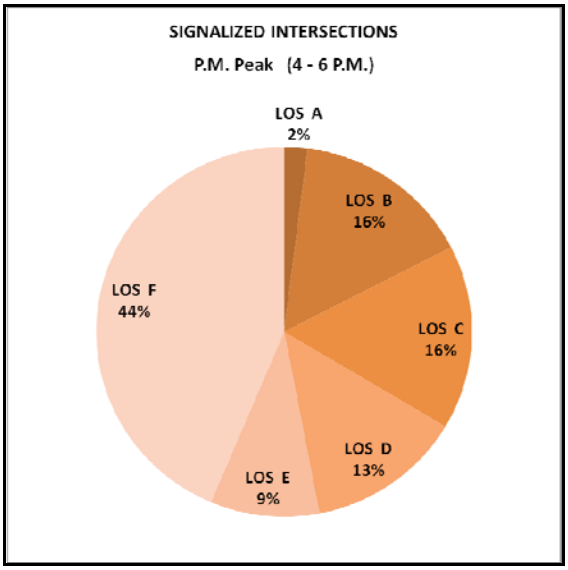


Figure 4-5: Study Sample's LOS Distribution for Signalized Intersections (P.M. Peak Period)

4.2.1.2.4.1 <u>Interpretation and Analysis</u>

Matching the data in Figure 4-5, and listed in descending order, the following LOS proportions were obtained for the P.M. Peak period:

Table 4-5: Study Sample's LOS Distribution for Signalized Intersections (P.M. Peak Period)

SIGNALIZED INTERSECTIONS					
P.M. Peak					
LOS	Delay (seconds)			Count (Intersections)	Proportion
	HCM 2000 Standards	Observed			
	Range	Min.	Max.		
A	≤ 10.0	9.0	9.5	3	2%
B	10.1 - 20.0	10.4	20.0	23	16%
C	20.1 - 35.0	20.2	34.8	24	16%
D	35.1 - 55.0	35.7	54.8	20	13%
E	55.1 - 80.0	57.7	79.0	14	9%
F	≥ 80.0	80.2	447.0	65	44%
				TOTAL **149**	**100%**

From the content displayed in Table 4-5, it can be noticed that the most frequent level is LOS F (excessive delay), corresponding to traffic delays $\geq$ 80.0 seconds. These statistics also seem to meet expectations since it is known that vehicular volume is very high during a regular weekday's afternoon peak hour (e.g. excessive traffic congestion when returning home after school and/or work, etc.)

Next in the scale, having frequencies within the 9-16% range, follow LOS B (minimal delay), LOS C (acceptable delay), LOS D (tolerable delay), and LOS E (significant delay). As for the A.M. Peak and Midday periods (refer to Tables 4-3 and 4-4), LOS E is again the 2nd least frequent level.

Finally, LOS A (insignificant delay) was the least frequent level. Only 3 of the 149 intersections in the sample reported a LOS A for this afternoon peak period; these outliers were:

- From Orange County
    - OC 206 = Colonial Dr. (SR 50) with CR-13
- From Hillsborough County
    - HC 1103 = Mission Hills Dr. with 56th St.
    - HC 1152 = Fowler Ave. with Riverhills Dr.

The intersections just listed are medium in size, and seem to be connected to some access points, located in residential/suburban areas; these factors can be attributed to the LOS A (i.e. delay $\leq$ 10.0 seconds) during the P.M. Peak period. This low frequency was also expected since a traffic delay $\leq$ 10.0 seconds is not that typical during the afternoon rush hour. For this reason, LOS A was considered to be an outlier level for the P.M. Peak period overall.

4.2.1.2.5 Late Evening Period

As shown in Figure 4-6, the LOS distribution that was obtained for the Late Evening period is composed by all LOS categories: LOS A, LOS B, LOS C, LOS D, LOS E and LOS F. The time frame considered for this period is from 8 P.M. to 10 P.M.

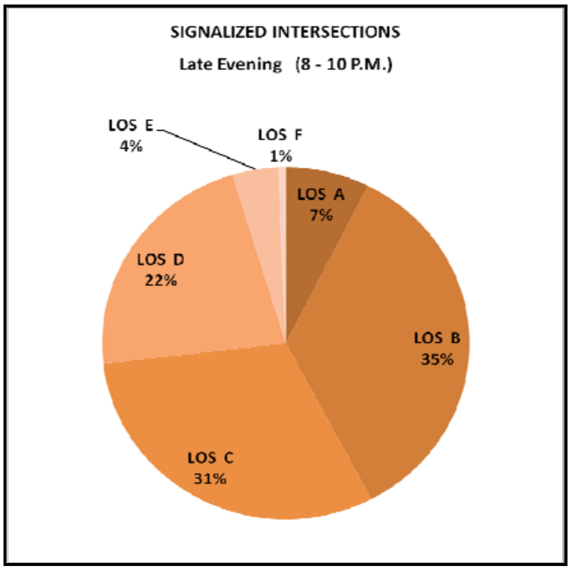


Figure 4-6: Study Sample's LOS Distribution for Signalized Intersections (Late Evening Period)

4.2.1.2.5.1 Interpretation and Analysis

Matching the data in Figure 4-6, and listed in descending order, the following LOS proportions were obtained for the Late Evening period:

Table 4-6: Study Sample's LOS Distribution for Signalized Intersections (Late Evening Period)

SIGNALIZED INTERSECTIONS					
Late Evening					
LOS	Delay (seconds)			Count (Intersections)	Proportion
	HCM 2000 Standards	Observed			
	Range	Min.	Max.		
A	< 10.0	4.1	9.9	11	7%
B	10.1 - 20.0	10.5	20	52	35%
C	20.1 - 35.0	20.3	34.8	46	31%
D	35.1 - 55.0	35.3	54.4	33	22%
E	55.1 - 80.0	55.8	70.5	6	4%
F	> 80.0	113.8		1	1%
				TOTAL **149**	**100%**

From the content displayed in Table 4-6, it can be noticed that the most frequent level is LOS B (minimal delay), corresponding to traffic delays within range of 10.1-20.0 seconds; recalling the Early Morning period, this is the second instance when LOS B has the highest frequency within the sample (refer to Table 4-2). These statistics also seem to match the author's expectations since it is commonly known that vehicular volume starts to decrease (i.e. traffic congestion dissipates) during the latest hours of a regular weekday.

Next in the scale is LOS C (acceptable delay), with a frequency almost as close as LOS B's; then are LOS D, LOS A and LOS E. As in the A.M. Peak, Midday and P.M. Peak periods (refer to Tables 4-3, 4-4 and 4-5), LOS E is the 2nd least frequent level.

Finally, LOS F (excessive delay) was the least frequent level. Only 1 of the 149 intersections in the sample reported a LOS F for the Late Evening period; this outlier was:

- From Orange County
    - OC 149 = Colonial Dr. (SR 50) with Chuluota Rd. (SR 419).

The intersection above is medium in size, and seems to be connected to some access points, located in residential/suburban areas; these factors can be attributed to the LOS F (i.e. delay ≥ 80.0 seconds) during the Late Evening period. This low frequency was also expected since very large delays, like the ones characterizing a rush hour, are not typical of the Late Evening hours of the day. For this reason, LOS F was considered to be an outlier for the Late Evening period overall.

4.2.1.3 Crash Frequency Distributions, by LOS

As mentioned earlier, the data for this study corresponded to the years 2000-2005. The following tables and figures denote the respective study sample's crash distributions by LOS.

Table 4-7: Study Sample's Crash Frequency Statistics for Signalized Intersections, by LOS (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS							
Study Sample's Crash Frequency Statistics (By LOS)							
Response Variable	LOS	No. of Intersections in Study Sample with the Indicated LOS	Crash Frequency				
			No. of Crashes	Mean	Min.	Max.	Std. Dev.
TOTAL CRASHES	A	42	63	1.5	0	8	1.9
	B	225	690	3.1	0	17	3.3
	C	170	1,073	6.3	0	29	4.7
	D	111	1,079	9.7	1	33	5.7
	E	56	696	12.4	1	36	7
	F	141	1,931	13.7	1	40	7.8
		TOTAL **745**	***TOTAL*** **5,532**				

Response Variable	LOS	No. of Intersections in Study Sample with the Indicated LOS	Crash Frequency				
			No. of Crashes	Mean	Min.	Max.	Std. Dev.
SEVERE CRASHES	**A**	42	12	0.3	0	2	0.6
	B	225	81	0.4	0	5	0.7
	C	170	106	0.6	0	4	0.9
	D	111	88	0.8	0	4	1
	E	56	50	0.9	0	4	1.1
	F	141	122	0.9	0	5	1.1
		TOTAL **745**	***TOTAL*** **459**				
REAR-END CRASHES	**A**	42	23	0.5	0	4	0.9
	B	225	254	1.1	0	8	1.5
	C	170	485	2.9	0	18	2.8
	D	111	512	4.6	0	17	3.5
	E	56	384	6.9	0	21	4.6
	F	141	1,075	7.6	0	29	5.6
		TOTAL **745**	***TOTAL*** **2,733**				
SIDESWIPE CRASHES	**A**	42	2	0	0	1	0.2
	B	225	19	0.1	0	2	0.3
	C	170	55	0.3	0	4	0.6
	D	111	65	0.6	0	3	0.8
	E	56	34	0.6	0	3	0.9
	F	141	153	1.1	0	5	1.3
		TOTAL **745**	***TOTAL*** **328**				
HEAD-ON CRASHES	**A**	42	1	0	0	1	0.2
	B	225	16	0.1	0	1	0.3
	C	170	22	0.1	0	3	0.4
	D	111	29	0.3	0	3	0.6
	E	56	6	0.1	0	1	0.3
	F	141	38	0.3	0	3	0.5
		TOTAL **745**	***TOTAL*** **112**				
ANGLE + LEFT-TURN CRASHES	**A**	42	23	0.5	0	4	1
	B	225	301	1.3	0	10	1.9
	C	170	382	2.2	0	13	2.1
	D	111	331	3	0	13	2.6
	E	56	189	3.4	0	12	2.6
	F	141	479	3.4	0	13	2.6
		TOTAL **745**	***TOTAL*** **1,705**				

Table 4-8: Study Sample's Crash Frequency Distribution for Signalized Intersections, by LOS (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS						
LOS	TOTAL CRASHES Frequency	SEVERE CRASHES Frequency	REAR-END CRASHES Frequency	SIDESWIPE CRASHES Frequency	HEAD-ON CRASHES Frequency	ANGLE + LEFT-TURN CRASHES Frequency
A	63	12	23	2	1	23
B	690	81	254	19	16	301
C	1,073	106	485	55	22	382
D	1,079	88	512	65	29	331
E	696	50	384	34	6	189
F	1,931	122	1,075	153	38	479
TOTAL	**5,532**	**459**	**2,733**	**328**	**112**	**1,705**
PROPORTIONS	**100.0%**	**8.3%**	**49.4%**	**5.9%**	**2.0%**	**30.8%**
	100.0%	**96.4%**				

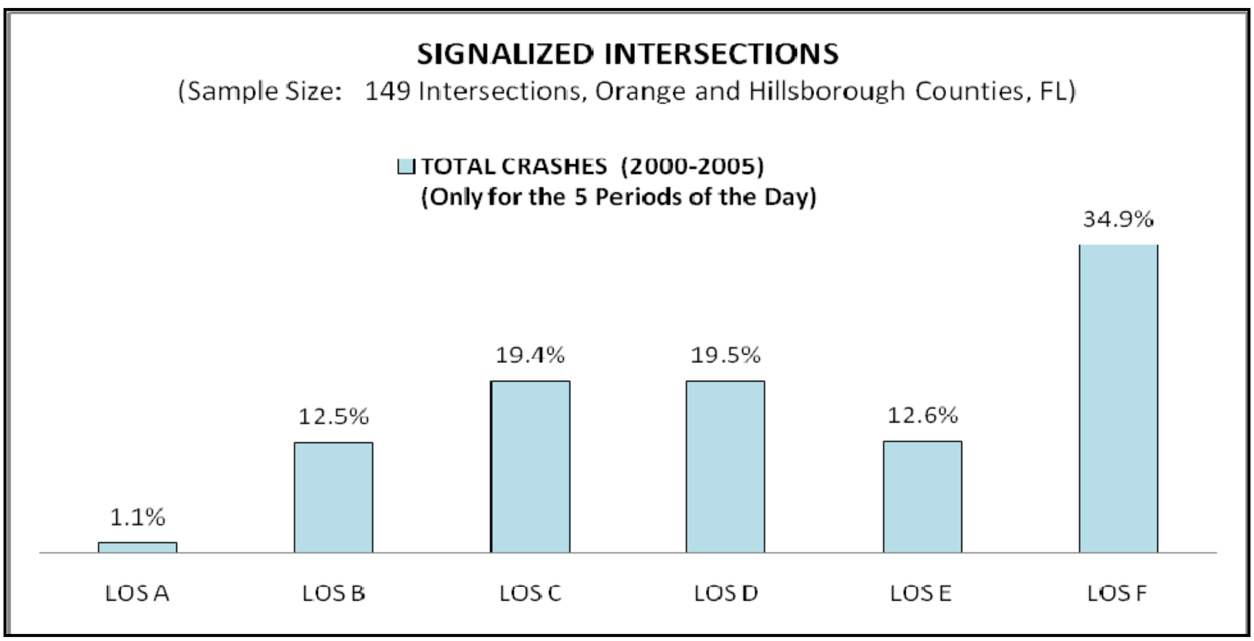


Figure 4-7: Study Sample's Distribution for "Total Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

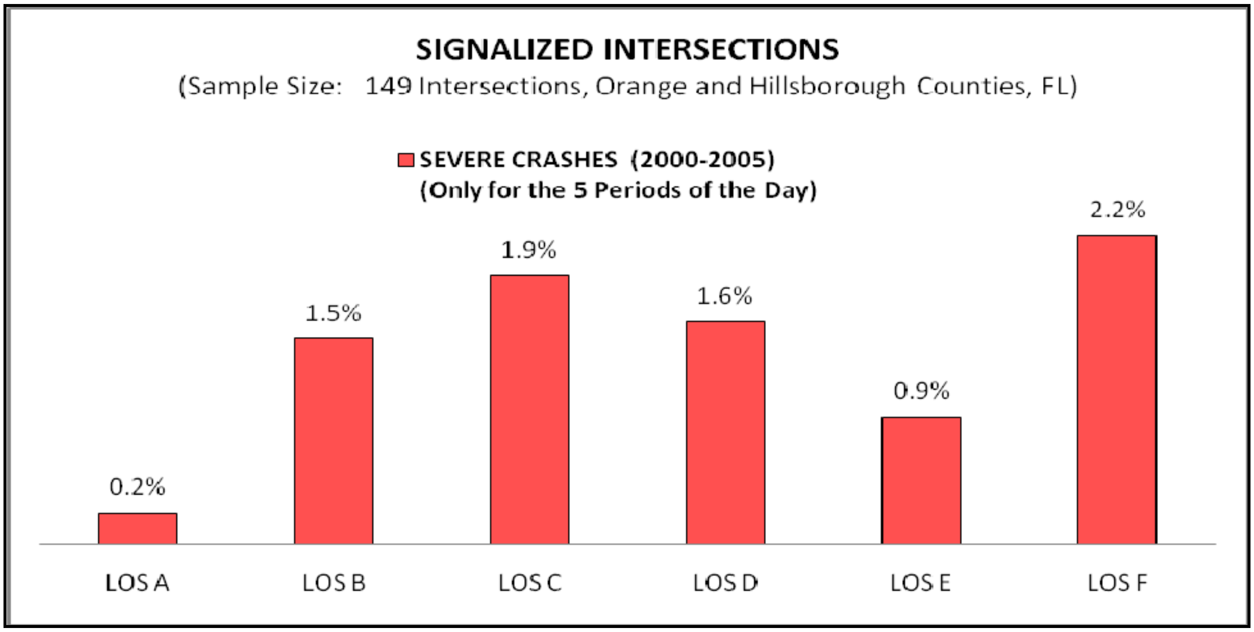


Figure 4-8: Study Sample's Distribution for "Severe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

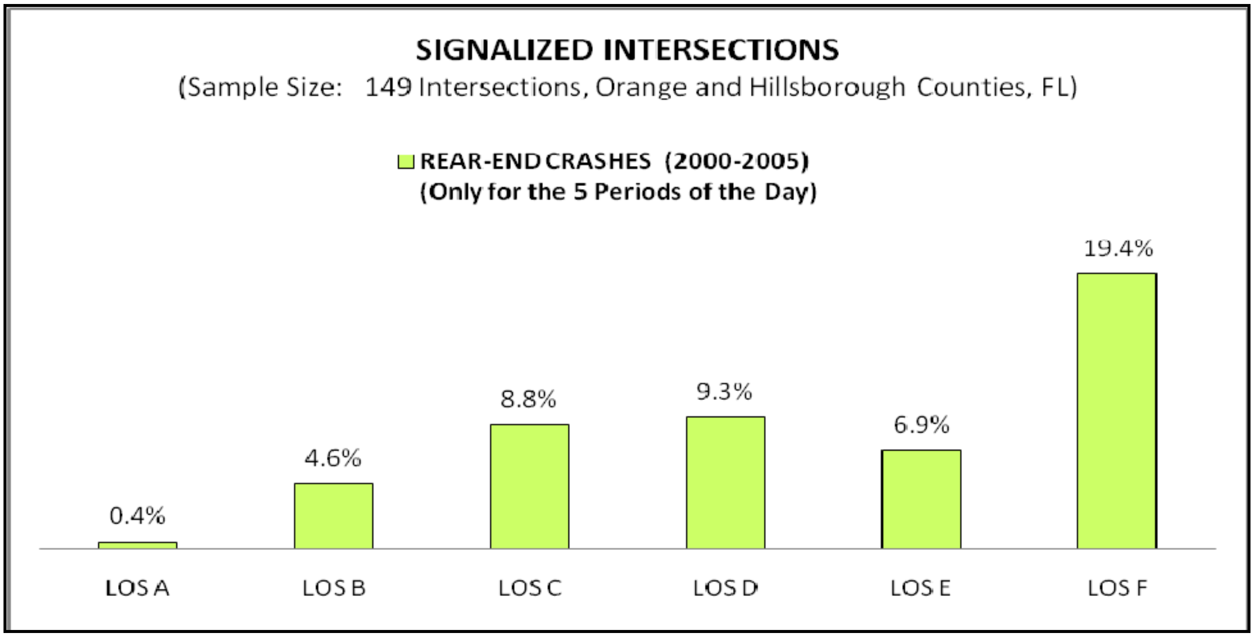


Figure 4-9: Study Sample's Distribution for "Rear-End Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

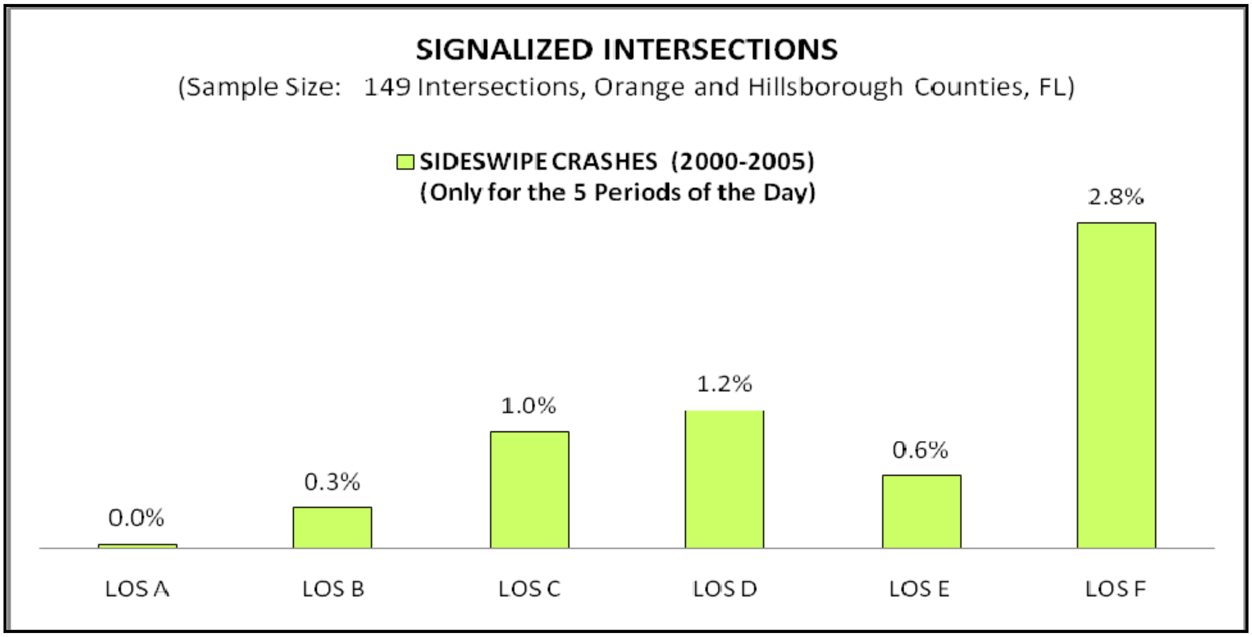


Figure 4-10: Study Sample's Distribution for "Sideswipe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

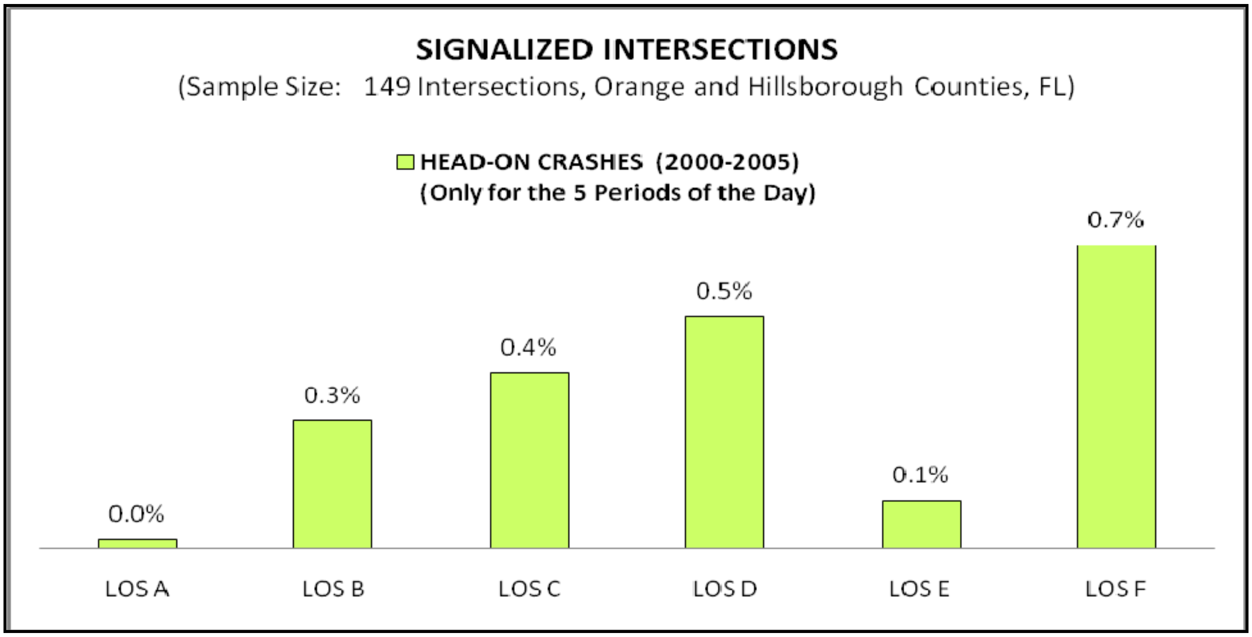


Figure 4-11: Study Sample's Distribution for "Head-On Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

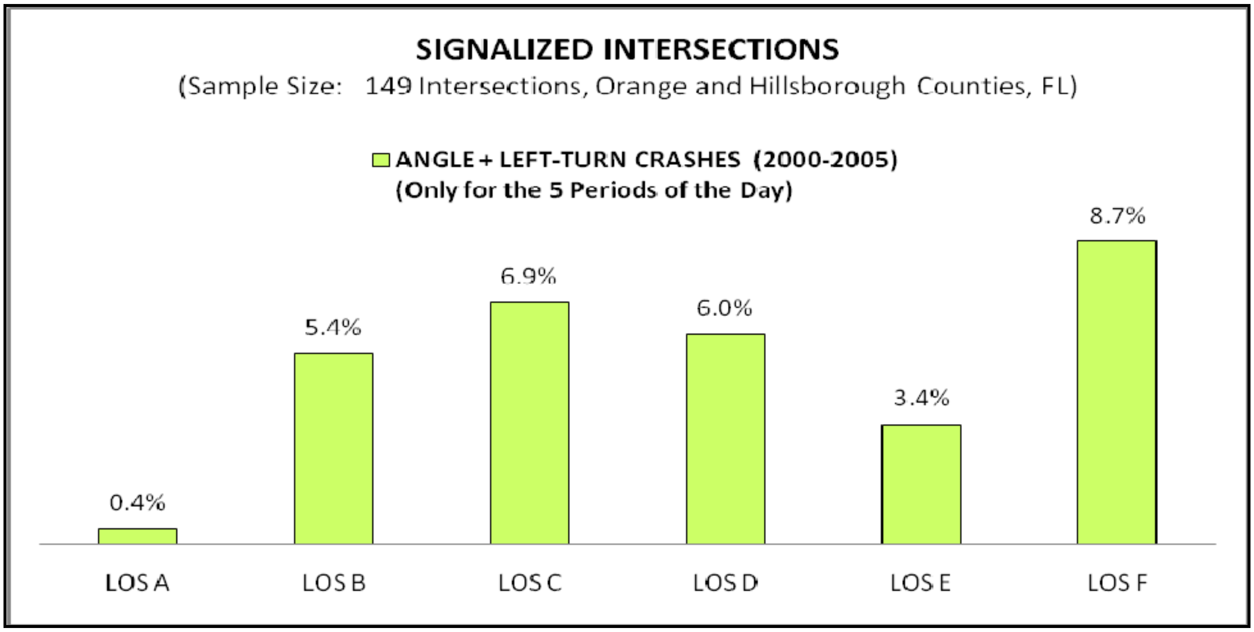


Figure 4-12: Study Sample's Distribution for "Angle and Left-Turn Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

4.2.1.4 Crash Frequency Distributions, by Period of the Day

General crash trends were also obtained by using the crash frequencies in the study's original dataset (see Figure 4-13). It can be seen that traffic volume is indeed linearly related to the average number of crashes per hour.

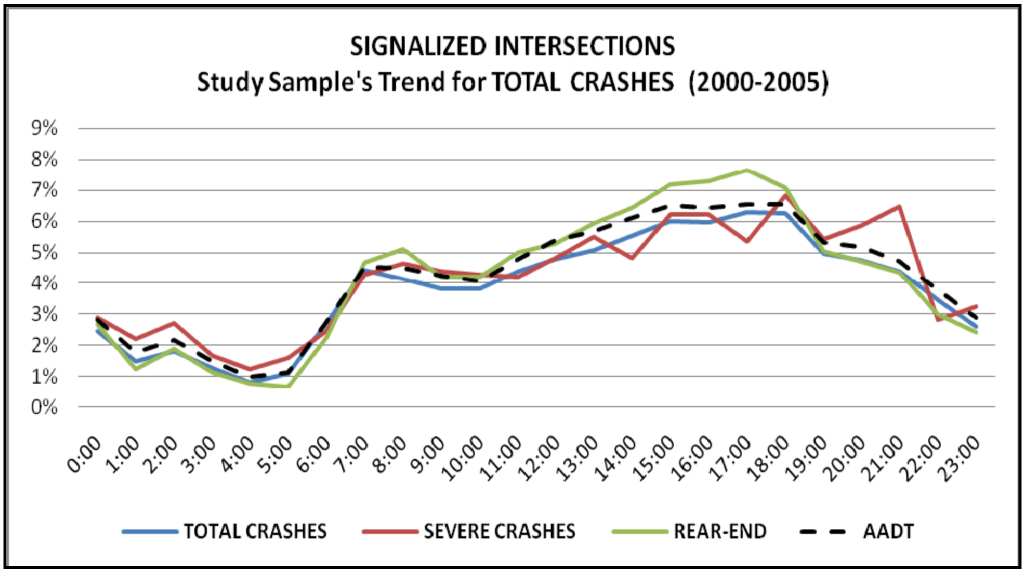


Figure 4-13: Sample Hourly Crash Distribution (2000-2005) for Total, Severe and Rear-End Crashes at Signalized Intersections, along with the AADT

Following is a similar graph, but based on the study sample's data (see Figure 4-14). It denotes the hours considered in the analysis, which belong to the 5 periods of the day considered.

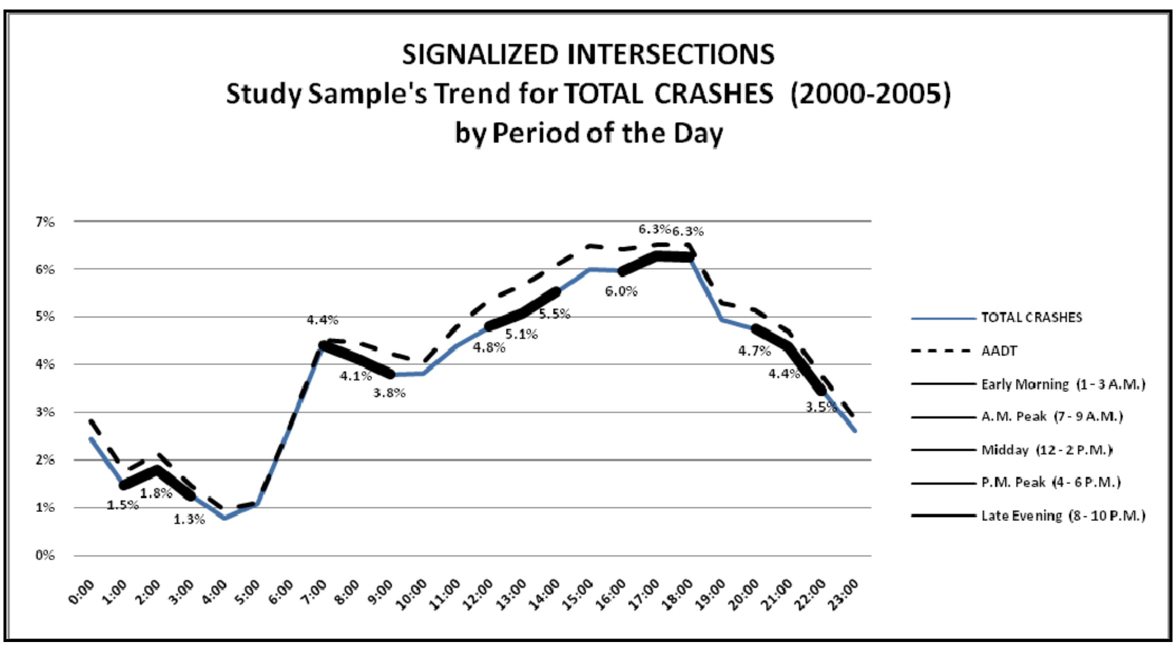


Figure 4-14: Study Sample's Hourly Crash Distributions (2000-2005) for Signalized Intersections (Considering the 5 Periods of the Day)

The data just shown denote the P.M. Peak period (4-6 P.M.) as the one having the highest average number of crashes per hour. Following are the Midday (12-2 P.M.), Late Evening (8-10 P.M.) and A.M. Peak (7-9 A.M.) periods, which show the 2nd, 3rd and 4th highest average number of crashes per hour; as expected, the Early Morning period (1-3 A.M.) has the lowest one (i.e. since traffic volume is very low at the earliest hours of the day, not many crashes can be recorded).

Having denoted these general crash trends over time, the following tables and figures denote the respective study sample's crash distributions by period of the day.

Table 4-9: Study Sample's Crash Frequency Statistics for Signalized Intersections, by Period of the Day

SIGNALIZED INTERSECTIONS								
Study Sample's Crash Frequency Statistics (By Period of the Day)								
Response Variable	Period of the Day	No. of Intersections in Study Sample for the Indicated Period of the Day		Crash Frequency				
				No. of Crashes	Mean	Min.	Max.	Std. Dev.
TOTAL CRASHES	EM	1 - 3 A.M.	149	238	1.6	0	17	2.6
	AM Peak	7 - 9 A.M.	149	1,245	8.4	0	38	6.4
	MD	12 - 2 P.M.	149	1,336	9	0	36	6.7
	PM Peak	4 - 6 P.M.	149	1,672	11.2	1	40	7.8
	LE	8 - 10 P.M.	149	1,041	7	0	23	5.5
		TOTAL	745	*TOTAL* 5,532				
SEVERE CRASHES	EM	1 - 3 A.M.	149	26	0.2	0	2	0.4
	AM Peak	7 - 9 A.M.	149	100	0.7	0	5	0.9
	MD	12 - 2 P.M.	149	103	0.7	0	4	1
	PM Peak	4 - 6 P.M.	149	117	0.8	0	5	1
	LE	8 - 10 P.M.	149	113	0.8	0	5	1.1
		TOTAL	745	*TOTAL* 459				

Response Variable	Period of the Day	No. of Intersections in Study Sample for the Indicated Period of the Day		Crash Frequency				
				No. of Crashes	Mean	Min.	Max.	Std. Dev.
REAR-END CRASHES	**EM**	**1 - 3 A.M.**	149	102	0.7	0	9	1.5
	AM Peak	**7 - 9 A.M.**	149	644	4.3	0	27	4.3
	MD	**12 - 2 P.M.**	149	683	4.6	0	24	4.2
	PM Peak	**4 - 6 P.M.**	149	873	5.9	0	29	5.3
	LE	**8 - 10 P.M.**	149	431	2.9	0	14	3.1
		TOTAL	**745**	***TOTAL*** **2,733**				
SIDESWIPE CRASHES	**EM**	**1 - 3 A.M.**	149	8	0.1	0	1	0.2
	AM Peak	**7 - 9 A.M.**	149	90	0.6	0	5	1
	MD	**12 - 2 P.M.**	149	73	0.5	0	4	0.8
	PM Peak	**4 - 6 P.M.**	149	105	0.7	0	5	1
	LE	**8 - 10 P.M.**	149	52	0.3	0	3	0.6
		TOTAL	**745**	***TOTAL*** **328**				
HEAD-ON CRASHES	**EM**	**1 - 3 A.M.**	149	4	0	0	1	0.2
	AM Peak	**7 - 9 A.M.**	149	20	0.1	0	2	0.4
	MD	**12 - 2 P.M.**	149	26	0.2	0	2	0.4
	PM Peak	**4 - 6 P.M.**	149	36	0.2	0	3	0.6
	LE	**8 - 10 P.M.**	149	26	0.2	0	2	0.4
		TOTAL	**745**	***TOTAL*** **112**				
ANGLE + LEFT-TURN CRASHES	**EM**	**1 - 3 A.M.**	149	72	0.5	0	5	0.9
	AM Peak	**7 - 9 A.M.**	149	362	2.4	0	10	2.2
	MD	**12 - 2 P.M.**	149	394	2.6	0	10	2.3
	PM Peak	**4 - 6 P.M.**	149	507	3.4	0	13	2.8
	LE	**8 - 10 P.M.**	149	370	2.5	0	13	2.4
		TOTAL	**745**	***TOTAL*** **1,705**				

(NOTE: EM=Early Morning, MD=Midday, LE=Late Evening).

4.3 LOS-Safety Study for Multilane High-Speed Arterial Corridors

4.3.1 Preliminary Descriptive Statistics and Distributions

4.3.1.1 Overall Analysis

Figure 4-15 denotes the LOS distribution among the road sections from the study sample. It has to be noted that these LOS data, provided by Hillsborough County, is specifically for the road segments in general (i.e. it is not based on signalized intersections operational conditions).

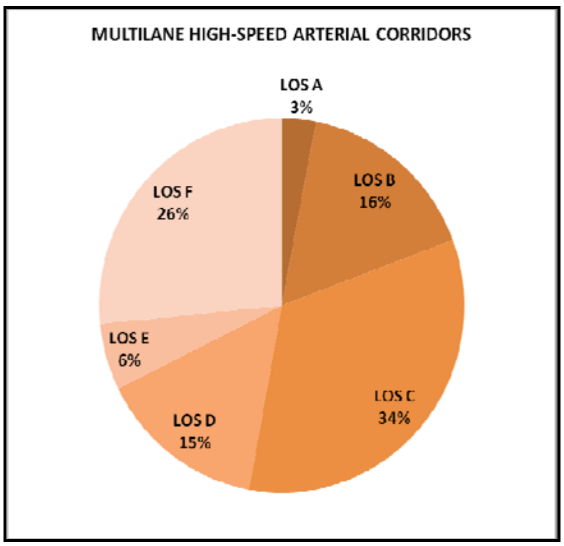


Figure 4-15: Study Sample's LOS Distribution for Multilane High-Speed Arterial Corridors

4.3.1.1.1 Interpretation and Analysis

Matching the data in Figure 4-15, and listed in descending order, the following LOS proportions were obtained:

Table 4-10: Study Sample's LOS Distribution for Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS		
LOS	**Count (Road Sections)**	**Proportion**
A	12	3%
B	65	16%
C	134	34%
D	59	15%
E	23	6%
F	106	26%
TOTAL	**399**	**100%**

As it can be observed, from the 399 road sections in the sample, 34% had LOS C (stable flow). These statistics suggest that most state roads in Hillsborough County in the stable flow range, with a high tendency towards having LOS F (forced or breakdown flow). Recall that these LOS data have been computed for the 100^{th} highest traffic hour of the year (refer to Section 3.3.3.1).

4.3.1.2 Crash Frequency Distribution, by LOS

As mentioned earlier, the data for this study corresponded to the years 2000-2001. The following tables and figures denote the respective study sample's crash distributions by LOS.

Table 4-11: Study Sample's Crash Frequency Statistics for Multilane High-Speed Arterial Corridors, by LOS

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS								
Study Sample's Crash Frequency Statistics (By LOS)								
Response Variable	LOS	No. of Road Sections in Study Sample with the Indicated LOS	Crash Frequency					
			No. of Crashes	Mean	Min.	Max.	Std. Dev.	
TOTAL CRASHES	A	12	209	17.4	3	40	12.6	
	B	65	1,897	29.2	0	99	21.1	
	C	134	4,101	30.6	1	464	43.5	
	D	59	2,611	44.3	1	133	33.2	
	E	23	987	42.9	0	174	40.4	
	F	106	4,534	42.8	1	157	31.3	
		TOTAL **399**	***TOTAL*** **14,339**					
SEVERE CRASHES	A	12	36	3	0	10	2.8	
	B	65	275	4.2	0	13	3.2	
	C	134	448	3.3	0	49	5.6	
	D	59	248	4.2	0	17	3.2	
	E	23	114	5	0	19	5.6	
	F	106	378	3.6	0	17	3.3	
		TOTAL **399**	***TOTAL*** **1,499**					
REAR-END CRASHES	A	12	50	4.2	1	14	4	
	B	65	681	10.5	0	49	10.4	
	C	134	1,561	11.6	0	216	20.2	
	D	59	1,120	19	0	77	17.3	
	E	23	504	21.9	0	55	18.4	
	F	106	2,390	22.5	0	106	18.9	
		TOTAL **399**	***TOTAL*** **6,306**					
SIDESWIPE CRASHES	A	12	20	1.7	0	7	2.2	
	B	65	125	1.9	0	8	2	
	C	134	233	1.7	0	24	2.8	
	D	59	170	2.9	0	12	3	
	E	23	74	3.2	0	16	3.6	
	F	106	339	3.2	0	15	2.9	
		TOTAL **399**	***TOTAL*** **961**					

Response Variable	LOS	Number of Road Sections in Study Sample with the Indicated LOS	Crash Frequency				
			No. of Crashes	Mean	Min.	Max.	Std. Dev.
HEAD-ON CRASHES	**A**	12	1	0.1	0	1	0.3
	B	65	24	0.4	0	2	0.7
	C	134	66	0.5	0	4	0.8
	D	59	37	0.6	0	4	0.9
	E	23	5	0.2	0	2	0.5
	F	106	57	0.5	0	4	0.8
		***TOTAL* 399**	***TOTAL* 190**				
ANGLE + LEFT-TURN CRASHES	**A**	12	65	5.4	0	16	4.6
	B	65	704	10.8	0	32	7.9
	C	134	1,629	12.2	0	167	16.7
	D	59	918	15.6	0	45	10.8
	E	23	283	12.3	0	84	17.4
	F	106	1,250	11.8	0	42	9.9
		***TOTAL* 399**	***TOTAL* 4,849**				

Table 4-12: Study Sample's Crash Frequency Distribution for Multilane High-Speed Arterial Corridors, by LOS

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS						
LOS	TOTAL CRASHES Frequency	SEVERE CRASHES Frequency	REAR-END CRASHES Frequency	SIDESWIPE CRASHES Frequency	HEAD-ON CRASHES Frequency	ANGLE + LEFT-TURN CRASHES Frequency
A	209	36	50	20	1	65
B	1,897	275	681	125	24	704
C	4,101	448	1,561	233	66	1,629
D	2,611	248	1,120	170	37	918
E	987	114	504	74	5	283
F	4,534	378	2,390	339	57	1,250
TOTAL	**14,339**	**1,499**	**6,306**	**961**	**190**	**4,849**
PROPORTIONS	**100.0%**	**10.5%**	**44.0%**	**6.7%**	**1.3%**	**33.8%**
	100.0%	**96.3%**				

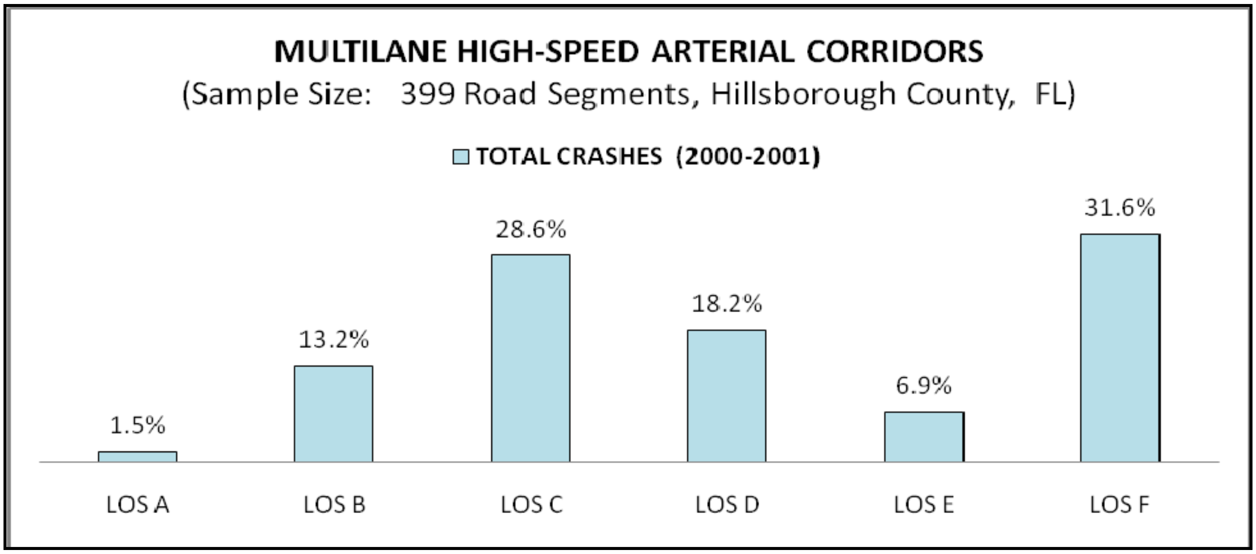


Figure 4-16: Study Sample's Distribution for "Total Crashes" at Multilane High-Speed Arterial Corridors

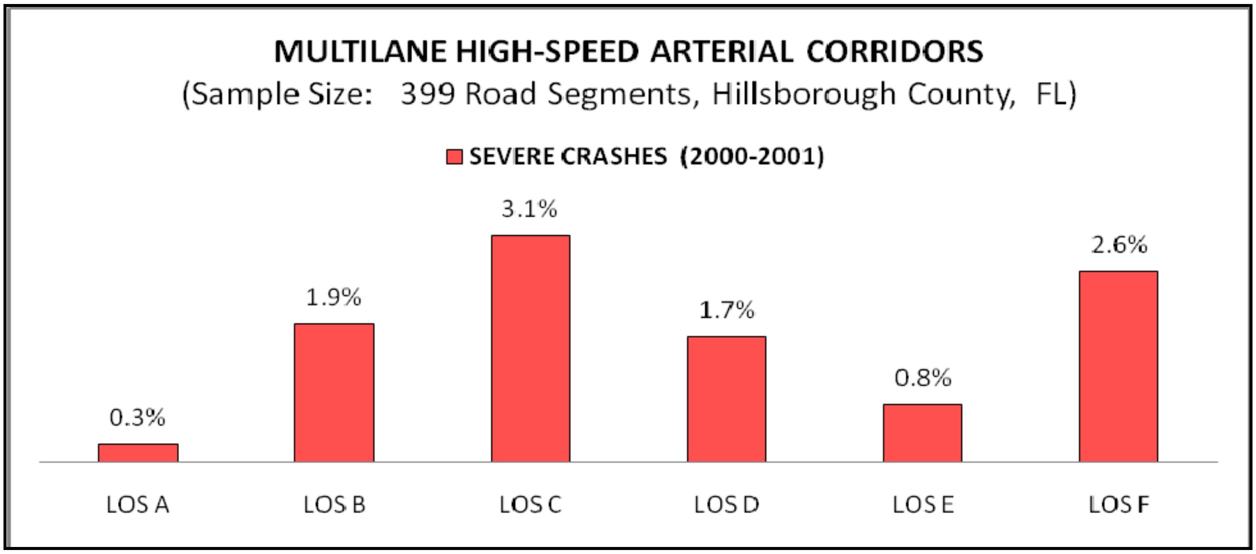


Figure 4-17: Study Sample's Distribution for "Severe Crashes" at Multilane High-Speed Arterial Corridors

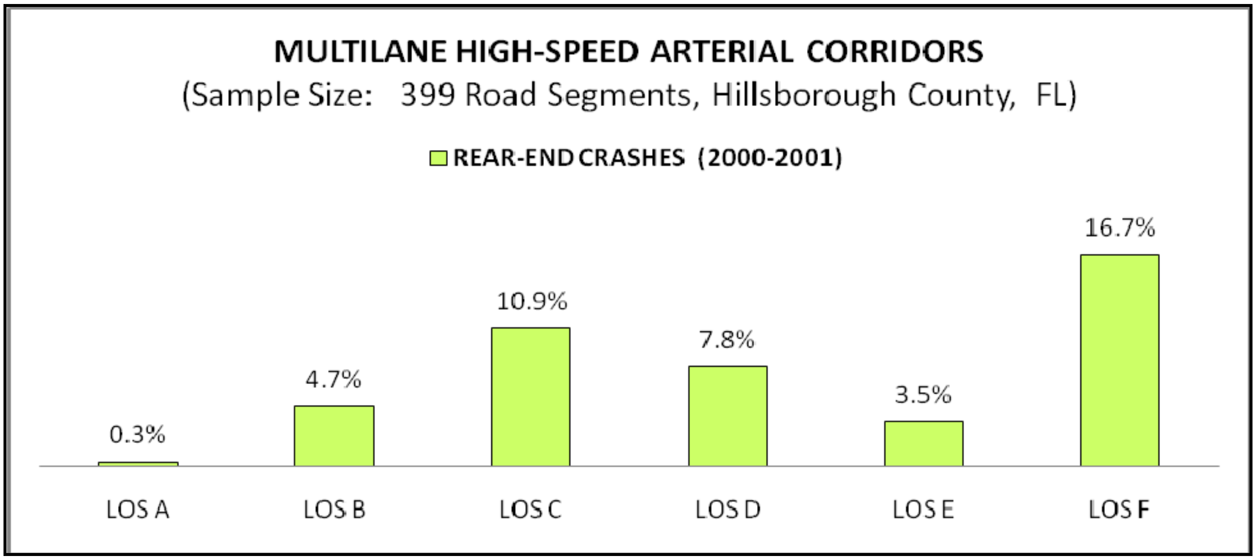


Figure 4-18: Study Sample's Distribution for "Rear-End Crashes" at Multilane High-Speed Arterial Corridors

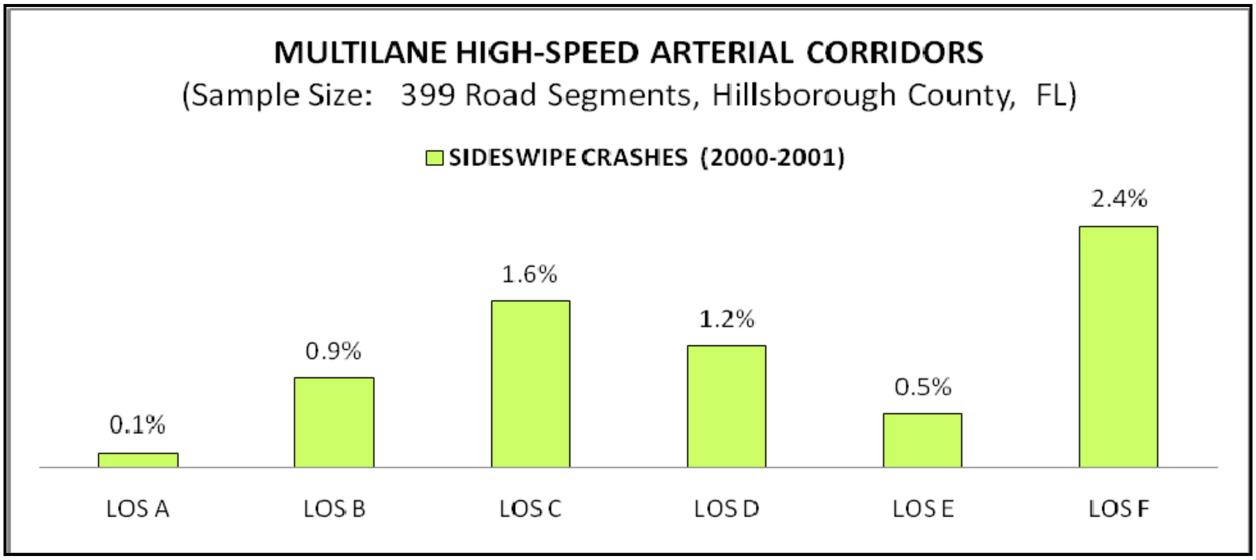


Figure 4-19: Study Sample's Distribution for "Sideswipe Crashes" at Multilane High-Speed Arterial Corridors

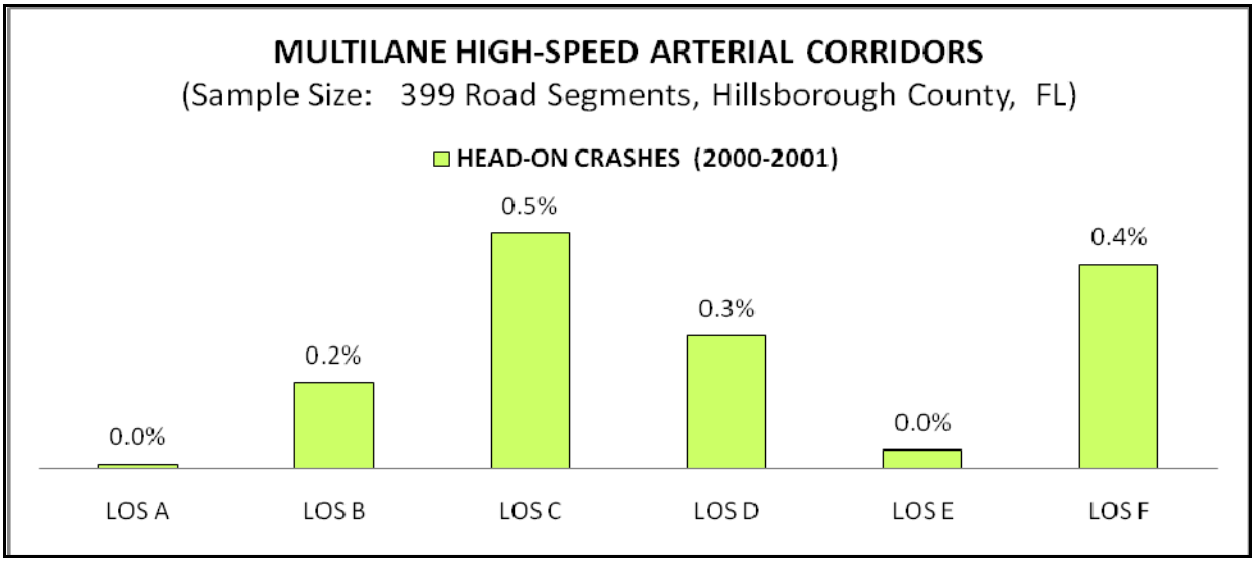


Figure 4-20: Study Sample's Distribution for "Head-On Crashes" at Multilane High-Speed Arterial Corridors

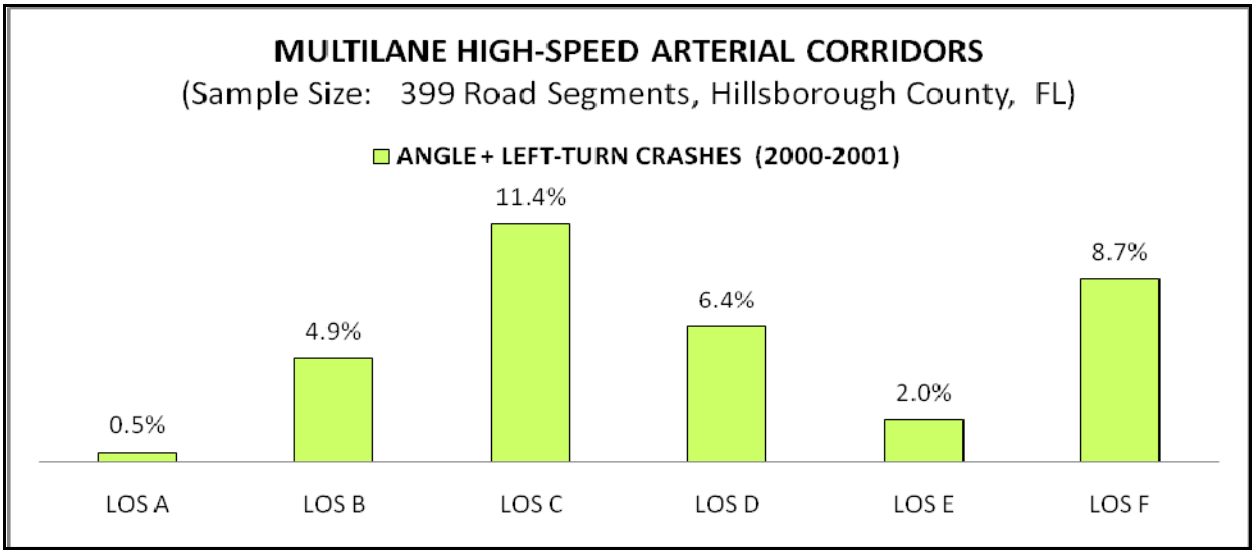


Figure 4-21: Study Sample's Distribution for "Angle and Left-Turn Crashes" at Multilane High-Speed Arterial Corridors

CHAPTER 5: MODELING APPROACH AND RESULTS

5.1 Overview

As observed from the preliminary trends shown in Chapter 4, the data from both studies comprising this research has a significant potential for modeling different case scenarios. With this in consideration, various statistical models were created for analyzing the LOS-Safety relationship based on the signalized intersections' and segments from multilane high-speed corridors' data in hand.

The current chapter discusses the 3 groups of models resulting from the analysis: Model Groups A and B (for Signalized Intersections' Aggregate and Disaggregate temporal analyses, respectively), and Model Group C (for Multilane High-Speed Arterial Corridors). Details on each model building procedure, as well as details on the statistical tools considered for the analysis and respective model assessment are also presented.

5.2 LOS-Safety Study for Signalized Intersections

5.2.1 Aggregate Analysis (Considering the 5 Periods of the Day)

5.2.1.1 Statistical Approach

5.2.1.1.1 Generalized Estimating Equations (GEEs)

The GEE technique provides an extension of the Generalized Linear Model (GLM). This method was used for the LOS-Safety aggregate analysis corresponding to signalized intersections –by aggregate referring to an overall analysis covering all periods of a regular day–. The choice of this statistical tool has been based on the type of data used and the criterion variable being studied: categorical data and frequency of crashes, respectively. For the purpose of studying the frequency of crashes per period of the day of each signalized intersection in the sample, a GEE analysis was performed using the Statistical Analysis Software (SAS).

An important aspect of the GEE modeling procedure in SAS was the computation of "Model-Based" standard error estimates for each of the different models' predictors; the MODELSE syntax was incorporated within the code in order to perform this task. To use the GEE's "Model-Based" estimates, instead of the "Empirical" ones which are provided as the default in the analysis, was deemed to be appropriate since a model-based analysis is better suited for medium-sized datasets like the one being used for this study. For this particular GEE estimation process, the scale parameter was computed by SAS as the square root of the normalized Pearson's chi-square (SAS Institute Inc., 2008).

One of the most important advantages from the GEE technique is that it accounts for the temporal correlation among repeated observations for each single subject composing a set of clustered data (Wang and Abdel-Aty, 2007); for this reason, the GEE facilitated the analysis of grouped count data from each signalized intersection in this study's sample, correlating their respective observations for five different time periods in a regular weekday.

Contrary to what is used for GLMs (e.g. regular Goodness-of-Fit criteria), the assessment of GEE models requires a more specific approach since their data are assumed to be correlated. For this purpose, the method of Cumulative Residuals was deemed to be appropriate for this study's model assessment; past analyses performed by Wang (2006) and Wang & Abdel-Aty (2007) of crashes occurring at signalized intersections corroborate the effectiveness of this method for evaluating GEEs. The Cumulative Residuals method used for this study is the one proposed by Lin, Wei and Ying (2002). Implying a quasi-likelihood function, this numerical and graphical statistical technique checks for a good fit of the link function of GEEs and the respective continuous covariates' functional form. In addition, this method can only work with balanced data (i.e. cannot work with underdispersed data).

5.2.1.2 Models

5.2.1.2.1 Total Crashes

5.2.1.2.1.1 Model Building Procedure

The analysis started by creating the first GEE model, which showed the frequency of "Total Crashes" as the dependent variable. On the other hand, the model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for "Total Crashes" included the following independent variables:

- *County (COUNTY)*: Having 2 levels (Orange and Hillsborough); Hillsborough County was treated as the base case.

- *Period of the day (APERIOD)*: having 5 levels (Early Morning, A.M. Peak, Midday, P.M. Peak and Late Evening); the Early Morning period was treated as the base case.
- *Logarithm of the Cycle Length (LOG_CL)*: Due to its continuous nature, this variable was processed having only one level.
- *LOS for the intersection as a whole (LOS_INT_NUM)*: Having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
- *Logarithm of the total traffic volume, for the 2-hour period of the day, on the Minor road (LOG_MIN_PER)*: Due to its continuous nature, this variable was processed having only one level.
- *Logarithm of the total left-turn traffic volume, for the 2-hour period of the day, on the Major road (LOG_MAJ_LTPER)*: Due to its continuous nature, this variable was processed having only one level.

From the group of covariates just described, only *County* and *Period of the day* were treated as dummy variables.

This “aggregate level analysis” (i.e. 5 periods of the day) for each of the final GEE models consisted in decomposing the analysis into four submodels, one per type of working correlation structure to be tried (independent, autoregressive, exchangeable and unstructured). The working correlation structure of all submodel types was based on a total of 149 clusters (number of signalized intersections), each having 5 as their maximum and minimum dimensions; this resulted in a 5x5 working correlation matrix per submodel. A total of 745 observations were read and used for each submodel, number that is five times the total number of signalized intersections in the study sample; this can be translated into the following: 149 intersections * 5 periods of the day = 745 observations. Tables 5-1 and 5-2 provide a summary of the model, its composition and respective model-based estimates for each correlation structure.

Table 5-1: General Model Information for "Total Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS	
Type III GEE Analysis - General Model Information	
MODEL 1A: "TOTAL CRASHES" (Considering the 5 Periods of the Day)	
Number of Clusters (i.e. Number of Signalized Intersections)	149
Cluster Size (i.e. Number of Continuous Periods of the Day)	5
Number of Observations	745
Number of Total Crashes	5,532

Table 5-2: Model-Based Standard Error Estimates for "Total Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS								
Type III GEE Analysis - Model-Based Standard Error Estimates								
MODEL 1A: "TOTAL CRASHES" (Considering the 5 Periods of the Day)								
PARAMETER	Working Correlation Structure							
	INDEPENDENT		AUTOREGRESSIVE		EXCHANGEABLE		UNSTRUCTURED	
	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)
Intercept	-3.9922	0.4739 (<.0001)	-3.3942	0.4755 (<.0001)	-3.1543	0.474 (<.0001)	-3.0494	0.4626 (<.0001)
County *(Dummy variable)*								
Orange	-0.3106	0.0528 (<.0001)	-0.2864	0.0686 (<.0001)	-0.3787	0.0766 (<.0001)	-0.2703	0.0794 (-0.0007)
Hillsborough	0	0 .	0	0 .	0	0 .	0	0 .

PARAMETER	Working Correlation Structure							
	INDEPENDENT		AUTOREGRESSIVE		EXCHANGEABLE		UNSTRUCTURED	
	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)
Period of the Day *(Dummy variable)*								
Late Evening	0.4197	0.1215 (-0.0006)	0.421	0.1285 (-0.0011)	0.496	0.1211 (<.0001)	0.4575	0.1171 (<.0001)
P.M. Peak	0.4254	0.1275 (-0.0009)	0.4949	0.1403 (-0.0004)	0.5919	0.1372 (<.0001)	0.5341	0.1416 (-0.0002)
Midday	0.4016	0.1243 (-0.0012)	0.4323	0.1318 (-0.001)	0.5167	0.13 (<.0001)	0.4645	0.1353 (-0.0006)
A.M. Peak	0.2126	0.1237 (-0.0857)	0.2903	0.1239 (-0.0192)	0.3811	0.1295 (-0.0032)	0.3282	0.137 (-0.0166)
Early Morning	0	0 .	0	0 .	0	0 .	0	0 .
Logarithm of the Cycle Length (for the Period of the Day)	0.5951	0.0975 (<.0001)	0.4552	0.0989 (<.0001)	0.4481	0.0986 (<.0001)	0.4001	0.0975 (<.0001)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)								
LOS F	0.7325	0.1965 (-0.0002)	0.6335	0.2001 (-0.0015)	0.5607	0.1905 (-0.0032)	0.5161	0.1847 (-0.0052)
LOS E	0.8459	0.1983 (<.0001)	0.7707	0.1986 (-0.0001)	0.6783	0.1899 (-0.0004)	0.6385	0.1846 (-0.0005)
LOS D	0.7707	0.1856 (<.0001)	0.7464	0.186 (<.0001)	0.6708	0.1768 (-0.0001)	0.5791	0.1707 (-0.0007)
LOS C	0.7734	0.1749 (<.0001)	0.7698	0.1736 (<.0001)	0.7009	0.1642 (<.0001)	0.626	0.1587 (<.0001)
LOS B	0.5709	0.1718 (-0.0009)	0.5165	0.1663 (-0.0019)	0.4823	0.1563 (-0.002)	0.4261	0.1496 (-0.0044)
LOS A	0	0 .	0	0 .	0	0 .	0	0 .
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	0.2209	0.0344 (<.0001)	0.2095	0.0416 (<.0001)	0.1646	0.0428 (-0.0001)	0.1905	0.0426 (<.0001)
Logarithm of Total Left-Turn Traffic Volume on Major Road (for the Period of the Day)	0.1902	0.0448 (<.0001)	0.2148	0.0524 (<.0001)	0.2303	0.0528 (<.0001)	0.2371	0.0507 (<.0001)
Dispersion	1.0982		1.0953		1.0966		1.0962	

5.2.1.2.1.2 Model Interpretation

The respective model interpretation, per correlation structure, is as follows:

- *Independent Correlation Structure*. As shown in Table 5-2, only the A.M. Peak period, second level for period of the day, was reported to be insignificant (P-value = 0.0857); in spite of this, the period of the day variable could be kept in this first submodel since the rest of its levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05). As an additional note, the delay variable was considered initially in the modeling process; however, this variable had to be removed since it did not contribute to the significance of the overall model; this could be attributed to the fact that delay is highly correlated to the intersection's LOS. As it will be seen in the following sections, this case applied to the rest of resulting models.

  A thorough evaluation of this first submodel for "Total Crashes" seems to suggest that the A.M. Peak period has the least estimated crash frequency, whereas the P.M. Peak period is the one with the highest estimated crash frequency; overall, more crashes seem to occur as it becomes later in the day. Regarding operational measures, a signalized intersection's Cycle Length was found to be significant by denoting an almost linear relationship with total crash frequency; this could be explained by saying that the longer the cycle length there is a bigger time frame open for crash occurrence. In terms of measures of exposure, the model predicts that traffic volume on the Minor road is also linearly related to the total crash frequency; this suggests that, at the same operational levels, a signalized intersection with very high vehicular volume on its Minor road will experience more crashes. Similarly, it was found that the higher the left-turn traffic volume on the Major road, the higher the chances for a crash to occur at a signalized intersection. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated crash frequency, whereas LOS E (significant delay) is the level with the highest estimated crash frequency for a signalized intersection; then, at LOS F (excessive delay) the crash rate seems to decrease again, being even less than the almost equal crash trends predicted for LOS C (acceptable delay) and LOS D (tolerable delay). Summarizing, as a signalized intersection becomes more congested, the higher the crash risk; however, as soon as there are jammed conditions (i.e. the intersection's capacity is exceeded), the crash risk's trend gets reversed (i.e. starts to decrease).

- *Autoregressive Correlation Structure*. As shown in Table 5-2, it can be observed that all predictors were found significant (P-value $\leq$ 0.05). These results indicate from the beginning that this model has a good fit.

  A thorough evaluation of this second submodel for "Total Crashes" seems to suggest that the A.M. Peak period has the least estimated crash frequency, whereas the P.M. Peak period is the one with the highest estimated crash frequency; overall, more crashes seem to occur as it becomes later in the day. Regarding operational measures, a signalized intersection's Cycle Length was found to be significant by denoting an almost linear relationship with total crash frequency; this could be explained by saying that the longer the cycle length there is a bigger time frame open for crash occurrence. In terms of measures of exposure, the model predicts that traffic volume on the Minor road is also linearly related to the total crash frequency; this suggests that, at the same operational levels, a signalized intersection with very high vehicular volume on its Minor road will experience more crashes. Similarly, it was found that the higher the left-turn traffic volume on the Major road, the higher the chances for a crash to occur at a signalized intersection. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated crash frequency, whereas LOS E (significant delay) is the level with the highest estimated crash frequency for a signalized intersection; then, at LOS F (excessive delay) the crash rate seems to decrease again, being even less than the crash trends predicted for LOS C (acceptable delay) and LOS D (tolerable delay). Summarizing, as a signalized intersection becomes more congested, the higher the crash risk; however, as soon as there are jammed conditions (i.e. the intersection's capacity is exceeded), the crash risk's trend gets reversed (i.e. starts to decrease).

- *Exchangeable Correlation Structure*. As shown in Table 5-2, it can be observed that all predictors were found significant (P-value $\leq$ 0.05). These results indicate from the beginning that this model has a good fit.

  A thorough evaluation of this third submodel for "Total Crashes" seems to suggest that the A.M. Peak period has the least estimated crash frequency, whereas the P.M. Peak period is the one with the highest estimated crash frequency; overall, more crashes seem to occur as it becomes later in the day.

Regarding operational measures, a signalized intersection's Cycle Length was found to be significant by denoting an almost linear relationship with total crash frequency; this could be explained by saying that the longer the cycle length there is a bigger time frame open for crash occurrence. In terms of measures of exposure, the model predicts that traffic volume on the Minor road is also linearly related to the total crash frequency; this suggests that, at the same operational levels, a signalized intersection with very high vehicular volume on its Minor road will experience more crashes. Similarly, it was found that the higher the left-turn traffic volume on the Major road, the higher the chances for a crash to occur at a signalized intersection. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated crash frequency, whereas LOS C (acceptable delay) is the level with the highest estimated crash frequency for a signalized intersection; then, the crash trend seems to decrease again with the almost equal predicted values for both LOS D (tolerable delay) and LOS E (significant delay), followed by LOS F (excessive delay). Summarizing, there is a higher probability for a crash to occur at a signalized intersection when it experiences delays between the acceptable and significant range; it is when delays become excessive (i.e. the intersection's capacity is exceeded) where the crash risk starts to decrease.

- *Unstructured Correlation Structure*. As shown in Table 5-2, it can be observed that all predictors were found significant (P-value $\leq$ 0.05). These results indicate from the beginning that this model has a good fit.

A thorough evaluation of this fourth and last submodel for “Total Crashes” seems to suggest that the A.M. Peak period has the least estimated crash frequency, whereas the P.M. Peak period is the one with the highest estimated crash frequency; overall, more crashes seem to occur as it becomes later in the day. Regarding operational measures, a signalized intersection’s Cycle Length was found to be significant by denoting an almost linear relationship with the total crash frequency; this could be explained by saying that the longer the cycle length there is a bigger time frame open for crash occurrence. In terms of measures of exposure, the model predicts that traffic volume on the Minor road is also linearly related to the total crash frequency; this suggests that, at the same operational levels, a signalized intersection with very high vehicular volume on its Minor road will experience more crashes. Similarly, it was found that the higher the left-turn traffic volume on the Major road, the higher the chances for a crash to occur at a signalized intersection. Finally, regarding the intersection’s LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated crash frequency, whereas LOS E (significant delay) is the level with the highest estimated crash frequency for a signalized intersection; then, at LOS F (excessive delay) the crash rate seems to decrease again, being even less than the crash trends predicted for LOS C (acceptable delay) and LOS D (tolerable delay). Summarizing, as a signalized intersection becomes more congested, the higher the crash risk; however, as soon as there are jammed conditions (i.e. the intersection’s capacity is exceeded), the crash risk’s trend gets reversed (i.e. starts to decrease).

5.2.1.2.1.3 Model Assessment

This section summarizes the respective model assessment performed through SAS when analyzing the GEE model designed for "Total Crashes" for the 5 periods of the day (refer to Section 5.2.1.2.1). Following are the plots that were generated through SAS' Output Delivery System (ODS) for each of the GEE model's type of working correlation structure (see Figure 5-1). When interpreting the plots, the heavy trend line represents the Cumulative Residuals whereas the light ones represent the simulated curves.

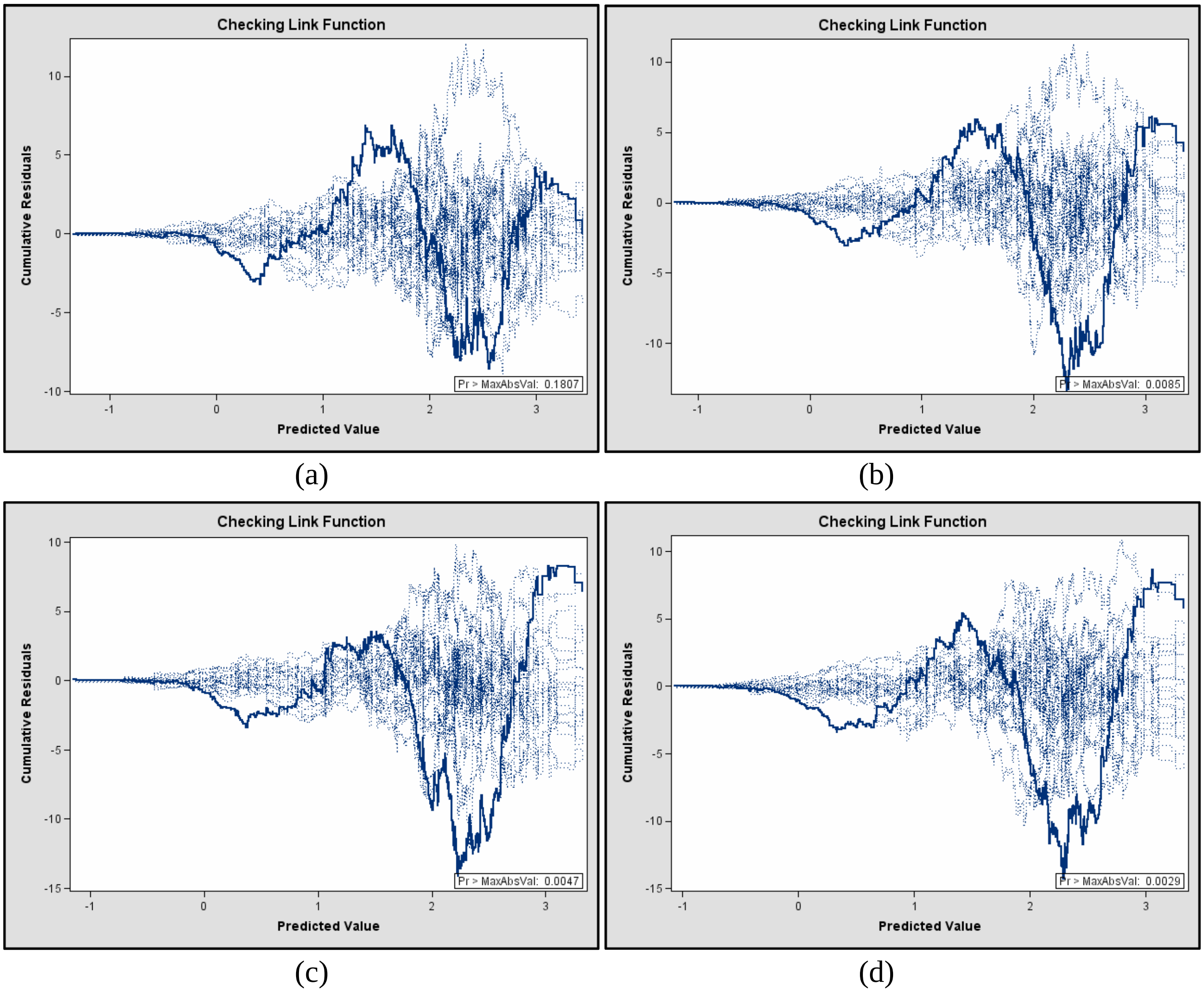


Figure 5-1: Model Assessment Plots (Cumulative Residuals for GEE Negative Binomial Analysis) for "Total Crashes" at Signalized Intersections (Considering the 5 Periods of the Day): (a) Independent Structure, (b) Autoregressive Structure, (c) Exchangeable Structure and (d) Unstructured

As it can be seen, these plots provide a P-value (Pr>MaxAbsVal) computed through a simulation of 10,000 residual paths; below is a summary of these computed values (see Table 5-3).

Table 5-3: Model Assessment Summary (Cumulative Residuals for GEE Negative Binomial Analysis) for "Total Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
GEE Analysis - Model Assessment Summary					
MODEL 1A: "TOTAL CRASHES" (Considering the 5 Periods of the Day)					
Working Correlation Structure	Assessment Variable	Max. Abs. Value	Replications	Seed	Pr > Max. Abs. Value
INDEPENDENT	Link Function	8.5706	10000	603708000	0.1807
AUTOREGRESSIVE	Link Function	13.4096	10000	603708000	0.0085
EXCHANGEABLE	Link Function	14.0946	10000	603708000	0.0047
UNSTRUCTURED	Link Function	14.3085	10000	603708000	0.0029

In addition, the GEE analysis procedure in SAS provided the model's respective score statistics based on a Type III analysis for each covariate in the model. It can be concluded from these results that, among all correlation structures, cycle length is the most significant of all covariates. Following is Table 5-4, which shows the corresponding Score Statistics for each covariate, listed in descending order of significance.

Table 5-4: GEE Score Statistics Summary for "Total Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Score Statistics					
MODEL 1A: "TOTAL CRASHES" (Considering the 5 Periods of the Day)					
COVARIATE	DF	Working Correlation Structure			
		INDEPENDENT	AUTOREGRESSIVE	EXCHANGEABLE	UNSTRUCTURED
		Chi-Sq. (P-value)	Chi-Sq. (P-value)	Chi-Sq. (P-value)	Chi-Sq. (P-value)
Logarithm of the Cycle Length (for the Period of the Day)	1	20.4 (<.0001)	16.8 (<.0001)	17.43 (<.0001)	15.29 (<.0001)
Logarithm of Total Left-Turn Traffic Volume on Major Road (for the Period of the Day)	1	9.47 (0.0021)	13.47 (0.0002)	19.1 (<.0001)	20.15 (<.0001)
County *(Dummy variable)*	1	15.89 (<.0001)	13.16 (0.0003)	20.79 (<.0001)	7.36 (0.0067)
Period of the Day *(Dummy variable)*	4	16.03 (0.003)	14.21 (0.0067)	16.75 (0.0022)	16.14 (0.0028)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)	5	14.01 (0.0156)	18.6 (0.0023)	16.52 (0.0055)	10.94 (0.0525)
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	1	13.27 (0.0003)	12.47 (0.0004)	9.49 (0.0021)	8.88 (0.0029)

The last outputs were the working correlation matrices for each structure type (see Table 5-5). As it can be seen, the Independent working correlation matrix was the result of a very naïve analysis procedure; the binary composition of this matrix reflects this fact. Regarding the Autoregressive working correlation matrix, it can be observed that it is characterized by correlation values that decrease over time. The Exchangeable working correlation matrix, however, has a defined and compound symmetry; the exchangeable working correlation had a reported value of 0.3322 for this model. Finally, it can be seen that the Unstructured working correlation matrix's composition has no particular specification.

Table 5-5: GEE Working Correlation Matrices for "Total Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Working Correlation Matrices					
MODEL 1A: "TOTAL CRASHES" (Considering the 5 Periods of the Day)					
INDEPENDENT	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0000	0.0000	0.0000	0.0000
A.M. Peak	0.0000	1.0000	0.0000	0.0000	0.0000
Midday	0.0000	0.0000	1.0000	0.0000	0.0000
P.M. Peak	0.0000	0.0000	0.0000	1.0000	0.0000
Late Evening	0.0000	0.0000	0.0000	0.0000	1.0000
AUTOREGRESSIVE	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.3471	0.1205	0.0418	0.0145
A.M. Peak	0.3471	1.0000	0.3471	0.1205	0.0418
Midday	0.1205	0.3471	1.0000	0.3471	0.1205
P.M. Peak	0.0418	0.1205	0.3471	1.0000	0.3471
Late Evening	0.0145	0.0418	0.1205	0.3471	1.0000
EXCHANGEABLE	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.3322	0.3322	0.3322	0.3322
A.M. Peak	0.3322	1.0000	0.3322	0.3322	0.3322
Midday	0.3322	0.3322	1.0000	0.3322	0.3322
P.M. Peak	0.3322	0.3322	0.3322	1.0000	0.3322
Late Evening	0.3322	0.3322	0.3322	0.3322	1.0000
UNSTRUCTURED	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.1272	0.1634	0.2188	0.4217
A.M. Peak	0.1272	1.0000	0.5113	0.5252	0.2608
Midday	0.1634	0.5113	1.0000	0.4910	0.2511
P.M. Peak	0.2188	0.5252	0.4910	1.0000	0.3825
Late Evening	0.4217	0.2608	0.2511	0.3825	1.0000

5.2.1.2.1.4 Conclusion

After comparing the results from all correlation structures, it may be concluded that the Independent correlation structure is the best among the other types, since it has the highest P-value (=0.1807). However, as it is known, the Independent correlation structure is the simplest of all, assuming no correlation among any of the observations within the study sample (Hardin and Hilbe, 2003). With the data in hand, and for the purpose of analyzing the LOS-Safety relationship throughout 5 sequentially ordered periods of the day, to choose the Independent correlation structure may not be the best decision.

Based on the previous statements, the Autoregressive correlation structure seems to be the best choice, not only for having the second highest P-value (=0.0085) (refer to Table 5-3) but also because it assumes that all observations in the sample being studied are temporally correlated (Wang, 2006); this seems to be appropriate since the data for this signalized intersections study is temporal in nature (i.e. it follows a natural order).

An appropriate interpretation of the results from the Autoregressive correlation structure, and recalling the basis of the Cumulative Residuals model assessment method, would be that "out of 10,000 realizations from the null distribution, 0.85% have maximum cumulative residuals greater than 13.4096". In addition, the 5x5 matrix for the Autoregressive correlation structure suggests that the repeated observations (i.e. 5 times) for each signalized intersection will become less correlated as the time-gap increases. Finally, the estimates obtained with the Autoregression correlation structure for each of the covariates in the model were all significant, corroborating the idea that the resulting model fits the data well. Finally, Table 5-6 lists the data summary statistics of the contributing factors included in the final model.

Table 5-6: Data Summary Statistics for Preferred Model for "Total Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Data Summary Statistics					
MODEL 1A: "TOTAL CRASHES" (Considering the 5 Periods of the Day)					
Preferred Correlation Structure: AUTOREGRESSIVE					
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**	**Chi-Sq. (P-value)**
Logarithm of the Cycle Length (for the Period of the Day) (seconds)	4.6	1.6	5.3	0.4	16.8 (<.0001)
Logarithm of Total Left-Turn Traffic Volume on Major Road (for the Period of the Day)	4.6	0.7	7.2	1.4	13.47 (0.0002)
County *(Dummy variable)* (Key: Orange=1, Hillsborough=2)	0.5	0	1	0.7	13.16 (0.0003)
Period of the Day *(Dummy variable)* (Key: Early Morning=1, A.M. Peak=2, Midday=3, P.M. Peak=4, Late Evening=5)	3.0	1	5	1.6	14.21 (0.0067)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9	18.6 (0.0023)
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	5.4	1.4	8.6	1.5	12.47 (0.0004)
Total Number of 5x5 Clusters (i.e. Signalized Intersections): 149					

5.2.1.2.2 Severe Crashes

5.2.1.2.2.1 Model Building Procedure

This analysis consisted in creating the second GEE model, which showed the frequency of "Severe Crashes" as the dependent variable.

The model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative

importance and significance these covariates could have for the model. Consequently, the best model for “Severe Crashes” included the following independent variables:

- *County (COUNTY)*: having 2 levels (Orange and Hillsborough); Hillsborough County was treated as the base case.
- *Period of the day (APERIOD)*: having 5 levels (Early Morning, A.M. Peak, Midday, P.M. Peak and Late Evening); the Early Morning period was treated as the base case.
- *LOS for the intersection as a whole (LOS_INT_NUM)*: Having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
- *Total Number of lanes on the Major road (LANMAJ_DAYC1)*: originally with 10 levels, this variable ended up having 5 levels after appropriate combinations were made (4 lanes, 5 with 6 lanes, 7 with 8 lanes, 9 with 10 lanes, and 11 combined with 12 and 14 lanes); the level corresponding to 4 lanes, the lowest number, was treated as the base case.

From the group of covariates just described, only *County* and *Period of the day* were treated as dummy variables.

This “aggregate level analysis” (i.e. 5 periods of the day) for each of the final GEE models consisted in decomposing the analysis into four submodels, one per type of working correlation structure to be tried (independent, autoregressive, exchangeable and unstructured). The working correlation structure of all submodel types was based on a total of 149 clusters (number of signalized intersections), each having 5 as their maximum and minimum dimensions; this resulted in a 5x5 working correlation matrix per submodel. A total of 745 observations were read and used for each submodel, number that is five times the total number of signalized intersections in the study sample; this can be translated into the following: 149 intersections * 5 periods of the day = 745 observations. Tables 5-7 and 5-8 provide a summary of the model, its composition and respective model-based estimates for each correlation structure.

Table 5-7: General Model Information for "Severe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS	
Type III GEE Analysis - Score Statistics	
MODEL 2A: "SEVERE CRASHES" (Considering the 5 Periods of the Day)	
Number of Clusters (i.e. Number of Signalized Intersections)	149
Cluster Size (i.e. Number of Continuous Periods of the Day)	5
Number of Observations	745
Number of Severe/Total Crashes	459/5,532

Table 5-8: Model-Based Standard Error Estimates for "Severe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS								
Type III GEE Analysis - Model-Based Standard Error Estimates								
MODEL 2A: "SEVERE CRASHES" (Considering the 5 Periods of the Day)								
PARAMETER	Working Correlation Structure							
	INDEPENDENT		AUTOREGRESSIVE		EXCHANGEABLE		UNSTRUCTURED	
	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)
Intercept	-2.7703	0.4359 (<.0001)	-2.7574	0.4475 (<.0001)	-2.7829	0.4583 (<.0001)	-2.795	0.465 (<.0001)
County *(Dummy variable)*								
Orange	-0.4683	0.1195 (<.0001)	-0.4604	0.1276 (-0.0003)	-0.462	0.1332 (-0.0005)	-0.4472	0.137 (-0.0011)
Hillsborough	0	0 .	0	0 .	0	0 .	0	0 .

PARAMETER	Working Correlation Structure							
	INDEPENDENT		AUTOREGRESSIVE		EXCHANGEABLE		UNSTRUCTURED	
	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)
Period of the Day *(Dummy variable)*								
Late Evening	1.2576	0.233 (<.0001)	1.2644	0.2332 (<.0001)	1.27	0.2271 (<.0001)	1.2653	0.2143 (<.0001)
P.M. Peak	1.0455	0.2477 (<.0001)	1.0685	0.2492 (<.0001)	1.0954	0.2439 (<.0001)	1.0783	0.2549 (<.0001)
Midday	1.0052	0.2419 (<.0001)	1.019	0.2422 (<.0001)	1.0345	0.2369 (<.0001)	1.0244	0.2525 (<.0001)
A.M. Peak	0.8991	0.249 (-0.0003)	0.9196	0.2429 (-0.0002)	0.9426	0.2447 (-0.0001)	0.931	0.2578 (-0.0003)
Early Morning	0	0 .	0	0 .	0	0 .	0	0 .
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)								
LOS F	0.9812	0.3364 (-0.0035)	0.9382	0.3439 (-0.0064)	0.9104	0.3496 (-0.0092)	0.9358	0.3518 (-0.0078)
LOS E	1.0258	0.3562 (-0.004)	0.9914	0.3613 (-0.0061)	0.9951	0.3662 (-0.0066)	0.9629	0.3684 (-0.009)
LOS D	0.8462	0.3355 (-0.0117)	0.8329	0.3396 (-0.0142)	0.8366	0.3445 (-0.0152)	0.8315	0.3458 (-0.0162)
LOS C	0.6699	0.3261 (-0.04)	0.6528	0.3295 (-0.0475)	0.665	0.3333 (-0.046)	0.6788	0.3354 (-0.043)
LOS B	0.2624	0.3252 (-0.4197)	0.257	0.3274 (-0.4324)	0.2806	0.3297 (-0.3947)	0.2819	0.3301 (-0.3933)
LOS A	0	0 .	0	0 .	0	0 .	0	0 .
Total Number of Lanes on Major Road								
11 ≤ lanes ≤ 14	0.6639	0.3102 (-0.0323)	0.6649	0.3303 (-0.0441)	0.6581	0.3461 (-0.0572)	0.7248	0.3564 (-0.042)
9 lanes or 10 lanes	0.9701	0.2889 (-0.0008)	0.9608	0.3077 (-0.0018)	0.9685	0.3218 (-0.0026)	0.988	0.3333 (-0.003)
7 lanes or 8 lanes	0.8365	0.2816 (-0.003)	0.8263	0.2997 (-0.0058)	0.8361	0.3134 (-0.0076)	0.8391	0.325 (-0.0098)
5 lanes or 6 lanes	0.9169	0.2942 (-0.0018)	0.9027	0.3135 (-0.004)	0.9156	0.3277 (-0.0052)	0.9035	0.3401 (-0.0079)
4 lanes	0	0 .	0	0 .	0	0 .	0	0 .
Dispersion	0.9923		0.992		0.9918		0.9918	

5.2.1.2.2.2 Model Interpretation

The respective model interpretation, per correlation structure, is as follows:

- *Independent Correlation Structure*. As shown in Table 5-8, only LOS B (minimal delay), second level of the LOS scale, was reported to be insignificant (P-value = 0.4197); in spite of this, the LOS for the intersection as a whole could be kept as a variable in this first submodel since the rest of its levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

  A thorough evaluation of this first submodel for "Severe Crashes" seems to suggest that the A.M. Peak period has the least estimated severe crash frequency, whereas the Late Evening period is the one with the highest estimated severe crash frequency; overall, more severe crashes seem to occur as it becomes later in the day. It can also be observed that a total number of lanes between 5 and 10 lanes on the Major road is associated with a very high severe crash risk, whereas for a total number of lanes of 11 or more this risk seems to decrease considerably. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated severe crash frequency, whereas LOS E (significant delay) is the level with the highest estimated severe crash frequency for a signalized intersection, immediately followed by LOS F (excessive delay) which is associated with the second highest severe crash frequency; on the other side of the scale, and in decreasing level of severe crash risk are LOS D (tolerable delay) in third place, followed by LOS C (acceptable delay) and LOS B (minimal delay). Summarizing, the probability for a severe crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

- *Autoregressive Correlation Structure*. As shown in Table 5-8, only LOS B (minimal delay), second level of the LOS scale, was reported to be insignificant (P-value = 0.4324); in spite of this, the LOS for the intersection as a whole could be kept as a variable in this second submodel since the rest of its levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

  A thorough evaluation of this second submodel for "Severe Crashes" seems to suggest that the A.M. Peak period has the least estimated severe crash frequency, whereas the Late Evening period is the one with the highest estimated severe crash frequency; overall, more severe crashes seem to occur as it becomes later in the day. It can also be observed that a total number of lanes between 5 and 10 lanes on the Major road is associated with a very high severe crash risk, whereas for a total number of lanes of 11 or more this risk seems to decrease considerably. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated severe crash frequency, whereas LOS E (significant delay) is the level with the highest estimated severe crash frequency for a signalized intersection, followed by LOS F (excessive delay) which is associated with the second highest severe crash frequency; on the other side of the scale, and in decreasing level of severe crash risk are LOS D (tolerable delay) in third place, followed by LOS C (acceptable delay) and LOS B (minimal delay). Summarizing, the probability for a severe crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

- *Exchangeable Correlation Structure*. As shown in Table 5-8, only LOS B (minimal delay) and a total number of 11 lanes or higher on the Major road were reported to be insignificant (P-values of 0.3947 and 0.0572, respectively); in spite of this, the respective variables could be kept in this third submodel since the rest of their levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

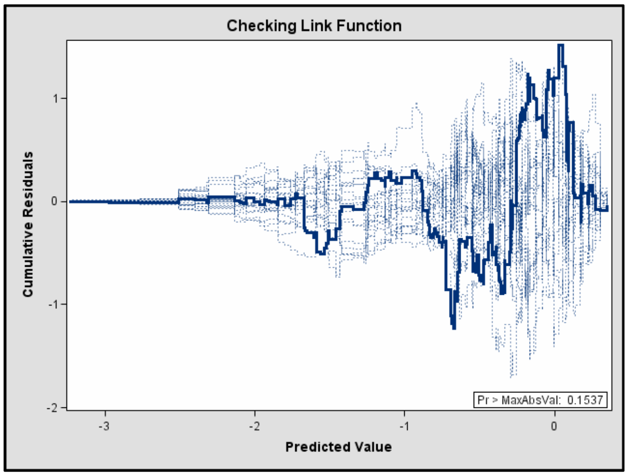
  A thorough evaluation of this third submodel for "Severe Crashes" seems to suggest that the A.M. Peak period has the least estimated severe crash frequency, whereas the Late Evening period is the one with the highest estimated severe crash frequency; overall, more severe crashes seem to occur as it becomes later in the day. It can also be observed that a total number of lanes between 5 and 10 lanes on the Major road is associated with a very high severe crash risk, whereas for a total number of lanes of 11 or more this risk seems to decrease considerably. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated severe crash frequency, whereas LOS E (significant delay) is the level with the highest estimated severe crash frequency for a signalized intersection, followed by LOS F (excessive delay) which is associated with the second highest severe crash frequency; on the other side of the scale, and in decreasing level of severe crash risk are LOS D (tolerable delay) in third place, followed by LOS C (acceptable delay) and LOS B (minimal delay). Summarizing, the probability for a severe crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

- *Unstructured Correlation Structure*. As shown in Table 5-8, only LOS B (minimal delay), second level of the LOS scale, was reported to be insignificant (P-value = 0.3933); in spite of this, the LOS for the intersection as a whole could be kept as a variable in this fourth submodel since the rest of its levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

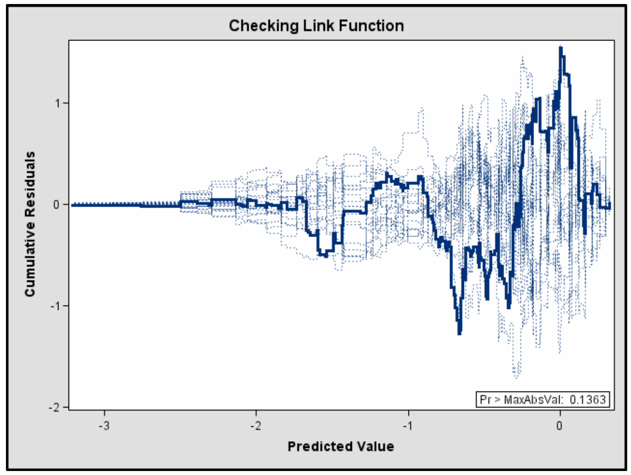
  A thorough evaluation of this fourth and last submodel for "Severe Crashes" seems to suggest that the A.M. Peak period has the least estimated severe crash frequency, whereas the Late Evening period is the one with the highest estimated severe crash frequency; overall, more severe crashes seem to occur as it becomes later in the day. It can also be observed that a total number of lanes between 5 and 10 lanes on the Major road is associated with a very high severe crash risk, whereas for a total number of lanes of 11 or more this risk seems to decrease considerably. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated severe crash frequency, whereas LOS E (significant delay) is the level with the highest estimated severe crash frequency for a signalized intersection, followed by LOS F (excessive delay) which is associated with the second highest estimated severe crash frequency; on the other side of the scale, and in decreasing level of severe crash risk are LOS D (tolerable delay) in third place, followed by LOS C (acceptable delay) and LOS B (minimal delay). Summarizing, the probability for a severe crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

5.2.1.2.2.3 Model Assessment

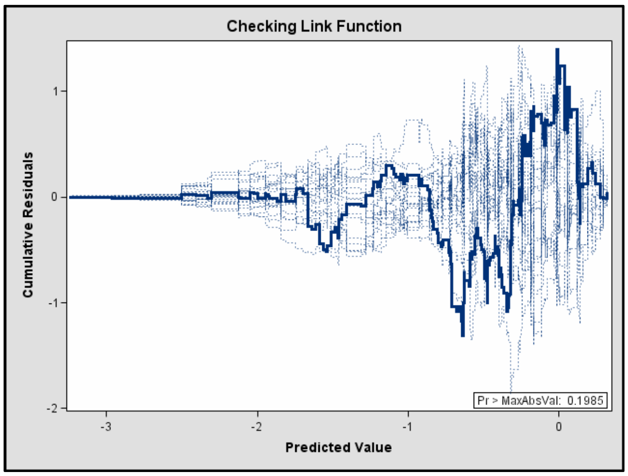
This section summarizes the respective model assessment performed through SAS when analyzing the GEE model designed for "Severe Crashes" for the 5 periods of the day. Following are the plots that were generated through SAS' Output Delivery System (ODS) for each of the GEE model's type of working correlation structure (see Figure 5-2). When interpreting the plots, the heavy trend line represents the Cumulative Residuals whereas the light ones represent the simulated curves.



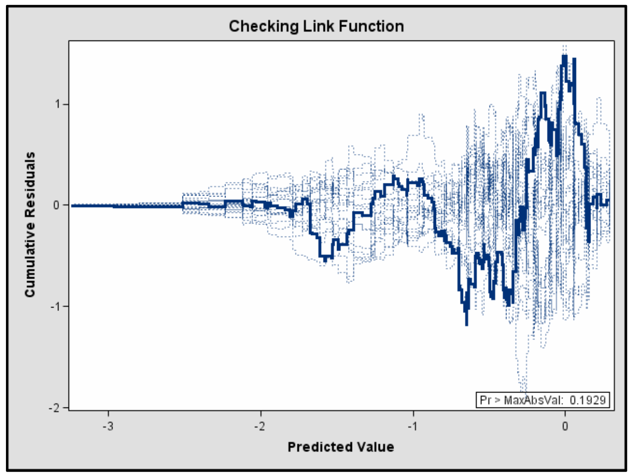
(a)



(b)



(c)



(d)

Figure 5-2: Model Assessment Plots (Cumulative Residuals for GEE Negative Binomial Analysis) for "Severe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day): (a) Independent Structure, (b) Autoregressive Structure, (c) Exchangeable Structure and (d) Unstructured

As it can be seen, these plots provide a P-value (Pr>MaxAbsVal) computed through a simulation of 10,000 residual paths; below is a summary of these computed values (see Table 5-9).

Table 5-9: Model Assessment Summary (Cumulative Residuals for GEE Negative Binomial Analysis) for "Severe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
GEE Analysis - Model Assessment Summary					
MODEL 2A: "SEVERE CRASHES" (Considering the 5 Periods of the Day)					
Working Correlation Structure	Assessment Variable	Max. Abs. Value	Replications	Seed	Pr > Max. Abs. Value
INDEPENDENT	Link Function	1.5146	10000	603708000	0.1537
AUTOREGRESSIVE	Link Function	1.5384	10000	603708000	0.1363
EXCHANGEABLE	Link Function	1.3985	10000	603708000	0.1985
UNSTRUCTURED	Link Function	1.4705	10000	603708000	0.1929

In addition, the GEE analysis procedure in SAS provided the model's respective score statistics based on a Type III analysis for each covariate in the model. It can be concluded from these results that, among all correlation structures, period of the day (dummy variable) is the most significant of all covariates. Following is Table 5-10, which shows the corresponding Score Statistics for each covariate, listed in descending order of significance.

Table 5-10: GEE Score Statistics Summary for "Severe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Score Statistics					
MODEL 2A: "SEVERE CRASHES" (Considering the 5 Periods of the Day)					
COVARIATE	DF	Working Correlation Structure			
		INDEPENDENT	AUTOREGRESSIVE	EXCHANGEABLE	UNSTRUCTURED
		Chi-Sq. (P-value)	Chi-Sq. (P-value)	Chi-Sq. (P-value)	Chi-Sq. (P-value)
Period of the Day *(Dummy variable)*	4	27.3 (<.0001)	27.83 (<.0001)	27.69 (<.0001)	26.88 (<.0001)
County *(Dummy variable)*	1	11.3 (0.0008)	10.83 (0.001)	10.69 (0.0011)	10.17 (0.0014)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)	5	16.56 (0.0054)	15.14 (0.0098)	13.04 (0.023)	11.98 (0.0351)
Total Number of Lanes on Major Road	4	8.27 (0.0821)	8.05 (0.0898)	8.4 (0.0781)	7.49 (0.1123)

The last outputs were the working correlation matrices for each structure type (see Table 5-11). As it can be seen, the Independent working correlation matrix was the result of a very naïve analysis procedure; the binary composition of this matrix reflects this fact. Regarding the Autoregressive working correlation matrix, it can be observed that it is characterized by correlation values that decrease over time. The Exchangeable working correlation matrix, however, has a defined and compound symmetry; the exchangeable working correlation had a reported value of 0.0674 for this model. Finally, it can be seen that the Unstructured working correlation matrix's composition has no particular specification.

Table 5-11: GEE Working Correlation Matrices for "Severe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Working Correlation Matrices					
MODEL 2A: "SEVERE CRASHES" (Considering the 5 Periods of the Day)					
INDEPENDENT	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0000	0.0000	0.0000	0.0000
A.M. Peak	0.0000	1.0000	0.0000	0.0000	0.0000
Midday	0.0000	0.0000	1.0000	0.0000	0.0000
P.M. Peak	0.0000	0.0000	0.0000	1.0000	0.0000
Late Evening	0.0000	0.0000	0.0000	0.0000	1.0000
AUTOREGRESSIVE	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0846	0.0072	0.0006	0.0001
A.M. Peak	0.0846	1.0000	0.0846	0.0072	0.0006
Midday	0.0072	0.0846	1.0000	0.0846	0.0072
P.M. Peak	0.0006	0.0072	0.0846	1.0000	0.0846
Late Evening	0.0001	0.0006	0.0072	0.0846	1.0000
EXCHANGEABLE	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0674	0.0674	0.0674	0.0674
A.M. Peak	0.0674	1.0000	0.0674	0.0674	0.0674
Midday	0.0674	0.0674	1.0000	0.0674	0.0674
P.M. Peak	0.0674	0.0674	0.0674	1.0000	0.0674
Late Evening	0.0674	0.0674	0.0674	0.0674	1.0000
UNSTRUCTURED	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	-0.0665	-0.1004	-0.0428	0.2031
A.M. Peak	-0.0665	1.0000	0.1018	0.2077	-0.0145
Midday	-0.1004	0.1018	1.0000	0.2390	0.1044
P.M. Peak	-0.0428	0.2077	0.2390	1.0000	0.0972
Late Evening	0.2031	-0.0145	0.1044	0.0972	1.0000

5.2.1.2.2.4 Conclusion

After comparing the results from all correlation structures, it may be concluded that the Exchangeable correlation structure is the best among the other types, since it has the highest P-value (=0.1985). As it is known, the Exchangeable correlation structure is a simple extension to the Independent correlation structure by imposing one additional association parameter within the GEE model; thus, the Exchangeable correlation structure assumes that there is a single common correlation between the observations within the study sample (Hardin and Hilbe, 2003). However, to choose the Exchangeable correlation structure may not be the best choice; not only does the structure ignores time dependence among the studied observations but it also allows any permutation of these, which contradicts the objective to analyze the LOS-Safety relationship throughout 5 sequentially ordered periods of the day.

Based on the previous statements, and based on the results for the model being discussed, the Unstructured correlation structure seems to be the best choice, not only for having the second highest P-value (=0.1929) (refer to Table 5-9) but also because it is the most general of all structures and it imposes no restriction with regards to the order of the studied observations (Hardin and Hilbe, 2003). The study conducted by Wang et al. (2009), which explored the temporal correlation of signalized intersection data, also supports the choice of the Unstructured correlation structure. Furthermore, recalling the Unstructured working correlation matrix obtained for the "Severe Crashes"' model (refer to Table 5-11), it can be observed that the correlation between the Midday and P.M. Peak periods and the one between the A.M. Peak and P.M. Peak periods have values of 0.2390 and 0.2077, respectively, which are the first and second highest of all; being the times of the day with similar traffic characteristics (i.e. highest traffic volumes, no "late" hours, etc.), the latter statement also justifies the choice of this structure.

An appropriate interpretation of the results from the Unstructured correlation structure, and recalling the basis of the Cumulative Residuals model assessment method, would be that "out of 10,000 realizations from the null distribution, 19.29% have maximum cumulative residuals greater than 1.4705". Finally, the estimates obtained with the Unstructured correlation structure for each of the covariates in the model were significant overall, corroborating the idea that the resulting model fits the data well. Finally, Table 5-12 lists the data summary statistics of the contributing factors included in the final model.

Table 5-12: Data Summary Statistics for Preferred Model for "Severe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Data Summary Statistics					
MODEL 2A: "SEVERE CRASHES" (Considering the 5 Periods of the Day)					
Preferred Correlation Structure: UNSTRUCTURED					
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**	**Chi-Sq. (P-value)**
Period of the Day *(Dummy variable)* (Key: Early Morning=1, A.M. Peak=2, Midday=3, P.M. Peak=4, Late Evening=5)	3.0	1	5	1.6	26.88 (<.0001)
County *(Dummy variable)* (Key: Orange=0, Hillsborough=1)	0.5	0	1	0.7	10.17 (0.0014)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9	11.98 (0.0351)
Total Number of Lanes on Major Road	7.9	4	14	2.1	7.49 (0.1123)
Total Number of 5x5 Clusters (i.e. Signalized Intersections): 149					

5.2.1.2.3 Crash Types: Rear-End Crashes

5.2.1.2.3.1 <u>Model Building Procedure</u>

This analysis consisted in creating the first of the last set of GEE models (i.e. models for different Crash Types) which shows the frequency of "Rear-End Crashes" as the dependent variable.

The model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for “Rear-End Crashes” included the following independent variables:

- *County (COUNTY)*: having 2 levels (Orange and Hillsborough); Hillsborough County was treated as the base case.
- *Period of the day (APERIOD)*: having 5 levels (Early Morning, A.M. Peak, Midday, P.M. Peak and Late Evening); the Early Morning period was treated as the base case.
- *LOS for the intersection as a whole (LOS_INT_NUM)*: resulting in 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
- *Logarithm of the total traffic volume, for the 2-hour period of the day, on the Major road (LOG_MAJ_PER)*: Due to its continuous nature, this variable was processed having only one level.
- *Logarithm of the total traffic volume, for the 2-hour period of the day, on the Minor road (LOG_MIN_PER)*: Due to its continuous nature, this variable was processed having only one level.

From the group of covariates just described, only *County* and *Period of the day* were treated as dummy variables.

This “aggregate level analysis” (i.e. 5 periods of the day) for each of the final GEE models consisted in decomposing the analysis into four submodels, one per type of working correlation structure (independent, autoregressive, exchangeable and unstructured). The working correlation structure of all submodel types was based on a total of 149 clusters (number of signalized intersections), each having 5 as their maximum and minimum dimensions; this resulted in a 5x5 working correlation matrix per submodel. A total of 745 observations were read and used for each submodel, number that is five times the total number of signalized intersections in the study sample; this can be translated into the following: 149 intersections * 5 periods of the day = 745 observations. Tables 5-13 and 5-14 provide a summary of the model, its composition and respective model-based estimates for each correlation structure.

Table 5-13: General Model Information for "Rear-End Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS	
Type III GEE Analysis - Score Statistics	
MODEL 3A: "REAR-END CRASHES" (Considering the 5 Periods of the Day)	
Number of Clusters (i.e. Number of Signalized Intersections)	149
Cluster Size (i.e. Number of Continuous Periods of the Day)	5
Number of Observations	745
Number of Rear-End/Total Crashes	2,733/5,532

Table 5-14: Model-Based Standard Error Estimates for "Rear-End Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS								
Type III GEE Analysis - Model-Based Standard Error Estimates								
MODEL 3A: "REAR-END CRASHES" (Considering the 5 Periods of the Day)								
PARAMETER	Working Correlation Structure							
	INDEPENDENT		AUTOREGRESSIVE		EXCHANGEABLE		UNSTRUCTURED	
	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)
Intercept	-5.3631	0.4584 (<.0001)	-5.3551	0.4983 (<.0001)	-5.0413	0.5082 (<.0001)	-5.2667	0.5091 (<.0001)

PARAMETER	Working Correlation Structure							
	INDEPENDENT		AUTOREGRESSIVE		EXCHANGEABLE		UNSTRUCTURED	
	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)
County *(Dummy variable)*								
Orange	-0.2357	0.0666 (-0.0004)	-0.2294	0.0809 (-0.0046)	-0.3194	0.0876 (-0.0003)	-0.2064	0.089 (-0.0204)
Hillsborough	0	0 .	0	0 .	0	0 .	0	0 .
Period of the Day *(Dummy variable)*								
Late Evening	-0.4992	0.1907 (-0.0088)	-0.4673	0.2053 (-0.0228)	-0.3224	0.202 (-0.1104)	-0.4036	0.1964 (-0.0399)
P.M. Peak	-0.5321	0.2093 (-0.011)	-0.453	0.2337 (-0.0526)	-0.2687	0.2356 (-0.254)	-0.3649	0.2414 (-0.1307)
Midday	-0.5285	0.2026 (-0.0091)	-0.4727	0.2216 (-0.0329)	-0.3017	0.2236 (-0.1771)	-0.3952	0.2326 (-0.0893)
A.M. Peak	-0.7597	0.2057 (-0.0002)	-0.6837	0.2202 (-0.0019)	-0.5068	0.2289 (-0.0268)	-0.6041	0.2373 (-0.0109)
Early Morning	0	0 .	0	0 .	0	0 .	0	0 .
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)								
LOS F	1.2476	0.2779 (<.0001)	1.1404	0.2909 (<.0001)	1.1418	0.2861 (<.0001)	1.0727	0.2753 (<.0001)
LOS E	1.3146	0.2812 (<.0001)	1.2471	0.2907 (<.0001)	1.2359	0.2871 (<.0001)	1.1935	0.2754 (<.0001)
LOS D	1.2129	0.269 (<.0001)	1.1577	0.2782 (<.0001)	1.16	0.2741 (<.0001)	1.088	0.2621 (<.0001)
LOS C	1.1401	0.2616 (<.0001)	1.1343	0.2681 (<.0001)	1.1233	0.2638 (<.0001)	1.0862	0.2517 (<.0001)
LOS B	0.7408	0.2637 (-0.005)	0.7309	0.2651 (-0.0058)	0.7707	0.2591 (-0.0029)	0.7208	0.2454 (-0.0033)
LOS A	0	0 .	0	0 .	0	0 .	0	0 .
Logarithm of Total Traffic Volume on Major Road (for the Period of the Day)	0.5254	0.0578 (<.0001)	0.5161	0.068 (<.0001)	0.477	0.0707 (<.0001)	0.501	0.0729 (<.0001)
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	0.3826	0.0389 (<.0001)	0.391	0.0444 (<.0001)	0.3655	0.0461 (<.0001)	0.3897	0.0465 (<.0001)
Dispersion	1.0995		1.0973		1.0942		1.0967	

5.2.1.2.3.2 <u>Model Interpretation</u>

The respective model interpretation, per correlation structure, is as follows:

- *Independent Correlation Structure*. As shown in Table 5-14, it can be observed that all predictors were found significant (P-value $\leq$ 0.05). These results indicate from the beginning that this model has a good fit.

  A thorough evaluation of this first submodel for "Rear-End Crashes" seems to suggest that the A.M. Peak period has the least estimated rear-end crash frequency, whereas the Late Evening period is the one with the highest estimated rear-end crash frequency; overall, more rear-end crashes seem to occur as it becomes later in the day. In terms of measures of exposure, the model predicts that both traffic volumes on the Major and Minor roads are linearly related to rear-end crash occurrence; it can be seen that traffic volume on the Major road is associated with a higher number of rear-end crashes than the Minor road. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated rear-end crash frequency, whereas LOS E (significant delay) is the level with the highest estimated rear-end crash frequency for a signalized intersection, followed by LOS F (excessive delay) which is associated with the second highest rear-end crash frequency; on the other side of the scale, and in decreasing level of rear-end crash risk are LOS D (tolerable delay) in third place, followed by LOS C (acceptable delay) and LOS B (minimal delay). Summarizing, the probability for a rear-end crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

- *Autoregressive Correlation Structure*. As shown in Table 5-14, only the P.M. Peak period, fourth level for period of the day, was reported to be insignificant (P-value = 0.0526); in spite of this, the period of the day variable could be kept in this second submodel since the rest of its levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

  A thorough evaluation of this second submodel for "Rear-End Crashes" seems to suggest that the A.M. Peak period has the least estimated rear-end crash frequency, whereas the P.M. Peak period is the one with the highest estimated rear-end crash frequency; overall, more rear-end crashes seem to occur as it becomes later in the day. In terms of measures of exposure, the model predicts that both traffic volumes on the Major and Minor roads are linearly related to rear-end crash occurrence; it can be seen that traffic volume on the Major road is associated with a higher number of rear-end crashes than the Minor road. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated rear-end crash frequency, whereas LOS E (significant delay) is the level with the highest estimated rear-end crash frequency for a signalized intersection; then are LOS D (tolerable delay) and LOS F (excessive delay), having the second and third highest rear-end crash risk, respectively, followed by LOS C (acceptable delay). Summarizing, the probability for a rear-end crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

- *Exchangeable Correlation Structure*. As shown in Table 5-14, the A.M. Peak, Midday and P.M. Peak periods, 3 of the levels for period of the day, were reported to be insignificant (P-values of 0.1771, 0.2540 and 0.1104, respectively); in spite of this, the period of the day variable could be kept in this third submodel since at least one of its levels was significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

  A thorough evaluation of this third submodel for "Rear-End Crashes" seems to suggest that the A.M. Peak period has the least estimated rear-end crash frequency, whereas the P.M. Peak period is the one with the highest estimated rear-end crash frequency; overall, more rear-end crashes seem to occur during the afternoon hours of the day (i.e. from noon to sunset). In terms of measures of exposure, the model predicts that both traffic volumes on the Major and Minor roads are linearly related to rear-end crash occurrence; it can be seen that traffic volume on the Major road is associated with a higher number of rear-end crashes than the Minor road. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated rear-end crash frequency, whereas LOS E (significant delay) is the level with the highest estimated rear-end crash frequency for a signalized intersection; then are LOS D (tolerable delay) and LOS F (excessive delay), having the second and third highest rear-end crash risk, respectively, followed by

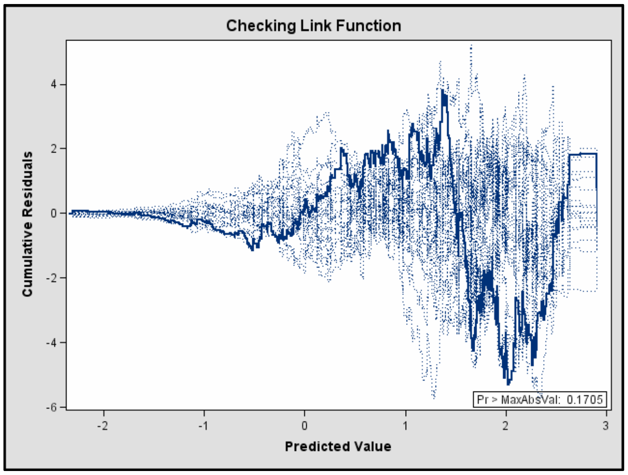
LOS C (acceptable delay). Summarizing, the probability for a rear-end crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

- *Unstructured Correlation Structure*. As shown in Table 5-14, the Midday and P.M. Peak periods, 2 of the levels for period of the day, were reported to be insignificant (P-values of 0.0893 and 0.1307, respectively); in spite of this, the period of the day variable could be kept in this fourth submodel since its other levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

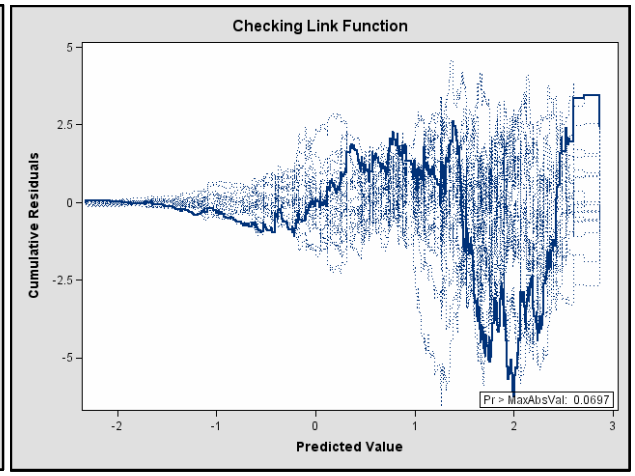
  A thorough evaluation of this fourth submodel for "Rear-End Crashes" seems to suggest that the A.M. Peak period has the least estimated rear-end crash frequency, whereas the P.M. Peak period is the one with the highest estimated rear-end crash frequency; overall, more rear-end crashes seem to occur during the afternoon hours of the day (i.e. from noon to sunset). In terms of measures of exposure, the model predicts that both traffic volumes on the Major and Minor roads are linearly related to rear-end crash occurrence; it can be seen that traffic volume on the Major road is associated with a higher number of rear-end crashes than the Minor road. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated rear-end crash frequency, whereas LOS E (significant delay) is the level with the highest estimated rear-end crash frequency for a signalized intersection; following are LOS D (tolerable delay) and LOS C (acceptable delay), having the second and third highest rear-end crash risk, respectively, followed by LOS F (excessive delay). Summarizing, the probability for a rear-end crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

5.2.1.2.3.3 Model Assessment

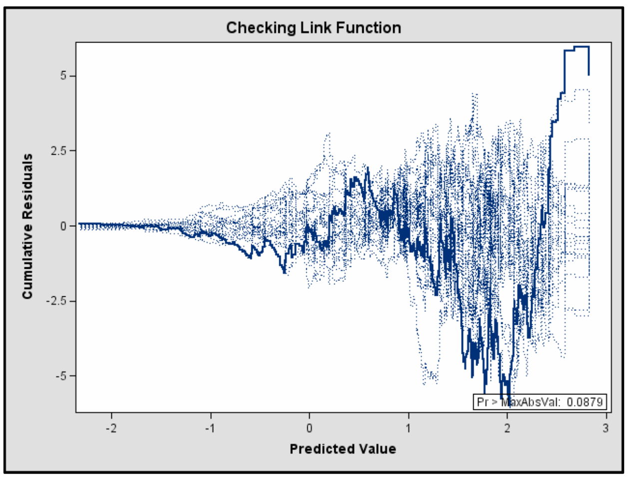
This section summarizes the respective model assessment performed through SAS when analyzing the GEE model designed for "Rear-End Crashes" for the 5 periods of the day. Following are the plots that were generated through SAS' Output Delivery System (ODS) for each of the GEE model's type of working correlation structure (see Figure 5-3). When interpreting the plots, the heavy trend line represents the Cumulative Residuals whereas the light ones represent the simulated curves.



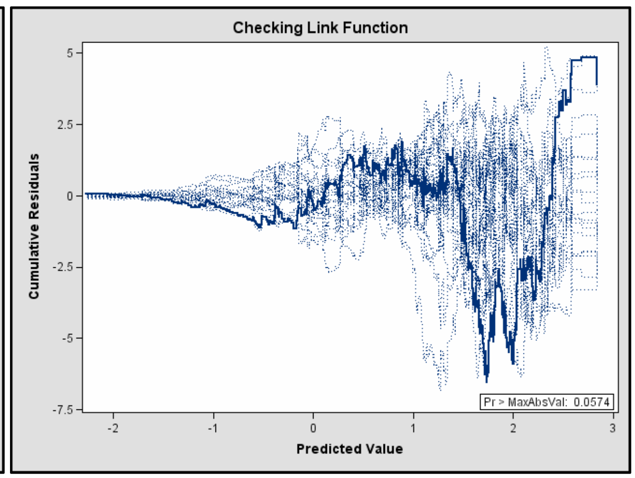
(a)



(b)



(c)



(d)

Figure 5-3: Model Assessment Plots (Cumulative Residuals for GEE Negative Binomial Analysis) for "Rear-End Crashes" at Signalized Intersections (Considering the 5 Periods of the Day): (a) Independent Structure, (b) Autoregressive Structure, (c) Exchangeable Structure and (d) Unstructured

As it can be seen, these plots provide a P-value (Pr>MaxAbsVal) computed through a simulation of 10,000 residual paths; below is a summary of these computed values (see Table 5-15).

Table 5-15: Model Assessment Summary (Cumulative Residuals for GEE Negative Binomial Analysis) for "Rear-End Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
GEE Analysis - Model Assessment Summary					
MODEL 3A: "REAR-END CRASHES" (Considering the 5 Periods of the Day)					
Working Correlation Structure	Assessment Variable	Max. Abs. Value	Replications	Seed	Pr > Max. Abs. Value
INDEPENDENT	Link Function	5.2945	10000	603708000	0.1705
AUTOREGRESSIVE	Link Function	6.2422	10000	603708000	0.0697
EXCHANGEABLE	Link Function	5.9969	10000	603708000	0.0879
UNSTRUCTURED	Link Function	6.5434	10000	603708000	0.0574

In addition, the GEE analysis procedure in SAS provided the model's respective score statistics based on a type III analysis for each covariate in the model. It can be concluded from these results that, among all correlation structures, traffic volume on the minor road is the most significant of all covariates. Following is Table 5-16, which shows the corresponding Score Statistics for each covariate, listed in descending order of significance.

Table 5-16: GEE Score Statistics Summary for "Rear-End Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Score Statistics					
MODEL 3A: "REAR-END CRASHES" (Considering the 5 Periods of the Day)					
COVARIATE	DF	Working Correlation Structure			
		INDEPENDENT	AUTOREGRESSIVE	EXCHANGEABLE	UNSTRUCTURED
		Chi-Sq. (P-value)	Chi-Sq. (P-value)	Chi-Sq. (P-value)	Chi-Sq. (P-value)
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	1	37.11 (<.0001)	38.88 (<.0001)	39.26 (<.0001)	30.54 (<.0001)
Logarithm of Total Traffic Volume on Major Road (for the Period of the Day)	1	26.42 (<.0001)	24.27 (<.0001)	20.34 (<.0001)	22.23 (<.0001)
Period of the Day *(Dummy variable)*	4	21.46 (0.0003)	20.08 (0.0005)	18.44 (0.001)	18.75 (0.0009)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)	5	18.43 (0.0025)	16.46 (0.0056)	14.51 (0.0127)	12.11 (0.0333)
County *(Dummy variable)*	1	7.62 (0.0058)	7.24 (0.0071)	12.77 (0.0004)	3.55 (0.0594)

The last outputs were the working correlation matrices for each structure type (see Table 5-17). As it can be seen, the Independent working correlation matrix was the result of a very naïve analysis procedure; the binary composition of this matrix reflects this fact. Regarding the Autoregressive working correlation matrix, it can be observed that it is characterized by correlation values that decrease over time. The Exchangeable working correlation matrix, however, has a defined and compound symmetry; the exchangeable working correlation had a reported value of 0.2163 for this model. Finally, it can be seen that the Unstructured working correlation matrix's composition has no particular specification.

Table 5-17: GEE Working Correlation Matrices for "Rear-End Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type 3 GEE Analysis - Working Correlation Matrices					
MODEL 3: "REAR-END CRASHES" (Considering the 5 Periods of the Day)					
INDEPENDENT	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0000	0.0000	0.0000	0.0000
A.M. Peak	0.0000	1.0000	0.0000	0.0000	0.0000
Midday	0.0000	0.0000	1.0000	0.0000	0.0000
P.M. Peak	0.0000	0.0000	0.0000	1.0000	0.0000
Late Evening	0.0000	0.0000	0.0000	0.0000	1.0000
AUTOREGRESSIVE	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.2528	0.0639	0.0162	0.0041
A.M. Peak	0.2528	1.0000	0.2528	0.0639	0.0162
Midday	0.0639	0.2528	1.0000	0.2528	0.0639
P.M. Peak	0.0162	0.0639	0.2528	1.0000	0.2528
Late Evening	0.0041	0.0162	0.0639	0.2528	1.0000
EXCHANGEABLE	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.2163	0.2163	0.2163	0.2163
A.M. Peak	0.2163	1.0000	0.2163	0.2163	0.2163
Midday	0.2163	0.2163	1.0000	0.2163	0.2163
P.M. Peak	0.2163	0.2163	0.2163	1.0000	0.2163
Late Evening	0.2163	0.2163	0.2163	0.2163	1.0000
UNSTRUCTURED	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0641	0.0339	0.1579	0.3958
A.M. Peak	0.0641	1.0000	0.4661	0.4034	0.0394
Midday	0.0339	0.4661	1.0000	0.4072	0.0334
P.M. Peak	0.1579	0.4034	0.4072	1.0000	0.1680
Late Evening	0.3958	0.0394	0.0334	0.1680	1.0000

5.2.1.2.3.4 Conclusion

After comparing the results from all correlation structures, it may be concluded that the Independent correlation structure is the best among the other types, since it has the highest P-value (=0.1705). However, as it is known, the Independent correlation structure is the simplest of all, assuming no correlation among any of the observations within the study sample (Hardin and Hilbe, 2003). With the data in hand, and for the purpose of analyzing the LOS-Safety relationship throughout 5 sequentially ordered periods of the day, to choose the Independent correlation structure may not be the best decision.

Based on the previous statements, the Autoregressive correlation structure seems to be the best choice; though this structure has the third highest P-value (=0.0697) (refer to Table 5-15), this structure is preferred over the Exchangeable correlation structure (i.e. second highest P-value) because it assumes that all observations in the sample being studied are temporally correlated (Wang, 2006); this seems to be appropriate since the data for this signalized intersections study is temporal in nature (i.e. it follows a natural order).

An appropriate interpretation of the results from the Autoregressive correlation structure, and recalling the basis of the Cumulative Residuals model assessment method, would be that "out of 10,000 realizations from the null distribution, 6.97% have maximum cumulative residuals greater than 6.2422". In addition, the 5x5 matrix for the Autoregressive correlation structure suggests that the repeated observations (i.e. 5 times) for each signalized intersection will become less correlated as the time-gap increases. Finally, the estimates obtained with the Autoregression correlation structure for each of the covariates in the model were all significant, corroborating the idea that the resulting model fits the data well. Finally, Table 5-18 lists the data summary statistics of the contributing factors included in the final model.

Table 5-18: Data Summary Statistics for Preferred Model for "Rear-End Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Data Summary Statistics					
MODEL 3A: "REAR-END CRASHES" (Considering the 5 Periods of the Day)					
Preferred Correlation Structure: AUTOREGRESSIVE					
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**	**Chi-Sq. (P-value)**
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	5.4	1.4	8.6	1.5	38.88 (<.0001)
Logarithm of Total Traffic Volume on Major Road (for the Period of the Day)	7.0	1.2	8.8	1.2	24.27 (<.0001)
Period of the Day *(Dummy variable)* (Key: Early Morning=1, A.M. Peak=2, Midday=3, P.M. Peak=4, Late Evening=5)	3.0	1	5	1.6	20.08 (0.0005)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9	16.46 (0.0056)
County *(Dummy variable)* (Key: Orange=0, Hillsborough=1)	0.5	0	1	0.7	7.24 (0.0071)
Total Number of 5x5 Clusters (i.e. Signalized Intersections): 149					

5.2.1.2.4 Crash Types: Sideswipe Crashes

5.2.1.2.4.1 Model Building Procedure

This analysis consisted in creating the second of the last set of GEE models (i.e. models for different Crash Types) which shows the frequency of "Sideswipe Crashes" as the dependent variable.

The model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative

importance and significance these covariates could have for the model. Consequently, the best model for "Sideswipe Crashes" included the following independent variables:

- *Period of the day (APERIOD)*: having 5 levels (Early Morning, A.M. Peak, Midday, P.M. Peak and Late Evening); the Early Morning period was treated as the base case.
- *Logarithm of the Cycle Length (LOG_CL)*: Due to its continuous nature, this variable was processed having only one level.
- *LOS for the intersection as a whole (LOS_INT_NUM)*: resulting in 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
- *Total Number of lanes on the Major road (LANMAJ_DAYC1)*: originally with 10 levels, this variable ended up having 5 levels after appropriate combinations were made (4 lanes, 5 with 6 lanes, 7 with 8 lanes, 9 with 10 lanes, and 11 combined with 12 and 14 lanes); the level corresponding to 4 lanes, the lowest number, was treated as the base case.

From the group of covariates just described, only *Period of the day* was treated as dummy variable. It has to be noted that, as in the rest of this study's models, the variable *County* was considered as a dummy variable; however, it had to be removed from this model since it was not found significant after many attempts.

This "aggregate level analysis" (i.e. 5 periods of the day) for each of the final GEE models consisted in decomposing the analysis into four submodels, one per type of working correlation structure to be tried (independent, autoregressive, exchangeable and unstructured). The working correlation structure of all submodel types was based on a total of 149 clusters (number of signalized intersections), each having 5 as their maximum and minimum dimensions; this resulted in a 5x5 working correlation matrix per submodel. A total of 745 observations were read and used for each submodel, number that is five times the total number of signalized intersections in the study sample; this can be translated into the following: 149 intersections * 5 periods of the day = 745 observations. Tables 5-19 and 5-20 provide a summary of the model, its composition and respective model-based estimates for each correlation structure.

Table 5-19: General Model Information for "Sideswipe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS	
Type III GEE Analysis - Score Statistics	
MODEL 4A: "SIDESWIPE CRASHES" (Considering the 5 Periods of the Day)	
Number of Clusters (i.e. Number of Signalized Intersections)	149
Cluster Size (i.e. Number of Continuous Periods of the Day)	5
Number of Observations	745
Number of Sideswipe/Total Crashes	328/5,532

Table 5-20: Model-Based Standard Error Estimates for "Sideswipe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS								
Type 3 GEE Analysis - Model-Based Standard Error Estimates								
MODEL 4A: "SIDESWIPE CRASHES" (Considering the 5 Periods of the Day)								
PARAMETER	Working Correlation Structure							
	INDEPENDENT		AUTOREGRESSIVE		EXCHANGEABLE		UNSTRUCTURED	
	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)
Intercept	-7.5063	1.589 (<.0001)	-7.4347	1.6027 (<.0001)	-7.4376	1.6042 (<.0001)	-7.4513	1.6278 (<.0001)

PARAMETER	Working Correlation Structure							
	INDEPENDENT		AUTOREGRESSIVE		EXCHANGEABLE		UNSTRUCTURED	
	Coef.	**S.E. (P-value)**	**Coef.**	**S.E. (P-value)**	**Coef.**	**S.E. (P-value)**	**Coef.**	**S.E. (P-value)**
Period of the Day *(Dummy variable)*								
Late Evening	1.1979	0.4156 (-0.0039)	1.1895	0.4145 (-0.0041)	1.2333	0.4102 (-0.0026)	1.1742	0.4211 (-0.0053)
P.M. Peak	1.1768	0.4272 (-0.0059)	1.1948	0.4272 (-0.0052)	1.2451	0.4249 (-0.0034)	1.1799	0.4252 (-0.0055)
Midday	1.1353	0.4177 (-0.0066)	1.1416	0.4167 (-0.0061)	1.1925	0.4137 (-0.0039)	1.1257	0.404 (-0.0053)
A.M. Peak	1.0723	0.4278 (-0.0122)	1.0851	0.4225 (-0.0102)	1.1359	0.4252 (-0.0076)	1.0681	0.4237 (-0.0117)
Early Morning	0	0 .	0	0 .	0	0 .	0	0 .
Logarithm of the Cycle Length (for the Period of the Day)	0.7131	0.3124 (-0.0224)	0.7061	0.3172 (-0.026)	0.7083	0.3196 (-0.0267)	0.7303	0.3249 (-0.0246)
LEVEL-OF-SERVICE **for the Intersection as a whole** (for the Period of the Day)								
LOS F	2.1398	0.7561 (-0.0047)	2.0723	0.7513 (-0.0058)	2.005	0.7354 (-0.0064)	2.0073	0.7397 (-0.0067)
LOS E	1.6137	0.7697 (-0.036)	1.5461	0.7647 (-0.0432)	1.4738	0.7489 (-0.0491)	1.4952	0.7526 (-0.047)
LOS D	1.5874	0.7556 (-0.0357)	1.562	0.749 (-0.037)	1.4841	0.7325 (-0.0428)	1.5004	0.7357 (-0.0414)
LOS C	1.3082	0.7525 (-0.0821)	1.2785	0.7449 (-0.0861)	1.2217	0.7269 (-0.0928)	1.2432	0.729 (-0.0881)
LOS B	0.5343	0.7714 (-0.4885)	0.4813	0.7627 (-0.528)	0.461	0.7413 (-0.5341)	0.3913	0.7502 (-0.602)
LOS A	0	0 .	0	0 .	0	0 .	0	0 .
Total Number of Lanes **on Major Road**								
11 $\leq$ lanes $\leq$ 14	1.1745	0.3794 (-0.002)	1.1766	0.3946 (-0.0029)	1.1634	0.4053 (-0.0041)	1.1367	0.4064 (-0.0052)
9 lanes or 10 lanes	1.0229	0.3726 (-0.006)	1.0273	0.3875 (-0.008)	1.0397	0.3974 (-0.0089)	1.0072	0.3986 (-0.0115)
7 lanes or 8 lanes	0.6621	0.3669 (-0.0711)	0.6623	0.3816 (-0.0826)	0.6767	0.3913 (-0.0837)	0.6366	0.3926 (-0.1049)
5 lanes or 6 lanes	0.0833	0.4207 (-0.8431)	0.085	0.438 (-0.8461)	0.09	0.4487 (-0.841)	0.053	0.4524 (-0.9068)
4 lanes	0	0 .	0	0 .	0	0 .	0	0 .
Dispersion	1.032		1.0301		1.0252		1.0336	

5.2.1.2.4.2 Model Interpretation

The respective model interpretation, per correlation structure, is as follows:

- *Independent Correlation Structure*. As shown in Table 5-20, LOS B, LOS C, as well as having between 5 and 8 as total number of lanes on the Major road were reported to be insignificant (P-values of 0.4885, 0.0821, 0.8431 and 0.0711, respectively); in spite of this, the respective variables could be kept in this first submodel since the rest of their levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

  A thorough evaluation of this first submodel for "Sideswipe Crashes" seems to suggest that the A.M. Peak period has the least estimated sideswipe crash frequency, whereas the Late Evening period is the one with the highest estimated sideswipe crash frequency; overall, more sideswipe crashes seem to occur as it becomes later in the day. Regarding operational measures, a signalized intersection's Cycle Length was found to be significant by denoting an almost linear relationship with the number of sideswipe crashes; this could be explained by saying that the longer the cycle length there is a bigger time frame open for sideswipe crash occurrence. It can also be observed that the higher the number of lanes on the Major road, the higher the probability for a sideswipe crash to happen. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated sideswipe crash frequency; then, going further along the scale, LOS F (excessive delay) is the level with the highest estimated sideswipe crash frequency for a signalized intersection. Summarizing, the probability for a sideswipe crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

- *Autoregressive Correlation Structure*. As shown in Table 5-20, LOS B and LOS C, as well as a total number of 5 or 6 lanes on the Major road were reported to be insignificant (P-values of 0.5280, 0.0861, 0.8461 and 0.0826, respectively); in spite of this, the respective variables could be kept in this second submodel since the rest of their levels were significant at the 0.05 confidence level or because they were necessary for being consistent throughout all models. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

  A thorough evaluation of this second submodel for "Sideswipe Crashes" seems to suggest that the A.M. Peak period has the least estimated sideswipe crash frequency, whereas the P.M. Peak period is the one with the highest estimated sideswipe crash frequency; overall, more sideswipe crashes seem to occur as it becomes later in the day. Regarding operational measures, a signalized intersection's Cycle Length was found to be significant by denoting an almost linear relationship with the number of sideswipe crashes; this could be explained by saying that the longer the cycle length there is a bigger time frame open for sideswipe crash occurrence. It can also be observed that the higher the number of lanes on the Major road, the higher the probability for a sideswipe crash to happen. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated sideswipe crash frequency; then, going further along the scale, the sideswipe crash trend increases until reaching LOS E (significant delay) where it has a very slight decrease in number of this particular crash type. Finally, LOS F (excessive delay) appears as the level with the highest estimated sideswipe crash frequency for a signalized intersection. Summarizing, the probability for a sideswipe crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

- *Exchangeable Correlation Structure*. As shown in Table 5-20, LOS B, LOS C, as well as having between 5 and 8 as total number of lanes on the Major road were reported to be insignificant (P-values of 0.5341, 0.0928, 0.8410 and 0.0837, respectively); in spite of this, the respective variables could be kept in this third submodel since the rest of their levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

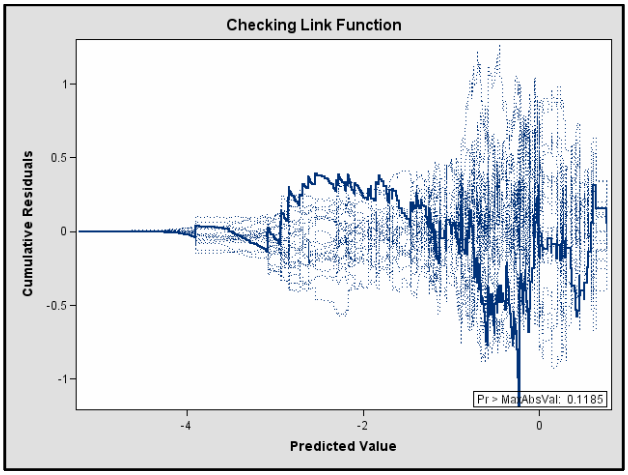
  A thorough evaluation of this third submodel for "Sideswipe Crashes" seems to suggest that the A.M. Peak period has the least estimated sideswipe crash frequency, whereas the P.M. Peak period is the one with the highest estimated sideswipe crash frequency; overall, more sideswipe crashes seem to occur during peak traffic hours and as it becomes later in the day. Regarding operational measures, a signalized intersection's Cycle Length was found to be significant by denoting an almost linear relationship with the number of sideswipe crashes; this could be explained by saying that the longer the cycle length there is a bigger time frame open for sideswipe crash occurrence. It can also be observed that the higher the number of lanes on the Major road, the higher the probability for a sideswipe crash to happen. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated sideswipe crash frequency; then, going further along the scale, the sideswipe crash trend increases until reaching LOS E (significant delay) where it has a very slight decrease in number of this particular crash type. Finally, LOS F (excessive delay) appears as the level with the highest estimated sideswipe crash frequency for a signalized intersection. Summarizing, the probability for a sideswipe crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

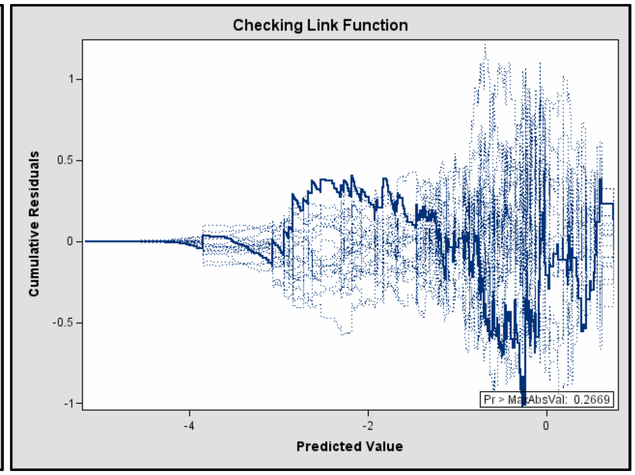
- *Unstructured Correlation Structure*. As shown in Table 5-20, LOS B, LOS C, as well as having between 5 and 8 as total number of lanes on the Major road were reported to be insignificant (P-values of 0.6020, 0.0881, 0.9068 and 0.1049, respectively); in spite of this, the respective variables could be kept in this fourth submodel since the rest of their levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

  A thorough evaluation of this fourth submodel for "Sideswipe Crashes" seems to suggest that the A.M. Peak period has the least estimated sideswipe crash frequency, whereas the P.M. Peak period is the one with the highest estimated sideswipe crash frequency; overall, more sideswipe crashes seem to occur during peak traffic hours and as it becomes later in the day. Regarding operational measures, a signalized intersection's Cycle Length was found to be significant by denoting an almost linear relationship with the number of sideswipe crashes; this could be explained by saying that the longer the cycle length there is a bigger time frame open for sideswipe crash occurrence. It can also be observed that the higher the number of lanes on the Major road, the higher the probability for a sideswipe crash to happen. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated sideswipe crash frequency; then, going further along the scale, the sideswipe crash trend increases until reaching LOS E (significant delay) where it has a sudden decrease in number of this particular crash type. Finally, LOS F (excessive delay) appears as the level with the highest estimated sideswipe crash frequency for a signalized intersection. Summarizing, the probability for a sideswipe crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

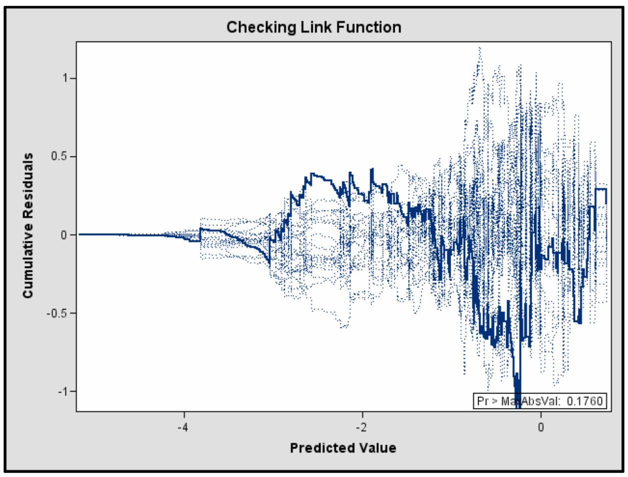
5.2.1.2.4.3 Model Assessment

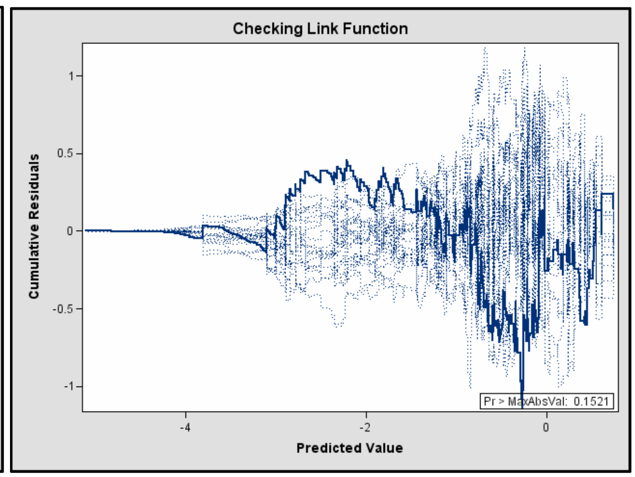
This section summarizes the respective model assessment performed through SAS when analyzing the GEE model designed for "Sideswipe Crashes" for the 5 periods of the day. Following are the plots that were generated through SAS' Output Delivery System (ODS) for each of the GEE model's type of working correlation structure (see Figure 5-4). When interpreting the plots, the heavy trend line represents the Cumulative Residuals whereas the light ones represent the simulated curves.





(a) (b)





(c) (d)

Figure 5-4: Model Assessment Plots (Cumulative Residuals for GEE Negative Binomial Analysis) for "Sideswipe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day): (a) Independent Structure, (b) Autoregressive Structure, (c) Exchangeable Structure and (d) Unstructured

As it can be seen, these plots provide a P-value (Pr>MaxAbsVal) computed through a simulation of 10,000 residual paths; below is a summary of these computed values (see Table 5-21).

Table 5-21: Model Assessment Summary (Cumulative Residuals for GEE Negative Binomial Analysis) for "Sideswipe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
GEE Analysis - Model Assessment Summary					
MODEL 4A: "SIDESWIPE CRASHES" (Considering the 5 Periods of the Day)					
Working Correlation Structure	Assessment Variable	Max. Abs. Value	Replications	Seed	Pr > Max. Abs. Value
INDEPENDENT	Link Function	1.1900	10000	603708000	0.1185
AUTOREGRESSIVE	Link Function	1.0221	10000	603708000	0.2669
EXCHANGEABLE	Link Function	1.1086	10000	603708000	0.1760
UNSTRUCTURED	Link Function	1.1452	10000	603708000	0.1521

In addition, the GEE analysis procedure in SAS provided the model's respective score statistics based on a type III analysis for each covariate in the model. It can be concluded from these results that, among all correlation structures, the LOS for the intersection as a whole is the most significant of all covariates. Following is Table 5-22, which shows the corresponding Score Statistics for each covariate, listed in descending order of significance.

Table 5-22: GEE Score Statistics Summary for Preferred Model for "Sideswipe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Score Statistics					
MODEL 4A: "SIDESWIPE CRASHES" (Considering the 5 Periods of the Day)					
COVARIATE	DF	Working Correlation Structure			
		INDEPENDENT	AUTOREGRESSIVE	EXCHANGEABLE	UNSTRUCTURED
		Chi-Sq. (P-value)	Chi-Sq. (P-value)	Chi-Sq. (P-value)	Chi-Sq. (P-value)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)	5	23.38 (0.0003)	23.28 (0.0003)	21.55 (0.0006)	23.01 (0.0003)
Total Number of Lanes on Major Road	4	18.33 (0.0011)	18.78 (0.0009)	18.57 (0.001)	17.85 (0.0013)
Period of the Day *(Dummy variable)*	4	13.55 (0.0089)	13.79 (0.008)	14.13 (0.0069)	12.88 (0.0119)
Logarithm of the Cycle Length (for the Period of the Day)	1	4.27 (0.0388)	3.91 (0.0479)	3.8 (0.0511)	3.57 (0.0588)

The last outputs were the working correlation matrices for each structure type (see Table 5-23). As it can be seen, the Independent working correlation matrix was the result of a very naïve analysis procedure; the binary composition of this matrix reflects this fact. Regarding the Autoregressive working correlation matrix, it can be observed that it is characterized by correlation values that decrease over time. The Exchangeable working correlation matrix, however, has a defined and compound symmetry; the exchangeable working correlation had a reported value of 0.0467 for this model. Finally, it can be seen that the Unstructured working correlation matrix's composition has no particular specification.

Table 5-23: GEE Working Correlation Matrices for "Sideswipe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type 3 GEE Analysis - Working Correlation Matrices					
MODEL 4A: "SIDESWIPE CRASHES" (Considering the 5 Periods of the Day)					
INDEPENDENT	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0000	0.0000	0.0000	0.0000
A.M. Peak	0.0000	1.0000	0.0000	0.0000	0.0000
Midday	0.0000	0.0000	1.0000	0.0000	0.0000
P.M. Peak	0.0000	0.0000	0.0000	1.0000	0.0000
Late Evening	0.0000	0.0000	0.0000	0.0000	1.0000
AUTOREGRESSIVE	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0544	0.0030	0.0002	0.0000
A.M. Peak	0.0544	1.0000	0.0544	0.0030	0.0002
Midday	0.0030	0.0544	1.0000	0.0544	0.0030
P.M. Peak	0.0002	0.0030	0.0544	1.0000	0.0544
Late Evening	0.0000	0.0002	0.0030	0.0544	1.0000
EXCHANGEABLE	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0467	0.0467	0.0467	0.0467
A.M. Peak	0.0467	1.0000	0.0467	0.0467	0.0467
Midday	0.0467	0.0467	1.0000	0.0467	0.0467
P.M. Peak	0.0467	0.0467	0.0467	1.0000	0.0467
Late Evening	0.0467	0.0467	0.0467	0.0467	1.0000
UNSTRUCTURED	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0538	0.1281	0.0392	-0.0490
A.M. Peak	0.0538	1.0000	0.0470	0.0842	-0.0628
Midday	0.1281	0.0470	1.0000	0.0501	0.1062
P.M. Peak	0.0392	0.0842	0.0501	1.0000	0.0822
Late Evening	-0.0490	-0.0628	0.1062	0.0822	1.0000

5.2.1.2.4.4 Conclusion

After comparing the results from all correlation structures, it may be concluded that the Autoregressive correlation structure is the best among the other types, since it has the highest P-value (=0.2669) (refer to Table 5-21). In addition, this structure assumes that all observations in the sample being studied are temporally correlated (Wang, 2006); this seems to be appropriate since the data for this signalized intersections study is temporal in nature (i.e. it follows a natural order).

An appropriate interpretation of the results from the Autoregressive correlation structure, and recalling the basis of the Cumulative Residuals model assessment method, would be that "out of 10,000 realizations from the null distribution, 26.69% have maximum cumulative residuals greater than 1.0221". In addition, the 5x5 matrix for the Autoregressive correlation structure suggests that the repeated observations (i.e. 5 times) for each signalized intersection will become less correlated as the time-gap increases. Finally, the estimates obtained with the Autoregression correlation structure for each of the covariates in the model were all significant, corroborating the idea that the resulting model fits the data well. Finally, Table 5-24 lists the data summary statistics of the contributing factors included in the final model.

Table 5-24: Data Summary Statistics for Preferred Model for "Sideswipe Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Data Summary Statistics					
Model 4A: "SIDESWIPE CRASHES" (Considering the 5 Periods of the Day)					
Preferred Correlation Structure: AUTOREGRESSIVE					
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**	**Chi-Sq. (P-value)**
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9	23.28 (0.0003)
Total Number of Lanes on Major Road	7.9	4	14	2.1	18.78 (0.0009)
Period of the Day *(Dummy variable)* (Key: Early Morning=1, A.M. Peak=2, Midday=3, P.M. Peak=4, Late Evening=5)	3.0	1	5	1.6	13.79 (0.008)
Logarithm of the Cycle Length (for the Period of the Day) (seconds)	4.6	1.6	5.3	0.4	3.91 (0.0479)
Total Number of 5x5 Clusters (i.e. Signalized Intersections): 149					

5.2.1.2.5 Crash Types: Head-On Crashes

5.2.1.2.5.1 Model Building Procedure

This analysis consisted in creating the third of the last set of GEE models (i.e. models for different Crash Types) which shows the frequency of "Head-On Crashes" as the dependent variable.

The model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for "Head-On" included the following independent variables:

- *County (COUNTY)*: having 2 levels (Orange and Hillsborough); Hillsborough County was treated as the base case.
- *Period of the day (APERIOD)*: having 5 levels (Early Morning, A.M. Peak, Midday, P.M. Peak and Late Evening); the Early Morning period was treated as the base case.
- *LOS for the intersection as a whole (LOS_INT_NUM)*: having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.

From the group of covariates just described, only "County" and "Period of the Day" were treated as dummy variables. In addition, this was the first "LOS only" model obtained, since no other covariates had a good fit in the model containing LOS.

This "aggregate level analysis" (i.e. 5 periods of the day) for each of the final GEE models consisted in decomposing the analysis into four submodels, one per type of working correlation structure to be tried (independent, autoregressive, exchangeable and unstructured). The working correlation structure of all submodel types was based on a total of 149 clusters (number of signalized intersections), each having 5 as their maximum and minimum dimensions; this resulted in a 5x5 working correlation matrix per submodel. A total of 745 observations were read and used for each submodel, number that is five times the total number of signalized intersections in the study sample; this can be translated into the following: 149 intersections * 5 periods of the day = 745 observations. Tables 5-25 and 5-26 provide a summary of the model, its composition and respective model-based estimates for each correlation structure.

Table 5-25: General Model Information for "Head-On Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS	
Type III GEE Analysis - Score Statistics	
MODEL 5A: "HEAD-ON CRASHES" (Considering the 5 Periods of the Day)	
Number of Clusters (i.e. Number of Signalized Intersections)	149
Cluster Size (i.e. Number of Continuous Periods of the Day)	5
Number of Observations	745
Number of Head-On/Total Crashes	112/5,532

Table 5-26: Model-Based Standard Error Estimates for "Head-On Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS								
Type III GEE Analysis - Model-Based Standard Error Estimates								
MODEL 5A: "HEAD-ON CRASHES" (Considering the 5 Periods of the Day)								
PARAMETER	Working Correlation Structure							
	INDEPENDENT		AUTOREGRESSIVE		EXCHANGEABLE		UNSTRUCTURED	
	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)
Intercept	-4.6246	1.146 (<.0001)	-4.6122	1.1406 (<.0001)	-4.6571	1.1594 (<.0001)	-4.6265	1.1603 (<.0001)
County *(Dummy variable)*								
Orange	-0.5356	0.2299 (-0.0198)	-0.5355	0.2314 (-0.0206)	-0.5345	0.2268 (-0.0185)	-0.538	0.2241 (-0.0164)
Hillsborough	0	0 .	0	0 .	0	0 .	0	0 .
Period of the Day *(Dummy variable)*								
Late Evening	1.4888	0.5782 (-0.01)	1.4906	0.578 (-0.0099)	1.4856	0.5799 (-0.0104)	1.4307	0.5451 (-0.0087)
P.M. Peak	1.4627	0.591 (-0.0133)	1.4652	0.5911 (-0.0132)	1.4578	0.5921 (-0.0138)	1.4468	0.5985 (-0.0156)
Midday	1.2934	0.5876 (-0.0277)	1.2956	0.5876 (-0.0275)	1.2887	0.5892 (-0.0287)	1.2867	0.596 (-0.0309)
A.M. Peak	0.9225	0.6081 (-0.1292)	0.9245	0.6065 (-0.1274)	0.9181	0.6095 (-0.132)	0.9067	0.6147 (-0.1402)
Early Morning	0	0 .	0	0 .	0	0 .	0	0 .
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)								
LOS F	2.2098	1.0747 (-0.0398)	2.1952	1.0698 (-0.0402)	2.2459	1.0877 (-0.0389)	2.213	1.0861 (-0.0416)
LOS E	1.3616	1.1397 (-0.2322)	1.3392	1.1354 (-0.2382)	1.4163	1.1508 (-0.2184)	1.4633	1.1451 (-0.2013)
LOS D	2.2037	1.0696 (-0.0394)	2.1904	1.0643 (-0.0396)	2.2388	1.0833 (-0.0388)	2.2697	1.0835 (-0.0362)
LOS C	1.5561	1.0718 (-0.1465)	1.5381	1.0664 (-0.1492)	1.5973	1.0857 (-0.1412)	1.5813	1.0887 (-0.1464)
LOS B	1.0981	1.0777 (-0.3082)	1.0871	1.0718 (-0.3104)	1.1302	1.0925 (-0.3009)	1.1056	1.0915 (-0.3111)
LOS A	0	0 .	0	0 .	0	0 .	0	0 .
Dispersion	1.0436		1.043		1.0448		1.0406	

5.2.1.2.5.2 Model Interpretation

The respective model interpretation, per correlation structure, is as follows:

- *Independent Correlation Structure*. As shown in Table 5-26, the A.M. Peak period, as well as LOS B, LOS C and LOS E, all were reported to be insignificant (P-values of 0.1292, as well as 0.3082, 0.1465 and 0.2322, respectively); in spite of this, the respective variables could still be kept in this first submodel since the rest of their levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

  A thorough evaluation of this first submodel for "Head-On Crashes" seems to suggest that the A.M. Peak period has the least estimated head-on crash frequency, whereas the Late Evening period is the one with the highest estimated head-on crash frequency; overall, more head-on crashes seem to occur as it becomes later in the day. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated head-on crash frequency; then, going further along the scale, the head-on crash trend increases until reaching LOS E (significant delay) where it has a sudden decrease in number of this particular crash type. Finally, LOS F (excessive delay) appears as the level with the highest estimated head-on crash frequency for at a signalized intersection, a number almost equal to that for LOS D (tolerable delay). Summarizing, the probability for a head-on crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

- *Autoregressive Correlation Structure*. As shown in Table 5-26, the A.M. Peak period, as well as LOS B, LOS C and LOS E, all were reported to be insignificant (P-values of 0.1274, as well as 0.3104, 0.1492 and 0.2382, respectively); in spite of this, the respective variables could still be kept in this second submodel since the rest of their levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

  A thorough evaluation of this second submodel for "Head-On Crashes" seems to suggest that the A.M. Peak period has the least estimated head-on crash frequency, whereas the Late Evening period is the one with the highest estimated head-on crash frequency; overall, more head-on crashes seem to occur as it becomes later in the day. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated head-on crash frequency; then, going further along the scale, the head-on crash trend increases until reaching LOS E (significant delay) where it has a sudden decrease in number of this particular crash type. Finally, LOS F (excessive delay) appears as the level with the highest estimated head-on crash frequency for a signalized intersection, a number almost equal to that of LOS D (tolerable delay). Summarizing, the probability for a head-on crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

- *Exchangeable Correlation Structure*. As shown in Table 5-26, the A.M. Peak period, as well as LOS B, LOS C and LOS E, all were reported to be insignificant (P-values of 0.1320, as well as 0.3009, 0.1412 and 0.2184, respectively); in spite of this, the respective variables could still be kept in this third submodel since the rest of their levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

  A thorough evaluation of this third submodel for "Head-On Crashes" seems to suggest that the A.M. Peak period has the least estimated head-on crash frequency, whereas the Late Evening period is the one with the highest estimated head-on crash frequency; overall, more head-on crashes seem to occur as it becomes later in the day. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated head-on crash frequency; then, going further along the scale, the head-on crash trend increases until reaching LOS E (significant delay) where it has a sudden decrease in number of this particular crash type. Finally, LOS F (excessive delay) appears as the level with the highest estimated head-on crash frequency for a signalized intersection, a number almost equal to that for LOS D (tolerable delay). Summarizing, the probability for a head-on crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

- *Unstructured Correlation Structure*. As shown in Table 5-26, the A.M. Peak period, as well as LOS B, LOS C and LOS E, all were reported to be insignificant

(P-values of 0.1402, as well as 0.3111, 0.1464 and 0.2013, respectively); in spite of this, the respective variables could still be kept in this fourth submodel since the rest of their levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

A thorough evaluation of this fourth submodel for "Head-On Crashes" seems to suggest that the A.M. Peak period has the least estimated head-on crash frequency, whereas both the P.M. Peak period and Late Evening periods are the ones with the first and second highest estimated head-on crash frequencies, respectively; overall, more head-on crashes seem to occur as it becomes later in the day. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated head-on crash frequency, whereas LOS D (tolerable delay) is the level with the highest estimated head-on crash frequency for a signalized intersection; then are LOS F (excessive delay), LOS C (acceptable delay) and LOS E (significant delay), which have the second, third and fourth highest head-on crash risk. Summarizing, the probability for a head-on crash to occur at a signalized intersection is highest while at a tolerable delay level.

5.2.1.2.5.3 Model Assessment

This section summarizes the respective model assessment performed through SAS when analyzing the GEE model designed for "Head-On Crashes" for the 5 periods of the day. Following are the plots that were generated through SAS' Output Delivery System (ODS) for each of the GEE model's type of working correlation structure (see Figure 5-5). When interpreting the plots, the heavy trend line represents the Cumulative Residuals whereas the light ones represent the simulated curves.

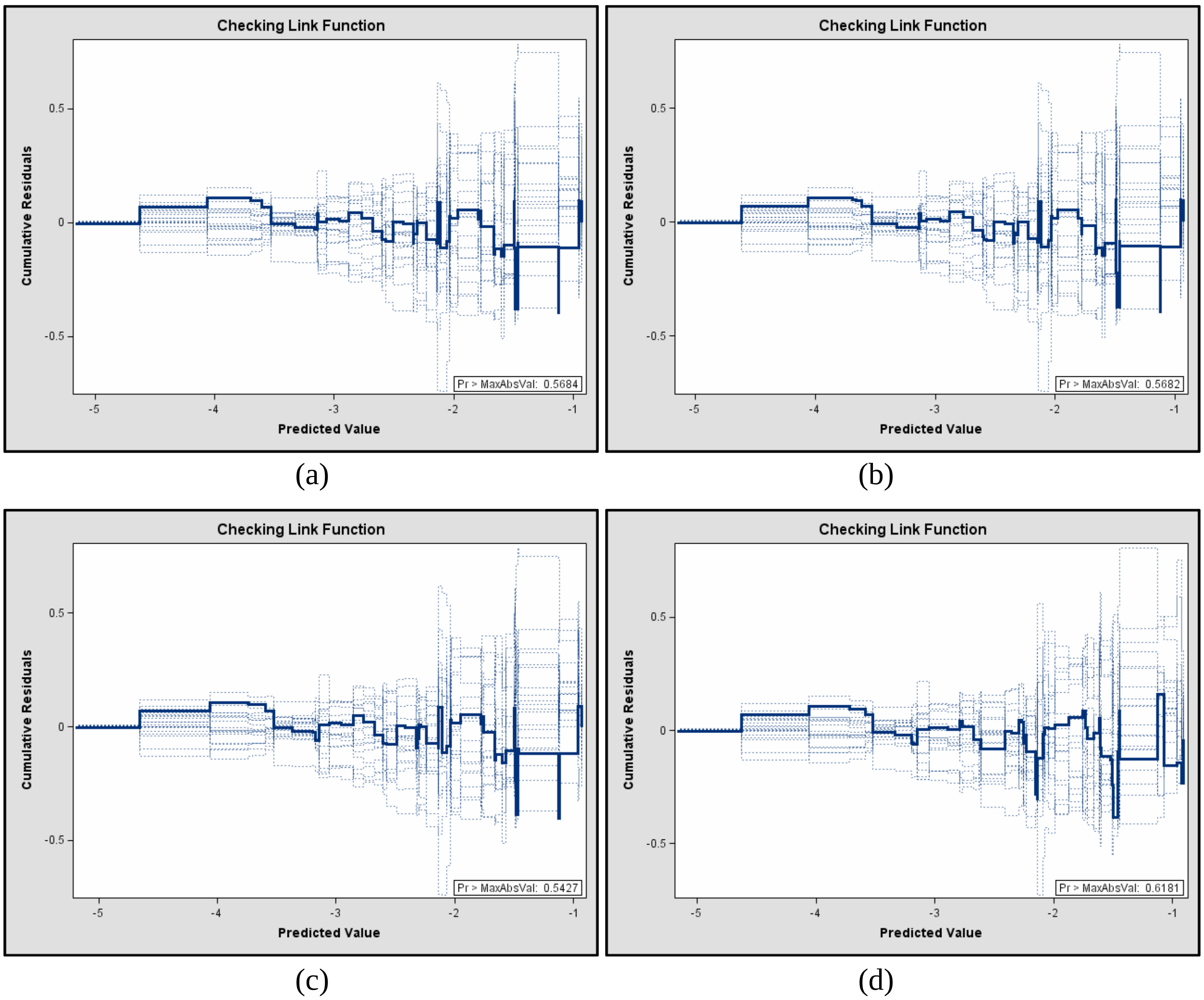


Figure 5-5: Model Assessment Plots (Cumulative Residuals for GEE Negative Binomial Analysis) for "Head-On Crashes" at Signalized Intersections (Considering the 5 Periods of the Day): (a) Independent Structure, (b) Autoregressive Structure, (c) Exchangeable Structure and (d) Unstructured

As it can be seen, these plots provide a P-value (Pr>MaxAbsVal) computed through a simulation of 10,000 residual paths; below is a summary of these computed values (see Table 5-27).

Table 5-27: Model Assessment Summary (Cumulative Residuals for GEE Negative Binomial Analysis) for "Head-On Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
GEE Analysis - Model Assessment Summary					
MODEL 5A: "HEAD-ON CRASHES" (Considering the 5 Periods of the Day)					
Working Correlation Structure	Assessment Variable	Max. Abs. Value	Replications	Seed	Pr > Max. Abs. Value
INDEPENDENT	Link Function	0.3950	10000	603708000	0.5684
AUTOREGRESSIVE	Link Function	0.3920	10000	603708000	0.5682
EXCHANGEABLE	Link Function	0.4031	10000	603708000	0.5427
UNSTRUCTURED	Link Function	0.3844	10000	603708000	0.6181

In addition, the GEE analysis procedure in SAS provided the model's respective score statistics based on a type III analysis for each covariate in the model. It can be concluded from these results that, among all correlation structures, the LOS for the intersection as a whole is the most significant of all covariates. Following is Table 5-28, which shows the corresponding Score Statistics for each covariate, listed in descending order of significance.

Table 5-28: GEE Score Statistics Summary for "Head-On Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS						
Type III GEE Analysis - Score Statistics						
MODEL 5A: "HEAD-ON CRASHES" (Considering the 5 Periods of the Day)						
COVARIATE	DF	Working Correlation Structure				
		INDEPENDENT	AUTOREGRESSIVE	EXCHANGEABLE	UNSTRUCTURED
		Chi-Sq. (P-value)	Chi-Sq. (P-value)	Chi-Sq. (P-value)	Chi-Sq. (P-value)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)	5	18.76 (-0.0021)	18.64 (-0.0022)	18.89 (-0.002)	19.74 (-0.0014)
Period of the Day *(Dummy variable)*	4	14.34 (-0.0063)	14.48 (-0.0059)	14.29 (-0.0064)	14 (-0.0073)
County *(Dummy variable)*	1	6.47 (-0.011)	6.47 (-0.011)	6.47 (-0.011)	6.29 (-0.0121)

The last outputs were the working correlation matrices for each structure type (see Table 5-29). As it can be seen, the Independent working correlation matrix was the result of a very naïve analysis procedure; the binary composition of this matrix reflects this fact. Regarding the Autoregressive working correlation matrix, it can be observed that it is characterized by correlation values that decrease over time. The Exchangeable working correlation matrix, however, has a defined and compound symmetry; the exchangeable working correlation had a reported value of -0.0079 for this model. Finally, it can be seen that the Unstructured working correlation matrix's composition has no particular specification.

Table 5-29: GEE Working Correlation Matrices for "Head-On Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Working Correlation Matrices					
MODEL 5A: "HEAD-ON CRASHES" (Considering the 5 Periods of the Day)					
INDEPENDENT	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0000	0.0000	0.0000	0.0000
A.M. Peak	0.0000	1.0000	0.0000	0.0000	0.0000
Midday	0.0000	0.0000	1.0000	0.0000	0.0000
P.M. Peak	0.0000	0.0000	0.0000	1.0000	0.0000
Late Evening	0.0000	0.0000	0.0000	0.0000	1.0000
AUTOREGRESSIVE	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0084	0.0001	0.0000	0.0000
A.M. Peak	0.0084	1.0000	0.0084	0.0001	0.0000
Midday	0.0001	0.0084	1.0000	0.0084	0.0001
P.M. Peak	0.0000	0.0001	0.0084	1.0000	0.0084
Late Evening	0.0000	0.0000	0.0001	0.0084	1.0000
EXCHANGEABLE	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	-0.0079	-0.0079	-0.0079	-0.0079
A.M. Peak	-0.0079	1.0000	-0.0079	-0.0079	-0.0079
Midday	-0.0079	-0.0079	1.0000	-0.0079	-0.0079
P.M. Peak	-0.0079	-0.0079	-0.0079	1.0000	-0.0079
Late Evening	-0.0079	-0.0079	-0.0079	-0.0079	1.0000
UNSTRUCTURED	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	-0.0668	-0.0815	-0.0918	0.1533
A.M. Peak	-0.0668	1.0000	0.1468	-0.1028	-0.0444
Midday	-0.0815	0.1468	1.0000	-0.0731	0.0507
P.M. Peak	-0.0918	-0.1028	-0.0731	1.0000	0.0298
Late Evening	0.1533	-0.0444	0.0507	0.0298	1.0000

5.2.1.2.5.4 Conclusion

After comparing the results from all correlation structures, it may be concluded that the Unstructured correlation structure is the best among the other types, since it has the highest P-value (=0.6181) (refer to Table 5-27). In addition, this is the most general of all structures and it imposes no restriction with regards to the order of the studied observations (Hardin and Hilbe, 2003). Recalling the Unstructured working correlation matrix obtained for the "Head-On Crashes"' model (refer to Table 5-29), it can be observed that the Early Morning and Late Evening periods, periods of the day with similar traffic characteristics (i.e. lowest traffic volumes, "late" hours, etc.), have the highest correlation (=0.1533) which justifies the choice of this structure. Furthermore, as it was mentioned before, the study conducted by Wang et al. (2009) also supports the choice of the Unstructured correlation structure. Apart from this, it has to be noted that the Autoregressive correlation structure was also considered in the selection process of the best correlation structure; however, this structure was disregarded for two reasons: 1) the Autoregressive working correlation matrix did not report any correlation for the later periods of the day (refer to Table 5-29), and 2) the Autoregressive structure did not count with stronger reasons for being chosen, when compared to the Unstructured correlation structure.

An appropriate interpretation of the results from the Unstructured correlation structure, and recalling the basis of the Cumulative Residuals model assessment method, would be that "out of 10,000 realizations from the null distribution, 61.81% have maximum cumulative residuals greater than 0.3844". Finally, the estimates obtained with the Unstructured correlation structure for each of the covariates in the model were significant, corroborating the idea that the resulting model fits the data well. Finally, Table 5-30 lists the data summary statistics of the contributing factors included in the final model.

Table 5-30: Data Summary Statistics for Preferred Model for "Head-On Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Data Summary Statistics					
MODEL 5A: "HEAD-ON CRASHES" (Considering the 5 Periods of the Day)					
Preferred Correlation Structure: UNSTRUCTURED					
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**	**Chi-Sq. (P-value)**
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9	19.74 (0.0014)
Period of the Day *(Dummy variable)* (Key: Early Morning=1, A.M. Peak=2, Midday=3, P.M. Peak=4, Late Evening=5)	3.0	1	5	1.6	14 (0.0073)
County *(Dummy variable)* (Key: Orange=0, Hillsborough=1)	0.5	0	1	0.7	6.29 (0.0121)
Total Number of 5x5 Clusters (i.e. Signalized Intersections): 149					

5.2.1.2.6 Crash Types: Angle and Left-Turn Crashes

5.2.1.2.6.1 Model Building Procedure

This analysis consisted in creating the model corresponding to the last set of GEE models (i.e. models for different Crash Types) which shows the frequency of "Angle and Left-Turn Crashes" combined as the dependent variable; this was done by adding the 2 types of crash frequencies together.

The model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative

importance and significance these covariates could have for the model. Consequently, the best model for "Angle and Left-Turn Crashes" included the following independent variables:

- *County (COUNTY)*: having 2 levels (Orange and Hillsborough); Hillsborough County was treated as the base case.
- *Period of the day (APERIOD)*: having 5 levels (Early Morning, A.M. Peak, Midday, P.M. Peak and Late Evening); the Early Morning period was treated as the base case.
- *LOS for the intersection as a whole (LOS_INT_NUM)*: Having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
- *Logarithm of the total traffic volume, for the 2-hour period of the day, on the Minor road (LOG_MIN_PER)*: Due to its continuous nature, this variable was processed having only one level.
- *Total Number of Through lanes on the Major road (TLANMAJ_DAYC1)*: originally with 7 levels, this variable ended up having 4 levels after appropriate combinations were made (2 lanes, 4 lanes, 5 with 6 and 7 lanes, as well as 8 combined with 9 lanes); the level corresponding to 2 lanes, the lowest number, was treated as the base case.
- *Total Number of Through lanes on the Minor road (TLANMIN_DAYC1)*: originally with 5 levels, this variable ended up having 3 levels after appropriate combinations were made (2 lanes, 3 lanes, as well as 4 combined with 5 and 6 lanes); the level corresponding to 2 lanes, the lowest number, was treated as the base case.

- *Total Number of Left-Turn lanes on the Major road (LLANMAJ_DAYC1)*: originally with 5 levels, this variable ended up having 3 levels after appropriate combinations were made (1 with 2 lanes, 3 with 4 lanes, as well as 5 lanes); the level corresponding to 1 combined with 2 lanes, the lowest numbers, was treated as the base case.

From the group of covariates just described, only *County* and *Period of the Day* were treated as dummy variables.

This "aggregate level analysis" (i.e. 5 periods of the day) for each of the final GEE models consisted in decomposing the analysis into four submodels, one per type of working correlation structure to be tried (independent, autoregressive, exchangeable and unstructured). The working correlation structure of all submodel types was based on a total of 149 clusters (number of signalized intersections), each having 5 as their maximum and minimum dimensions; this resulted in a 5x5 working correlation matrix per submodel. A total of 745 observations were read and used for each submodel, number that is five times the total number of signalized intersections in the study sample; this can be translated into the following: 149 intersections * 5 periods of the day = 745 observations. Tables 5-31 and 5-32 provide a summary of the model, its composition and respective model-based estimates for each correlation structure.

Table 5-31: General Model Information for "Angle and Left-Turn Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS	
Type III GEE Analysis - Score Statistics	
MODEL 6A: "ANGLE + LEFT-TURN CRASHES" (Considering the 5 Periods of the Day)	
Number of Clusters (i.e. Number of Signalized Intersections)	149
Cluster Size (i.e. Number of Continuous Periods of the Day)	5
Number of Observations	745
Number of Angle and Left-Turn/Total Crashes	1,705/5,532

Table 5-32: Model-Based Standard Error Estimates for "Angle and Left-Turn Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS								
Type 3 GEE Analysis - Model-Based Standard Error Estimates								
MODEL 6A: "ANGLE + LEFT-TURN CRASHES" (Considering the 5 Periods of the Day)								
PARAMETER	Working Correlation Structure							
	INDEPENDENT		AUTOREGRESSIVE		EXCHANGEABLE		UNSTRUCTURED	
	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)
Intercept	-2.5058	0.3232 (<.0001)	-2.3979	0.336 (<.0001)	-2.2448	0.3394 (<.0001)	-2.3505	0.3406 (<.0001)
County *(Dummy variable)*								
Orange	-0.4648	0.0767 (<.0001)	-0.4557	0.0922 (<.0001)	-0.4873	0.0997 (<.0001)	-0.4123	0.1032 (<.0001)
Hillsborough	0	0 .	0	0 .	0	0 .	0	0 .
Period of the Day *(Dummy variable)*								
Late Evening	1.13	0.1689 (<.0001)	1.1411	0.175 (<.0001)	1.1872	0.1684 (<.0001)	1.1712	0.1762 (<.0001)
P.M. Peak	1.2422	0.1784 (<.0001)	1.2747	0.1889 (<.0001)	1.3478	0.1857 (<.0001)	1.2955	0.2033 (<.0001)
Midday	1.0677	0.1736 (<.0001)	1.0895	0.1796 (<.0001)	1.1501	0.177 (<.0001)	1.117	0.1897 (<.0001)
A.M. Peak	0.9245	0.1786 (<.0001)	0.9525	0.1769 (<.0001)	1.0179	0.1835 (<.0001)	0.9745	0.1972 (<.0001)
Early Morning	0	0 .	0	0 .	0	0 .	0	0 .

PARAMETER	Working Correlation Structure							
	INDEPENDENT		AUTOREGRESSIVE		EXCHANGEABLE		UNSTRUCTURED	
	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)	Coef.	S.E. (P-value)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)								
LOS F	1.0616	0.2682 (<.0001)	0.9888	0.2772 (-0.0004)	0.8895	0.271 (-0.001)	0.8825	0.275 (-0.0013)
LOS E	1.1921	0.2734 (<.0001)	1.1499	0.2785 (<.0001)	1.0458	0.2735 (-0.0001)	1.0586	0.2756 (-0.0001)
LOS D	1.0966	0.2592 (<.0001)	1.0559	0.2641 (<.0001)	0.9729	0.2579 (-0.0002)	0.9029	0.2599 (-0.0005)
LOS C	1.0431	0.2497 (<.0001)	1.0168	0.2522 (<.0001)	0.9416	0.2454 (-0.0001)	0.8665	0.2474 (-0.0005)
LOS B	0.8554	0.2481 (-0.0006)	0.7879	0.2465 (-0.0014)	0.7279	0.2386 (-0.0023)	0.6866	0.2392 (-0.0041)
LOS A	0	0 .	0	0 .	0	0 .	0	0 .
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	0.2119	0.0469 (<.0001)	0.2013	0.0538 (-0.0002)	0.1833	0.056 (-0.0011)	0.1973	0.0582 (-0.0007)
Total Number of Through Lanes on Major Road								
8 lanes or 9 lanes	0.6148	0.1764 (-0.0005)	0.594	0.2113 (-0.0049)	0.5939	0.2252 (-0.0084)	0.5734	0.2421 (-0.0178)
5 ≤ lanes ≤ 7	0.2494	0.1146 (-0.0295)	0.2096	0.1371 (-0.1262)	0.2178	0.1464 (-0.1367)	0.2354	0.1553 (-0.1295)
4 lanes	0.3036	0.0999 (-0.0024)	0.2727	0.1194 (-0.0224)	0.2663	0.1277 (-0.037)	0.291	0.1355 (-0.0317)
2 lanes	0	0 .	0	0 .	0	0 .	0	0 .
Total Number of Through Lanes on Minor Road								
4 ≤ lanes ≤ 6	0.1986	0.105 (-0.0585)	0.2105	0.1255 (-0.0936)	0.1844	0.1351 (-0.1723)	0.2984	0.14 (-0.033)
3 lanes	0.3321	0.1163 (-0.0043)	0.3464	0.1393 (-0.0129)	0.3498	0.1498 (-0.0196)	0.3527	0.1575 (-0.0252)
2 lanes	0	0 .	0	0 .	0	0 .	0	0 .
Total Number of Left-Turn Lanes on Major Road								
5 lanes	-0.6563	0.3977 (-0.0989)	-0.6553	0.4852 (-0.1768)	-0.9241	0.5683 (-0.1039)	-0.3577	0.5039 (-0.4778)
3 lanes or 4 lanes	-0.267	0.0981 (-0.0065)	-0.2551	0.118 (-0.0306)	-0.2734	0.1279 (-0.0326)	-0.1994	0.1315 (-0.1296)
1 lanes or 2 lanes	0	0 .	0	0 .	0	0 .	0	0 .
Dispersion	1.0246		1.025		1.0286		1.0287	

5.2.1.2.6.2 Model Interpretation

The respective model interpretation, per correlation structure, is as follows:

- *Independent Correlation Structure*. As shown in Table 5-32, having between 4 and 6 lanes for total number of through lanes on the Minor road, and having 5 or more as total number of left-turn lanes on the Major road, were reported to be insignificant (P-values of 0.0585 and 0.0989, respectively); in spite of this, the respective variables could still be kept in this first submodel since the rest of their levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

  A thorough evaluation of this first submodel for "Angle and Left-Turn Crashes" seems to suggest that the A.M. Peak period has the least estimated frequency of angle and left-turn crashes, whereas the P.M. Peak period is the one with the highest estimated frequency; overall, more angle and left-turn crashes seem to occur as it becomes later in the day. In terms of measures of exposure, the model predicts that traffic volume on the Minor road is also linearly related to the number of angle and left-turn crashes; this suggests that, at the same operational levels, a signalized intersection with very high vehicular volume on its Minor road will experience more crashes. Regarding total number of through lanes on the Major road, it can be observed that a total of 5, 6 or 7 through lanes on the Major road is associated with the least number of angle and left-turn crashes, whereas a total of 8 or more through lanes on this road has the highest crash risk. In addition, it can be observed that a total of 4 or more through lanes on the Minor road is associated with the least number of angle and left-turn crashes, whereas a total of 3 through lanes on this road has the highest crash risk. In similar terms, the higher the number of exclusive left-turn lanes on the Major road, the less the probability for an angle and left-turn crash to occur. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated frequency of angle and left-turn crashes, whereas LOS E (significant delay) is the level having the highest occurrence of this particular crash type; then, at LOS F (excessive delay) the crash rate seems to decrease again, being even less than the crash trend predicted for LOS D (tolerable delay). Summarizing, as a signalized intersection becomes more congested, the higher the crash risk; however, as soon as there are jammed conditions (i.e. the intersection's capacity is exceeded), the crash risk's trend gets reversed (i.e. starts to decrease).

- *Autoregressive Correlation Structure*. As shown in Table 5-32, having between 5 and 7 lanes for total number of through lanes on the Major road, having between 4 and 6 lanes for total number of through lanes on the Minor road, and having 5 or more as total number of left-turn lanes on the Major road, all were reported to be insignificant (P-values of 0.1262, 0.0936 and 0.1768, respectively); in spite of this, the respective variables could still be kept in this second submodel since the rest of their levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

A thorough evaluation of this second submodel for “Angle and Left-Turn Crashes” seems to suggest that the A.M. Peak period has the least estimated frequency of angle and left-turn crashes, whereas the P.M. Peak period is the one with the highest crash frequency; overall, more angle and left-turn crashes seem to occur as it becomes later in the day. In terms of measures of exposure, the model predicts that traffic volume on the Minor road is also linearly related to the number of angle and left-turn crashes; this suggests that, at the same operational levels, a signalized intersection with very high vehicular volume on its Minor road will experience more crashes. Regarding total number of through lanes on the Major road, it can be observed that a total of 5, 6 or 7 through lanes on the Major road is associated with the least number of angle and left-turn crashes, whereas a total of 8 or more through lanes on this road has the highest crash risk. In addition, it can be observed that a total of 4 or more through lanes on the Minor road is associated with the least number of angle and left-turn crashes, whereas a total of 3 through lanes on this road has the highest crash risk. In similar terms, the higher the number of exclusive left-turn lanes on the Major road, the less the probability for an angle and left-turn crash to occur. Finally, regarding the intersection’s LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated frequency of angle and left-turn crashes, whereas LOS E (significant delay) is the level having the highest occurrence of this particular crash type; then, at LOS F (excessive delay) the crash rate seems to decrease again, being even less than the crash trends predicted for LOS C (acceptable delay) and LOS D (tolerable delay). Summarizing, as a signalized intersection becomes more congested, the higher the crash risk; however, as soon as there are jammed conditions (i.e. the intersection's capacity is exceeded), the crash risk's trend gets reversed (i.e. starts to decrease).

- *Exchangeable Correlation Structure*. As shown in Table 5-32, having between 5 and 7 lanes for total number of through lanes on the Major road, having between 4 and 6 lanes for total number of through lanes on the Minor road, and having 5 or more as total number of left-turn lanes on the Major road, all were reported to be insignificant (P-values of 0.1367, 0.1723 and 0.1039, respectively); in spite of this, the respective variables could still be kept in this third submodel since the rest of their levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

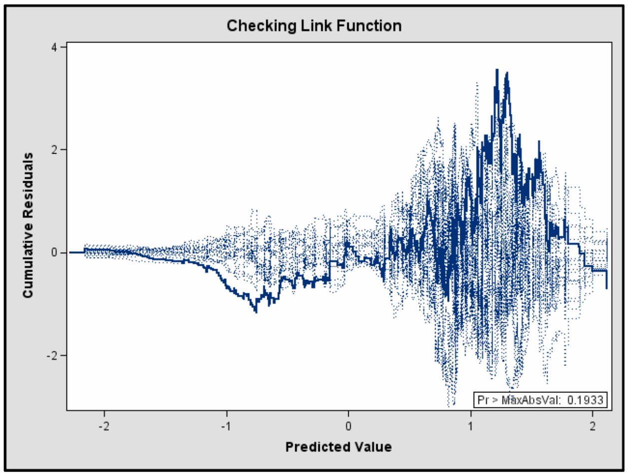
  A thorough evaluation of this third submodel for "Angle and Left Turn Crashes" seems to suggest that the A.M. Peak period has the least estimated frequency of angle and left-turn crashes, whereas the P.M. Peak period is the one with the highest crash frequency; overall, more angle and left-turn crashes seem to occur as it becomes later in the day. In terms of measures of exposure, the model predicts that traffic volume on the Minor road is also linearly related to the number of angle and left-turn crashes; this suggests that, at the same operational levels, a signalized intersection with very high vehicular volume on its Minor road will experience more crashes. Regarding total number of through lanes on the Major road, it can be observed that a total of 5, 6 or 7 through lanes on the Major road is associated with the least number of angle and left-turn crashes, whereas a total of 8 or more through lanes on this road has the highest crash risk. In addition, it can be observed that a total of 4 or more through lanes on the Minor road is associated with the least number of angle and left-turn crashes, whereas a total of 3 through lanes on this road has the highest crash risk. In similar terms, the higher the number of exclusive left-turn lanes on the Major road, the less the probability for an angle and left-turn crash to occur. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated frequency of angle and left-turn crashes, whereas LOS E (significant delay) is the level having the highest occurrence of this particular crash type; then, at LOS F (excessive delay) the crash rate seems to decrease again, being even less than the crash trends predicted for LOS C (acceptable delay) and LOS D (tolerable delay). Summarizing, as a signalized intersection becomes more congested, the higher the crash risk; however, as soon as there are jammed conditions (i.e. the intersection's capacity is exceeded), the crash risk's trend gets reversed (i.e. starts to decrease).

- *Unstructured Correlation Structure*. As shown in Table 5-32, having between 5 and 7 lanes for total number of through lanes on the Major road, as well as having between 3 or more as a total number of left-turn lanes on the Major road, were reported to be insignificant (P-values of 0.1295, 0.1296 and 0.4778, respectively); in spite of this, the respective variables could still be kept in this fourth submodel since the rest of their levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

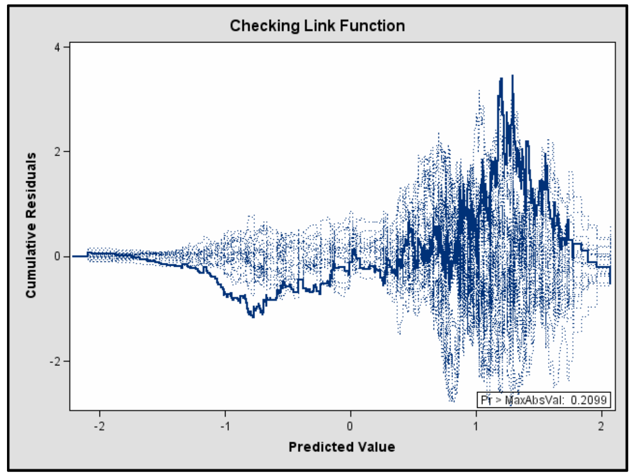
A thorough evaluation of this fourth submodel for “Angle and Left Turn Crashes” seems to suggest that the A.M. Peak period has the least estimated frequency of angle and left-turn crashes, whereas the P.M. Peak period is the one with the highest crash frequency; overall, more angle and left-turn crashes seem to occur as it becomes later in the day. In terms of measures of exposure, the model predicts that traffic volume on the Minor road is also linearly related to the number of angle and left-turn crashes; this suggests that, at the same operational levels, a signalized intersection with very high vehicular volume on its Minor road will experience more crashes. Regarding total number of through lanes on the Major road, it can be observed that a total of 5, 6 or 7 through lanes on the Major road is associated with the least number of angle and left-turn crashes, whereas a total of 8 or more through lanes on this road has the highest crash risk. In addition, it can be observed that a total of 4 or more through lanes on the Minor road is associated with the least number of angle and left-turn crashes, whereas a total of 3 through lanes on this road has the highest crash risk. In similar terms, the higher the number of exclusive left-turn lanes on the Major road, the less the probability for an angle and left-turn crash to occur. Finally, regarding the intersection’s LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with the least estimated frequency of angle and left-turn crashes, whereas LOS E (significant delay) is the level having the highest occurrence of this particular crash type; then, at LOS F (excessive delay) the crash rate seems to decrease again, being even less than the crash trend predicted for LOS D (tolerable delay). Summarizing, as a signalized intersection becomes more congested, the higher the crash risk; however, as soon as there are jammed conditions (i.e. the intersection's capacity is exceeded), the crash risk's trend gets reversed (i.e. starts to decrease).

5.2.1.2.6.3 Model Assessment

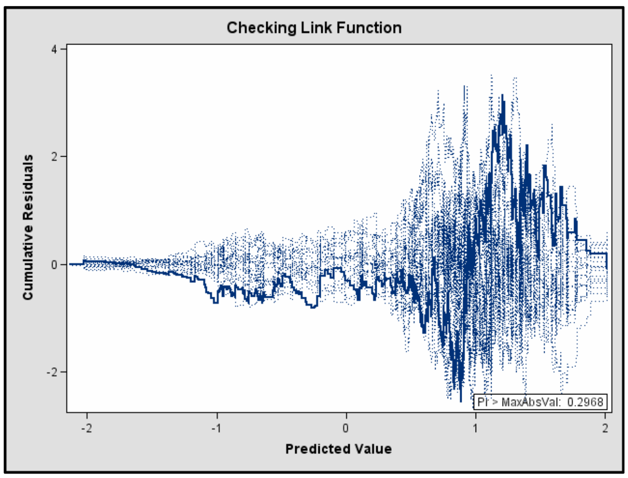
This section summarizes the respective model assessment performed through SAS when analyzing the GEE model designed for "Angle and Left-Turn Crashes" for the 5 periods of the day. Following are the plots that were generated through SAS' Output Delivery System (ODS) for each of the GEE model's type of working correlation structure (see Figure 5-6). When interpreting the plots, the heavy trend line represents the Cumulative Residuals whereas the light ones represent the simulated curves.



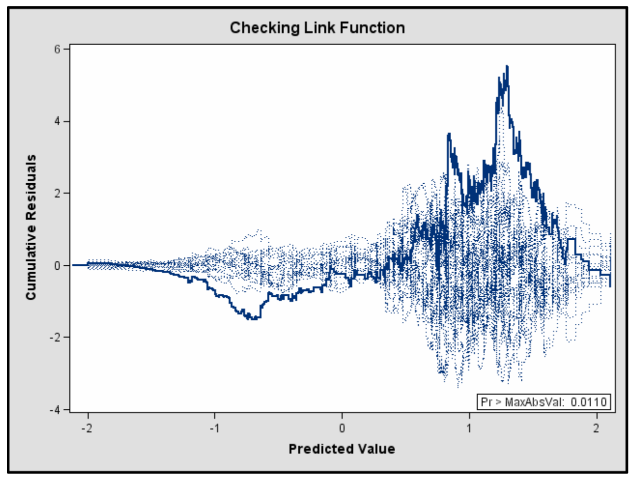
(a)



(b)



(c)



(d)

Figure 5-6: Model Assessment Plots (Cumulative Residuals for GEE Negative Binomial Analysis) for "Angle and Left-Turn Crashes" at Signalized Intersections (Considering the 5 Periods of the Day): (a) Independent Structure, (b) Autoregressive Structure, (c) Exchangeable Structure and (d) Unstructured

As it can be seen, these plots provide a P-value (Pr>MaxAbsVal) computed through a simulation of 10,000 residual paths; below is a summary of these computed values (see Table 5-33).

Table 5-33: Model Assessment Summary (Cumulative Residuals for GEE Negative Binomial Analysis) for "Angle and Left-Turn Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
GEE Analysis - Model Assessment Summary					
MODEL 6A: "ANGLE + LEFT-TURN CRASHES" (Considering the 5 Periods of the Day)					
Working Correlation Structure	Assessment Variable	Max. Abs. Value	Replications	Seed	Pr > Max. Abs. Value
INDEPENDENT	Link Function	3.5519	10000	603708000	0.1933
AUTOREGRESSIVE	Link Function	3.4426	10000	603708000	0.2099
EXCHANGEABLE	Link Function	3.1381	10000	603708000	0.2968
UNSTRUCTURED	Link Function	5.5092	10000	603708000	0.0110

In addition, the GEE analysis procedure in SAS provided the model's respective score statistics based on a type III analysis for each covariate in the model. It can be concluded from these results that, among all correlation structures, period of the day is the most significant of all covariates. Following is Table 5-34, which shows the corresponding Score Statistics for each covariate, listed in descending order of significance.

Table 5-34: GEE Score Statistics Summary for "Angle and Left-Turn Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type 3 GEE Analysis - Score Statistics					
MODEL 6A: "ANGLE + LEFT-TURN CRASHES" (Considering the 5 Periods of the Day)					
COVARIATE	DF	Working Correlation Structure			
		INDEPENDENT	AUTOREGRESSIVE	EXCHANGEABLE	UNSTRUCTURED
		Chi-Sq. (P-value)	Chi-Sq. (P-value)	Chi-Sq. (P-value)	Chi-Sq. (P-value)
Period of the Day *(Dummy variable)*	4	35.67 (<.0001)	36.15 (<.0001)	35.44 (<.0001)	38.92 (<.0001)
County *(Dummy variable)*	1	15.95 (<.0001)	15.24 (<.0001)	16.47 (<.0001)	12.17 (0.0005)
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	1	12.65 (0.0004)	11.16 (0.0008)	9.52 (0.002)	9.34 (0.0022)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)	5	14.42 (0.0131)	18.11 (0.0028)	17.35 (0.0039)	14.73 (0.0116)
Total Number of Through Lanes on Major Road	3	7.69 (0.0528)	6.63 (0.0846)	6.17 (0.1035)	6.96 (0.0731)
Total Number of Through Lanes on Minor Road	2	4.77 (0.0921)	5.19 (0.0747)	4.67 (0.0967)	6.46 (0.0395)
Total Number of Left-Turn Lanes on Major Road	2	3.9 (0.1422)	3.4 (0.1829)	3.94 (0.1391)	2.11 (0.3485)

The last outputs were the working correlation matrices for each structure type (see Table 5-35). As it can be seen, the Independent working correlation matrix was the result of a very naïve analysis procedure; the binary composition of this matrix reflects this fact. Regarding the Autoregressive working correlation matrix, it can be observed that it is characterized by correlation values that decrease over time. The Exchangeable working correlation matrix, however, has a defined and compound symmetry; the exchangeable working correlation had a reported value of 0.1924 for this model. Finally, it can be seen that the Unstructured working correlation matrix's composition has no particular specification.

Table 5-35: GEE Working Correlation Matrices for "Angle and Left-Turn Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type 3 GEE Analysis - Working Correlation Matrices					
MODEL 6A: "ANGLE + LEFT-TURN CRASHES" (Considering the 5 Periods of the Day)					
INDEPENDENT	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.0000	0.0000	0.0000	0.0000
A.M. Peak	0.0000	1.0000	0.0000	0.0000	0.0000
Midday	0.0000	0.0000	1.0000	0.0000	0.0000
P.M. Peak	0.0000	0.0000	0.0000	1.0000	0.0000
Late Evening	0.0000	0.0000	0.0000	0.0000	1.0000
AUTOREGRESSIVE	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.2356	0.0555	0.0131	0.0031
A.M. Peak	0.2356	1.0000	0.2356	0.0555	0.0131
Midday	0.0555	0.2356	1.0000	0.2356	0.0555
P.M. Peak	0.0131	0.0555	0.2356	1.0000	0.2356
Late Evening	0.0031	0.0131	0.0555	0.2356	1.0000
EXCHANGEABLE	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	0.1924	0.1924	0.1924	0.1924
A.M. Peak	0.1924	1.0000	0.1924	0.1924	0.1924
Midday	0.1924	0.1924	1.0000	0.1924	0.1924
P.M. Peak	0.1924	0.1924	0.1924	1.0000	0.1924
Late Evening	0.1924	0.1924	0.1924	0.1924	1.0000
UNSTRUCTURED	Early Morning	A.M. Peak	Midday	P.M. Peak	Late Evening
Early Morning	1.0000	-0.0103	0.0000	-0.1168	0.0863
A.M. Peak	-0.0103	1.0000	0.2983	0.4077	0.2430
Midday	0.0000	0.2983	1.0000	0.4444	0.2180
P.M. Peak	-0.1168	0.4077	0.4444	1.0000	0.3330
Late Evening	0.0863	0.2430	0.2180	0.3330	1.0000

5.2.1.2.6.4 Conclusion

After comparing the results from all correlation structures, it may be concluded that the Exchangeable correlation structure is the best among the other types, since it has the highest P-value (=0.2968). As it is known, the Exchangeable correlation structure is a simple extension to the Independent correlation structure by imposing one additional association parameter within the GEE model; thus, the Exchangeable correlation structure assumes that there is a single common correlation between the observations within the study sample (Hardin and Hilbe, 2003). However, to choose the Exchangeable correlation structure may not be the best choice; not only does the structure ignores time dependence among the studied observations but it also allows any permutation of these, which contradicts the objective to analyze the LOS-Safety relationship throughout 5 sequentially ordered periods of the day.

Based on the previous statements, the Autoregressive correlation structure seems to be the best choice, not only for having the second highest P-value (=0.2099) (refer to Table 5-33) but also because it assumes that all observations in the sample being studied are temporally correlated (Wang, 2006); this seems to be appropriate since the data for this signalized intersections study is temporal in nature (i.e. it follows a natural order).

An appropriate interpretation of the results from the Autoregressive correlation structure, and recalling the basis of the Cumulative Residuals model assessment method, would be that "out of 10,000 realizations from the null distribution, 20.99% have maximum cumulative residuals greater than 3.4426". In addition, the 5x5 matrix for the Autoregressive correlation structure suggests that the repeated observations (i.e. 5 times) for each signalized intersection will become less correlated as the time-gap increases. Finally, the estimates obtained with the Autoregressive correlation structure for each of the covariates in the model were significant overall (refer to Table 5-32), corroborating the idea that the resulting model fits the data well. Finally, Table 5-36 lists the data summary statistics of the contributing factors included in the final model.

Table 5-36: Data Summary Statistics for Preferred Model for "Angle and Left-Turn Crashes" at Signalized Intersections (Considering the 5 Periods of the Day)

SIGNALIZED INTERSECTIONS					
Type III GEE Analysis - Data Summary Statistics					
MODEL 6A: "ANGLE + LEFT=TURN CRASHES" (Considering the 5 Periods of the Day)					
Preferred Correlation Structure: AUTOREGRESSIVE					
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**	**Chi-Sq. (P-value)**
Period of the Day *(Dummy variable)* (Key: Early Morning=1, A.M. Peak=2, Midday=3, P.M. Peak=4, Late Evening=5)	3.0	1	5	1.6	36.15 (<.0001)
County *(Dummy variable)* (Key: Orange=0, Hillsborough=1)	0.5	0	1	0.7	15.24 (<.0001)
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	5.4	1.4	8.6	1.5	11.16 (0.0008)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9	18.11 (0.0028)
Total Number of Through Lanes on Major Road	4.4	2	9	1.6	6.63 (0.0846)
Total Number of Through Lanes on Minor Road	2.6	2	6	1.0	5.19 (0.0747)
Total Number of Left-Turn Lanes on Major Road	2.4	1	5	0.7	3.4 (0.1829)
Total Number of 5x5 Clusters (i.e. Signalized Intersections): 149					

5.2.2 Disaggregate Analysis (Considering one Period of the Day at a time)

5.2.2.1 Statistical Approach

5.2.2.1.1 Negative Binomial

The Negative Binomial technique was used for the LOS-safety disaggregate analysis corresponding to signalized intersections –by disaggregate referring to a more specific analysis, having each period of a day analyzed independently–. This statistical tool was chosen since it is appropriate for analyzing cross-sectional count data, as is the case of crash frequencies.

5.2.2.2 Models

5.2.2.2.1 Early Morning Period

5.2.2.2.1.1 Model Building Procedure

The analysis started by creating the Negative Binomial model for the Early Morning period (1-3 A.M.), which showed the frequency of "Total Crashes" occurring at this period of the day as the dependent variable. On the other hand, the model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for the Early Morning period with "Total Crashes" included the following independent variables:

- *LOS for the intersection as a whole (LOS_INT_NUM)*: Having 4 levels (A, B, C and D); LOS A, the most optimum level within the LOS scale, was treated as the base case.

  (Note: LOS E and LOS F did not appear at any of the observations corresponding to this period of the day; this is due to the fact that at roads are not highly congested at those times of the day).

- *Logarithm of the total traffic volume, for the 2-hour period of the day, on the Major road (LOG_MAJ_PER)*: Due to its continuous nature, this variable was processed having only one level.

- *Lighting Conditions (LIGHT_EMC1)*: originally with 3 levels, an appropriate combination was made in order to obtain only 2 levels (lack of lighting and presence of lighting –partial lighting combined with full lighting conditions–); the former level, lack of lighting, was treated as the base case.

This “disaggregate level analysis” consisted in studying the 149 signalized intersections in the sample with regular Negative Binomial analysis, having the Early Morning period as the only time frame of interest. Consequently, a total of 149 observations were read and used for this first model.

For this first model, the algorithm successfully converged during the analysis in SAS; no errors were reported. Following is Table 5-37, which displays the predictors and estimates corresponding to the best model for the Early Morning period with “Total Crashes”. It can be observed that all predictors were found significant (P-value $\leq$ 0.05). Also, from all the covariates in the model, LOS C and LOS D were found to be the most significant (P-value < 0.0001).

Table 5-37: Parameter Estimates for "Total Crashes" at Signalized Intersections (Early Morning Period)

SIGNALIZED INTERSECTIONS					
Negative Binomial Analysis - Parameter Estimates					
MODEL 1B: "TOTAL CRASHES" (Early Morning)					
PARAMETER	**DF**	**Coef.**	**Chi-Sq.**	**S.E. (P-value)**
Intercept	1	-4.7755	15.85	1.1995 (<.0001)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)				
LOS D	1	2.9181	17.54	0.6967 (<.0001)
LOS C	1	2.0445	19.61	0.4617 (<.0001)
LOS B	1	0.8892	4.04	0.4423 (0.0444)
LOS A	0	0	.	0 .
Logarithm of Total Traffic Volume on Major Road (for the Period of the Day)	1	0.7848	13.34	0.2148 (0.0003)
Lighting Conditions				
Partial or Full	1	-0.5555	4.65	0.2578 (0.0311)
None	0	0	.	0 .
Dispersion	1	0.8009		0.2111

5.2.2.2.1.1 Model Interpretation

A thorough evaluation of this first model seems to suggest that traffic volume "for the Early Morning period" (i.e. 2 hours) along the Minor road forming the intersection is linearly related to the total crash frequency; this suggests that, at the same operational levels, a signalized intersection with very high vehicular volume on its Minor road will experience more crashes within the 1-3 A.M. time frame. It can also be observed that lighting conditions play a significant role with regards to crash occurrence at signalized intersections; presence of light at the intersection, either in partial or full mode, seems to reduce the total crash frequency (estimate = -0.5555). Finally, regarding the intersection's LOS, the main parameter, the model shows that the most optimum LOS levels –as represented by LOS A (insignificant delay), LOS B (minimal delay) and LOS C (acceptable delay) in this case– are associated with a low crash frequency, whereas LOS D (tolerable delay) is the level with the highest crash frequency for a signalized intersection at very early hours of the day. Summarizing, the more congested the signalized intersection within the 1-3 A.M. time frame, the higher the probability for a crash to occur at that location.

5.2.2.2.1.2 Model Assessment

Table 5-38 below contains the goodness-of-fit criteria obtained through SAS when analyzing the Negative Binomial model designed for the Early Morning period (refer to Section 5.2.2.2.1).

Table 5-38: Goodness-of-Fit Criteria for "Total Crashes" at Signalized Intersections (Early Morning Period)

SIGNALIZED INTERSECTIONS			
Negative Binomial Analysis - Criteria for Assessing Goodness-of-Fit			
MODEL 1B: "TOTAL CRASHES" (Early Morning)			
Criterion	DF	Value	Value/DF
Deviance	143	146.8724	1.0271
Scaled Deviance	143	146.8724	1.0271
Pearson Chi-Square	143	151.1231	1.0568
Scaled Pearson X2	143	151.1231	1.0568
Log Likelihood		-10.5021	

5.2.2.2.1.1 Conclusion

Overall, the results corroborate the idea that this model fits the data well. Apart from having all covariates significant at the 0.05 confidence level, the Dispersion parameter was found to be greater than zero (=0.8009), which indicates that the response variable, "Total Crashes" during the Early Morning period (1-3 A.M.), is somewhat over-dispersed; this supports the use of the Negative Binomial regression in the analysis (UCLA, 2008). Furthermore, the obtained ratio of the Deviance to Degrees of Freedom (Dev/DF) was close to 1 (=1.0271), value that also supports the use of this model for obtaining insights on the LOS-Safety relationship for the Early Morning period. Finally, Table 5-39 lists the data summary statistics of the contributing factors included in the model.

Table 5-39: Data Summary Statistics for Final Model for "Total Crashes" at Signalized Intersections (Early Morning Period)

SIGNALIZED INTERSECTIONS				
Negative Binomial Analysis - Data Summary Statistics				
MODEL 1B: "TOTAL CRASHES" (Early Morning)				
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9
Logarithm of Total Traffic Volume on Major Road (for the Period of the Day)	7.0	1.2	8.8	1.2
Lighting Conditions (Key: None=0, Partial or Full=1)	0.5	0	1	0.7
Total Number of Observations (i.e. Signalized Intersections): 149				

5.2.2.2.2 A.M. Peak Period

5.2.2.2.2.1 Model Building Procedure

The analysis started by creating the Negative Binomial model for the A.M. Peak period (7-9 A.M.), which showed the frequency of "Total Crashes" occurring at this period of the day as the dependent variable. On the other hand, the model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for the A.M. Peak period with "Total Crashes" included the following independent variables:

- *Logarithm of the Cycle Length (LOG_CL)*: due to its continuous nature, this variable was processed having only one level.
- *LOS for the intersection as a whole (LOS_INT_NUM)*: having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
- *Logarithm of the total Left-Turn traffic volume, for the 2-hour period of the day, on the Major road (LOG_MAJ_LTPER)*: Due to its continuous nature, this variable was processed having only one level.
- *Total Number of lanes on the Major road (LANMAJ_AMC2)*: originally with 10 levels, this variable ended up having 5 levels after appropriate combinations were made (4 lanes, 5 with 6 lanes, 7 with 8 lanes, 9 with 10 lanes, and 11 combined with 12 and 14 lanes); the level corresponding to 4 lanes, the lowest number, was treated as the base case.

This "disaggregate level analysis" consisted in studying the 149 signalized intersections in the sample with regular Negative Binomial analysis, having the A.M. Peak period as the only time frame of interest. Consequently, a total of 149 observations were read and used for this second model.

For this second model, the algorithm successfully converged during the analysis in SAS; no errors were reported. Following is Table 5-40, which displays the predictors and estimates corresponding to the best model for the A.M. Peak period with "Total Crashes". It can be observed that 5 and 6 lanes, as well as 7 and 8 lanes for total number of through lanes on the Major road were reported to be insignificant (P-values of 0.0534 and 0.0605, respectively); in spite of this, the respective variable could still be kept in this second model since its other levels were significant at the 0.05 confidence level. On the other hand, the rest of the model's predictors were found to be significant (P-value $\leq$ 0.05).

Table 5-40: Parameter Estimates for "Total Crashes" at Signalized Intersections (A.M. Peak Period)

SIGNALIZED INTERSECTIONS				
Negative Binomial Analysis - Parameter Estimates				
MODEL 2B: "TOTAL CRASHES" (A.M. Peak)				
PARAMETER	**DF**	**Coef.**	**Chi-Sq.**	**S.E. (P-value)**
Intercept	1	-4.1	11.01	1.2359 (0.0009)
Logarithm of the Cycle Length (for the Period of the Day)	1	0.7532	12.04	0.2171 (0.0005)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)				
LOS F	1	1.204	7.85	0.4298 (0.0051)
LOS E	1	1.1581	6.95	0.4392 (0.0084)
LOS D	1	1.0672	6.06	0.4334 (0.0138)
LOS C	1	0.8662	4.07	0.4296 (0.0438)
LOS B	1	1.0433	5.5	0.4448 (0.019)
LOS A	0	0	.	0 .
Logarithm of Total Left-Turn Traffic Volume on Major Road (for the Period of the Day)	1	0.1922	7.51	0.0701 (0.0061)
Total Number of Lanes On Major Road				
11 ≤ lanes ≤ 14	1	0.7087	8.14	0.2485 (0.0043)
9 lanes or 10 lanes	1	0.6024	7.01	0.2275 (0.0081)
7 lanes or 8 lanes	1	0.4027	3.52	0.2145 (0.0605)
5 lanes or 6 lanes	1	0.453	3.73	0.2345 (0.0534)
4 lanes	0	0	.	0 .
Dispersion	1	0.1955		0.0401

5.2.2.2.2.2 Model Interpretation

A thorough evaluation of this second model, and regarding operational measures, a signalized intersection's Cycle Length was found to be significant by denoting an almost linear relationship with the total crash frequency; this could be explained by saying that the longer the cycle length there is a bigger time frame open for crash occurrence. Also, it seems that left-turn traffic volume "for the A.M. Peak period" (i.e. 2 hours) along the Major road forming the intersection is linearly related to the total crash frequency; this suggests that, at the same operational levels, a signalized intersection with very high left-turning vehicular volume on its Major road will experience more crashes within the 7-9 A.M. time frame. It can also be observed that a total number of lanes of 11 or greater on the Major road is associated with a very high risk of crash occurrence. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS C (acceptable delay) is associated with a low crash frequency, whereas LOS F (excessive delay) is the level with the highest crash frequency for a signalized intersection; then are LOS E (significant delay), LOS D (tolerable delay) and LOS B (minimal delay), which are associated with the second, third and fourth highest crash frequencies, respectively. Summarizing, the more congested the signalized intersection within the 7-9 A.M. time frame, the higher the probability for a crash to occur at that location.

5.2.2.2.2.3 Model Assessment

Table 5-41 below contains the goodness-of-fit criteria obtained through SAS when analyzing the Negative Binomial model designed for the A.M. Peak period (refer to Section 5.2.2.2.2).

Table 5-41: Goodness-of-Fit Criteria for "Total Crashes" at Signalized Intersections (A.M. Peak Period)

SIGNALIZED INTERSECTIONS			
Negative Binomial Analysis - Criteria for Assessing Goodness-of-Fit			
MODEL 2B: "TOTAL CRASHES" (A.M. Peak)			
Criterion	DF	Value	Value/DF
Deviance	137	166.6697	1.2166
Scaled Deviance	137	166.6697	1.2166
Pearson Chi-Square	137	161.4901	1.1788
Scaled Pearson X2	137	161.4901	1.1788
Log Likelihood		1591.0101	

5.2.2.2.2.4 Conclusion

Overall, the results corroborate the idea that this model fits the data well. Apart from having all covariates significant at the 0.05 confidence level, the Dispersion parameter was found to be greater than zero (=0.1955), which indicates that the response variable, "Total Crashes" during the A.M. Peak period (7-9 A.M.), is over-dispersed; this supports the use of the Negative Binomial regression in the analysis (UCLA, 2008). Furthermore, the obtained ratio of the Deviance to Degrees of Freedom (Dev/DF) was somewhat close to 1 (=1.2166), value that also supports the use of this model for obtaining insights on the LOS-Safety relationship for the A.M. Peak period. Finally, Table 5-42 lists the data summary statistics of the contributing factors included in the model.

Table 5-42: Data Summary Statistics for Final Model for "Total Crashes" at Signalized Intersections (A.M. Peak Period)

SIGNALIZED INTERSECTIONS				
Negative Binomial Analysis - Data Summary Statistics				
MODEL 2B: "TOTAL CRASHES" (A.M. Peak)				
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**
Logarithm of the Cycle Length (for the Period of the Day) (seconds)	4.6	1.6	5.3	0.4
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9
Logarithm of Total Left-Turn Traffic Volume on Major Road (for the Period of the Day)	4.6	0.7	7.2	1.4
Total Number of Lanes on Major Road	7.9	4	14	2.1
Total Number of Observations (i.e. Signalized Intersections): 149				

5.2.2.2.3 Midday Period

5.2.2.2.3.1 Model Building Procedure

The analysis started by creating the Negative Binomial model for the Midday period (12-2 P.M.), which showed the frequency of "Total Crashes" occurring at this period of the day as the dependent variable. On the other hand, the model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for the Midday period with "Total Crashes" included the following independent variables:

- *LOS for the intersection as a whole (LOS_INT_NUM)*: having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
- *Logarithm of the total traffic volume, for the 2-hour period of the day, on the Major road (LOG_MAJ_PER)*: Due to its continuous nature, this variable was processed having only one level.
- *Logarithm of the total traffic volume, for the 2-hour period of the day, on the Minor road (LOG_MIN_PER)*: Due to its continuous nature, this variable was processed having only one level.

This "disaggregate level analysis" consisted in studying the 149 signalized intersections in the sample with regular Negative Binomial analysis, having the Midday period as the only time frame of interest. Consequently, a total of 149 observations were read and used for this second model.

For this third model, the algorithm successfully converged during the analysis in SAS; no errors were reported. Following is Table 5-43, which displays the predictors and estimates corresponding to the best model for the Midday period with "Total Crashes". Only LOS B (minimal delay), second level of the LOS scale, was reported to be insignificant (P-value = 0.0851); in spite of this, the LOS for the intersection as a whole could be kept as a variable in this first submodel since the rest of its levels were significant at the 0.05 confidence level. On the other hand, the rest of this submodel's predictors were found to be significant (P-value $\leq$ 0.05).

Table 5-43: Parameter Estimates for "Total Crashes" at Signalized Intersections (Midday Period)

SIGNALIZED INTERSECTIONS				
Negative Binomial Analysis - Parameter Estimates				
MODEL 3B: "TOTAL CRASHES" (Midday)				
PARAMETER	**DF**	**Coef.**	**Chi-Sq.**	**S.E. (P-value)**
Intercept	1	-3.0233	16.02	0.7554 (<.0001)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)				
LOS F	1	0.9508	8.32	0.3296 (0.0039)
LOS E	1	0.9282	7.69	0.3347 (0.0055)
LOS D	1	0.7357	5.33	0.3186 (0.021)
LOS C	1	0.7597	5.93	0.3121 (0.0149)
LOS B	1	0.5541	2.96	0.3218 (0.0851)
LOS A	0	0	.	0 .
Logarithm of Total Traffic Volume on Major Road (for the Period of the Day)	1	0.3739	23.31	0.0774 (<.0001)
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	1	0.262	26.74	0.0507 (<.0001)
Dispersion	1	0.1356		0.0324

5.2.2.2.3.2 <u>Model Interpretation</u>

A thorough evaluation of this third model seems to suggest that both traffic volumes "for the Midday period" (i.e. 2 hours) along the Major and Minor roads forming the intersection are linearly related to the total crash frequency; this suggests that, at the same operational levels, a signalized intersection with very high vehicular volume on its Major and Minor roads will experience more crashes within the 12-2 P.M. time frame. In addition, it can be noted that traffic volume on the Major road is associated with a higher crash risk when compared to the one on the Minor road. Finally, regarding the intersection's LOS, the main parameter, the model shows that LOS B (minimal delay) is associated with a low crash frequency; then, going further along the scale, the crash trend increases until reaching LOS D (tolerable delay) where it has a sudden and slight decrease in number of crashes. Then, the crash trend increases again until reaching LOS F (excessive delay), which is the level with the highest crash frequency for a signalized intersection. Summarizing, the probability for a crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

5.2.2.2.3.3 Model Assessment

Table 5-44 below contains the goodness-of-fit criteria obtained through SAS when analyzing the Negative Binomial model designed for the Midday period (refer to Section 5.2.2.2.3).

Table 5-44: Goodness-of-Fit Criteria for "Total Crashes" at Signalized Intersections (Midday Period)

SIGNALIZED INTERSECTIONS			
Negative Binomial Analysis - Criteria for Assessing Goodness-of-Fit			
MODEL 3B: "TOTAL CRASHES" (Midday)			
Criterion	DF	Value	Value/DF
Deviance	141	170.4068	1.2086
Scaled Deviance	141	170.4068	1.2086
Pearson Chi-Square	141	152.4525	1.0812
Scaled Pearson X2	141	152.4525	1.0812
Log Likelihood		1804.0927	

5.2.2.2.3.4 Conclusion

Overall, the results corroborate the idea that this model fits the data well. Apart from having all covariates significant at the 0.05 confidence level, the Dispersion parameter was found to be greater than zero (=0.1356), which indicates that the response variable, "Total Crashes" during the Midday period (12-2 P.M.), is over-dispersed; this supports the use of the Negative Binomial regression in the analysis (UCLA, 2008). Furthermore, the obtained ratio of the Deviance to Degrees of Freedom (Dev/DF) was somewhat close to 1 (=1.2086), value that also supports the use of this model for obtaining insights on the LOS-Safety relationship for the Midday period. Finally, Table 5-45 lists the data summary statistics of the contributing factors included in the model.

Table 5-45: Data Summary Statistics for Final Model for "Total Crashes" at Signalized Intersections (Midday Period)

SIGNALIZED INTERSECTIONS				
Negative Binomial Analysis - Data Summary Statistics				
MODEL 3B: "TOTAL CRASHES" (Midday)				
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9
Logarithm of Total Traffic Volume on Major Road (for the Period of the Day)	7.0	1.2	8.8	1.2
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	5.4	1.4	8.6	1.5
Total Number of Observations (i.e. Signalized Intersections): 149				

5.2.2.2.4 P.M. Peak Period

5.2.2.2.4.1 Model Building Procedure

The analysis started by creating the Negative Binomial model for the P.M. Peak period (4-6 P.M.), which showed the frequency of "Total Crashes" occurring at this period of the day as the dependent variable. On the other hand, the model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for the P.M. Peak period with "Total Crashes" included the following independent variables:

- *Logarithm of the Cycle Length (LOG_CL)*: due to its continuous nature, this variable was processed having only one level.
- *LOS for the intersection as a whole (LOS_INT_NUM)*: having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
- *Logarithm of the total Left-Turn traffic volume, for the 2-hour period of the day, on the Major road (LOG_MAJ_LTPER)*: Due to its continuous nature, this variable was processed having only one level.
- *Total Number of lanes on the Major road (LANMAJ_PMC1)*: originally with 10 levels, this variable ended up having 5 levels after appropriate combinations were made (4 lanes, 5 with 6 lanes, 7 with 8 lanes, 9 with 10 lanes, and 11 combined with 12 and 14 lanes); the level corresponding to 4 lanes, the lowest number, was treated as the base case.

This “disaggregate level analysis” consisted in studying the 149 signalized intersections in the sample with regular Negative Binomial analysis, having the P.M. Peak period as the only time frame of interest. Consequently, a total of 149 observations were read and used for this second model.

For this fourth model, the algorithm successfully converged during the analysis in SAS; no errors were reported. Following is Table 5-46, which displays the predictors and estimates corresponding to the best model for the P.M. Peak period with “Total Crashes”. It can be observed that all predictors were found significant (P-value $\leq$ 0.05).

Table 5-46: Parameter Estimates for "Total Crashes" at Signalized Intersections (P.M. Peak Period)

SIGNALIZED INTERSECTIONS				
Negative Binomial Analysis - Parameter Estimates				
MODEL 4B: "TOTAL CRASHES" (P.M. Peak)				
PARAMETER	**DF**	**Coef.**	**Chi-Sq.**	**S.E. (P-value)**
Intercept	1	-2.8935	7.62	1.0482 (0.0058)
Logarithm of the Cycle Length (for the Period of the Day)	1	0.4015	4.24	0.1951 (0.0396)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)				
LOS F	1	1.5251	8.92	0.5106 (0.0028)
LOS E	1	1.495	8.26	0.5203 (0.0041)
LOS D	1	1.2164	5.55	0.5165 (0.0185)
LOS C	1	1.3687	7.31	0.5062 (0.0069)
LOS B	1	1.2588	6.26	0.503 (0.0123)
LOS A	0	0	.	0 .
Logarithm of Total Left-Turn Traffic Volume on Major Road (for the Period of the Day)	1	0.247	10.95	0.0747 (0.0009)
Total Number of Lanes on Major Road				
11 ≤ lanes ≤ 14	1	0.6444	7.38	0.2372 (0.0066)
9 lanes or 10 lanes	1	0.6382	8.48	0.2192 (0.0036)
7 lanes or 8 lanes	1	0.5101	6.59	0.1988 (0.0103)
5 lanes or 6 lanes	1	0.7164	11.27	0.2134 (0.0008)
4 lanes	0	0	.	0 .
Dispersion	1	0.1938		0.0348

5.2.2.2.4.2 Model Interpretation

A thorough evaluation of this fourth model, and regarding operational measures, a signalized intersection's Cycle Length was found to be significant by denoting an almost linear relationship with the total crash frequency; this could be explained by saying that the longer the cycle length there is a bigger time frame open for crash occurrence. Also, it seems that left-turn traffic volume "for the P.M. Peak period" (i.e. 2 hours) along the Major road forming the intersection is linearly related to the total crash frequency; this suggests that, at the same operational levels, a signalized intersection with very high left-turning vehicular volume on its Major road will experience more crashes within the 4-6 P.M. time frame. It can also be observed that having 5 or 6 lanes as total number of lanes (i.e. not the highest number of lanes) on the Major road is associated with a very high risk of crash occurrence; this is different to the A.M. Peak period, where the total number of crashes is at its highest when having the most number of lanes on the Major road. Finally, regarding the intersection's LOS, the main parameter, the model shows that the total crash frequency increases as operation conditions deteriorate until reaching LOS D (tolerable delay), which is associated with the lowest crash frequency; then, going further along the scale, the crash trend increases again until reaching LOS F (excessive delay), level with the highest crash frequency for a signalized intersection. Summarizing, the probability for a crash to occur at a signalized intersection increases as it approaches undesirable traffic operation conditions.

5.2.2.2.4.3 Model Assessment

Table 5-47 below contains the goodness-of-fit criteria obtained through SAS when analyzing the Negative Binomial model designed for the P.M. Peak period (refer to Section 5.2.2.2.4).

Table 5-47: Goodness-of-Fit Criteria for "Total Crashes" at Signalized Intersections (P.M. Peak Period)

SIGNALIZED INTERSECTIONS			
Negative Binomial Analysis - Criteria for Assessing Goodness-of-Fit			
MODEL 4B: "TOTAL CRASHES" (P.M. Peak)			
Criterion	DF	Value	Value/DF
Deviance	137	154.5218	1.1279
Scaled Deviance	137	154.5218	1.1279
Pearson Chi-Square	137	149.7829	1.0933
Scaled Pearson X2	137	149.7829	1.0933
Log Likelihood		2599.5451	

5.2.2.2.4.4 Conclusion

Overall, the results corroborate the idea that this model fits the data well. Apart from having all covariates significant at the 0.05 confidence level, the Dispersion parameter was found to be greater than zero (=0.1938), which indicates that the response variable, "Total Crashes" during the P.M. Peak period (4-6 P.M.), is over-dispersed; this supports the use of the Negative Binomial regression in the analysis (UCLA, 2008). Furthermore, the obtained ratio of the Deviance to Degrees of Freedom (Dev/DF) was somewhat close to 1 (=1.1279), value that also supports the use of this model for obtaining insights on the LOS-Safety relationship for the P.M.

Peak period. Finally, Table 5-48 lists the data summary statistics of the contributing factors included in the model.

Table 5-48: Data Summary Statistics for Final Model for "Total Crashes" at Signalized Intersections (P.M. Peak Period)

SIGNALIZED INTERSECTIONS				
Negative Binomial Analysis - Data Summary Statistics				
MODEL 4B: "TOTAL CRASHES" (P.M. Peak)				
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**
Logarithm of the Cycle Length (for the Period of the Day) (seconds)	4.6	1.6	5.3	0.4
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9
Logarithm of Total Left-Turn Traffic Volume on Major Road (for the Period of the Day)	4.6	0.7	7.2	1.4
Total Number of Lanes on Major Road	7.9	4	14	2.1
Total Number of Observations (i.e. Signalized Intersections): 149				

5.2.2.2.5 Late Evening Period

5.2.2.2.5.1 Model Building Procedure

For this last model for signalized intersections, the analysis started by creating the Negative Binomial model for the Late Evening period (8-10 P.M.), which showed the frequency of "Total Crashes" occurring at this period of the day as the dependent variable. On the other hand, the model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for the Late Evening period with "Total Crashes" included the following independent variables:

- *LOS for the intersection as a whole (LOS_INT_NUM)*: having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
- *Land Use (LAND_LEC2)*: originally with 3 levels, an appropriate combination was made in order to obtain only 2 levels (Rural area, and Urban area –Suburban and Urban areas combined–); the former level, Rural area, was treated as the base case.
- *Logarithm of the total traffic volume, for the 2-hour period of the day, on the Major road (LOG_MAJ_PER)*: Due to its continuous nature, this variable was processed having only one level.
- *Logarithm of the total traffic volume, for the 2-hour period of the day, on the Minor road (LOG_MIN_PER)*: Due to its continuous nature, this variable was processed having only one level.

This "disaggregate level analysis" consisted in studying the 149 signalized intersections in the sample with regular Negative Binomial analysis, having the Late Evening period as the only time frame of interest. Consequently, a total of 149 observations were read and used for this first model.

For this fifth and last model, the algorithm successfully converged during the analysis in SAS; no errors were reported. Following is Table 5-49, which displays the predictors and estimates corresponding to the best model for the Late Evening period with "Total Crashes". Only LOS F (excessive delay), sixth and last level of the LOS scale, was reported to be insignificant (P-value = 0.8290); in spite of this, the LOS for the intersection as a whole could be kept as a variable in this first submodel since the rest of its levels were significant at the 0.05 confidence level. On the other hand, it can be observed that the rest of predictors were found significant, considering a significance level of 0.05; among these, traffic volumes “for the Late Evening period” (i.e. 2 hours) along the Major and Minor roads were found to be the most significant (P-value < 0.0001). As an additional note, lighting conditions was considered initially in the modeling process; however, this variable had to be removed since it did not contribute to the significance of the overall model (contrary to the case of the Early Morning period).

Table 5-49: Parameter Estimates for "Total Crashes" at Signalized Intersections (Late Evening Period)

SIGNALIZED INTERSECTIONS				
Negative Binomial Analysis - Parameter Estimates				
MODEL 5B: "TOTAL CRASHES" (Late Evening)				
PARAMETER	DF	Coef.	Chi-Sq.	S.E. (P-value)
Intercept	1	-2.9931	16.36	0.7399 (<.0001)
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day)				
LOS F	1	0.1328	0.05	0.615 (0.829)
LOS E	1	0.9452	8.95	0.316 (0.0028)
LOS D	1	0.9779	13.11	0.27 (0.0003)
LOS C	1	0.805	10.65	0.2467 (0.0011)
LOS B	1	0.556	5.27	0.2423 (0.0217)
LOS A	0	0	.	0 .
Land Use				
Suburban or Urban	1	-0.4939	6.71	0.1907 (0.0096)
Rural	0	0	.	0 .
Logarithm of Total Traffic Volume on Major Road (for the Period of the Day)	1	0.3471	15.84	0.0872 (<.0001)
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	1	0.3753	41.31	0.0584 (<.0001)
Dispersion	1	0.0977		0.0319

5.2.2.2.5.2 Model Interpretation

A thorough evaluation of this fifth and last model seems to suggest that land use plays a significant role with regards to crash occurrence at signalized intersections; a signalized intersection located in either a suburban or urban area seems to have a smaller probability of crash occurrence (estimate = -0.4939). Also, it can be seen that both traffic volumes “for the Late Evening period” (i.e. 2 hours) along the Major and Minor roads forming the intersection are linearly related to the total crash frequency; this suggests that, at the same operational levels, a signalized intersection with very high vehicular volume on its Major and Minor roads will experience more crashes within the 8-10 P.M. time frame. In addition, it can be noted that traffic volume on the Minor road is associated with a higher crash risk when compared to the one on the Major road. Finally, regarding the intersection’s LOS, the main parameter, the model shows that LOS F (excessive delay) is associated with a low crash frequency, whereas LOS D (tolerable delay) is the level with the highest crash frequency for a signalized intersection. Summarizing, the probability for a crash to occur at a signalized intersection increases as it gets close to tolerable operation conditions, point where it reaches its highest; on the other hand, this probability is at its lowest when the signalized intersection experiences excessive delays.

5.2.2.2.5.3 Model Assessment

Finally, Table 5-50 below contains the goodness-of-fit criteria obtained through SAS when analyzing the Negative Binomial model designed for the Late Evening period (refer to Section 5.2.2.2.5).

Table 5-50: Goodness-of-Fit Criteria for “Total Crashes” at Signalized Intersections (Late Evening Period)

SIGNALIZED INTERSECTIONS			
Negative Binomial Analysis - Criteria for Assessing Goodness-of-Fit			
MODEL 5B: "TOTAL CRASHES" (Late Evening)			
Criterion	DF	Value	Value/DF
Deviance	140	175.0629	1.2504
Scaled Deviance	140	175.0629	1.2504
Pearson Chi-Square	140	171.4485	1.2246
Scaled Pearson X2	140	171.4485	1.2246
Log Likelihood		1174.5464	

5.2.2.2.5.4 Conclusion

Overall, the results corroborate the idea that this model fits the data well. Apart from having all covariates significant at the 0.05 confidence level, the Dispersion parameter was found to be somewhat greater than zero (=0.0977), which indicates that the response variable, “Total Crashes” during the Late Evening period (8-10 P.M.), is over-dispersed; this supports the use of the Negative Binomial regression in the analysis (UCLA, 2008). Furthermore, the obtained ratio of the Deviance to Degrees of Freedom (Dev/DF) was somewhat close to 1 (=1.2504), value that also supports the use of this model for obtaining insights on the LOS-Safety relationship for the Late Evening period. Finally, Table 5-51 lists the data summary statistics of the contributing factors included in the model.

Table 5-51: Data Summary Statistics for Final Model for "Total Crashes" at Signalized Intersections (Late Evening Period)

SIGNALIZED INTERSECTIONS				
Negative Binomial Analysis - Data Summary Statistics				
MODEL 5B: "TOTAL CRASHES" (Late Evening)				
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**
LEVEL-OF-SERVICE for the Intersection as a whole (for the Period of the Day) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9
Land Use (Key: Rural=0, Suburban or Urban=1)	0.5	0	1	0.7
Logarithm of Total Traffic Volume on Major Road (for the Period of the Day)	7.0	1.2	8.8	1.2
Logarithm of Total Traffic Volume on Minor Road (for the Period of the Day)	5.4	1.4	8.6	1.5
Total Number of Observations (i.e. Signalized Intersections): 149				

5.3 Multilane High-Speed Arterial Corridors

5.3.1 Overall Analysis

5.3.1.1 Statistical Approach

5.3.1.1.1 Negative Binomial

Similar to the "per period" analysis for the previous study (i.e. Signalized Intersections), the Negative Binomial technique was used for analyzing the LOS-safety relationship of multilane high-speed arterial corridors. This statistical tool facilitated the study of the cross-sectional count data in hand.

5.3.1.2 Models

5.3.1.2.1 Total Crashes

5.3.1.2.1.1 Model Building Procedure

The analysis started by creating a Negative Binomial model applicable to the safety evaluation of multilane high-speed arterial corridors. Since temporal correlation is not taken into account in this study, the response variable was frequency of "Total Crashes" within the study area, corresponding to the years 2000 and 2001; all days of the week have been considered, along with a 24-hour time frame. The model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for "Total Crashes" at Multilane High-Speed Arterial Corridors included the following independent variables:

- *LOS for the road section as a whole (LOS_SECT_NUM)*: having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
  - It has to be noted that this LOS has been calculated for the 100^{th} highest traffic hour of the year, per road section; this is based on FDOT's ARTPLAN standards detailed in the Quality/Level-of-Service Handbook.

- *Total Number of lanes (for left, through and right traffic flows) in a single direction (LANES_CORRSTOTC1)*: originally with 7 levels, this variable ended up having 4 levels after appropriate combinations were made ("2 lanes in a single direction", "3 or 4 lanes in a single direction", 5 or 6 lanes in a single direction, and 7 or 8 lanes in a single direction); the level corresponding to 2 lanes, the lowest number, was treated as the base case.
- *Speed limit on the road section (MAXSPEED_CORRSTOTC1)*: originally with 6 levels, this variable ended up having 3 levels after appropriate combinations were made (40 with 45 mph, 50 with 55 mph, as well as 60 combined with 65 mph); the level corresponding to 40 with 45 mph, the lowest speed limit range, was treated as the base case.
- *Total Length (in miles) of the road section (TOT_LENGTH_CORRS)*: due to its continuous nature, this variable was processed having only one level.

This analysis consisted in studying the 399 road sections in the sample with regular Negative Binomial analysis, considering that the respective LOS data were correspond to the 100^{th} highest traffic hour of the year.

For this model, the algorithm successfully converged during the analysis in SAS; no errors were reported. Following are Tables 5-52 and 5-53; the latter displays the predictors and estimates corresponding to the best model for "Total Crashes" at Multilane High-Speed Arterial Corridors. Only the speed limits of 50 and 55 mph were reported to be quite insignificant (P-value = 0.0561); in spite of this, speed limit on the road section could be kept as a variable in this first model since the rest of its levels were significant at the 0.05 confidence level. On the other hand, it can be observed that the rest of predictors were found significant, considering a significance level of 0.05; among these, LOS F, total number of lanes in a single direction, as well as a road section's total length all were found to be the most significant parameters (P-value < 0.0001).

Table 5-52: General Model Information for "Total Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS	
Negative Binomial Analysis - General Model Information	
MODEL 1C: "TOTAL CRASHES"	
Number of Observations (i.e. Number of Road Sections)	399
Number of Total Crashes	14,339

Table 5-53: Parameter Estimates for "Total Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
Negative Binomial Analysis - Parameter Estimates				
MODEL 1C: "TOTAL CRASHES"				
PARAMETER	**DF**	**Coef.**	**Chi-Sq.**	**S.E. (P-value)**
Intercept	1	1.774	43.1	0.2702 (<.0001)
LEVEL-OF-SERVICE for the Road Section (for the 100th highest traffic hour of the year)				
LOS F	1	0.9948	15.83	0.25 (<.0001)
LOS E	1	0.9039	9.92	0.287 (0.0016)
LOS D	1	1.0082	15.11	0.2594 (0.0001)
LOS C	1	0.7254	8.87	0.2436 (0.0029)
LOS B	1	0.5057	4.33	0.2431 (0.0375)
LOS A	0	0	.	0 .
Total Number of Lanes in a Single Direction of the Road Section				
7 lanes or 8 lanes	1	1.1869	40.42	0.1867 (<.0001)
5 lanes or 6 lanes	1	0.8819	52.1	0.1222 (<.0001)
3 lanes or 4 lanes	1	0.5411	22.74	0.1135 (<.0001)
2 lanes	0	0	.	0 .
Posted Speed Limit for the Road Section				
60 mph or 65 mph	1	-0.6734	7.08	0.2531 (0.0078)
50 mph or 55 mph	1	-0.1613	3.65	0.0844 (0.0561)
40 mph or 45 mph	0	0	.	0 .
Total Length of the Road Section	1	0.5142	59.24	0.0668 (<.0001)
Dispersion	1	0.4602		0.0345

5.3.1.2.1.2 Model Interpretation

A thorough evaluation of this model suggests that the higher the number of lanes along the road, the more probability for crashes to occur. In addition, the results indicate that more crashes take place at lower speed limits (40 or 45 mph) than at higher ones (60 or 65 mph). Also, the length of the road section was found to be significant by denoting an almost linear relationship with the total crash frequency. Finally, regarding the road section's LOS, the main parameter, the model shows that LOS B (reasonably free flow) is associated with a small crash frequency, whereas LOS D (approaching unstable flow) is the level with the highest crash frequency; then are LOS F (forced or breakdown flow), LOS E (unstable flow) and LOS C (stable flow), which are associated with the second, third and fourth highest crash frequencies, respectively. Summarizing, the probability for a crash to occur at a multilane high-speed arterial corridor increases as it approaches undesirable traffic flow and/or operation conditions.

5.3.1.2.1.3 Model Assessment

Table 5-54 below contains the goodness-of-fit criteria obtained through SAS when analyzing the Negative Binomial model designed for "Total Crashes" on multilane high-speed arterial corridors (refer to Section 5.3.1.2.1).

Table 5-54: Goodness-of-Fit Criteria for "Total Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS			
Negative Binomial Analysis - Criteria for Assessing Goodness-of-Fit			
MODEL 1C: "TOTAL CRASHES"			
Criterion	**DF**	**Value**	**Value/DF**
Deviance	387	438.5928	1.1333
Scaled Deviance	387	438.5928	1.1333
Pearson Chi-Square	387	335.9005	0.868
Scaled Pearson X2	387	335.9005	0.868
Log Likelihood		41200.3851	

5.3.1.2.1.4 Conclusion

The results for this model corroborate the idea that it fits the road segments' data well. Apart from having all covariates significant at the 0.05 confidence level, the Dispersion parameter was found to be greater than zero (=0.4602), which indicates that the response variable, "Total Crashes", is over-dispersed; this supports the use of the Negative Binomial regression in the analysis (UCLA, 2008). Furthermore, the obtained ratio of the Deviance to Degrees of Freedom (Dev/DF) was somewhat close to 1 (=1.1333), value that also supports the use of this model for obtaining insights on the LOS-Safety relationship for the multilane high-speed arterial corridors. Finally, Table 5-55 lists the data summary statistics of the contributing factors included in the model.

Table 5-55: Data Summary Statistics for Final Model for "Total Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
Negative Binomial Analysis - Data Summary Statistics				
MODEL 1C: "TOTAL CRASHES"				
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**
LEVEL-OF-SERVICE for the Road Section (for the 100th highest traffic hour of the year) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9
Total Number of Lanes in a Single Direction on the Road Section	4.4	2	8	1.6
Posted Speed Limit for the Road Section (mph)	46.9	40	65	5.5
Total Length of the Road Section (mi)	0.73	0.02	7.97	0.91
Total Number of Observations (i.e. Road Sections): 399				

5.3.1.2.2 Severe Crashes

5.3.1.2.2.1 Model Building Procedure

The analysis started by creating a Negative Binomial model applicable to the safety evaluation of multilane high-speed arterial corridors; the response variable was frequency of "Severe Crashes" within the study area, corresponding to the years 2000 and 2001; all days of the week have been considered, along with a 24-hour time frame. The model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for "Severe Crashes" at Multilane High-Speed Arterial Corridors included the following independent variables:

- *LOS for the road section as a whole (LOS_SECT_NUM)*: having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
    - It has to be noted that this LOS has been calculated for the 100th highest traffic hour of the year, per road section; this is based on FDOT's ARTPLAN standards detailed in the Quality/Level-of-Service Handbook.
- *Total Number of lanes (for left, through and right traffic flows) in a single direction (LANES_CORRSTOTC1)*: originally with 7 levels, this variable ended up having 4 levels after appropriate combinations were made ("2 lanes in a single direction", "3 or 4 lanes in a single direction", 5 or 6 lanes in a single direction, and 7 or 8 lanes in a single direction); the level corresponding to 2 lanes, the lowest number, was treated as the base case.
- *Speed limit on the road section (MAXSPEED_CORRSTOTC1)*: originally with 6 levels, this variable ended up having 3 levels after appropriate combinations were made (40 with 45 mph, 50 with 55 mph, as well as 60 combined with 65 mph); the level corresponding to 40 with 45 mph, the lowest speed limit range, was treated as the base case.
- *Total Length (in miles) of the road section (TOT_LENGTH_CORRS)*: due to its continuous nature, this variable was processed having only one level.
- *Pavement surface type of the road section (SURFNUM_SECTION)*: having 3 levels (Unknown/Other, Portland Cement Concrete and Asphaltic Concrete); the Unknown/Other category was treated as the base case.

- *Inside Shoulder Width (in feet) of the road section (ISLDWDTH_SECTION)*: due to its continuous nature, this variable was processed having only one level.

This analysis consisted in studying the 399 road sections in the sample with regular Negative Binomial analysis, considering that the respective LOS data were correspond to the 100th highest traffic hour of the year.

For this model, the algorithm successfully converged during the analysis in SAS; no errors were reported. Following are Tables 5-56 and 5-57; the latter displays the predictors and estimates corresponding to the best model for "Severe Crashes" at Multilane High-Speed Arterial Corridors. It can be observed that LOS C, as well as the speed limits of 50 and 55 mph were reported to be insignificant (P-values of 0.0606 and 0.5700, respectively); in spite of this, the respective variables could still be kept in this second model since their other levels were significant at the 0.05 confidence level. On the other hand, it can be observed that the rest of predictors were found significant, considering a significance level of 0.05; among these, having between 3 and 6 as a total number of lanes in a single direction, as well as a road section's total length were found to be the most significant parameters (P-value < 0.0001).

Table 5-56: General Model Information for "Severe Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS	
Negative Binomial Analysis - General Model Information	
MODEL 2C: "SEVERE CRASHES"	
Number of Observations (i.e. Number of Road Sections)	399
Number of Severe/Total Crashes	1,499/14,339

Table 5-57: Parameter Estimates for "Severe Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
Negative Binomial Analysis - Parameter Estimates				
MODEL 2C: "SEVERE CRASHES"				
PARAMETER	DF	Coef.	Chi-Sq.	S.E. (P-value)
Intercept	1	-1.3241	6.51	0.5189 (-0.0107)
LEVEL-OF-SERVICE for the Road Section (for the 100th highest traffic hour of the year)				
LOS F	1	0.7888	5.94	0.3235 (-0.0148)
LOS E	1	0.9727	7.32	0.3594 (-0.0068)
LOS D	1	0.9579	8.35	0.3315 (-0.0039)
LOS C	1	0.5897	3.52	0.3142 (-0.0606)
LOS B	1	0.7152	5.2	0.3136 (-0.0226)
LOS A	0	0	.	0 .
Total Number of Lanes in a Single Direction of the Road Section				
7 lanes or 8 lanes	1	0.903	13.81	0.243 (-0.0002)
5 lanes or 6 lanes	1	0.8311	28.83	0.1548 (<.0001)
3 lanes or 4 lanes	1	0.57	15.23	0.1461 (<.0001)
2 lanes	0	0	.	0 .
Posted Speed Limit for the Road Section				
60 mph or 65 mph	1	-0.7379	5.38	0.3181 (-0.0203)
50 mph or 55 mph	1	-0.0588	0.32	0.1035 (-0.57)
40 mph or 45 mph	0	0	.	0 .
Total Length of the Road Section	1	0.5631	57.24	0.0744 (<.0001)

PARAMETER	DF	Coef.	Chi-Sq.	S.E. (P-value)
Pavement Surface Type for the Road Section				
Asphaltic Concrete	1	0.9615	5.9	0.3958 (-0.0151)
Portland Cement Concrete	1	0.2832	0.37	0.4633 (-0.5411)
Unknown/Other	0	0	.	0 .
Inside Shoulder Width of the Road Section (ft)	1	-0.1064	4.94	0.0479 (-0.0263)
Dispersion	1	0.4159		0.0527

5.3.1.2.2.2 Model Interpretation

A thorough evaluation of this model suggests that the higher the number of lanes along the road, the more probability for severe crashes to occur. In addition, the results indicate that more severe crashes take place at lower speed limits (40 or 45 mph) than at higher ones (60 or 65 mph). Also, the length of the road section was found to be significant by denoting an almost linear relationship with the severe crash frequency. In terms of road surface type, asphaltic concrete is associated with a higher severe crash frequency when compared to Portland cement concrete. Finally, regarding the road section's LOS, the main parameter, the model shows that LOS C (stable flow) is associated with a low severe crash frequency, whereas LOS E (unstable flow) is the level with the highest severe crash frequency; then are LOS D (approaching unstable flow), LOS F (forced or breakdown flow) and LOS B (reasonably free flow), which are associated with the second, third and fourth highest severe crash frequencies, respectively. Summarizing, the probability for a severe crash to occur at a multilane high-speed arterial corridor increases as it approaches undesirable traffic flow and/or operation conditions.

5.3.1.2.2.3 Model Assessment

Table 5-58 contains the goodness-of-fit criteria obtained through SAS when analyzing the Negative Binomial model designed for "Severe Crashes" on multilane high-speed arterial corridors (refer to Section 5.3.1.2.1).

Table 5-58: Goodness-of-Fit Criteria for "Severe Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS			
Negative Binomial Analysis - Criteria for Assessing Goodness-of-Fit			
MODEL 2C: "SEVERE CRASHES"			
Criterion	**DF**	**Value**	**Value/DF**
Deviance	384	443.0422	1.1538
Scaled Deviance	384	443.0422	1.1538
Pearson Chi-Square	384	380.6852	0.9914
Scaled Pearson X2	384	380.6852	0.9914
Log Likelihood		822.0382	

5.3.1.2.2.4 Conclusion

The results for this model corroborate the idea that it fits the road segments' data well. Apart from having all covariates significant at the 0.05 confidence level, the Dispersion parameter was found to be greater than zero (=0.4159), which indicates that the response variable, "Severe Crashes", is over-dispersed; this supports the use of the Negative Binomial regression in the analysis (UCLA, 2008). Furthermore, the obtained ratio of the Deviance to Degrees of Freedom (Dev/DF) was somewhat close to 1 (=1.1538), value that also supports the use of this model for obtaining insights on the LOS-Safety relationship for the multilane high-speed arterial corridors. Finally, Table 5-59 lists the data summary statistics of the contributing factors included in the model.

Table 5-59: Data Summary Statistics for Final Model for "Severe Crashes" at Multilane High Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
Negative Binomial Analysis - Data Summary Statistics				
MODEL 2C: "SEVERE CRASHES"				
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**
LEVEL-OF-SERVICE for the Road Section (for the 100th highest traffic hour of the year) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9
Total Number of Lanes in a Single Direction on the Road Section	4.4	2	8	1.6
Posted Speed Limit for the Road Section (mph)	46.9	40	65	5.5
Total Length of the Road Section (mi)	0.73	0.02	7.97	0.91
Pavement Surface Type for the Road Section (Key: Unknown or Other=1, Portland Cement Concrete=2, Asphaltic Concrete=3	1.0	0	2	1.0
Inside Shoulder Width of the Road Section (ft)	0.6	0	9.4	1.1
Total Number of Observations (i.e. Road Sections): 399				

5.3.1.2.3 Crash Types: Rear-End Crashes

5.3.1.2.3.1 Model Building Procedure

The analysis started by creating a Negative Binomial model applicable to the safety evaluation of multilane high-speed arterial corridors; the response variable was frequency of "Rear-End Crashes" within the study area, corresponding to the years 2000 and 2001; all days of the week have been considered, along with a 24-hour time frame. The model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for "Rear-End Crashes" at Multilane High-Speed Arterial Corridors included the following independent variables:

- *LOS for the road section as a whole (LOS_SECT_NUM)*: having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
    - It has to be noted that this LOS has been calculated for the 100th highest traffic hour of the year, per road section; this is based on FDOT's ARTPLAN standards detailed in the Quality/Level-of-Service Handbook.
- *Total Length (in miles) of the road section (TOT_LENGTH_CORRS)*: due to its continuous nature, this variable was processed having only one level.
- *Logarithm of the weighted ADT (in number of vehicle passenger cars equivalent) for the road section (LOG_ADTSECTION)*: Due to its continuous nature, this variable was processed having only one level.

This analysis consisted in studying the 399 road sections in the sample with regular Negative Binomial analysis, considering that the respective LOS data were correspond to the 100th highest traffic hour of the year.

For this model, the algorithm successfully converged during the analysis in SAS; no errors were reported. Following are Tables 5-60 and 5-61; the latter displays the predictors and estimates corresponding to the best model for "Rear-End Crashes" at Multilane High-Speed Arterial Corridors. It can be observed that LOS B and LOS F were reported to be insignificant

(P-values of 0.1395 and 0.0667, respectively); in spite of this, the LOS for the road section could be kept as a variable in this third model since the rest of its levels were considered to be significant at the 0.05 confidence level. On the other hand, it can be observed that the rest of predictors were found significant, considering a significance level of 0.05; among these, a road section's total length and its traffic flow (i.e. weighted ADT) were found to be the most significant parameters (P-value < 0.0001).

Table 5-60: General Model Information for "Rear-End Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS	
Negative Binomial Analysis - General Model Information	
MODEL 3C: "REAR-END CRASHES"	
Number of Observations (i.e. Number of Road Sections)	399
Number of Rear-End/Total Crashes	6,306/14,339

Table 5-61: Parameter Estimates for "Rear-End Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
Negative Binomial Analysis - Parameter Estimates				
MODEL 3C: "REAR-END CRASHES"				
PARAMETER	**DF**	**Coef.**	**Chi-Sq.**	**S.E. (P-value)**
Intercept	1	-10.0473	119.8	0.9179 (<.0001)
LEVEL-OF-SERVICE for the Road Section (for the 100th highest traffic hour of the year)				
LOS F	1	0.5736	3.36	0.3128 (0.0667)
LOS E	1	0.6907	4.07	0.3422 (0.0436)
LOS D	1	0.8705	8	0.3079 (0.0047)
LOS C	1	0.6044	4.45	0.2864 (0.0349)
LOS B	1	0.4317	2.18	0.2922 (0.1395)
LOS A	0	0	.	0 .
Total Length of the Road Section (ft)	1	0.4505	58.06	0.0591 (<.0001)
Logarithm of the Weighted ADT for the Road Section (vehicle passenger cars equivalent)	1	1.14	154.79	0.0916 (<.0001)
Dispersion	1	0.5345		0.0454

5.3.1.2.3.2 Model Interpretation

A thorough evaluation of this model shows that both the length of the road section as well as its traffic flow (i.e. weighted ADT) are significant by denoting an almost linear relationship with the rear-end crash frequency. Regarding the road section's LOS, the main parameter, the model shows that LOS B (reasonably free flow) is associated with a low rear-end crash frequency, whereas LOS D (approaching unstable flow) is the level with the highest rear-end crash frequency; then are LOS E (unstable flow), LOS C (stable flow) and LOS F (forced or breakdown flow), which are associated with the second, third and fourth highest rear-end crash frequencies, respectively. Summarizing, the probability for a rear-end crash to occur at a multilane high-speed arterial corridor is at its highest when the traffic flow and/or operation conditions are quite stable.

5.3.1.2.3.3 Model Assessment

Table 6-62 below contains the goodness-of-fit criteria obtained through SAS when analyzing the Negative Binomial model designed for "Rear-End Crashes" on multilane high-speed arterial corridors (refer to Section 5.3.1.2.1).

Table 5-62: Goodness-of-Fit Criteria for "Rear-End Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS			
Negative Binomial Analysis - Criteria for Assessing Goodness-of-Fit			
MODEL 3C: "REAR-END CRASHES"			
Criterion	**DF**	**Value**	**Value/DF**
Deviance	391	454.2376	1.1617
Scaled Deviance	391	454.2376	1.1617
Pearson Chi-Square	391	380.7584	0.9738
Scaled Pearson X2	391	380.7584	0.9738
Log Likelihood		13541.7696	

5.3.1.2.3.4 Conclusion

The results for this model corroborate the idea that it fits the road segments' data well. Apart from having all covariates significant at the 0.05 confidence level, the Dispersion parameter was found to be greater than zero (=0.5345), which indicates that the response variable, “Rear-End Crashes”, is over-dispersed; this supports the use of the Negative Binomial regression in the analysis (UCLA, 2008). Furthermore, the obtained ratio of the Deviance to Degrees of Freedom (Dev/DF) was somewhat close to 1 (=1.1617), value that also supports the use of this model for obtaining insights on the LOS-Safety relationship for the multilane high-speed arterial corridors. Finally, Table 5-63 lists the data summary statistics of the contributing factors included in the model.

Table 5-63: Data Summary Statistics for Final Model for “Rear-End Crashes” at Multilane High Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
Negative Binomial Analysis - Data Summary Statistics				
MODEL 3C: "REAR-END CRASHES"				
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**
LEVEL-OF-SERVICE for the Road Section (for the 100th highest traffic hour of the year) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9
Total Length of the Road Section (mi)	0.73	0.02	7.97	0.91
Logarithm of the Weighted ADT for the Road Section (vehicle passenger cars equivalent)	8.8	6.7	10.1	0.4
Total Number of Observations (i.e. Road Sections): 399				

5.3.1.2.4 Crash Types: Sideswipe Crashes

5.3.1.2.4.1 Model Building Procedure

The analysis started by creating a Negative Binomial model applicable to the safety evaluation of multilane high-speed arterial corridors; the response variable was frequency of

“Sideswipe Crashes” within the study area, corresponding to the years 2000 and 2001; all days of the week have been considered, along with a 24-hour time frame. The model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for “Sideswipe Crashes” at Multilane High-Speed Arterial Corridors included the following independent variables:

- *LOS for the road section as a whole (LOS_SECT_NUM)*: having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
    - It has to be noted that this LOS has been calculated for the 100th highest traffic hour of the year, per road section; this is based on FDOT’s ARTPLAN standards detailed in the Quality/Level-of-Service Handbook.
- *Total Number of lanes (for left, through and right traffic flows) in a single direction (LANES_CORRSTOTC1)*: originally with 7 levels, this variable ended up having 4 levels after appropriate combinations were made (“2 lanes in a single direction”, “3 or 4 lanes in a single direction”, 5 or 6 lanes in a single direction, and 7 or 8 lanes in a single direction); the level corresponding to 2 lanes, the lowest number, was treated as the base case.
- *Total Length (in miles) of the road section (TOT_LENGTH_CORRS)*: due to its continuous nature, this variable was processed having only one level.

This analysis consisted in studying the 399 road sections in the sample with regular Negative Binomial analysis, considering that the respective LOS data were correspond to the 100th highest traffic hour of the year.

For this model, the algorithm successfully converged during the analysis in SAS; no errors were reported. Following are Tables 5-64 and 5-65; the latter displays the predictors and estimates corresponding to the best model for "Sideswipe Crashes" at Multilane High-Speed Arterial Corridors. It can be observed that LOS B and LOS C were reported to be insignificant (P-values of 0.4677 and 0.2051, respectively); in spite of this, the respective variable could still be kept in this fourth model since its other levels were significant at the 0.05 confidence level. On the other hand, it can be observed that the rest of predictors were found significant, considering a significance level of 0.05; among these, having between 5 and 8 as a total number of lanes in a single direction, as well as a road section's total length were found to be the most significant parameters (P-value < 0.0001).

Table 5-64: General Model Information for "Sideswipe Crashes" at Multilane High-Speed Arterial Corridors for the 100th Highest Traffic Hour of the Year

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS	
Negative Binomial Analysis - General Model Information	
MODEL 4C: "SIDESWIPE CRASHES"	
Number of Observations (i.e. Number of Road Sections)	399
Number of Sideswipe/Total Crashes	961/14,339

Table 5-65: Parameter Estimates for "Sideswipe Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
Negative Binomial Analysis - Parameter Estimates				
MODEL 4C: "SIDESWIPE CRASHES"				
PARAMETER	DF	Coef.	Chi-Sq.	S.E. (P-value)
Intercept	1	-0.9095	6.13	0.3672 (0.0133)
LEVEL-OF-SERVICE for the Road Section (for the 100th highest traffic hour of the year)				
LOS F	1	0.9186	7.68	0.3315 (0.0056)
LOS E	1	0.8282	4.98	0.3711 (0.0256)
LOS D	1	0.8341	6	0.3406 (0.0143)
LOS C	1	0.4123	1.61	0.3254 (0.2051)
LOS B	1	0.2409	0.53	0.3317 (0.4677)
LOS A	0	0	.	0 .
Total Number of Lanes in a Single Direction of the Road Section				
7 lanes or 8 lanes	1	1.5052	41.83	0.2327 (<.0001)
5 lanes or 6 lanes	1	1.1108	39.81	0.1761 (<.0001)
3 lanes or 4 lanes	1	0.5994	12.52	0.1694 (0.0004)
2 lanes	0	0	.	0 .
Total Length of the Road Section (mi)	1	0.4251	51.21	0.0594 (<.0001)
Dispersion	1	0.3617		0.0605

5.3.1.2.4.2 Model Interpretation

A thorough evaluation of this model suggests that the higher the number of lanes along the road, the more probability for sideswipe crashes to occur. Also, the length of the road section was found to be significant by denoting an almost linear relationship with the sideswipe crash frequency. Finally, regarding the road section's LOS, the main parameter, the model shows that LOS B (reasonably free flow) is associated with a low sideswipe crash frequency; then, going further along the scale, the crash trend increases again until reaching LOS F (forced or breakdown flow), level with the highest sideswipe crash frequency. Summarizing, the probability for a sideswipe crash to occur at a multilane high-speed arterial corridor increases as it approaches undesirable traffic flow and/ or operation conditions.

5.3.1.2.4.3 Model Assessment

Table 5-66 below contains the goodness-of-fit criteria obtained through SAS when analyzing the Negative Binomial model designed for "Sideswipe Crashes" on multilane high-speed arterial corridors (refer to Section 5.3.1.2.1).

Table 5-66: Goodness-of-Fit Criteria for "Sideswipe Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS			
Negative Binomial Analysis - Criteria for Assessing Goodness-of-Fit			
MODEL 4C: "SIDESWIPE CRASHES"			
Criterion	DF	Value	Value/DF
Deviance	389	431.6041	1.1095
Scaled Deviance	389	431.6041	1.1095
Pearson Chi-Square	389	384.2355	0.9878
Scaled Pearson X2	389	384.2355	0.9878
Log Likelihood		103.1642	

5.3.1.2.4.4 Conclusion

The results for this model corroborate the idea that it fits the road segments' data well. Apart from having all covariates significant at the 0.05 confidence level, the Dispersion parameter was found to be greater than zero (=0.3617), which indicates that the response variable, "Sideswipe Crashes", is over-dispersed; this supports the use of the Negative Binomial regression in the analysis (UCLA, 2008). Furthermore, the obtained ratio of the Deviance to Degrees of Freedom (Dev/DF) was somewhat close to 1 (=1.1095), value that also supports the use of this model for obtaining insights on the LOS-Safety relationship for the multilane high-speed arterial corridors. Finally, Table 5-67 lists the data summary statistics of the contributing factors included in the model.

Table 5-67: Data Summary Statistics for Final Model for "Sideswipe Crashes" at Multilane High Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
Negative Binomial Analysis - Data Summary Statistics				
MODEL 4C: "SIDESWIPE CRASHES"				
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**
LEVEL-OF-SERVICE for the Road Section (for the 100th highest traffic hour of the year) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9
Total Number of Lanes in a Single Direction on the Road Section	4.4	2	8	1.6
Total Length of the Road Section (mi)	0.73	0.02	7.97	0.91
Total Number of Observations (i.e. Road Sections): 399				

5.3.1.2.5 Crash Types: Head-On Crashes

5.3.1.2.5.1 Model Building Procedure

The analysis started by creating a Negative Binomial model applicable to the safety evaluation of multilane high-speed arterial corridors; the response variable was frequency of

"Head-On Crashes" within the study area, corresponding to the years 2000 and 2001; all days of the week have been considered, along with a 24-hour time frame. The model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for "Head-On Crashes" at Multilane High-Speed Arterial Corridors included the following independent variables:

- *LOS for the road section as a whole (LOS_SECT_NUM)*: having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
  - It has to be noted that this LOS has been calculated for the 100th highest traffic hour of the year, per road section; this is based on FDOT's ARTPLAN standards detailed in the Quality/Level-of-Service Handbook.
- *Total Number of lanes (for left, through and right traffic flows) in a single direction (LANES_CORRSTOTC1)*: originally with 7 levels, this variable ended up having 4 levels after appropriate combinations were made ("2 lanes in a single direction", "3 or 4 lanes in a single direction", 5 or 6 lanes in a single direction, and 7 or 8 lanes in a single direction); the level corresponding to 2 lanes, the lowest number, was treated as the base case.
- *Speed limit on the road section (MAXSPEED_CORRSTOTC1)*: originally with 6 levels, this variable ended up having 3 levels after appropriate combinations were made (40 with 45 mph, 50 with 55 mph, as well as 60 combined with 65 mph); the level corresponding to 40 with 45 mph, the lowest speed limit range, was treated as the base case.

- *Median Width (in feet) of the road section (MEDWIDTH_SECTION)*: due to its continuous nature, this variable was processed having only one level.
- *Inside Shoulder Width (in feet) of the road section (ISLDWDTH_SECTION)*: due to its continuous nature, this variable was processed having only one level.

This analysis consisted in studying the 399 road sections in the sample with regular Negative Binomial analysis, considering that the respective LOS data were correspond to the 100th highest traffic hour of the year.

For this model, the algorithm successfully converged during the analysis in SAS; no errors were reported. Following are Tables 5-68 and 5-69; the latter displays the predictors and estimates corresponding to the best model for "Head-On Crashes" at Multilane High-Speed Arterial Corridors. It can be observed that LOS B and LOS E, as well as having between 3 and 6 as total number of lanes in a single direction, were reported to be insignificant (P-values of 0.0819 and 0.2023, as well as 0.7772 and 0.4661, respectively); in spite of this, the respective variables could still be kept in this fifth model since their other levels were significant at the 0.05 confidence level. On the other hand, it can be observed that the rest of predictors were found significant, considering a significance level of 0.05.

Table 5-68: General Model Information for "Head-On Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS	
Negative Binomial Analysis - General Model Information	
MODEL 5C: "HEAD-ON CRASHES"	
Number of Observations (i.e. Number of Road Sections)	399
Number of Head-On/Total Crashes	190/14,339

Table 5-69: Parameter Estimates for "Head-On Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
Negative Binomial Analysis - Parameter Estimates				
MODEL 5C: "HEAD-ON CRASHES"				
PARAMETER	DF	Coef.	Chi-Sq.	S.E. (P-value)
Intercept	1	-2.7971	6.82	1.0712 (0.009)
LEVEL-OF-SERVICE for the Road Section (for the 100th highest traffic hour of the year)				
LOS F	1	2.4848	5.42	1.0671 (0.0199)
LOS E	1	1.4766	1.63	1.1581 (0.2023)
LOS D	1	2.6037	5.88	1.0742 (0.0154)
LOS C	1	2.0943	3.91	1.0591 (0.048)
LOS B	1	1.8449	3.03	1.0604 (0.0819)
LOS A	0	0	.	0 .
Total Number of Lanes in a Single Direction of the Road Section				
7 lanes or 8 lanes	1	1.2155	7.38	0.4474 (0.0066)
5 lanes or 6 lanes	1	0.2313	0.53	0.3174 (0.4661)
3 lanes or 4 lanes	1	0.0806	0.08	0.2849 (0.7772)
2 lanes	0	0	.	0 .
Posted Speed Limit for the Road Section				
60 mph or 65 mph	1	1.0166	4.17	0.4978 (0.0411)
50 mph or 55 mph	1	0.6725	9.93	0.2135 (0.0016)
40 mph or 45 mph	0	0	.	0 .
Median Width of the Road Section (ft)	1	-0.0277	10.77	0.0084 (0.001)
Inside Shoulder Width of the Road Section (ft)	1	-0.195	4.09	0.0965 (0.0432)
Dispersion	1	0.3449		0.185

5.3.1.2.5.2 Model Interpretation

A thorough evaluation of this model suggests that the higher the number of lanes along the road, the more probability for head-on crashes to occur. In addition, the results indicate that more head-on crashes take place at higher speed limits (60 or 65 mph) than at lower ones. Also, both the median and inside shoulder widths of a road section were found to be significant; note their negative coefficients (-0.0277 and -0.1950, respectively) indicating a reduction in the number of head-on crashes. Finally, regarding the road section's LOS, our main parameter, the model shows that LOS E (unstable flow) is associated with a low head-on crash frequency, whereas LOS D (approaching unstable flow) is the level with the highest head-on crash frequency; then are LOS F (forced or breakdown flow), LOS C (stable flow) and LOS B (reasonably free flow), which are associated with the second, third and fourth highest head-on crash frequencies, respectively. Summarizing, the probability for a head-on crash to occur at a multilane high-speed arterial corridor is at its highest either when the traffic flow and/ or operation conditions are quite stable or when the road reaches the highest level of congestion.

5.3.1.2.5.3 Model Assessment

Table 5-70 below contains the goodness-of-fit criteria obtained through SAS when analyzing the Negative Binomial model designed for "Head-On Crashes" on multilane high-speed arterial corridors (refer to Section 5.3.1.2.1).

Table 5-70: Goodness-of-Fit Criteria for "Head-On Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS			
Negative Binomial Analysis - Criteria for Assessing Goodness-of-Fit			
MODEL 5C: "HEAD-ON CRASHES"			
Criterion	DF	Value	Value/DF
Deviance	386	343.2414	0.8892
Scaled Deviance	386	343.2414	0.8892
Pearson Chi-Square	386	381.4368	0.9882
Scaled Pearson X2	386	381.4368	0.9882
Log Likelihood		-302.1739	

5.3.1.2.5.4 Conclusion

The results for this model corroborate the idea that it fits the road segments' data well. Apart from having all covariates significant at the 0.05 confidence level, the Dispersion parameter was found to be greater than zero (=0.3449), which indicates that the response variable, "Head-On Crashes", is over-dispersed; this supports the use of the Negative Binomial regression in the analysis (UCLA, 2008). Furthermore, the obtained ratio of the Deviance to Degrees of Freedom (Dev/DF) was somewhat close to 1 (=0.8892), value that also supports the use of this model for obtaining insights on the LOS-Safety relationship for the multilane high-speed arterial corridors. Finally, Table 5-71 lists the data summary statistics of the contributing factors included in the model.

Table 5-71: Data Summary Statistics for Final Model for "Head-On Crashes" at Multilane High Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
Negative Binomial Analysis - Data Summary Statistics				
MODEL 5C: "HEAD-ON CRASHES"				
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**
LEVEL-OF-SERVICE for the Road Section (for the 100th highest traffic hour of the year) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9
Total Number of Lanes in a Single Direction on the Road Section	4.4	2	8	1.6
Posted Speed Limit for the Road Section (mph)	46.9	40	65	5.5
Median Width of the Road Section (ft)	20.9	0.0	116.4	14.6
Inside Shoulder Width of the Road Section (ft)	0.6	0	9.4	1.1
Total Number of Observations (i.e. Road Sections): 399				

5.3.1.2.6 Crash Types: Angle and Left-Turn Crashes

5.3.1.2.6.1 <u>Model Building Procedure</u>

The analysis started by creating a Negative Binomial model applicable to the safety evaluation of multilane high-speed arterial corridors; the response variable was frequency of "Angle and Left-Turn Crashes" within the study area, corresponding to the years 2000 and 2001; all days of the week have been considered, along with a 24-hour time frame. The model calibration (i.e. the act of checking or adjusting, by comparison with a standard value) and independent variable selection processes were based on the relative importance and significance these covariates could have for the model. Consequently, the best model for "Angle and Left-

Turn Crashes" at Multilane High-Speed Arterial Corridors included the following independent variables:

- *LOS for the road section as a whole (LOS_SECT_NUM)*: having 6 levels (A, B, C, D, E and F); LOS A, the most optimum level within the LOS scale, was treated as the base case.
    - It has to be noted that this LOS has been calculated for the 100th highest traffic hour of the year, per road section; this is based on FDOT's ARTPLAN standards detailed in the Quality/Level-of-Service Handbook.
- *Total Number of lanes (for left, through and right traffic flows) in a single direction (LANES_CORRSTOTC1)*: originally with 7 levels, this variable ended up having 4 levels after appropriate combinations were made ("2 lanes in a single direction", "3 or 4 lanes in a single direction", 5 or 6 lanes in a single direction, and 7 or 8 lanes in a single direction); the level corresponding to 2 lanes, the lowest number, was treated as the base case.
- *Speed limit on the road section (MAXSPEED_CORRSTOTC1)*: originally with 6 levels, this variable ended up having 3 levels after appropriate combinations were made (40 with 45 mph, 50 with 55 mph, as well as 60 combined with 65 mph); the level corresponding to 40 with 45 mph, the lowest speed limit range, was treated as the base case.
- *Total Length (in miles) of the road section (TOT_LENGTH_CORRS)*: due to its continuous nature, this variable was processed having only one level.
- *Inside Shoulder Width (in feet) of the road section (ISLDWDTH_SECTION)*: due to its continuous nature, this variable was processed having only one level.

This analysis consisted in studying the 399 road sections in the sample with regular Negative Binomial analysis, considering that the respective LOS data were correspond to the 100th highest traffic hour of the year.

For this model, the algorithm successfully converged during the analysis in SAS; no errors were reported. Following are Tables 5-72 and 5-73; the latter displays the predictors and estimates corresponding to the best model for “Angle and Left-Turn Crashes” at Multilane High-Speed Arterial Corridors. It can be observed that LOS B and LOS E were reported to be insignificant (P-values of 0.0524 and 0.0803, respectively); in spite of this, the respective variable could still be kept in this last model since its other levels were significant at the 0.05 confidence level. On the other hand, it can be observed that the rest of predictors were found significant, considering a significance level of 0.05; among these, total number of lanes in a single direction as well as a road section’s total length were found to be the most significant parameters (P-value < 0.0001).

Table 5-72: General Model Information for “Angle and Left-Turn Crashes” at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS	
Negative Binomial Analysis - General Model Information	
MODEL 6C: "ANGLE + LEFT-TURN CRASHES"	
Number of Observations (i.e. Number of Road Sections)	399
Number of Angle and Left-Turn/Total Crashes	4,849/14339

Table 5-73: Parameter Estimates for "Angle and Left-Turn Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
Negative Binomial Analysis - Parameter Estimates				
MODEL 6C: "ANGLE + LEFT-TURN CRASHES"				
PARAMETER	**DF**	**Coef.**	**Chi-Sq.**	**S.E. (P-value)**
Intercept	1	0.7969	5.81	0.3307 (0.016)
LEVEL-OF-SERVICE for the Road Section (for the 100th highest traffic hour of the year)				
LOS F	1	0.701	5.1	0.3103 (0.0239)
LOS E	1	0.6109	3.06	0.3493 (0.0803)
LOS D	1	0.9509	8.83	0.32 (0.003)
LOS C	1	0.7969	6.84	0.3047 (0.0089)
LOS B	1	0.5911	3.76	0.3047 (0.0524)
LOS A	0	0	.	0 .
Total Number of Lanes in a Single Direction of the Road Section				
7 lanes or 8 lanes	1	1.2501	28.62	0.2337 (<.0001)
5 lanes or 6 lanes	1	0.9355	41.84	0.1446 (<.0001)
3 lanes or 4 lanes	1	0.6973	27.31	0.1334 (<.0001)
2 lanes	0	0	.	0 .
Posted Speed Limit for the Road Section				
60 mph or 65 mph	1	-1.0997	12.29	0.3137 (0.0005)
50 mph or 55 mph	1	-0.2009	4.37	0.0961 (0.0366)
40 mph or 45 mph	0	0	.	0 .
Total Length of the Road Section (mi)	1	0.4983	45.11	0.0742 (<.0001)
Inside Shoulder Width of the Road Section (ft)	1	-0.132	9.61	0.0426 (0.0019)
Dispersion	1	0.5542		0.0476

5.3.1.2.6.2 Model Interpretation

A thorough evaluation of this model suggests that the higher the number of lanes along the road, the more probability for angle and left-turn crashes to occur. In addition, the results indicate that more angle and left-turn crashes take place at lower speed limits (40 or 45 mph) than at higher ones (60 or 65 mph). Also, the length of the road section was found to be significant by denoting an almost linear relationship with the frequency of angle and left-turn crashes. The inside shoulder width was also found significant; note its negative coefficient (= -0.1320) indicating a reduction in the frequency of angle and left-turn crashes. Finally, regarding the road section's LOS, the main parameter, the model shows that LOS B (reasonably free flow) is associated with a low frequency of angle and left-turn crashes, whereas LOS D (approaching unstable flow) is the level with the highest frequency of angle and left-turn crashes; then are LOS C (stable flow), LOS F (forced or breakdown flow) and LOS E (unstable flow), which are associated with the second, third and fourth highest frequencies, respectively. Summarizing, the probability for a Head-On crash to occur at a multilane high-speed arterial corridor is at its highest when the traffic flow and/ or operation conditions are stable.

5.3.1.2.6.3 Model Assessment

Table 5-74 below contains the goodness-of-fit criteria obtained through SAS when analyzing the Negative Binomial model designed for "Angle and Left-Turn Crashes" on multilane high-speed arterial corridors (refer to Section 5.3.1.2.1).

Table 5-74: Goodness-of-Fit Criteria for "Angle and Left-Turn Crashes" at Multilane High-Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS			
Negative Binomial Analysis - Criteria for Assessing Goodness-of-Fit			
MODEL 6C: "ANGLE + LEFT-TURN CRASHES"			
Criterion	**DF**	**Value**	**Value/DF**
Deviance	386	458.5708	1.188
Scaled Deviance	386	458.5708	1.188
Pearson Chi-Square	386	360.4243	0.9337
Scaled Pearson X2	386	360.4243	0.9337
Log Likelihood		8600.9715	

5.3.1.2.6.4 Conclusion

The results for this model corroborate the idea that it fits the road segments' data well. Apart from having all covariates significant at the 0.05 confidence level, the Dispersion parameter was found to be greater than zero (=0.5542), which indicates that the response variable, "Total Crashes", is over-dispersed; this supports the use of the Negative Binomial regression in the analysis (UCLA, 2008). Furthermore, the obtained ratio of the Deviance to Degrees of Freedom (Dev/DF) was close to 1 (=1.1880), value that also supports the use of this model for obtaining insights on the LOS-Safety relationship for the multilane high-speed arterial corridors. Finally, Table 5-75 lists the data summary statistics of the contributing factors included in this last model.

Table 5-75: Data Summary Statistics for Final Model for "Angle and Left-Turn Crashes" at Multilane High Speed Arterial Corridors

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
Negative Binomial Analysis - Data Summary Statistics				
MODEL 6C: "ANGLE + LEFT=TURN CRASHES"				
Predictor	**Mean**	**Min.**	**Max.**	**Std. Dev.**
LEVEL-OF-SERVICE for the Road Section (for the 100th highest traffic hour of the year) (Key: A=1, B=2, C=3, D=4, E=5, F=6)	3.5	1	6	1.9
Total Number of Lanes in a Single Direction on the Road Section	4.4	2	8	1.6
Posted Speed Limit for the Road Section (mph)	46.9	40	65	5.5
Total Length of the Road Section (mi)	0.73	0.02	7.97	0.91
Inside Shoulder Width of the Road Section (ft)	0.6	0	9.4	1.1
Total Number of Observations (i.e. Road Sections): 399				

CHAPTER 6: RESEARCH SUMMARY AND CONTRIBUTIONS

6.1 Overview

As the final chapter of this thesis, a comprehensive summary of the research conducted, including details on its implications and results, is presented in the following sections. Overall, the two studies performed constitute a good reference for transportation engineers and planners in order to see that the LOS can be used as a significant factor in the prediction of the safety performance of signalized intersections and multilane high-speed arterial corridors. Due to the existing potential for future research in this subject, a series of recommendations is also presented.

6.2 LOS-Safety Study for Signalized Intersections

6.2.1 Aggregate Analysis (Considering the 5 Periods of the Day)

6.2.1.1 Models' Summary and Implications

Recalling the corresponding modeling approach described in Chapter 5, six final models (i.e. Model Group A) resulted after using the Type III GEE technique with Negative Binomial link function; this method was used for analyzing the relationship between the LOS and crash occurrence taking all 5 periods of the day into account (i.e. aggregate analysis). From the four types of GEE correlation structures considered, only the Autoregressive and Unstructured correlation structures were selected as the “best” or most appropriate based on the data in hand; this selection was based on past studies and theoretical background on the GEE technique.

The LOS for the intersection as a whole was found to be a significant predictor in every model, a finding that meets the purpose of the research conducted. In order to better understand these findings, following are the implications of the most important aspects of this investigation.

Following is Table 6-1, which contains the summary of all the resulting models, detailing their corresponding correlation structure and significant factors.

Table 6-1: Summary of Signalized Intersections' Final Models: Model Group A (Type III GEE Analysis)

SIGNALIZED INTERSECTIONS	
Models' Summary	
Type III GEE Analysis (Considering the 5 Periods of the Day)	
MODEL (BEST CORRELATION STRUCTURE)	**SIGNIFICANT FACTORS** (α = 0.05) **(Listed from Highest to Lowest Degree of Relative Significance)**
Model 1A TOTAL CRASHES (AUTOREGRESSIVE)	**Log. of the Cycle Length** (for the Period of the Day) **Log. of the Total Left-Turn Traffic Volume on the Major Road** (for the Period of the Day) **County** *(Dummy variable)* **Period of the Day** *(Dummy variable)* **LEVEL-OF-SERVICE for the Intersection as a whole** (for the Period of the Day) **Log. of the Total Traffic Volume on the Minor Road** (for the Period of the Day)
Model 2A SEVERE CRASHES (UNSTRUCTURED)	**Period of the Day** *(Dummy variable)* **County** *(Dummy variable)* **LEVEL-OF-SERVICE for the Intersection as a whole** (for the Period of the Day) **Total Number of Lanes on the Major Road**
Model 3A REAR-END CRASHES (AUTOREGRESSIVE)	**Log. of the Total Traffic Volume on the Minor Road** (for the Period of the Day) **Log. of the Total Traffic Volume on the Major Road** (for the Period of the Day) **Period of the Day** *(Dummy variable)* **LEVEL-OF-SERVICE for the Intersection as a whole** (for the Period of the Day) **County** *(Dummy variable)*
Model 4A SIDESWIPE CRASHES (AUTOREGRESSIVE)	**LEVEL-OF-SERVICE for the Intersection as a whole** (for the Period of the Day) **Total Number of Lanes on the Major Road** **Period of the Day** *(Dummy variable)* **Log. of the Cycle Length** (for the Period of the Day)
Model 5A HEAD-ON CRASHES (UNSTRUCTURED)	**LEVEL-OF-SERVICE for the Intersection as a whole** (for the Period of the Day) **Period of the Day** *(Dummy variable)* **County** *(Dummy variable)*
Model 6A ANGLE + LEFT-TURN CRASHES (AUTOREGRESSIVE)	**Period of the Day** *(Dummy variable)* **County** *(Dummy variable)* **Log. of the Total Traffic Volume on the Minor Road** (for the Period of the Day) **LEVEL-OF-SERVICE for the Intersection as a whole** (for the Period of the Day) **Total Number of Through Lanes on the Major Road** **Total Number of Through Lanes on the Minor Road** **Total Number of Left-Turn Lanes on the Major Road**

6.2.1.1.1 Significant Factors

Overall, crash occurrence at signalized intersections can be attributed to the following factor categories:

- Geometric design features (i.e. number of lanes)
- Location (i.e. County) (dummy variable)
- Time (i.e. Period of the Day) (dummy variable)
- Traffic characteristics (i.e. traffic volumes)
- Traffic control and operational measures (i.e. cycle length)
- Traffic and operation conditions as a whole (i.e. LOS).

As has been mentioned throughout this thesis, the LOS has been the predictor of interest in the modeling process; when applicable, also the period of the day. Considering LOS A (insignificant delay) as the base case, following are plots derived from Model Group A's GEE estimates (i.e. coefficients) for the LOS variable (see Figures 6-1 through 6-6). Similarly, considering the Early Morning period (1-3 A.M.) as the base case, following are plots derived from Model Group A's GEE estimates (i.e. coefficients) for the period of the day variable (see Figures 6-7 through 6-12). The observations obtained from these plots were considered for the preparation of the safety guidelines detailed in the following section.

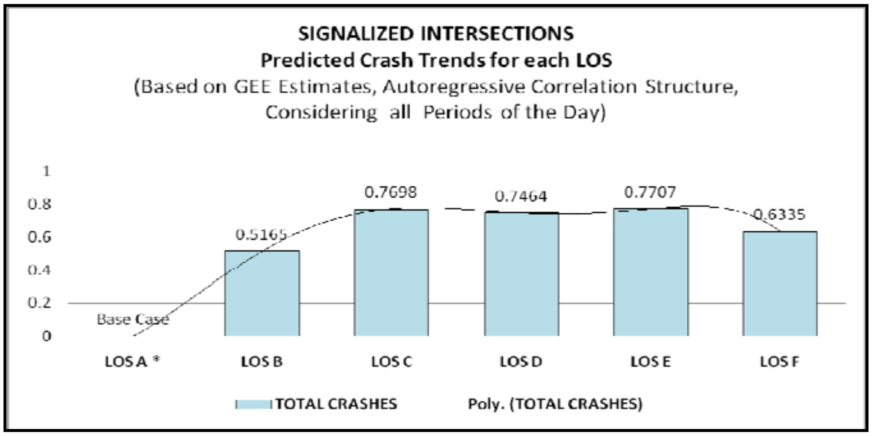


Figure 6-1: Predicted Trend for Model 1A ("Total Crashes") Based on Estimates for LOS for the Intersection as a Whole

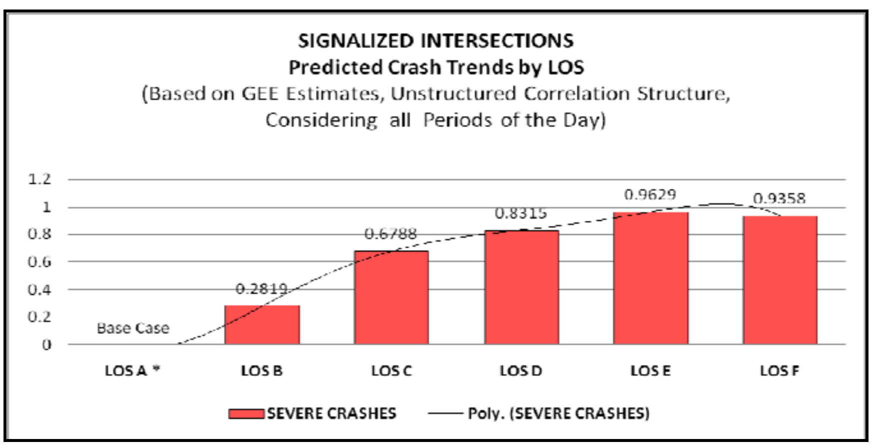


Figure 6-2: Predicted Trend for Model 2A ("Severe Crashes") Based on Estimates for LOS for the Intersection as a Whole

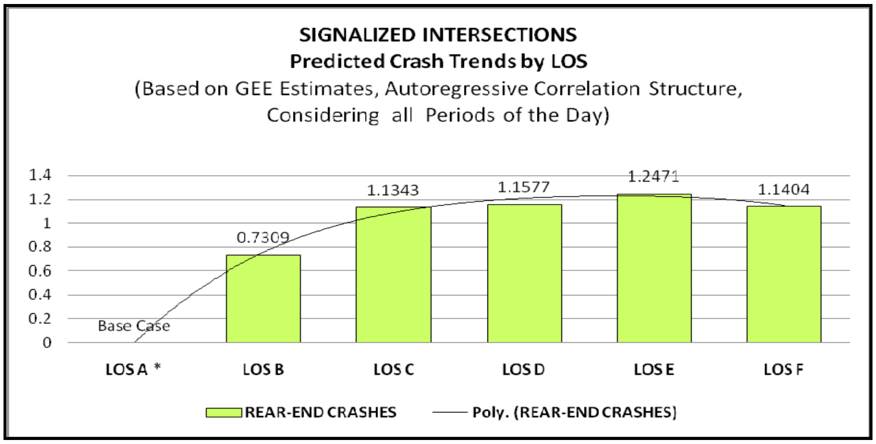


Figure 6-3: Predicted Trend for Model 3A ("Rear-End Crashes") Based on Estimates for LOS for the Intersection as a Whole

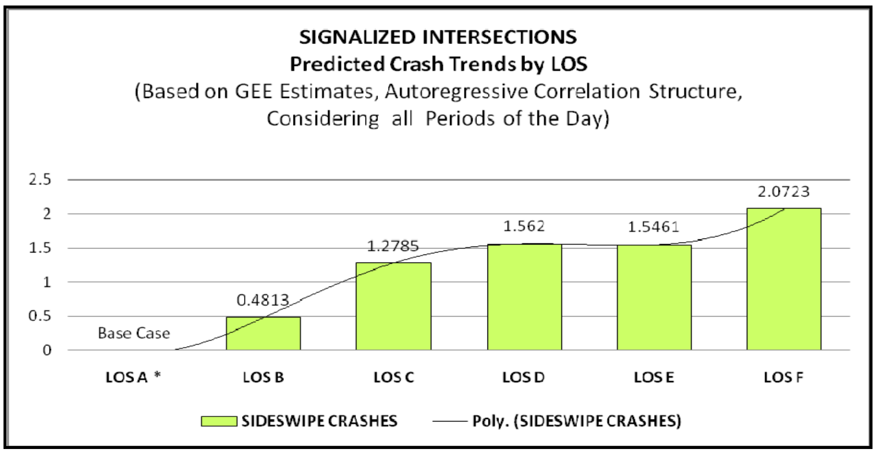


Figure 6-4: Predicted Trend for Model 4A ("Sideswipe Crashes") Based on Estimates for LOS for the Intersection as a Whole

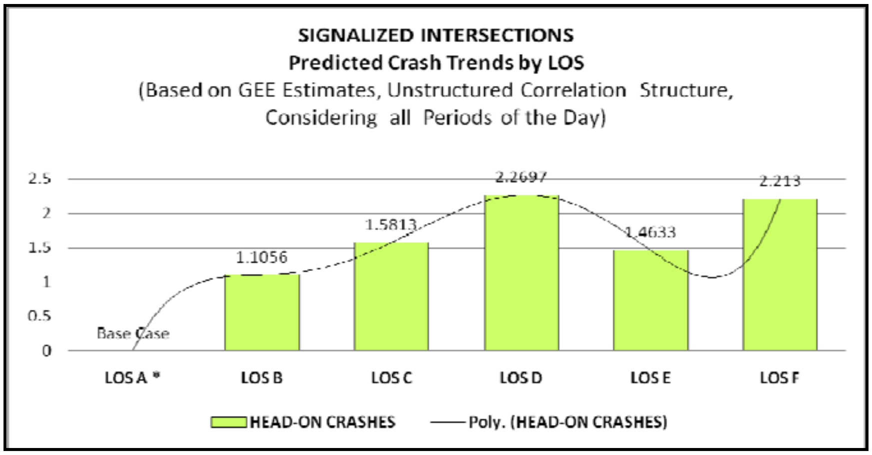


Figure 6-5: Predicted Trend for Model 5A ("Head-On Crashes") Based on Estimates for LOS for the Intersection as a Whole

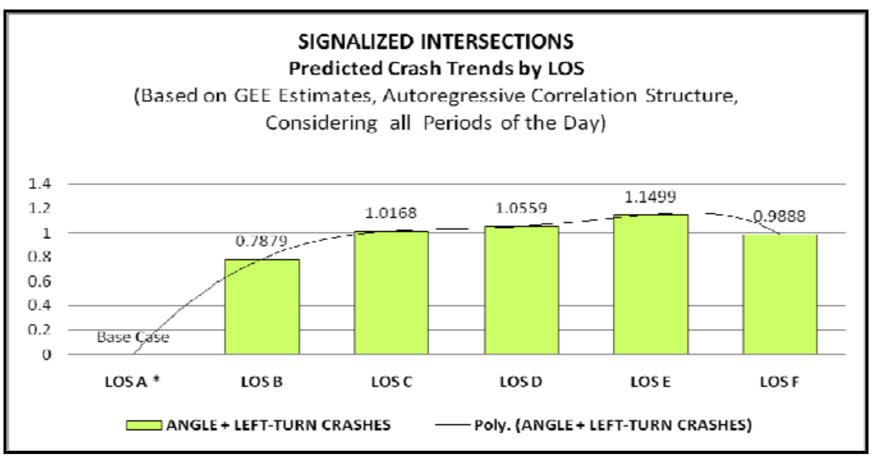


Figure 6-6: Predicted Trend for Model 6A ("Angle and Left-Turn Crashes") Based on Estimates for LOS for the Intersection as a Whole

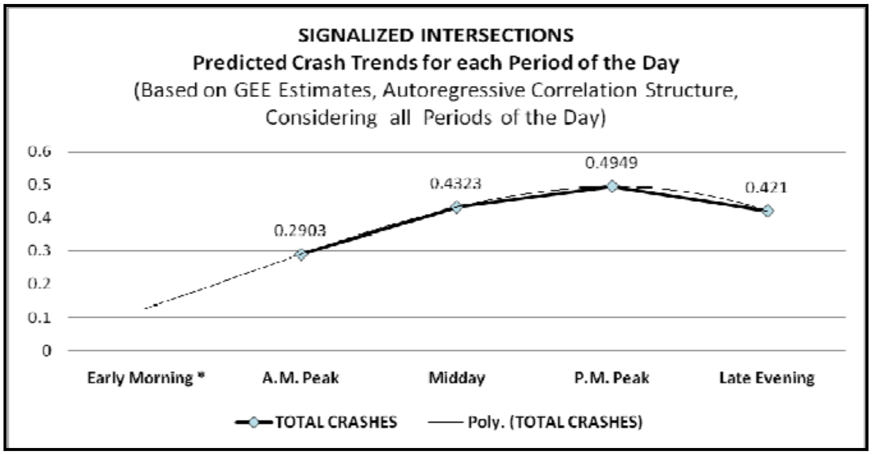


Figure 6-7: Predicted Trend for Model 1A ("Total Crashes") Based on Estimates for Period of the Day Variable

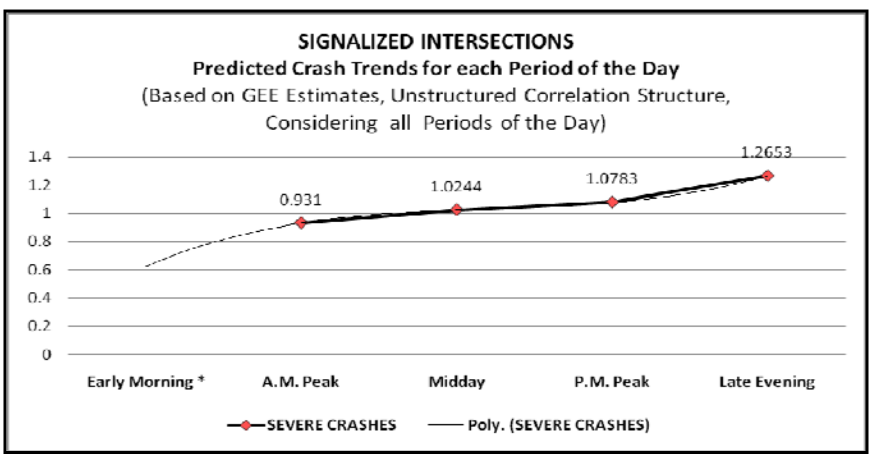


Figure 6-8: Predicted Trend for Model 2A ("Severe Crashes") Based on Estimates for Period of the Day Variable

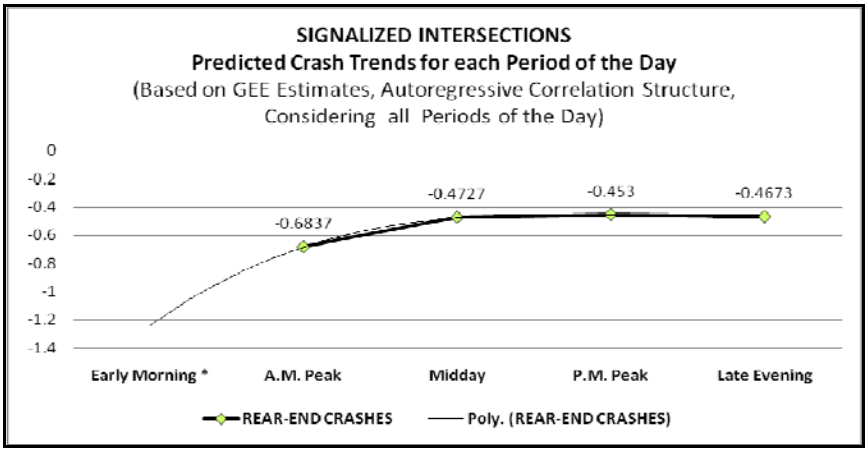


Figure 6-9: Predicted Trend for Model 3A ("Rear-End Crashes") Based on Estimates for Period of the Day Variable

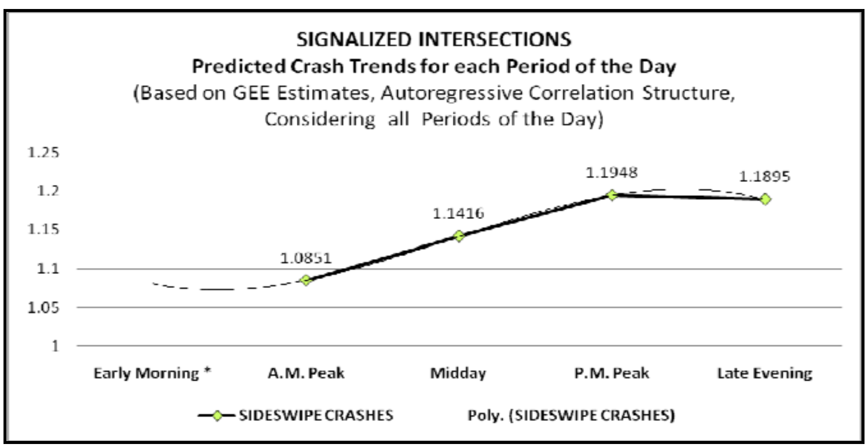


Figure 6-10: Predicted Trend for Model 4A ("Sideswipe Crashes") Based on Estimates for Period of the Day Variable

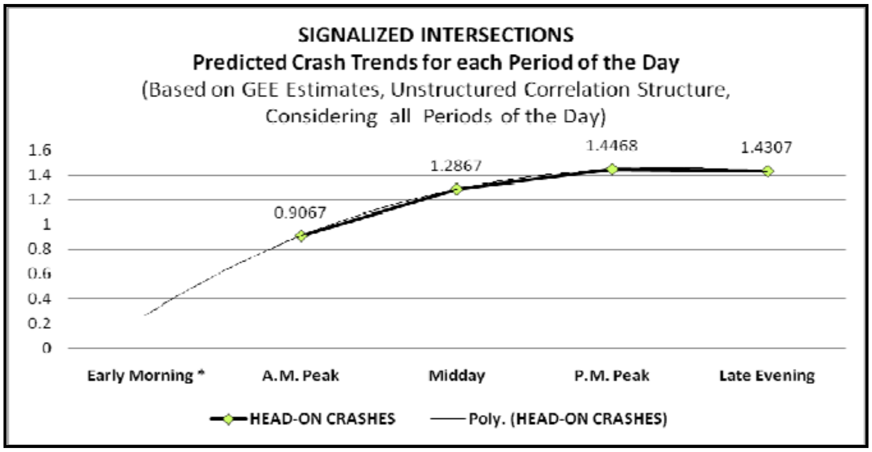


Figure 6-11: Predicted Trend for Model 5A ("Head-On Crashes") Based on Estimates for Period of the Day Variable

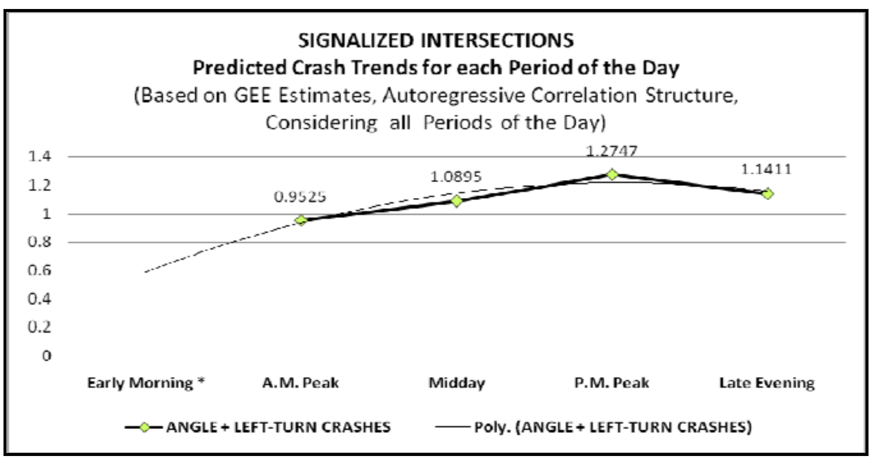


Figure 6-12: Predicted Trend for Model 6A ("Angle and Left-Turn Crashes") Based on Estimates for Period of the Day Variable

6.2.1.2 Applications and Contributions

6.2.1.2.1 Traffic Safety Guidelines

The results obtained have contributed to the preparation of a summary of "Traffic Safety Guidelines" based on the significant factors just presented. Overall, these have been organized by contributing factor and its potential effect on crash occurrence, as well as by crash type and risk associated to the respective factor; a reference to the corresponding models is also provided. The uniqueness of these guidelines relies on the fact that they are derived from the LOS-based models discussed throughout this thesis, which all have been statistically tested; in addition, they are clear and concise.

6.2.1.2.1.1 "Intersection Features-Based" Traffic Safety Guideline

The first set of traffic safety guidelines are derived from Model Group A, which are the models from the analysis of Signalized Intersections considering the 5 periods of the day (refer to Section 5.2.1.2). Following is Table 6-2 with the set of such guidelines, based on the significant factors related to the overall features of signalized intersections.

Table 6-2: "Intersection Features-Based" Traffic Safety Guideline for Signalized Intersections, Derived from Model Group A

SIGNALIZED INTERSECTIONS					
MODEL GROUP A: TRAFFIC SAFETY GUIDELINE					
Based on Signalized Intersections' Overall Features					
Approach(es)	FACTOR		CRASH OCCURRENCE		Reference
	Name	Level	Type	Risk or Effect	
Major & Minor	Cycle Length		TOTAL	Increase	MODEL 1A
			SIDESWIPE	Increase	MODEL 4A
Major	Left-Turn Traffic Volume		TOTAL	Increase	MODEL 1A
Minor	Total (Left+Through+Right) Traffic Volume		TOTAL	Increase	MODEL 1A
			REAR-END	Increase	MODEL 3A
			ANGLE + LEFT-TURN	Increase	MODEL 6A
Major	Total (Left+Through+Right) Traffic Volume		REAR-END	Increase	MODEL 3A
Major	Total (Left+Through+Right) Number of Lanes	9 or 10 lanes	SEVERE	HIGH	MODEL 2A
		5 or 6 lanes	SEVERE	MED-HIGH	MODEL 2A
		7 or 8 lanes	SEVERE	MEDIUM	MODEL 2A
		11 ≤ lanes ≤ 14	SEVERE	LOW	MODEL 2A
		4 lanes	SEVERE	(Base Case)	MODEL 2A
		11 ≤ lanes ≤ 14	SIDESWIPE	HIGH	MODEL 4A
		9 or 10 lanes	SIDESWIPE	MED-HIGH	MODEL 4A
		7 or 8 lanes	SIDESWIPE	MEDIUM	MODEL 4A
		5 or 6 lanes	SIDESWIPE	LOW	MODEL 4A
		4 lanes	SIDESWIPE	(Base Case)	MODEL 4A
Major	Total (Through) Number of Lanes	8 or 9 lanes	ANGLE + LEFT-TURN	HIGH	MODEL 6A
		4 lanes	ANGLE + LEFT-TURN	MEDIUM	MODEL 6A
		5 ≤ lanes ≤ 7	ANGLE + LEFT-TURN	LOW	MODEL 6A
		2 lanes	ANGLE + LEFT-TURN	(Base Case)	MODEL 6A
Minor	Total (Through) Number of Lanes	3 lanes	ANGLE + LEFT-TURN	HIGH	MODEL 6A
		4 ≤ lanes ≤ 6	ANGLE + LEFT-TURN	LOW	MODEL 6A
		2 lanes	ANGLE + LEFT-TURN	(Base Case)	MODEL 6A
Major	Total (Left) Number of Lanes	3 or 4 lanes	ANGLE + LEFT-TURN	HIGH	MODEL 6A
		5 lanes	ANGLE + LEFT-TURN	LOW	MODEL 6A
		1 or 2 lanes	ANGLE + LEFT-TURN	(Base Case)	MODEL 6A

Based on the table just presented, the following conclusions can be made with regards to safety conditions of signalized intersections in general:

- Cycle length contributes to the increase of total and sideswipe crashes. This could be explained by the fact that longer cycle lengths constitute a bigger time frame for a crash to occur; drivers, knowing a cycle length is long enough, feel with more time and/or freedom to pass other vehicles or do other maneuvers which, if not done properly, could result in a crash.
- Left-turn traffic volume on the major road contributes to the increase of total crashes. This could be explained by the fact that left-turn traffic at signalized intersections is usually large, reason for which more vehicles are exposed to a possible crash event.
- Total (left, through and right) traffic volume on the minor road contributes to the increase of total, rear-end, as well as angle and left-turn crashes. This could be explained by the fact that vehicles traveling on the major road have flow priority when compared to vehicles traveling on the minor road, reason for which the latter tend to be more exposed to a possible crash event.
- Total (left, through and right) traffic volume on the major road contributes to the increase of rear-end crashes. This could be explained by the fact that vehicles traveling on the major road usually travel at higher speeds, reason for which at the moment to stop (i.e. red phase) these are exposed to a crash of this type.
- Total (left, through and right) number of lanes on the major road contributes to the increase of severe crashes.

- More specifically, and considering 4 lanes as the base case, the results suggest that the risk for severe crashes is higher while having 9 or 10 as total number of lanes on the major road; on the other hand, the results also suggest that the risk for severe crashes is lower while having 11 $\leq$ lanes $\leq$ 14 on the major road.

The same factor, total (left, through and right) number of lanes on the major road, also contributes to the increase of sideswipe crashes.

- More specifically, and considering 4 lanes as the base case, the results suggest that the risk for sideswipe crashes is higher while having 11 $\leq$ lanes $\leq$ 14 on the major road; on the other hand, the results also suggest that the risk for sideswipe crashes is lower while having 5 or 6 lanes on the major road.

- Total number of through lanes on the major road contributes to the increase of angle and left-turn crashes.
  - More specifically, and considering 2 through lanes as the base case, the results suggest that the risk for angle and left-turn crashes is higher while having 8 or 9 through lanes on the major road; on the other hand, the results also suggest that the risk for angle and left-turn crashes is lower while having 5 $\leq$ through lanes $\leq$ 7 on the major road.
- Total number of through lanes on the minor road contributes to the increase of angle and left-turn crashes.
  - More specifically, and considering 2 through lanes as the base case, the results suggest that the risk for angle and left-turn crashes is higher while

having 3 through lanes on the minor road; on the other hand, the results also suggest that the risk for angle and left-turn crashes is lower while having 4 $\leq$ through lanes $\leq$ 6 on the minor road.

- Total number of left-turn lanes on the major road contributes to the increase of angle and left-turn crashes.
    - More specifically, and considering 1 or 2 lanes as the base case, the results suggest that the risk for angle and left-turn crashes is higher while having 3 or 4 left-turn lanes on the major road; on the other hand, the results also suggest that the risk for angle and left-turn crashes is lower while having 5 left-turn lanes on the major road.

The reader may refer to Table 6-2 for more accurate details and a complete review of the guidelines just presented.

6.2.1.2.1.2 "Time-Based" Traffic Safety Guideline

A second set of traffic safety guidelines were derived from Model Group A (refer to Section 6.2.1.2.1.1). One of the features of this second set of guidelines is that they add the time factor (i.e. period of the day) into this safety framework. Following is Table 6-3 with the set of such guidelines, based on the "period of the day" factor.

Table 6-3: “Time-Based” Traffic Safety Guideline for Signalized Intersections, Derived from Model Group A

SIGNALIZED INTERSECTIONS				
MODEL GROUP A: "TIME-BASED" TRAFFIC SAFETY GUIDELINE				
FACTOR		CRASH		Reference
Period of the Day		OCCURRENCE		
Name	Hours	Risk	Type	
P.M. Peak	4 - 6 P.M.	HIGH	TOTAL	MODEL 1A
Midday	12 - 2 P.M.	MED-HIGH	TOTAL	MODEL 1A
Late Evening	8 - 10 P.M.	MEDIUM	TOTAL	MODEL 1A
A.M. Peak	7 - 9 A.M.	LOW	TOTAL	MODEL 1A
Early Morning	1 - 3 A.M.	(Base Case)	TOTAL	MODEL 1A
Late Evening	8 - 10 P.M.	HIGH	SEVERE	MODEL 2A
P.M. Peak	4 - 6 P.M.	MED-HIGH	SEVERE	MODEL 2A
Midday	12 - 2 P.M.	MEDIUM	SEVERE	MODEL 2A
A.M. Peak	7 - 9 A.M.	LOW	SEVERE	MODEL 2A
Early Morning	1 - 3 A.M.	(Base Case)	SEVERE	MODEL 2A
P.M. Peak	4 - 6 P.M.	HIGH	REAR-END	MODEL 3A
Late Evening	8 - 10 P.M.	MED-HIGH	REAR-END	MODEL 3A
Midday	12 - 2 P.M.	MEDIUM	REAR-END	MODEL 3A
A.M. Peak	7 - 9 A.M.	LOW	REAR-END	MODEL 3A
Early Morning	1 - 3 A.M.	(Base Case)	REAR-END	MODEL 3A
P.M. Peak	4 - 6 P.M.	HIGH	SIDESWIPE	MODEL 4A
Late Evening	8 - 10 P.M.	MED-HIGH	SIDESWIPE	MODEL 4A
Midday	12 - 2 P.M.	MEDIUM	SIDESWIPE	MODEL 4A
A.M. Peak	7 - 9 A.M.	LOW	SIDESWIPE	MODEL 4A
Early Morning	1 - 3 A.M.	(Base Case)	SIDESWIPE	MODEL 4A
P.M. Peak	4 - 6 P.M.	HIGH	HEAD-ON	MODEL 5A
Late Evening	8 - 10 P.M.	MED-HIGH	HEAD-ON	MODEL 5A
Midday	12 - 2 P.M.	MEDIUM	HEAD-ON	MODEL 5A
A.M. Peak	7 - 9 A.M.	LOW	HEAD-ON	MODEL 5A
Early Morning	1 - 3 A.M.	(Base Case)	HEAD-ON	MODEL 5A
P.M. Peak	4 - 6 P.M.	HIGH	ANGLE + LEFT-TURN	MODEL 6A
Late Evening	8 - 10 P.M.	MED-HIGH	ANGLE + LEFT-TURN	MODEL 6A
Midday	12 - 2 P.M.	MEDIUM	ANGLE + LEFT-TURN	MODEL 6A
A.M. Peak	7 - 9 A.M.	LOW	ANGLE + LEFT-TURN	MODEL 6A
Early Morning	1 - 3 A.M.	(Base Case)	ANGLE + LEFT-TURN	MODEL 6A

Based on the “time-based” table just presented, the following conclusions can be made with regards to safety conditions of signalized intersections in general:

- With regards to total crashes, and considering the Early Morning period (1-3 A.M.) as the base case, the results suggest that this crash risk is higher during the P.M. Peak period (4-6 P.M.); on the other hand, the results suggest that this crash risk is lower during the earlier periods of the day.
- With regards to severe crashes, and considering the Early Morning period (1-3 A.M.) as the base case, the results suggest that this crash risk is higher during the Late Evening period (8-10 P.M.); on the other hand, the results suggest that this crash risk is lower during the earlier periods of the day.
- With regards to rear-end, sideswipe, head-on, as well as angle and left-turn crashes, and considering the Early Morning period (1-3 A.M.) as the base case, the results suggest that these types of crash risk are higher during the P.M. Peak period (4-6 P.M.); on the other hand, the results suggest that these types of crash risk are lower during the earlier periods of the day.

The reader may refer to Table 6-3 for more accurate details and a complete review of the guidelines just presented.

6.2.1.2.1.3 “LOS-Based” Traffic Safety Guideline

A third set of traffic safety guidelines were derived from Model Group A (refer to Section 6.2.1.1). One of the features of this third set of guidelines is that they add the LOS indicator as a contributing factor into this safety framework. Following is Table 6-4 with the set of such guidelines, based on the “LOS” factor.

Table 6-4: "LOS-Based" Traffic Safety Guideline for Signalized Intersections, Derived from Model Group A

SIGNALIZED INTERSECTIONS				
MODEL GROUP A: "LOS-BASED" TRAFFIC SAFETY GUIDELINE				
FACTOR		CRASH		Reference
LOS		OCCURRENCE		
Level	Delay	Risk	Type	
E	Significant	HIGH	TOTAL	MODEL 1A
C	Acceptable	MED-HIGH	TOTAL	MODEL 1A
D	Tolerable	MEDIUM	TOTAL	MODEL 1A
F	Excessive	MED-LOW	TOTAL	MODEL 1A
B	Minimal	LOW	TOTAL	MODEL 1A
A	Insignificant	(Base Case)	TOTAL	MODEL 1A
E	Significant	HIGH	SEVERE	MODEL 2A
F	Excessive	MED-HIGH	SEVERE	MODEL 2A
D	Tolerable	MEDIUM	SEVERE	MODEL 2A
C	Acceptable	MED-LOW	SEVERE	MODEL 2A
B	Minimal	LOW	SEVERE	MODEL 2A
A	Insignificant	(Base Case)	SEVERE	MODEL 2A
E	Significant	HIGH	REAR-END	MODEL 3A
D	Tolerable	MED-HIGH	REAR-END	MODEL 3A
F	Excessive	MEDIUM	REAR-END	MODEL 3A
C	Acceptable	MED-LOW	REAR-END	MODEL 3A
B	Minimal	LOW	REAR-END	MODEL 3A
A	Insignificant	(Base Case)	REAR-END	MODEL 3A
F	Excessive	HIGH	SIDESWIPE	MODEL 4A
D	Tolerable	MED-HIGH	SIDESWIPE	MODEL 4A
E	Significant	MEDIUM	SIDESWIPE	MODEL 4A
C	Acceptable	MED-LOW	SIDESWIPE	MODEL 4A
B	Minimal	LOW	SIDESWIPE	MODEL 4A
A	Insignificant	(Base Case)	SIDESWIPE	MODEL 4A
D	Tolerable	HIGH	HEAD-ON	MODEL 5A
F	Excessive	MED-HIGH	HEAD-ON	MODEL 5A
C	Acceptable	MEDIUM	HEAD-ON	MODEL 5A
E	Significant	MED-LOW	HEAD-ON	MODEL 5A
B	Minimal	LOW	HEAD-ON	MODEL 5A
A	Insignificant	(Base Case)	HEAD-ON	MODEL 5A
E	Significant	HIGH	ANGLE + LEFT-TURN	MODEL 6A
D	Tolerable	MED-HIGH	ANGLE + LEFT-TURN	MODEL 6A
C	Acceptable	MEDIUM	ANGLE + LEFT-TURN	MODEL 6A
F	Excessive	MED-LOW	ANGLE + LEFT-TURN	MODEL 6A
B	Minimal	LOW	ANGLE + LEFT-TURN	MODEL 6A
A	Insignificant	(Base Case)	ANGLE + LEFT-TURN	MODEL 6A

Based on the "LOS-based" table just presented, the following conclusions can be made with regards to safety conditions of signalized intersections in general:

- Considering LOS A (insignificant delay) as the base case, the results suggest that the risk for total, severe, rear-end, as well as angle and left-turn crashes is higher for signalized intersections that have a LOS E (significant delay).
- Considering LOS A (insignificant delay) as the base case, the results suggest that the risk for sideswipe crashes is higher for signalized intersections that have a LOS F (excessive delay).
- Considering LOS A (insignificant delay) as the base case, the results suggest that the risk for head-on crashes is higher for signalized intersections that have a LOS D (tolerable delay).

The reader may refer to Table 6-4 for more accurate details and a complete review of the guidelines just presented.

6.2.2 Disaggregate Analysis (Considering one Period of the Day at a time)

6.2.2.1 Models' Summary and Implications

Recalling the corresponding modeling approach described in Chapter 5, five final models (i.e. Model Group B) resulted after using the Negative Binomial technique; this method was used for analyzing the relationship between the LOS and crash occurrence for each period of the day independently (i.e. disaggregate analysis). The results of this approach would provide more specific insights for this study. Following is Table 6-5, which contains the summary of all the resulting models, detailing the corresponding significant factors.

Table 6-5: Summary of Signalized Intersections' Models (from Negative Binomial Analysis)

SIGNALIZED INTERSECTIONS	
Models' Summary	
Negative Binomial Analysis (Considering Each Period of the Day, Individually)	
MODEL	**SIGNIFICANT FACTORS** (α = 0.05) **(Listed from Highest to Lowest Degree of Significance)**
Model 1B **TOTAL** **CRASHES** **Early Morning** **(1-3 A.M.)**	**LEVEL-OF-SERVICE for the Intersection as a whole** (for the Period of the Day) **Log. of the Total Traffic Volume on the Major Road (for the Period of the Day)** **Lighting Conditions**
Model 2B **TOTAL** **CRASHES** **A.M. Peak** **(7-9 A.M.)**	**Log. of the Cycle Length** (for the Period of the Day) **LEVEL-OF-SERVICE for the Intersection as a whole** (for the Period of the Day) **Log. of the Total Left-Turn Traffic Volume on the Major Road** (for the Period of the Day) **Total Number of Lanes on the Major Road**
Model 3B **TOTAL** **CRASHES** **Midday** **(12-2 P.M.)**	**Log. of the Total Traffic Volume on the Major Road (for the Period of the Day)** **Log. of the Total Traffic Volume on the Minor Road** (for the Period of the Day) **LEVEL-OF-SERVICE for the Intersection as a whole** (for the Period of the Day)
Model 4B **TOTAL** **CRASHES** **P.M. Peak** **(4-6 P.M.)**	**Total Number of Lanes on the Major Road** **Log. of the Total Left-Turn Traffic Volume on the Major Road** (for the Period of the Day) **LEVEL-OF-SERVICE for the Intersection as a whole** (for the Period of the Day) **Log. of the Cycle Length** (for the Period of the Day)
Model 5B **TOTAL** **CRASHES** **Late Evening** **(8-10 P.M.)**	**Log. of the Total Traffic Volume on the Major Road (for the Period of the Day)** **Log. of the Total Traffic Volume on the Minor Road** (for the Period of the Day) **LEVEL-OF-SERVICE for the Intersection as a whole** (for the Period of the Day) **Land Use**

As it can be seen, for this part of the investigation the LOS was also found to be a significant predictor in every model. In order to better understand the results, the following sections will detail, per model, the most important findings and contributions.

6.2.2.1.1 Significant Factors

As mentioned before, crash occurrence at signalized intersections can be generally attributed to the following categories of contributing factors:

- Geometric design features (i.e. number of lanes)
- Location (i.e. County) (dummy variable)
- Time (i.e. Period of the Day) (dummy variable)
- Traffic characteristics (i.e. traffic volumes)
- Traffic control and operational measures (i.e. cycle length)
- Traffic and operation conditions as a whole (i.e. LOS).

Being LOS the predictor of interest in the modeling process, and considering LOS A (insignificant delay) as the base case, following are plots derived from Model Group B's estimates (i.e. coefficients) for this predictor (see Figures 6-13 through 6-18); these were used in the preparation of the safety guidelines detailed in the following section.

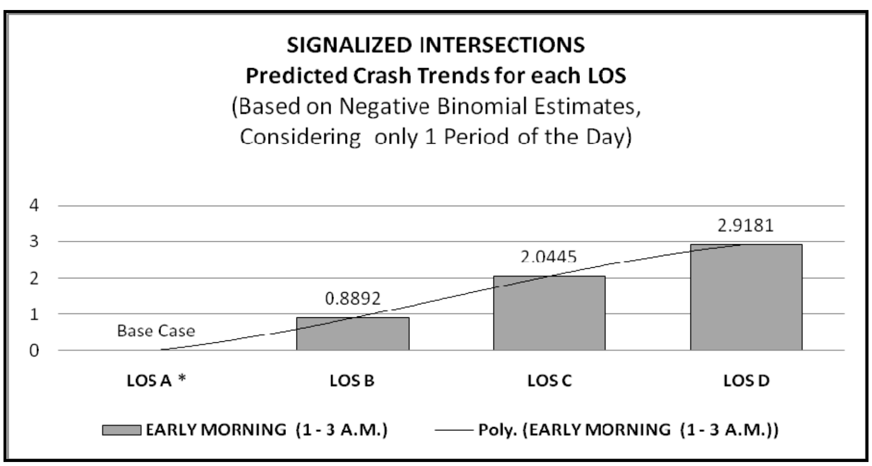


Figure 6-13: Predicted Trend for Model 1B ("Total Crashes", Considering only the Early Morning Period) Based on Estimates for LOS for the Intersection as a Whole

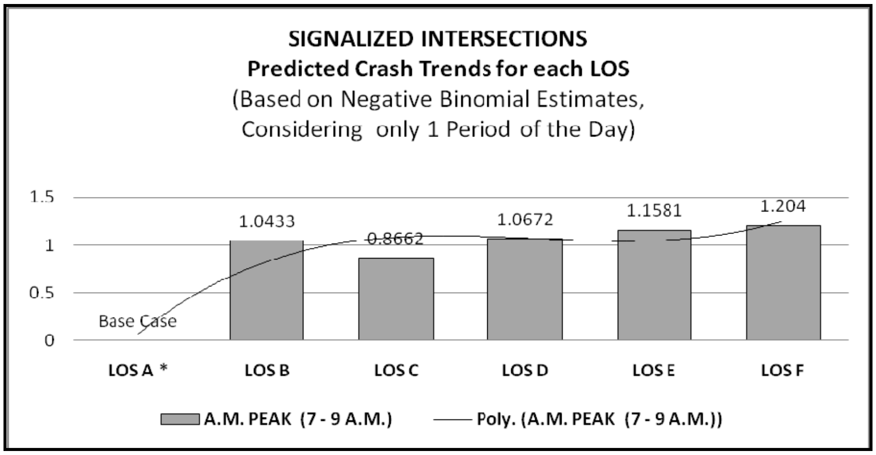


Figure 6-14: Predicted Trend for Model 2B ("Total Crashes", Considering only the A.M. Peak Period) Based on Estimates for LOS for the Intersection as a Whole

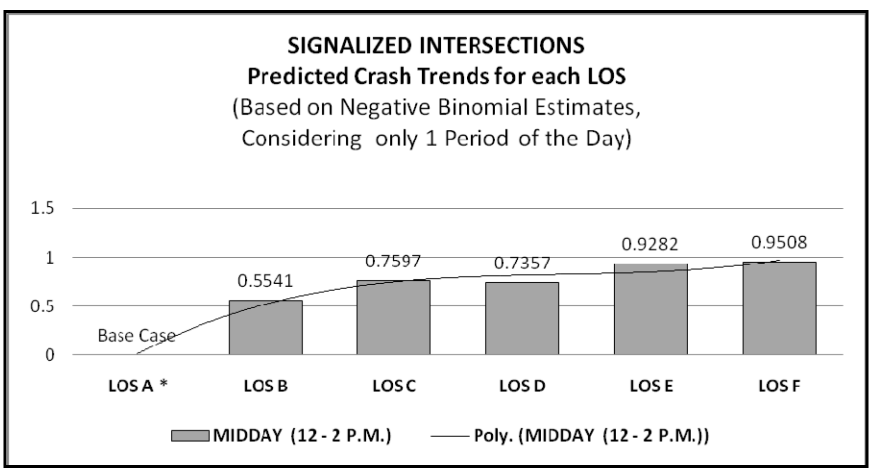


Figure 6-15: Predicted Trend for Model 3B ("Total Crashes", Considering only the Midday Period) Based on Estimates for LOS for the Intersection as a Whole

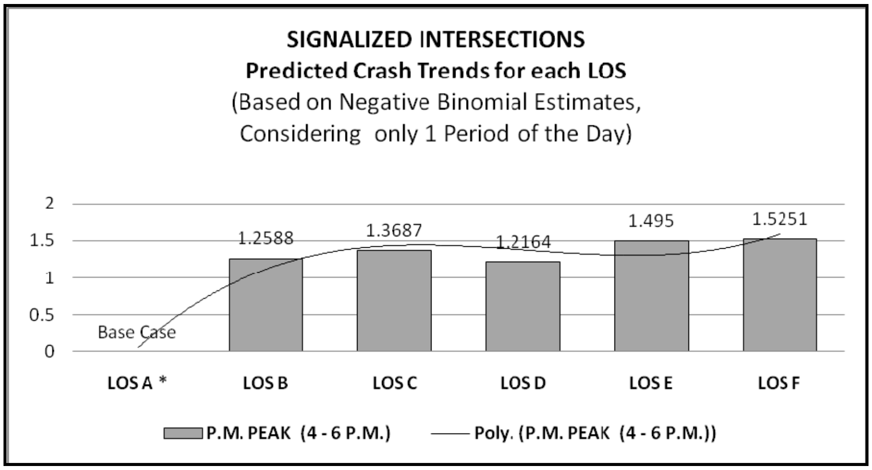


Figure 6-16: Predicted Trend for Model 4B ("Total Crashes", Considering only the P.M. Peak Period) Based on Estimates for LOS for the Intersection as a Whole

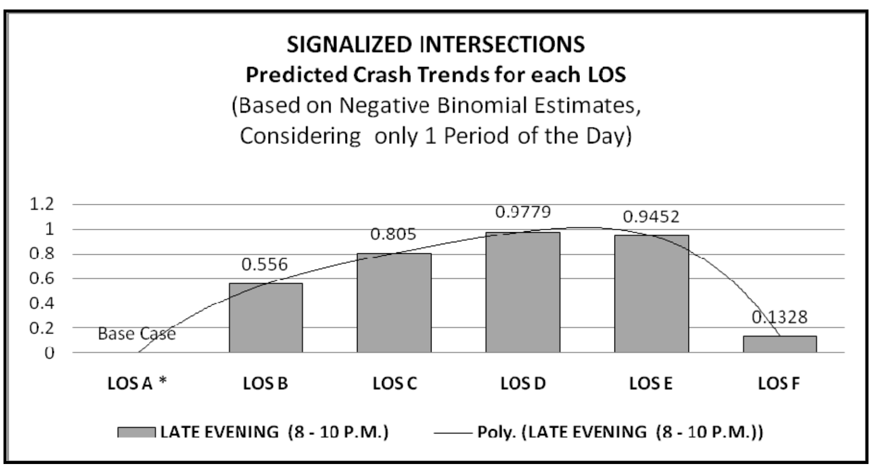


Figure 6-17: Predicted Trend for Model 5B ("Total Crashes", Considering only the Late Evening Period) Based on Estimates for LOS for the Intersection as a Whole

6.2.2.2 Applications and Contributions

6.2.2.2.1 Traffic Safety Guidelines

The results obtained have contributed to the preparation of a second set of "Traffic Safety Guidelines" based on the significant factors just presented. These guidelines have been organized as mentioned in Section 6.2.1.2.1. One of the features of this second set of guidelines is that they add the time factor (i.e. period of the day) into the table.

6.2.2.2.1.1 "Intersection Features-Based" Traffic Safety Guideline

Following is the first set of traffic safety guidelines derived from Model Group B, constituted by the models from the analysis of Signalized Intersections considering one period of the day at a time (refer to Section 5.2.2.2). Following is Table 6-6 with the set of such guidelines, based on the significant factors related to the overall features of signalized intersections.

Table 6-6: "Intersection Features-Based" Traffic Safety Guideline for Signalized Intersections, Derived from Model Group B

SIGNALIZED INTERSECTIONS						
MODEL GROUP B: TRAFFIC SAFETY GUIDELINE						
Based on Signalized Intersections' Overall Features						
Approach(es)	**FACTOR**		**CRASH OCCURRENCE**			**Reference**
	Name	**Level**	**Type**	**Hours**	**Risk or Effect**	
Major	**Total (Left+Through+Right) Traffic Volume**		TOTAL	1 - 3 A.M.	Increase	MODEL 1B
			TOTAL	12 - 2 P.M.	Increase	MODEL 3B
			TOTAL	8 - 10 P.M.	Increase	MODEL 5B
Minor	**Total (Left+Through+Right) Traffic Volume**		TOTAL	12 - 2 P.M.	Increase	MODEL 3B
			TOTAL	8 - 10 P.M.	Increase	MODEL 5B
Major	**Left-Turn Traffic Volume**		TOTAL	7 - 9 A.M.	Increase	MODEL 2B
			TOTAL	4 - 6 P.M.	Increase	MODEL 4B
Major & Minor	**Cycle Length**		TOTAL	7 - 9 A.M.	Increase	MODEL 2B
			TOTAL	4 - 6 P.M.	Increase	MODEL 4B
Major	**Total (Left+Through+Right) Number of Lanes**	11 ≤ lanes ≤ 14	TOTAL	7 - 9 A.M.	HIGH	MODEL 2B
		9 or 10 lanes	TOTAL	7 - 9 A.M.	MED-HIGH	MODEL 2B
		5 or 6 lanes	TOTAL	7 - 9 A.M.	MEDIUM	MODEL 2B
		7 or 8 lanes	TOTAL	7 - 9 A.M.	LOW	MODEL 2B
		4 lanes	TOTAL	7 - 9 A.M.	(Base Case)	MODEL 2B
		5 or 6 lanes	TOTAL	4 - 6 P.M.	HIGH	MODEL 4B
		11 ≤ lanes ≤ 14	TOTAL	4 - 6 P.M.	MED-HIGH	MODEL 4B
		9 or 10 lanes	TOTAL	4 - 6 P.M.	MEDIUM	MODEL 4B
		7 or 8 lanes	TOTAL	4 - 6 P.M.	LOW	MODEL 4B
		4 lanes	TOTAL	4 - 6 P.M.	(Base Case)	MODEL 4B
Major & Minor	**Lighting Conditions**	Partial or Full	TOTAL	1 - 3 A.M.	Reduction	MODEL 1B
		None	TOTAL	1 - 3 A.M.	(Base Case)	MODEL 1B
Major & Minor	**Land Use**	Suburban or Urban	TOTAL	8 - 10 P.M.	Reduction	MODEL 5B
		Rural	TOTAL	8 - 10 P.M.	(Base Case)	MODEL 5B

Based on the table just presented, the following conclusions can be made with regards to safety conditions of signalized intersections in general:

- Total (left, through and right) traffic volume on the major road contributes to the increase of total crashes; this applies to the Early Morning (1-3 A.M.), Midday (12-2 P.M.) and Late Evening (8-10 P.M.) models. This could be explained by the fact that vehicles traveling on the major road at these periods of the day can usually travel at higher speeds, reason for which at the moment to stop (i.e. red phase) these are exposed to a possible crash event.
- Total (left, through and right) traffic volume on the minor road contributes to the increase of total crashes; this applies to the Midday (12-2 P.M.) and Late Evening (8-10 P.M.) models. This could be explained by the fact that vehicles traveling on the minor road at these periods of the day can usually travel at higher speeds, reason for which at the moment to stop (i.e. red phase) these are exposed to a possible crash event.
- Left-turn traffic volume on the major road contributes to the increase of total crashes; this applies to the A.M. Peak (7-9 A.M.) and P.M. Peak (4-6 P.M.) models. This could be explained by the fact that left-turn traffic at signalized intersections is usually large during these periods of the day, reason for which more vehicles are exposed to a possible crash event.
- Cycle length contributes to the increase of total crashes; this applies to the A.M. Peak (7-9 A.M.) and P.M. Peak (4-6 P.M.) models. This could be explained by the fact that longer cycle lengths constitute a bigger time frame for a crash to occur; drivers, knowing a cycle length is long enough during these periods of the day, feel with more time and/or freedom to maneuver which, if not done properly, could result in a crash.

- Total (left, through and right) number of lanes on the major road contributes to the increase of total crashes; this applies to the A.M. Peak (7-9 A.M.) and P.M. Peak (4-6 P.M.) models.
    - More specifically, focusing on the A.M. Peak period (7-9 A.M.) model, and considering 4 lanes as the base case, the results suggest that the risk for total crashes is higher while having $11 \leq$ lanes ≤ 14 on the major road; on the other hand, the results also suggest that the risk for total crashes is lower while having 7 or 8 lanes on the major road.
    - Furthermore, focusing on the P.M. Peak period (4-6 P.M.) model, and considering 4 lanes as the base case, the results suggest that the risk for total crashes is higher while having 5 or 6 lanes on the major road; on the other hand, the results also suggest that the risk for total crashes is lower while having 7 or 8 lanes on the major road.
- Lighting conditions appear to contribute to the reduction of total crashes; this applies to the Early Morning period (1-3 A.M.) model.
    - More specifically, focusing on the Early Morning period (1-3 A.M.) model, and considering lack of lighting as the base case, the results suggest that the risk for total crashes is reduced while having partially or fully illuminated signalized intersections.
- The land use, an indicator of the location of signalized intersections, appears to contribute to the reduction of total crashes; this applies to the Late Evening period (8-10 P.M.) model.

- More specifically, focusing on the Late Evening period (8-10 P.M.) model, and considering rural areas as the base case, the results suggest that signalized intersections located in suburban or urban areas are associated with a lower risk for total crashes.

The reader may refer to Table 6-6 for more accurate details and a complete review of the guidelines just presented.

6.2.2.2.1.2 “LOS- and Time-Based” Traffic Safety Guideline

A second set of traffic safety guidelines were derived from Model Group B (refer to Section 6.2.2.2.1.1). One of the features of this second set of guidelines is that they combine both the LOS indicator and time factor into this safety framework. Following is Table 6-7 with the set of such guidelines, based on both the “LOS” and “period of the day” factors.

Table 6-7: "LOS- and Time-Based" Traffic Safety Guideline for Signalized Intersections, Derived from Model Groups A & B

SIGNALIZED INTERSECTIONS						
MODEL GROUPS A & B: "LOS- & TIME-BASED" TRAFFIC SAFETY GUIDELINE						
FACTORS				CRASH OCCURRENCE		Reference
Period of the Day		LOS				
Risk	Hours	Level	Delay	Risk	Type	
HIGH	4 - 6 P.M.	F	Excessive	HIGH	TOTAL	MODEL 4B
	4 - 6 P.M.	E	Significant	MED-HIGH	TOTAL	MODEL 4B
	4 - 6 P.M.	C	Acceptable	MEDIUM	TOTAL	MODEL 4B
	4 - 6 P.M.	B	Minimal	MED-LOW	TOTAL	MODEL 4B
	4 - 6 P.M.	D	Tolerable	LOW	TOTAL	MODEL 4B
	4 - 6 P.M.	A	Insignificant	(Base Case)	TOTAL	MODEL 4B
MED-HIGH	12 - 2 P.M.	F	Excessive	HIGH	TOTAL	MODEL 3B
	12 - 2 P.M.	E	Significant	MED-HIGH	TOTAL	MODEL 3B
	12 - 2 P.M.	C	Acceptable	MEDIUM	TOTAL	MODEL 3B
	12 - 2 P.M.	D	Tolerable	MED-LOW	TOTAL	MODEL 3B
	12 - 2 P.M.	B	Minimal	LOW	TOTAL	MODEL 3B
	12 - 2 P.M.	A	Insignificant	(Base Case)	TOTAL	MODEL 3B
MEDIUM	8 - 10 P.M.	D	Tolerable	HIGH	TOTAL	MODEL 5B
	8 - 10 P.M.	E	Significant	MED-HIGH	TOTAL	MODEL 5B
	8 - 10 P.M.	C	Acceptable	MEDIUM	TOTAL	MODEL 5B
	8 - 10 P.M.	B	Minimal	MED-LOW	TOTAL	MODEL 5B
	8 - 10 P.M.	F	Excessive	LOW	TOTAL	MODEL 5B
	8 - 10 P.M.	A	Insignificant	(Base Case)	TOTAL	MODEL 5B
LOW-MED	7 - 9 A.M.	F	Excessive	HIGH	TOTAL	MODEL 2B
	7 - 9 A.M.	E	Significant	MED-HIGH	TOTAL	MODEL 2B
	7 - 9 A.M.	D	Tolerable	MEDIUM	TOTAL	MODEL 2B
	7 - 9 A.M.	B	Minimal	MED-LOW	TOTAL	MODEL 2B
	7 - 9 A.M.	C	Acceptable	LOW	TOTAL	MODEL 2B
	7 - 9 A.M.	A	Insignificant	(Base Case)	TOTAL	MODEL 2B
LOW	1 - 3 A.M.	D	Tolerable	HIGH	TOTAL	MODEL 1B
	1 - 3 A.M.	C	Acceptable	MEDIUM	TOTAL	MODEL 1B
	1 - 3 A.M.	B	Minimal	LOW	TOTAL	MODEL 1B
	1 - 3 A.M.	A	Insignificant	(Base Case)	TOTAL	MODEL 1B

Based on the table just presented, the following conclusions can be made with regards to safety conditions of signalized intersections in general:

- Considering the Early Morning period (1-3 A.M.) as the base case, the results suggest that the risk for total crashes is at its highest during the P.M. Peak period (4-6 P.M.).
    - Within this period of the day, and considering LOS A (insignificant delay) as the base case, the results suggest that the risk for total crashes is the highest for signalized intersections that have a LOS F (excessive delay).
    - Also, within this period of the day, and considering LOS A (insignificant delay) as the base case, the results suggest that the risk for total crashes is the 2nd highest for signalized intersections that have a LOS E (significant delay).
- Considering the Early Morning period (1-3 A.M.) as the base case, the results suggest that the risk for total crashes is 2nd highest during the Midday period (12-2 P.M.).
    - Within this period of the day, and considering LOS A (insignificant delay) as the base case, the results suggest that the risk for total crashes is the highest for signalized intersections that have a LOS F (excessive delay).
    - Also, within this period of the day, and considering LOS A (insignificant delay) as the base case, the results suggest that the risk for total crashes is the 2nd highest for signalized intersections that have a LOS E (significant delay).

The reader may refer to Table 6-7 for more accurate details and a complete review of the guidelines just presented.

6.3 LOS-Safety Study for Multilane High-Speed Arterial Corridors

6.3.1 Overall Analysis

6.3.1.1 Models' Summary and Implications

Recalling the corresponding modeling approach described in Chapter 5, six final models (i.e. Model Group C) resulted after using the Negative Binomial technique; this method was used for analyzing the relationship between the LOS and crash occurrence at multilane high-speed arterial corridors. Following is Table 6-8, which contains the summary of all the resulting models, detailing the corresponding significant factors.

Table 6-8: Summary of Multilane High-Speed Arterial Corridors' Models (from Negative Binomial Analysis)

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS	
Models' Summary	
Negative Binomial Analysis	
MODEL	**SIGNIFICANT FACTORS** (α = 0.05) **(Listed from Highest to Lowest Degree of Significance)**
Model 1C TOTAL CRASHES	**Total Number of Lanes in a Single Direction of the Road Section** **Total Length of the Road Section** **LEVEL-OF-SERVICE for the Road Section** (for the 100th highest traffic hour of the year) **Posted Speed Limit for the Road Section**
Model 2C SEVERE CRASHES	**Total Length of the Road Section** **Total Number of Lanes in a Single Direction of the Road Section** **LEVEL-OF-SERVICE for the Road Section** (for the 100th highest traffic hour of the year) **Pavement Surface Type of the Road Section** **Posted Speed Limit for the Road Section** **Inside Shoulder Width for the Road Section**
Model 3C REAR-END CRASHES	**Log. of the ADT for the Road Section** **Total Length of the Road Section** **LEVEL-OF-SERVICE for the Road Section** (for the 100th highest traffic hour of the year)
Model 4C SIDESWIPE CRASHES	**Total Length of the Road Section** **Total Number of Lanes in a Single Direction of the Road Section** **LEVEL-OF-SERVICE for the Road Section** (for the 100th highest traffic hour of the year)
Model 5C HEAD-ON CRASHES	**Median Width of the Road Section** **Posted Speed Limit for the Road Section** **Total Number of Lanes in a Single Direction of the Road Section** **LEVEL-OF-SERVICE for the Road Section** (for the 100th highest traffic hour of the year) **Inside Shoulder Width for the Road Section**
Model 6C ANGLE + LEFT-TURN CRASHES	**Total Number of Lanes in a Single Direction of the Road Section** **Total Length of the Road Section** **Posted Speed Limit for the Road Section** **Inside Shoulder Width for the Road Section** **LEVEL-OF-SERVICE for the Road Section** (for the 100th highest traffic hour of the year)

As it can be seen, for this part of the investigation the LOS was also found to be a significant predictor in every model. In order to better understand the results, the following sections will detail, per model, the most important findings and contributions.

6.3.1.1.1 Significant Factors

Based on the results obtained for this study, crash occurrence at multilane high-speed arterial corridors can be attributed to the following categories of contributing factors:

- Control and Geometric design features (i.e. posted speed limit, number of lanes, length of the road section, median width, inside shoulder width)
- Surface characteristics (i.e. pavement surface type)
- Traffic characteristics (i.e. ADT)
- Traffic and operating conditions as a whole (i.e. LOS).

Being LOS the predictor of interest in the modeling process for this study, and considering LOS A (free flow) as the base case, following are plots derived from Model Group C's estimates (i.e. coefficients) for this predictor (see Figures 6-18 through 6-23); these were used in the preparation of the safety guidelines detailed in the following section.

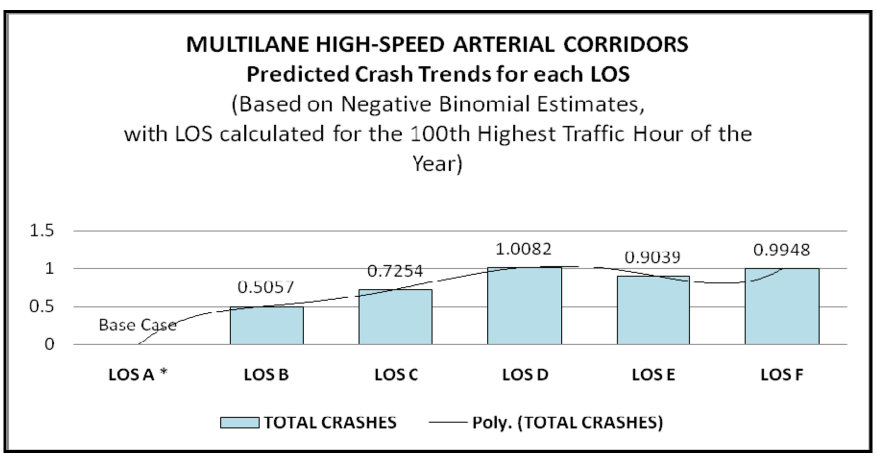


Figure 6-18: Predicted Trend for Model 1 ("Total Crashes") Based on Estimates for LOS for the Road Section

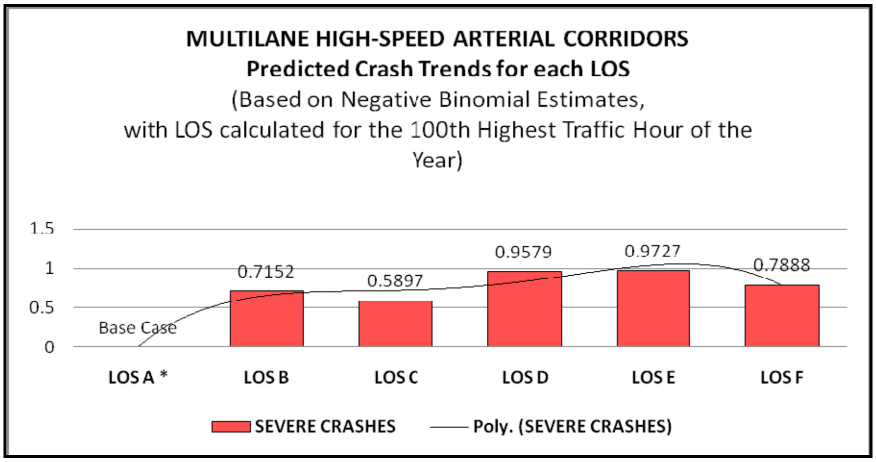


Figure 6-19: Predicted Trend for Model 2 ("Severe Crashes") Based on Estimates for LOS for the Road Section

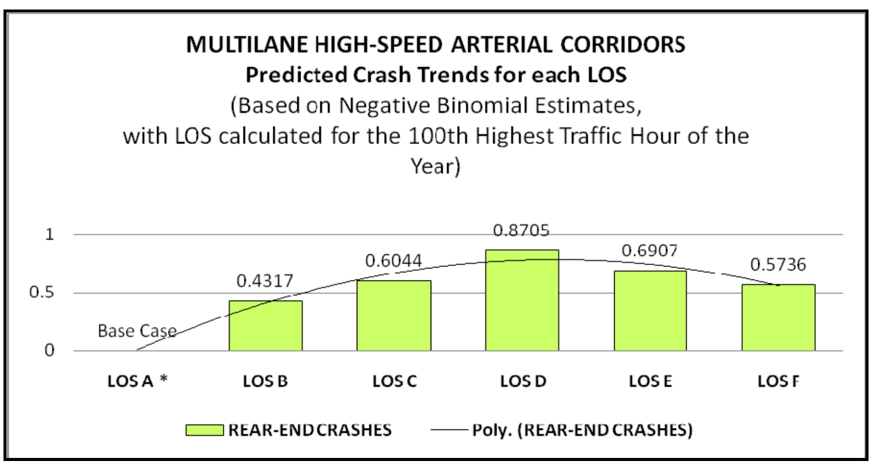


Figure 6-20: Predicted Trend for Model 3 ("Rear-End Crashes") Based on Estimates for LOS for the Road Section

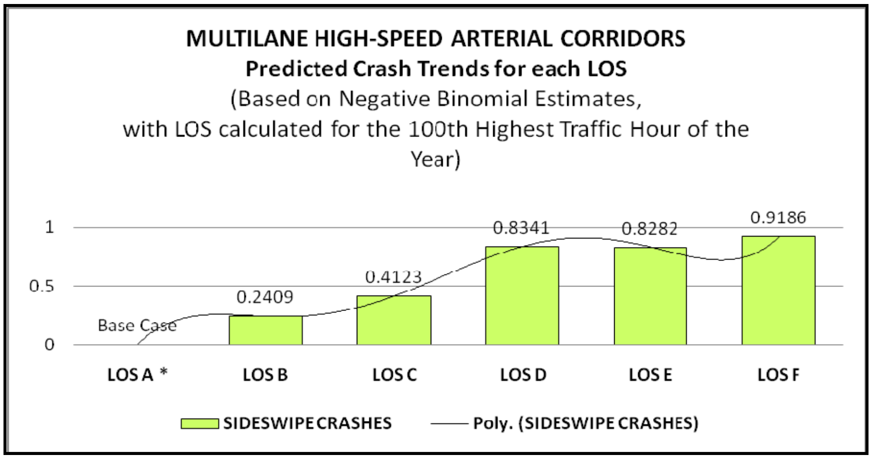


Figure 6-21: Predicted Trend for Model 4 ("Sideswipe Crashes") Based on Estimates for LOS for the Road Section

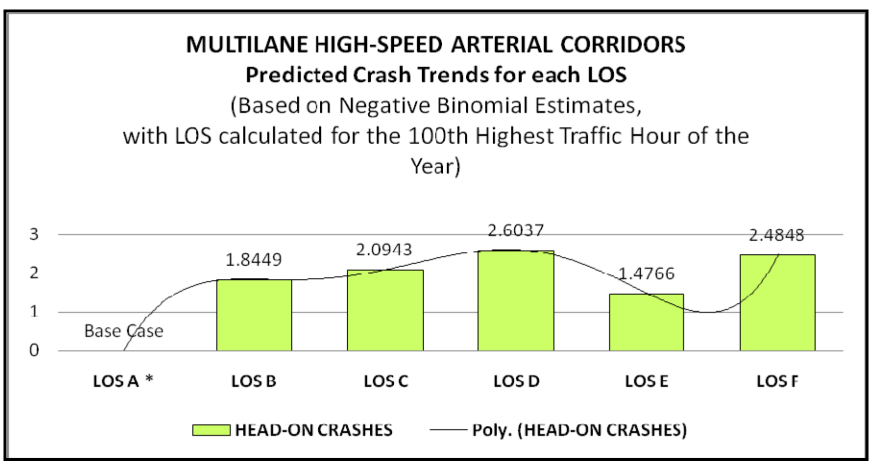


Figure 6-22: Predicted Trend for Model 5 ("Head-On Crashes") Based on Estimates for LOS for the Road Section

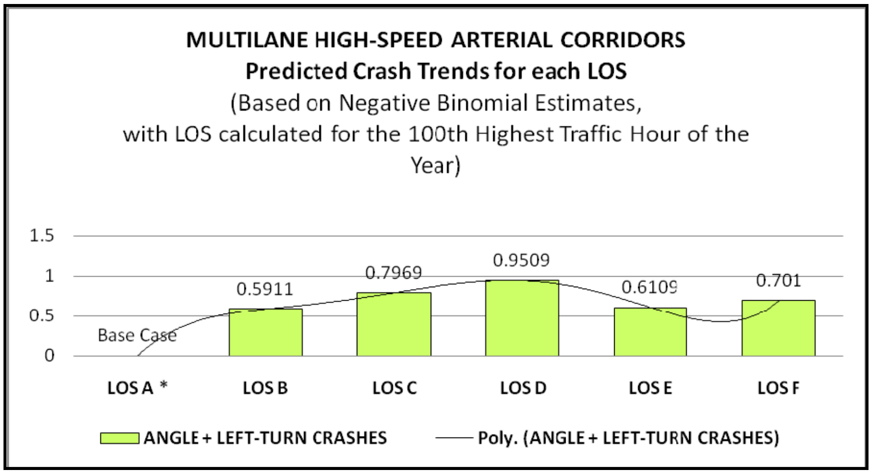


Figure 6-23: Predicted Trend for Model 6 ("Angle and Left-Turn Crashes") Based on Estimates for LOS for the Road Section

6.3.1.2 Applications and Contributions

6.3.1.2.1 Traffic Safety Guidelines

The results obtained have contributed to the preparation of a third set of "Traffic Safety Guidelines" based on the significant factors just presented. These guidelines have been organized as mentioned in Section 6.2.1.2.1.

6.3.1.2.1.1 "Road Features-Based" Traffic Safety Guideline

Following is this first set of traffic safety guidelines derived from Model Group C, constituted by the models from the analysis of Multilane High-Speed Arterial Corridors (refer to Section 5.3.1.2). Following is Table 6-9 with the set of such guidelines, based on the significant factors related to the overall features of this type of transportation facilities.

Table 6-9: "Road Features-Based" Traffic Safety Guideline for Multilane High-Speed Arterial Corridors, Derived from Model Group C

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
MODEL GROUP C: TRAFFIC SAFETY GUIDELINE				
FACTOR		CRASH OCCURRENCE		Reference
Name	Level	Type	Risk or Effect	
Total (Left+Through+Right) Number of Lanes In a Single Direction	7 or 8 lanes	TOTAL	HIGH	MODEL 1C
	5 or 6 lanes	TOTAL	MEDIUM	MODEL 1C
	3 or 4 lanes	TOTAL	LOW	MODEL 1C
	2 lanes	TOTAL	(Base Case)	MODEL 1C
	7 or 8 lanes	SEVERE	HIGH	MODEL 2C
	5 or 6 lanes	SEVERE	MEDIUM	MODEL 2C
	3 or 4 lanes	SEVERE	LOW	MODEL 2C
	2 lanes	SEVERE	(Base Case)	MODEL 2C
	7 or 8 lanes	SIDESWIPE	HIGH	MODEL 4C
	5 or 6 lanes	SIDESWIPE	MEDIUM	MODEL 4C
	3 or 4 lanes	SIDESWIPE	LOW	MODEL 4C
	2 lanes	SIDESWIPE	(Base Case)	MODEL 4C
	7 or 8 lanes	HEAD-ON	HIGH	MODEL 5C
	5 or 6 lanes	HEAD-ON	MEDIUM	MODEL 5C
	3 or 4 lanes	HEAD-ON	LOW	MODEL 5C
	2 lanes	HEAD-ON	(Base Case)	MODEL 5C
	7 or 8 lanes	ANGLE + LEFT-TURN	HIGH	MODEL 6C
	5 or 6 lanes	ANGLE + LEFT-TURN	MEDIUM	MODEL 6C
	3 or 4 lanes	ANGLE + LEFT-TURN	LOW	MODEL 6C
	2 lanes	ANGLE + LEFT-TURN	(Base Case)	MODEL 6C
Average Daily Traffic (ADT)		REAR-END	Increase	MODEL 3C
Posted Speed Limit	50 or 55 mph	TOTAL	HIGH	MODEL 1C
	60 or 65 mph	TOTAL	LOW	MODEL 1C
	40 or 45 mph	TOTAL	(Base Case)	MODEL 1C
	50 or 55 mph	SEVERE	HIGH	MODEL 2C
	60 or 65 mph	SEVERE	LOW	MODEL 2C
	40 or 45 mph	SEVERE	(Base Case)	MODEL 2C
	60 or 65 mph	HEAD-ON	HIGH	MODEL 5C
	50 or 55 mph	HEAD-ON	LOW	MODEL 5C
	40 or 45 mph	HEAD-ON	(Base Case)	MODEL 5C
	50 or 55 mph	ANGLE + LEFT-TURN	HIGH	MODEL 6C
	60 or 65 mph	ANGLE + LEFT-TURN	LOW	MODEL 6C
	40 or 45 mph	ANGLE + LEFT-TURN	(Base Case)	MODEL 6C

FACTOR		CRASH OCCURRENCE		Reference
Name	Level	Type	Risk or Effect	
Road Section's Length		TOTAL	Increase	MODEL 1C
		SEVERE	Increase	MODEL 2C
		REAR-END	Increase	MODEL 3C
		SIDESWIPE	Increase	MODEL 4C
		ANGLE + LEFT-TURN	Increase	MODEL 6C
Pavement Surface Type	Asphaltic Concrete	SEVERE	HIGH	MODEL 2C
	Portland Cement Concrete	SEVERE	LOW	MODEL 2C
	Other/Unknown	SEVERE	(Base Case)	MODEL 2C
Inside Shoulder Width		SEVERE	Reduction	MODEL 2C
		HEAD-ON	Reduction	MODEL 5C
		ANGLE + LEFT-TURN	Reduction	MODEL 6C
Median Width		HEAD-ON	Reduction	MODEL 5C

Based on the table just presented, the following conclusions can be made with regards to safety conditions of multilane high-speed arterial corridors in general:

- Total (left, through and right) number of lanes, in a single direction, of the road section contributes to the increase of total, severe, sideswipe, head-on, as well as angle and left-turn crashes.
    - More specifically, and considering 2 lanes as the base case, the results suggest that the risk for total, severe, sideswipe, head-on, as well as angle and left-turn crashes is higher while having a total of 7 or 8 lanes in a single direction of the road section; on the other hand, the results also suggest that the risk for all of these crash types is lower while having a total of 3 or 4 lanes in a single direction of the road section.
- Average Daily Traffic (ADT) (i.e. traffic volume) for the road section contributes to the increase of rear-end crashes. This could be explained by the fact that as

more vehicles are traveling on the road these are more exposed to a possible crash event (i.e. vehicle proximity).

- The posted speed limit contributes to the increase of total, severe, head-on, as well as angle and left-turn crashes.
    - More specifically, and considering 40 or 45 mph as the base case, the results suggest that the risk for total, severe, as well as angle and left-turn crashes is higher while having speed limits of 50 or 55 mph on the road section; on the other hand, the results also suggest that the risk for all of these crash types is lower while having speed limits of 60 or 65 mph on the road section.
    - Also, and considering 40 or 45 mph as the base case, the results suggest that the risk for head-on crashes is higher while having speed limits of 60 or 65 mph on the road section; on the other hand, the results also suggest that the risk for head-on crashes is lower while having speed limits of 50 or 55 mph on the road section.
- The length of the road section appears to be linearly related to total, severe, rear-end, sideswipe, as well as angle and left-turn crash occurrence.
- Pavement surface type seems to contribute to severe crash occurrence.
    - More specifically, and having other/unknown pavement surface type as the base case, the results suggest that the risk for severe crashes is higher while having a road section made of asphaltic concrete; on the other hand, the results also suggest that the risk for severe crashes is lower while having a road section made of Portland cement concrete (i.e. better skid

resistance, brighter surface at night, etc.) according to the FHWA (Aven, 2008).

- The inside shoulder width of the road section contributes to the decrease of severe, head-on, as well as angle and left-turn crashes.
- The median width of the road section contributes to the decrease of head-on crashes.

The reader may refer to Table 6-9 for more accurate details and a complete review of the guidelines just presented.

6.3.1.2.1.2 “LOS-Based” Traffic Safety Guideline

A second set of traffic safety guidelines were derived from Model Group C (refer to Section 6.3.1.2.1.1); this is also the last set of guidelines that resulted from the overall investigation. One of the features of this second set of guidelines is that they add the LOS indicator as a factor into this safety framework. Following is Table 6-10 with the set of such guidelines, based on the “LOS” factor.

Table 6-10: "LOS-Based" Traffic Safety Guideline for Multilane High-Speed Arterial Corridors, Derived from Model Group C

MULTILANE HIGH-SPEED ARTERIAL CORRIDORS				
MODEL GROUP C: "LOS-BASED" TRAFFIC SAFETY GUIDELINE				
FACTOR		CRASH		Reference
LOS		OCCURRENCE		
Level	Traffic Flow	Risk	Type	
D	Approaching Unstable	HIGH	TOTAL	MODEL 1C
F	Forced/Breakdown	MED-HIGH	TOTAL	MODEL 1C
E	Unstable	MEDIUM	TOTAL	MODEL 1C
C	Stable	MED-LOW	TOTAL	MODEL 1C
B	Reasonably Free	LOW	TOTAL	MODEL 1C
A	Free	(Base Case)	TOTAL	MODEL 1C
E	Unstable	HIGH	SEVERE	MODEL 2C
D	Approaching Unstable	MED-HIGH	SEVERE	MODEL 2C
F	Forced/Breakdown	MEDIUM	SEVERE	MODEL 2C
B	Reasonably Free	MED-LOW	SEVERE	MODEL 2C
C	Stable	LOW	SEVERE	MODEL 2C
A	Free	(Base Case)	SEVERE	MODEL 2C
D	Approaching Unstable	HIGH	REAR-END	MODEL 3C
E	Unstable	MED-HIGH	REAR-END	MODEL 3C
C	Stable	MEDIUM	REAR-END	MODEL 3C
F	Forced/Breakdown	MED-LOW	REAR-END	MODEL 3C
B	Reasonably Free	LOW	REAR-END	MODEL 3C
A	Free	(Base Case)	REAR-END	MODEL 3C
F	Forced/Breakdown	HIGH	SIDESWIPE	MODEL 4C
D	Approaching Unstable	MED-HIGH	SIDESWIPE	MODEL 4C
E	Unstable	MEDIUM	SIDESWIPE	MODEL 4C
C	Stable	MED-LOW	SIDESWIPE	MODEL 4C
B	Reasonably Free	LOW	SIDESWIPE	MODEL 4C
A	Free	(Base Case)	SIDESWIPE	MODEL 4C
D	Approaching Unstable	HIGH	HEAD-ON	MODEL 5C
F	Forced/Breakdown	MED-HIGH	HEAD-ON	MODEL 5C
C	Stable	MEDIUM	HEAD-ON	MODEL 5C
B	Reasonably Free	MED-LOW	HEAD-ON	MODEL 5C
E	Unstable	LOW	HEAD-ON	MODEL 5C
A	Free	(Base Case)	HEAD-ON	MODEL 5C
D	Approaching Unstable	HIGH	ANGLE + LEFT-TURN	MODEL 6C
C	Stable	MED-HIGH	ANGLE + LEFT-TURN	MODEL 6C
F	Forced/Breakdown	MEDIUM	ANGLE + LEFT-TURN	MODEL 6C
E	Unstable	MED-LOW	ANGLE + LEFT-TURN	MODEL 6C
B	Reasonably Free	LOW	ANGLE + LEFT-TURN	MODEL 6C
A	Free	(Base Case)	ANGLE + LEFT-TURN	MODEL 6C

Based on the table just presented, the following conclusions can be made with regards to safety conditions of multilane high-speed arterial corridors in general:

- Considering LOS A (free flow) as the base case, the results suggest that the risk for total, rear-end, head-on, as well as angle and left-turn crashes is higher for road sections from multilane high-speed arterial corridors that have a LOS D (approaching unstable flow) .
- Considering LOS A (free flow) as the base case, the results suggest that the risk for severe crashes is higher for road sections from multilane high-speed arterial corridors that have a LOS E (unstable flow).
- Considering LOS A (free flow) as the base case, the results suggest that the risk for sideswipe crashes is higher for road sections from multilane high-speed arterial corridors that have a LOS F (forced or breakdown flow).

The reader may refer to Table 6-10 for more accurate details and a complete review of the guidelines just presented.

6.4 Research Summary

Based on the results from the studies just presented, it can be concluded that the LOS can be used as an explicit indicator of traffic safety conditions at both signalized intersections and multilane high-speed arterial corridors. Following are the most important findings of this investigation:

- The LOS can be used as predictor in crash frequency models. Based on the results, the LOS appears to perform better in terms of significance when interacting with other factors and/or characteristics of the transportation facility being studied. This denotes that the LOS alone is not sufficient for predicting safety performance of transportation facilities and that operational conditions have a different effect per crash type.

- In order to compare the significance of the delay parameter (i.e. quantitative) and the LOS indicator (i.e. qualitative), both were introduced in the crash frequency models. It was observed that either one of these parameters has to be included when modeling crash frequencies in order to find it to be significant. Apart from this, it was observed that the LOS had a better significance than the delay parameter.
- The period of the day can also be used as predictor in crash frequency models. The significance of this factor indicates that each crash type has a different mechanism depending the time of day.
- Regarding significant factors for crashes at signalized intersections, the final models included the LOS for the intersection as a whole, cycle length, lighting conditions, land use, traffic volume (major and minor roads), left-turn traffic volume (major road only), posted speed limit (major and minor roads), total number of through lanes (major and minor roads), overall total and total number of left-turn lanes (major road only), as well as county and period of the day (dummy variables). Among these factors, partial or full lighting conditions was found to be associated with a reduction in overall crash frequency; this applies to the Early Morning period (1-3 A.M.) only. On the other hand, suburban and urban land uses were found to be associated with a reduction in the overall crash frequency; this applies to the Late Evening period (8-10 P.M.) only. The safety guidelines in which all these factors have been incorporated are summarized in Tables 6-2 and 6-6.

- Regarding significant factors for crashes at multilane high-speed arterial corridors, the final models included the LOS for the road section, average daily traffic (ADT), total number of through lanes in a single direction, total length of the road section, pavement surface type, as well as median and inside shoulder widths. Among these factors, the inside shoulder width was found to be associated with a reduction in severe, head-on, as well as angle and left-turn crash frequency, whereas the median width was found to be associated with a reduction in head-on crashes only. The safety guidelines in which all these factors have been incorporated are summarized in Table 6-9.
- For signalized intersections, and considering the Early Morning period (1-3 A.M.) as the base case, crash occurrence generally increases as it becomes later in the day; however, these results varied per crash frequency model. For example, the P.M. Peak period (4-6 P.M.) was associated with the highest risk for total, rear-end, sideswipe, head-on, as well as angle and left-turn crashes; and the Late Evening period (8-10 P.M.) was associated with the highest risk for severe crashes. On the other hand, the A.M. Peak period (7-9 A.M.) (i.e. earlier hours of the day) was consistently associated with the lowest risk in all crash frequency models (refer to Table 6-3). Overall, these results contradict the consensus that less congested roads are safer (i.e. less crash occurrence), since generally the P.M. Peak period is one of the most congested ones throughout a regular weekday.
- For signalized intersections, and considering LOS A (insignificant delay) as the base case, crash occurrence generally increases as operational conditions deteriorate; however, these results varied per crash frequency model. For example, LOS E (significant delay) was associated with the highest total, severe, rear-end, as well as angle and left-turn crash occurrence; LOS F (excessive delay) was associated with the highest sideswipe crash

occurrence; and LOS D (tolerable delay) was associated with the highest head-on crash occurrence. On the other hand, LOS B (minimal delay) was consistently associated with the lowest risk in all crash frequency models (refer to Table 6-4). Overall, these results again contradict the consensus that less congested roads are safer (i.e. less crash occurrence). Based on these results, and considering planning practices, to achieve a LOS C or better is recommended in order to provide desirable safety conditions at signalized intersections.

- For signalized intersections, and considering both the Early Morning period (1-3 A.M.) and LOS A (insignificant delay) as base cases, crash occurrence generally increases as becomes later in the day and as operational conditions deteriorate; however, these results varied per crash frequency model. For example, the P.M. Peak period (4-6 P.M.) has been found to have the highest risk overall in all crash frequency models, and within this time frame this risk is at its highest while having a LOS F (excessive delay) and is lower while having a LOS D (tolerable delay); in addition, the Midday period (12-2 P.M.) has been found to have the 2nd highest risk overall in all crash frequency models, and within this time frame this risk is at its highest also while having a LOS F (excessive delay) and is lower while having a LOS B (minimal delay). On the other hand, the A.M. Peak period (7-9 A.M.) (i.e. earlier hours of the day) was found to have the lowest risk overall in all crash frequency models, and within this time frame this risk is at its highest while having a LOS F (excessive delay) and is lower while having a LOS C (acceptable delay) (refer to Table 6-7). Once again, these results contradict the consensus that less congested roads are safer (i.e. less crash occurrence). Again, considering planning practices, to achieve a

LOS C or better is recommended in order to provide desirable safety conditions at signalized intersections.

- For multilane high-speed arterial corridors, and considering LOS A (free flow) as the base case, it also seems that crash occurrence increases as operational conditions deteriorate; however, these results varied per crash frequency model. For example, LOS D (approaching unstable flow) was associated with the highest total, rear-end, head-on, as well as angle and left-turn crash occurrence; LOS E (unstable flow) was associated with the highest severe crash occurrence; and LOS F (forced or breakdown flow) was associated with the highest sideswipe crash occurrence. On the other hand, LOS B (minimal delay) was associated with the lowest total, rear-end, sideswipe, as well as angle and left-turn crash occurrence; LOS C (stable flow) was associated with the lowest severe crash occurrence; and LOS E (unstable flow) was associated with the lowest head-on crash occurrence (refer to Table 6-10). Same as for signalized intersections, these results also contradict the consensus that less congested roads are safer (i.e. less crash occurrence). Based on these results, and considering planning practices, to achieve a LOS C or better is recommended in order to provide desirable safety conditions at these facilities.

Finally, it has to be emphasized that though these results are the result of a careful and extensive exploratory analysis, these models serve as guidelines that would need to be tested and/or evaluated. An analysis based on professional standards would have to be considered to complement these findings, so that an effective selection of countermeasures and policy decisions can be made.

6.5 Limitations and Extensions

After reviewing the summary of the results just presented, it can be observed that a thoughtful process had to be applied from beginning to end. The aforementioned process, however, also consisted in overcoming a series of limitations.

Regarding the LOS-Safety studies for signalized intersections, the first limitation encountered was that of being knowledgeable on the use and interpretation of the signal timing plans; this was a critical issue since most of the LOS calculation process in HCS involved entering data coming from these plans (refer to Section 3.2.3.2). Fortunately, this problem could be overcome by contacting transportation professionals and/or researchers in the field with experience on the matter. However, if not having the right contacts this problem would have been hard to overcome, especially since a guide on how to use this type of plans is not readily available or accessible to the public; in general, knowing how to use the signal timing data is based on experience.

Another limitation is with regards to the capabilities of the software used. As it is known, HCS is the most trusted highway capacity analysis software in many parts around the world. Still, one of the limitations related to the use of this software was that of entering the intersection-related data and extracting the LOS-related results for each of the signalized intersections in the sample (i.e. 149 intersections), one file per period of the day at a time (i.e. 5 periods of the day). Had the software a more advanced or interactive (i.e. faster, efficient, etc.) interface, the LOS calculation process (refer to Section 3.2.3.2) would have been completed in a shorter amount of time; because of this, extra attention had to be put for such meticulous process. Still, despite this limitation, the LOS-related data provided by the software proved to be of good quality and acceptable to HCM 2000 standards.

Apart from this, if not a data preparation issue, references from similar research were very limited. As mentioned in Chapter 1, few researchers have conducted studies on the relationship between the LOS and traffic safety or similar. It is true that innovation is inherent to any type of research, but it always helps to count with different insights on the topic of interest in order to better refine the thinking process. Overall, this was not a big limitation; in fact, it expanded this investigation's framework in general.

On the other hand, the LOS-Safety study of multilane high-speed arterial corridors had a very big limitation from the start: data availability. As stated in Section 3.3.2.1, the LOS data for state road sections provided by Hillsborough County's planning authorities did not have any variable(s) in common with the crash records available in the FDOT's CAR database; this was a very critical issue since crash-related data is very important for traffic studies and these have to be linked in some way to data in hand. Fortunately, the use of appropriate GIS software (refer to Section 3.3.2.2) aided in the retrieval of roadway ids and mile points, the most important variables within the CAR database used for data merging processes.

Furthermore, with regards to the aforementioned road sections' LOS data, another issue was its compatibility with the crash and road features data available. As mentioned in Section 3.3.2.2.1, the road sections' LOS data corresponded to the year 1999, reason for which no crash- or road features-related information could be obtained for that year since the respective databases only contain data from recent years. In the end, this could be overcome by using crash and road features data for the closest years to 1999 (2000-2001 and 2002, respectively) and by contacting other researchers in the field who had already downloaded the aforementioned data.

6.6 Future Directions

This thesis has covered an area of transportation that not only is of interest to researchers in academia but that also appeals to professionals in planning and current practice. With this being said, potential exists for more studies of the LOS-Safety relationship within the transportation and traffic engineering framework.

Signalized intersections, being entities involving many factors, could be studied in the future using a similar methodology to the one presented in this thesis. The transferability of this methodology can be applicable to an approach-level analysis, which would use the respective approach-level LOS (i.e. Eastbound, Westbound, Northbound and Southbound) also provided by HCS; to perform such analysis would provide more detailed insights into how the LOS and operational conditions interact with the mechanisms of different crash types. Furthermore, provided the respective data types are available, a similar study for unsignalized intersections would be an extension of the study presented here and complement these results; there is the advantage that HCS also computes the LOS for unsignalized intersections. Last but not least, if performing a temporal analysis for the two types of studies just described, it would be good to find a way to incorporate the flashing patterns present, for example, during the Early Morning period (1-3 A.M.) since these are frequent over regular weekdays.

Regarding the study performed for multilane high-speed arterial corridors, it is recommended to incorporate intersection-related data for the respective study area. If this can be achieved, it has to be emphasized the importance of having all data types (e.g. corridor- and intersection-related LOS, geometric features and signalization plans) current and matching in terms of years considered for the study; this would need to be considered for the crash data to be used as well. In addition to exploring the LOS-Safety relationship, the use of GIS technology in a study of that nature would be a powerful tool for identifying the spatial correlation between corridors and intersections as well; with this in consideration, including socioeconomic data into the spatial analysis would provide more valuable insights (Stamatiadis and Puccini, 1999).

Finally, and based on these recommendations, it can be observed that counting with a complete and representative dataset is the key to meaningful results. The author also recommends for any future LOS-Safety study to incorporate data from more counties in the State. In this way, a more comprehensive set of LOS-based guidelines can be prepared for the areas with the highest crash risk, which would also constitute a good tool for transportation planning in general.

APPENDIX A:
STUDY SAMPLE'S LIST OF SIGNALIZED INTERSECTIONS

ORANGE COUNTY			
COUNT	**INTERSECTION ID**	**MAJOR Road**	**MINOR Road**
1	**OC97**	Colonial Dr (SR-50)	Semoran Blvd (SR-436)
2	**OC95**	Research Pkwy	Alafaya Tr (SR-434)
3	**OC93**	Colonial Dr (SR-50)	Alafaya Tr (SR-434)
4	**OC92**	University Blvd (SR-436A)	Alafaya Tr (SR-434)
5	**OC90**	Scarlet Dr	Semoran Blvd (SR-436)
6	**OC89**	Aloma Ave (SR-426)	Semoran Blvd (SR-436)
7	**OC438**	Colonial Dr (SR-50)	Vineland Rd
8	**OC436**	Sadler Rd	Orange Blossom Tr (US-441)
9	**OC405**	Colonial Dr (SR-50)	Blackwood Ave
10	**OC404**	Colonial Dr (SR-50)	Apopka Vineland Rd (SR-535)
11	**OC327**	Silver Star Rd (SR-438)	Ocoee-Apopka (SR-437)
12	**OC303**	Silver Star Rd (SR-438)	Bluford Ave
13	**OC290**	Silver Star Rd (SR-438)	Powers Dr
14	**OC289**	Silver Star Rd (SR-438)	Mercy Dr
15	**OC249**	Lake Underhill Rd	Goldenrod Rd (SR-551)
16	**OC231**	Lake Margaret Dr	Conway Rd (SR-15)
17	**OC230**	Anderson Ave (SR-15)	Conway Rd (SR-15)
18	**OC228**	Colonial Dr (SR-50)	Powers Dr
19	**OC227**	Colonial Dr (SR-50)	Paul St
20	**OC225**	Colonial Dr (SR-50)	Ninth St
21	**OC224**	Colonial Dr (SR-50)	Maguire Rd
22	**OC220**	Colonial Dr (SR-50)	Good Homes Rd
23	**OC219**	Colonial Dr (SR-50)	Dorscher Rd
24	**OC218**	Colonial Dr (SR-50)	Bluford Ave
25	**OC215**	Colonial Dr (SR-50)	Avalon Rd
26	**OC213**	Colonial Dr (SR-50)	Rouse Rd
27	**OC212**	Colonial Dr (SR-50)	Murdock Blvd
28	**OC211**	Colonial Dr (SR-50)	Lake Pickett
29	**OC210**	Colonial Dr (SR-50)	Forsyth Rd
30	**OC209**	Colonial Dr (SR-50)	Econlockhatchee Tr
31	**OC207**	Colonial Dr (SR-50)	Pebble Beach Blvd
32	**OC206**	Colonial Dr (SR-50)	CR-13
33	**OC205**	Colonial Dr (SR-50)	Bonneville Rd
34	**OC189**	Old Cheney Highway	Semoran Blvd (SR-436)
35	**OC171**	Colonial Dr (SR-50)	Kirkman Rd (SR-435)
36	**OC166**	Silver Star Rd (SR-438)	Pine Hills Rd
37	**OC157**	Pershing Ave	Goldenrod Rd (SR-551)
38	**OC156**	Curry Ford Rd (SR-552)	Goldenrod Rd (SR-551)
39	**OC154**	Colonial Dr (SR-50)	Pine Hills Rd

40	**OC153**	Colonial Dr (SR-50)	Hiawassee Rd
41	**OC152**	Colonial Dr (SR-50)	Goldenrod Rd (SR-551)
42	**OC151**	Colonial Dr (SR-50)	Daniels St
43	**OC150**	Colonial Dr (SR-50)	Dean Rd
44	**OC149**	Colonial Dr (SR-50)	Chuluota Rd (SR-419)
45	**OC139**	McCulloch Rd	Alafaya Tr (SR-434)
46	**OC135**	Westinghouse	Alafaya Tr (SR-434)
47	**OC123**	Silver Star Rd (SR-438)	Dardanelle Dr
48	**OC122**	Banchory Rd	Semoran Blvd (SR-436)
49	**OC120**	Lee Rd (SR423)	Wymore Rd
50	**OC111**	Silver Star Rd (SR-438)	Sheringham Rd
51	**OC107**	University Blvd (SR-436A)	Goldenrod Rd (SR-551)
52	**OC101**	Lee Rd (SR423)	Edgewater Dr (SR-424)

HILLSBOROUGH COUNTY

COUNT	INTERSECTION ID	MAJOR Road	MINOR Road
1	**HC1001**	Hillsborough Ave(US 92)	Falkenburg Rd
2	**HC1002**	Hillsborough Ave(US 92)	Williams Rd
3	**HC1003**	Hillsborough Ave(US 92)	Mango Rd(CR 579)
4	**HC1004**	Hillsborough Ave(US 92)	Peach St-School
5	**HC1005**	Hillsborough Ave(US 92)	Pine St-School
6	**HC1006**	Hillsborough Ave(US 92)	Parsons Ave
7	**HC1007**	Hillsborough Ave(US 92)	Kingsway Rd
8	**HC1008**	MLKing Blvd(SR 574)	Falkenburg Rd
9	**HC1010**	MLKing Blvd(SR 574)	Williams Rd
10	**HC1011**	MLKing Blvd(SR 574)	Lakewood Dr
11	**HC1013**	MLKing Blvd(SR 574)	Mango Rd(CR 579)
12	**HC1014**	MLKing Blvd(SR 574)	Pine St
13	**HC1015**	MLKing Blvd(SR 574)	Parsons Ave
14	**HC1016**	MLKing Blvd(SR 574)	Kingsway Ave
15	**HC1027**	Brandon Blvd(SR 60)	BrandonCrossings Entr
16	**HC1028**	Brandon Blvd(SR 60)	Falkenburg Rd
17	**HC1031**	Brandon Blvd(SR 60)	GrandRegency Blvd
18	**HC1033**	Brandon Blvd(SR 60)	Lakewood Dr
19	**HC1034**	Brandon Blvd(SR 60)	Hilltop Rd
20	**HC1036**	Brandon Blvd(SR 60)	Kings Ave
21	**HC1037**	Brandon Blvd(SR 60)	Parsons Ave
22	**HC1038**	Brandon Blvd(SR 60)	Lithia-Pinecrest Rd
23	**HC1039**	Brandon Blvd(SR 60)	Kingsway Rd
24	**HC1040**	Brandon Blvd(SR 60)	Ridgewood Ave
25	**HC1045**	Adamo Dr(SR 60)	US 301
26	**HC1046**	PalmRiver Rd	US 301
27	**HC1048**	Causeway Blvd	US 301

28	**HC1064**	Bloomingdale Ave	US 301
29	**HC1066**	Riverview Dr	US 301
30	**HC1068**	Gibsonton Dr	US 301
31	**HC1083**	Adamo Dr(SR 60)	78th St
32	**HC1085**	Causeway Blvd	78th St
33	**HC1086**	Brandon Blvd(SR 60)	GorntoLake Rd
34	**HC1088**	MLKing Blvd(SR 574)	Highview Rd
35	**HC1101**	Fowler Ave	56th St
36	**HC1102**	Whiteway Dr	56th St
37	**HC1103**	MissionHills Dr	56th St
38	**HC1104**	Busch Blvd	56th St
39	**HC1105**	Riverhills Dr	56th St
40	**HC1106**	Puritan Rd	56th St
41	**HC1107**	Sligh Ave	56th St
42	**HC1108**	Hanna Ave	56th St
43	**HC1109**	Hillsborough Ave	56th St
44	**HC1112**	Hillsborough Ave(US 92)	Harney Rd
45	**HC1114**	Hillsborough Ave (SR 600)	Orient Road
46	**HC1120**	Fowler Ave	Gillette Ave
47	**HC1121**	Fowler Ave	Riverhills Dr
48	**HC1126**	Fletcher Ave	Florida Ave
49	**HC1130**	Bearss Ave	Florida Ave
50	**HC1133**	Crenshaw-Whitaker Rd	US 41/SR45
51	**HC1135**	Sunset Lane	US 41/SR45
52	**HC1143**	Skipper Rd	Nebraska Ave
53	**HC1144**	Fletcher Ave	Nebraska Ave
54	**HC1152**	131st St	Nebraska Ave
55	**HC1162**	Fletcher Ave	DaleMabry Hwy
56	**HC1166**	Linebaugh Ave	DaleMabry Hwy
57	**HC1181**	Hillsborough Ave(SR 580)	Hoover Blvd
58	**HC1184**	Hillsborough Ave(SR 580)	George Rd
59	**HC1188**	Hillsborough Ave(SR 580)	Town-N-Country Blvd
60	**HC1189**	Hillsborough Ave(SR 580)	Webb Rd
61	**HC1221**	Bearss Ave-Ehrlich Rd	DaleMabry Hwy
62	**HC1222**	Northdale Blvd	DaleMabry Hwy
63	**HC1225**	S. Entr. N Lakeview Dr	DaleMabry Hwy
64	**HC1288**	19th Ave(Ruskin)	US 41/SR45
65	**HC1289**	College Ave E(SR 674)	21st St SE
66	**HC1301**	BigBend Rd	US 301
67	**HC1302**	BigBend Rd	US 41/SR45
68	**HC1309**	College Ave(SR 674)	US 41/SR45
69	**HC1318**	VanDyke Rd	DaleMabry Hwy

70	**HC1321**	MLKing Blvd(SR 574)	Forbes Rd
71	**HC1322**	MLKing Blvd(SR 574)	McIntosh Rd
72	**HC1325**	MLKing Blvd(SR 574)	SydneyDover Rd
73	**HC1326**	MLKing Blvd(SR 574)	TurkeyCreek Rd
74	**HC1328**	MLKing Blvd(SR 574)	Valrico Rd
75	**HC1332**	Falkenburg Rd	US 301
76	**HC1348**	Hillsborough Ave(SR 580)	W. Longboat Blvd
77	**HC1356**	KnightsGriffin Rd	Paul Buchman Hwy(SR 39)
78	**HC1362**	Madison Ave	US 41/SR45
79	**HC1365**	Hillsborough Ave(US 92)	McIntosh Rd
80	**HC1366**	Brandon Blvd(SR 60)	Miller Rd
81	**HC1367**	Brandon Blvd(SR 60)	MtCarmel Rd
82	**HC1369**	Palm Ave	US 41/SR45
83	**HC1371**	Sun City Center Blvd(SR 674)	PebbleBeach Blvd
84	**HC1375**	Riverview Dr	US 41/SR45
85	**HC1376**	Sabal-Industrial Blvd	US 301
86	**HC1377**	SamAllen Rd	Paul Buchman Hwy(SR 39)
87	**HC1378**	ShellPoint Rd	US 41/SR45
88	**HC1380**	Brandon Blvd(SR 60)	James Redman Pkwy(SR 39)
89	**HC1381**	Trapnell Rd	James Redman Pkwy(SR 39)
90	**HC1382**	Brandon Blvd(SR 60)	SouthDover Rd
91	**HC1385**	Brandon Blvd(SR 60)	Valrico Rd
92	**HC1386**	Sun City Center Blvd(SR 674)	TrinityLakes Dr
93	**HC1387**	Sun City Center Blvd(SR 674)	US 301
94	**HC1388**	Sun City Center Blvd(SR 674)	ValleyForge-Kings Blvd
95	**HC1394**	College Ave E(SR 674)	30th St
96	**HC1397**	TempleHeights Rd	56th St
97	**HC1422**	Brandon Blvd(SR 60)	Mulrennen Rd

APPENDIX B:
SAMPLE SIGNAL TIMING PLANS

ORANGE COUNTY TRAFFIC SIGNAL TIMING

Intersection: ALAFAYA TR & E. COLONIAL DR **Node:** 3 **Address:** 3B9
Equipment: EAGLE **Date:** 05/07/04

BASIC TIMING

Phase	1	2	3	4	5	6	7	8
Direction	EBL	WB	SBL	NB	WBL	EB	NBL	SB
Min Green (sec)	5	15	5	15	5	15	5	15
Vehicle Gap (sec)	1.8	3.0	1.8	3.0	1.8	3.0	1.8	3.0
Max Green 1 (sec)	25	50	25	30	25	50	25	30
Max Green 2 (sec)	25	50	25	30	25	50	25	30
Yellow (sec)	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3
All-Red (sec)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Walk (sec)		5		5		5		5
Flash Don't Walk (sec)		34		30		36		30
Recall/Memory	LK	SF/LK	LK	LK	LK	SF/LK	LK	LK
Detector Delay (sec)								
Dual Entry		Y		Y		Y		Y
Overlap								
Flash	R	R	R	R	R	R	R	R
Speed (mph)	45	45	45	45	45	45	45	45
Crossing Distance (ft)		159.0		143.0		164.0		143.0
Ped Clearence (sec)		39.8		35.8		41.0		35.8

COORDINATION PLANS

Coordination Pattern	1/1/1	2/1/1	3/1/1	4/1/1	5/1/1	Day	Time	Pattern
Cycle	150	130	140			1	0:00	FREE
Split 1	25	23	22			1	9:00	2/1/1
Split 2	54	45	45			1	23:00	FREE
Split 3	21	23	33			2	0:00	FREE
Split 4	50	39	40			2	6:00	1/1/1
Split 5	20	23	21			2	9:00	2/1/1
Split 6	59	45	46			2	15:00	3/1/1
Split 7	21	23	22			2	19:00	2/1/1
Split 8	50	39	51			2	23:30	FREE
Offset	23	0	65			7	0:00	FREE
Lagging Phases	0/3/5/0	0/0/5/0	0/0/5/7			7	7:30	2/1/1
Coord Implemented						7	23:30	FREE
Source Day	**Equate 1**	**Equate 2**	**Equate 3**	**Equate 4**	**Equate 5**			
2	3	4	5	6				

Notes: FLASH ALL RED, START UP TIME = 6 SEC IN RED STATE
COORD MODE = PYL, YIELD = 1 SECOND, SHORTWAY +
ALT SEQ 04 = REVERSE PHASE 5 & 6
ALT SEQ 06 = REVERSE PHASE 3 & 4 AND 5 & 6
ALT SEQ 12 = REVERSE PHASE 5 & 6 AND 7 & 8

HILLSBOROUGH COUNTY TRAFFIC ENGINEERING

SIGNAL TIMING SHEET

Location : Dale Mabry Hwy @ Fletcher Ave

Intersection # Mist # 162 System: DaleMabry_Hwy

Node # 87 File Name: Fletcher_Ave

Flash	
Max II	ON WITH COORDINATION

MOVEMENT	Ø	MIN	EXT	CLR	RED	MAX I	MAX II	WLK	FDW	RECALL	LOCK/CNA
NBLT Lead	1	5	2.5	3.0	2.0	15	40				LOCK
SB	2	15	3	5.0	2.0	35	65	4	18	MAX	Coord/CNA
EBLT	3	5	2.5	3.0	2.0	15	35				LOCK
WB	4	7	3	4.0	2.5	20	50	4	32		
SBLT Lag	5	5	2.5	5.0	2.0	15	40	4		MAX	LOCK/CNA
NB	6	15	3	5.0	2.0	35	120	4	18	MAX	LOCK
WBLT	7	5	2.5	3.0	2.0	15	35				LOCK
EB	8	7	3	4.0	2.5	20	50	4	30		

Comments:

Overlap Calls

OLA = Ø N/A

OLB = Ø N/A

OLC = Ø N/A

OLD = Ø N/A

*** Set all Perms to 0 - 5 Seconds ***

* Intersection Configured for NBLT Leading (Ø1) & SBLT lagging (Ø5) Opertion.

* Ø2 & Ø5 Coordinated Phases * CKT 25 on with System

* Adjusted 4-1-1 HC/CT 04/08/98

* Ensured Ø5 MIN obligation was met. REW 10/09/98

**Enable CKT 56 - WRM, found under Time Clock CKT Mode

**Recall Ø2 & Ø6 WRM, found under Memory/Recall/CNA

Prepared By: R. E. Wood Date: Approved By: M. McCarthy P.E. Date:

Reviewed By: G. Bassett Date: Implemented By: Date:

File:G:\TS\Systems\DaleMabry_Hwy\Fletcher_Ave Date of Print: 05/14/04

HILLSBOROUGH COUNTY TRAFFIC ENGINEERING

Location : Dale Mabry Hwy @ Fletcher Ave
Intersection # Mist # 162

Ø #	Movement
Ø1=	NBLT
Ø2=	SB
Ø3=	EBLT
Ø4=	WB
Ø5=	SBLT
Ø6=	NB
Ø7=	WBLT
Ø8=	EB
OLA=	Ø N/A
OLB=	Ø N/A
OLC=	Ø N/A
OLD=	Ø N/A

Day Plan # 1

Day	Time	Cycle	Split	Offset
M-F	00:00	CKT	13	OFF
M-F	00:00	CKT	25	OFF
M-F	06:00	CKT	13	ON
M-F	06:00	CKT	25	ON
M-F	06:00	1	1	1
M-F	09:00	4	1	1
M-F	15:00	3	1	1
M-F	19:00	4	1	1
M-F	22:00	5	1	1

Day Plan # 2

Day	Time	Cycle	Split	Offset
S-S	00:00	CKT	13	ON
S-S	00:00	CKT	25	ON
S-S	00:00	4	1	1
S-S	02:00	CKT	13	OFF
S-S	02:00	CKT	25	OFF
S-S	07:00	CKT	13	ON
S-S	07:00	CKT	25	ON
S-S	07:00	5	1	1
S-S	10:00	4	1	1
S-S	12:00	2	1	1
S-S	19:00	4	1	1

Cycle Lengths

1	2	3	4	5
180	180	200	130	110

Offsets

SECS	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
Off 1	149	168	24	114	25
Off 2					
Off 3					
%	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5
Off 1	83%	93%	12%	88%	23%
Off 2					
Off 3					

Phase Splits

	CYCLE 1=	180			CYCLE 2=	180			CYCLE 3=	200			CYCLE 4=	130			CYCLE 5=	110		
Ø	SPLIT 1		SPLIT 2		SPLIT 1		SPLIT 2		SPLIT 1		SPLIT 2		SPLIT 1		SPLIT 2		SPLIT 1		SPLIT 2	
	%	sec	%	sec	%	sec	%	sec	%	sec	%	sec	%	sec	%	sec	%	sec	%	sec
1	14%	25.2			17%	30.6			20%	40.0			14%	18.2			16%	17.6		
2	44%	79.2			40%	72.0			39%	78.0			45%	58.5			47%	51.7		
3	20%	36.0			17%	30.6			18%	36.0			18%	23.4			15%	16.5		
4	22%	39.6			26%	46.8			23%	46.0			23%	29.9			22%	24.2		
5	20%	36.0			14%	25.2			16%	32.0			20%	26.0			26%	28.6		
6	38%	68.4			43%	77.4			43%	86.0			39%	50.7			37%	40.7		
7	20%	36.0			17%	30.6			18%	36.0			18%	23.4			15%	16.5		
8	22%	39.6			26%	46.8			23%	46.0			23%	29.9			22%	24.2		

Force Offs

	CYCLE 1				CYCLE 2				CYCLE 3				CYCLE 4				CYCLE 5			
Ø	SPLIT 1		SPLIT 2		SPLIT 1		SPLIT 2		SPLIT 1		SPLIT 2		SPLIT 1		SPLIT 2		SPLIT 1		SPLIT 2	
	%	sec	%	sec	%	sec	%	sec	%	sec	%	sec	%	sec	%	sec	%	sec	%	sec
1	67%	121			71%	128			71%	142			70%	92			71%	78		
2	0%	0			0%	0			0%	0			0%	0			0%	0		
3	31%	56			28%	51			28%	56			33%	43			33%	37		
4	52%	94			53%	96			50%	101			55%	72			54%	59		
5	0%	0			0%	0			0%	0			0%	0			0%	0		
6	90%	162			96%	173			93%	186			94%	122			90%	99		
7	31%	56			28%	51			28%	56			33%	43			33%	37		
8	52%	94			53%	96			50%	101			55%	72			54%	59		

Page 2 of File: File:G:\TS\Systems\Dale Mabry_Hwy\Fletcher_Ave Date of Print: 05/14/04

APPENDIX C:
SAMPLE TRAFFIC COUNTS

Accurate Traffic Counts, Inc

920 Kerwood Circle
Oviedo, Fl 32765
Phone/Fax: (407) 359-0962

Counter: 1893/0677
Counted By: Ney/Jose
Weather: Clear
Site:Alafaya Trail@Colonial Dr

File Name : or1893-06770811
Site Code : 00001893
Start Date : 8/11/04
Page No : 1

Groups Printed- General Traffic

	Alafaya Trail Southbound				Colonial Dr Westbound				Alafaya Trail Northbound				Colonial Dr Eastbound				
Start Time	Left	Thru	Right	App. Total	Left	Thru	Right	App. Total	Left	Thru	Right	App. Total	Left	Thru	Right	App. Total	Int. Total
Factor	1.0	1.0	1.0		1.0	1.0	1.0		1.0	1.0	1.0		1.0	1.0	1.0		
07:00	82	106	35	223	53	193	139	385	35	148	30	213	56	111	28	195	1016
07:15	66	107	46	219	54	223	166	443	61	191	39	291	91	134	47	272	1225
07:30	83	133	47	263	61	207	188	456	43	249	44	336	96	169	64	329	1384
07:45	46	133	52	231	61	225	197	483	41	317	52	410	137	150	61	348	1472
Total	277	479	180	936	229	848	690	1767	180	905	165	1250	380	564	200	1144	5097
08:00	76	140	35	251	65	166	145	376	49	256	48	353	90	141	62	293	1273
08:15	72	137	41	250	62	214	189	465	40	267	25	332	98	137	47	282	1329
08:30	47	142	51	240	76	187	178	441	30	228	55	313	95	137	75	307	1301
08:45	87	140	33	260	95	205	145	445	45	175	49	269	83	150	55	288	1262
Total	282	559	160	1001	298	772	657	1727	164	926	177	1267	366	565	239	1170	5165
11:30	112	306	82	500	95	171	77	343	91	179	85	355	100	163	111	374	1572
11:45	123	279	101	503	90	181	80	351	111	178	89	378	87	163	132	382	1614
Total	235	585	183	1003	185	352	157	694	202	357	174	733	187	326	243	756	3186
12:00	150	289	107	546	103	178	114	395	96	212	81	389	101	182	116	399	1729
12:15	126	269	90	485	103	188	98	389	91	250	88	429	109	228	132	469	1772
12:30	138	250	101	489	80	173	119	372	92	292	99	483	108	209	120	437	1781
12:45	135	280	94	509	105	183	119	407	107	299	77	483	102	210	108	420	1819
Total	549	1088	392	2029	391	722	450	1563	386	1053	345	1784	420	829	476	1725	7101
13:00	155	268	75	498	118	181	105	404	117	283	98	498	83	207	105	395	1795
13:15	137	273	88	498	85	182	105	372	83	320	104	507	104	270	162	536	1913
Total	292	541	163	996	203	363	210	776	200	603	202	1005	187	477	267	931	3708
16:00	117	256	79	452	77	165	65	307	107	187	92	386	67	263	136	466	1611
16:15	168	285	60	513	78	165	76	319	95	166	107	368	90	181	115	386	1586
16:30	188	345	81	614	101	192	83	376	83	230	112	425	66	173	139	378	1793
16:45	176	334	101	611	111	205	69	385	86	167	102	355	87	285	161	533	1884
Total	649	1220	321	2190	367	727	293	1387	371	750	413	1534	310	902	551	1763	6874
17:00	182	425	139	746	101	194	109	404	105	250	130	485	87	244	133	464	2099
17:15	147	341	106	594	99	225	88	412	102	175	85	362	100	286	152	538	1906
17:30	166	431	64	661	87	189	78	354	117	220	88	425	81	275	135	491	1931
17:45	182	398	69	649	98	203	97	398	115	191	128	434	105	260	110	475	1956
Total	677	1595	378	2650	385	811	372	1568	439	836	431	1706	373	1065	530	1968	7892
Grand Total	2961	6067	1777	10805	2058	4595	2829	9482	1942	5430	1907	9279	2223	4728	2506	9457	39023
Apprch %	27.4	56.1	16.4		21.7	48.5	29.8		20.9	58.5	20.6		23.5	50.0	26.5		
Total %	7.6	15.5	4.6	27.7	5.3	11.8	7.2	24.3	5.0	13.9	4.9	23.8	5.7	12.1	6.4	24.2	

	Alafaya Trail Southbound				Colonial Dr Westbound				Alafaya Trail Northbound				Colonial Dr Eastbound				
Start Time	Left	Thru	Right	App. Total	Left	Thru	Right	App. Total	Left	Thru	Right	App. Total	Left	Thru	Right	App. Total	Int. Total
Peak Hour From 07:00 to 08:45 - Peak 1 of 1																	
Intersection	07:30																
Volume	277	543	175	995	249	812	719	1780	173	1089	169	1431	421	597	234	1252	5458
Percent	27.8	54.6	17.6		14.0	45.6	40.4		12.1	76.1	11.8		33.6	47.7	18.7		
07:45 Volume	46	133	52	231	61	225	197	483	41	317	52	410	137	150	61	348	1472
Peak Factor																	0.927
High Int.	07:30				07:45				07:45				07:45				
Volume	83	133	47	263	61	225	197	483	41	317	52	410	137	150	61	348	
Peak Factor				0.946				0.921				0.873				0.899	
Peak Hour From 11:30 to 13:15 - Peak 1 of 1																	
Intersection	12:30																
Volume	565	1071	358	1994	388	719	448	1555	399	1194	378	1971	397	896	495	1788	7308
Percent	28.3	53.7	18.0		25.0	46.2	28.8		20.2	60.6	19.2		22.2	50.1	27.7		
13:15 Volume	137	273	88	498	85	182	105	372	83	320	104	507	104	270	162	536	1913
Peak Factor																	0.955
High Int.	12:45				12:45				13:15				13:15				
Volume	135	280	94	509	105	183	119	407	83	320	104	507	104	270	162	536	
Peak Factor				0.979				0.955				0.972				0.834	

Accurate Traffic Counts, Inc
920 Kerwood Circle
Oviedo, Fl 32765
Phone/Fax: (407) 359-0962

File Name : or1893-06770811
Site Code : 00001893
Start Date : 8/11/04

	Alafaya Trail Southbound				Colonial Dr Westbound				Alafaya Trail Northbound				Colonial Dr Eastbound				
Start Time	Left	Thru	Right	App. Total	Left	Thru	Right	App. Total	Left	Thru	Right	App. Total	Left	Thru	Right	App. Total	Int. Total
Peak Hour From 16:00 to 17:45 - Peak 1 of 1																	
Intersection	17:00																
Volume	677	1595	378	2650	385	811	372	1568	439	836	431	1706	373	1065	530	1968	7892
Percent	25.5	60.2	14.3		24.6	51.7	23.7		25.7	49.0	25.3		19.0	54.1	26.9		
17:00 Volume	182	425	139	746	101	194	109	404	105	250	130	485	87	244	133	464	2099
Peak Factor																	0.940
High Int.	17:00				17:15				17:00				17:15				
Volume	182	425	139	746	99	225	88	412	105	250	130	485	100	286	152	538	
Peak Factor				0.888				0.951				0.879				0.914	

APPENDIX D:
STUDY SAMPLE'S LIST OF ROAD SECTIONS ALONG MULTILANE HIGH-SPEED ARTERIAL CORRIDORS

HILLSBOROUGH COUNTY					
COUNT	ROAD SECTION ID	MAIN ROAD SECTION	LOS	Intersecting Road 1	Intersecting Road 2
1	0	DALE MABRY HWY	C	LUTZ LAKE FERN	COUNTYLINE RD
2	1	DALE MABRY HWY	C	GERACI RD	LUTZ LAKE FERN
3	3	DALE MABRY HWY	C	VAN DYKE RD	VETERAN'S EXPWY
4	7	DALE MABRY HWY	C	N NORTH LAKEVIE	VAN DYKE RD
5	8	US HWY 41	B	NEBRASKA/FLORIA	CRENSHAW LAKE R
6	9	DALE MABRY HWY	D	S NORTH LAKEVIE	N NORTH LAKEVIE
7	12	DALE MABRY HWY	D	NORTHDALE BLVD	S NORTH LAKEVIE
8	14	BEARSS AVE	F	FLORIDA AVE	I-275
9	16	FLORIDA AVE	C	FLETCHER AVE	BEARSS AVE
10	18	NEBRASKA AVE	B	FLETCHER AVE	SKIPPER RD
11	19	DALE MABRY HWY	F	FLETCHER AVE	EHRLICH RD
12	20	US HWY 301	C	HARNEY RD	CR 579
13	21	FLORIDA AVE	C	FOWLER AVE	124TH ST
14	22	NEBRASKA AVE	D	FOWLER AVE	131ST AVE
15	23	FOWLER AVE	F	I-275	NEBRASKA AVE
16	24	FOWLER AVE	D	15TH ST	22ND ST
17	25	FOWLER AVE	D	22ND ST	30TH ST
18	26	FOWLER AVE	F	MCKINLEY BLVD	46TH ST
19	27	FOWLER AVE	F	46TH ST	50TH ST
20	28	US HWY 301	C	MAIN ST	HARNEY RD
21	30	US HWY 301	C	FOWLER AVE	MAIN ST
22	31	FOWLER AVE	B	I-75	US 301
23	32	DALE MABRY HWY	F	HUDSON	FLETCHER AVE
24	33	US HWY 301	C	WILLIAMS ROAD	FOWLER AVE
25	35	56TH ST	F	WHITEWAY DR	FOWLER AVE
26	37	56TH ST	F	SERENA	WHITEWAY DR
27	39	FLORIDA AVE	C	LINEBAUGH AVE	BOUGAINVILLEA A
28	40	DALE MABRY HWY	F	LINEBAUGH AVE	HUDSON
29	41	56TH ST	F	MISSION HILLS D	SERENA
30	42	FLORIDA AVE	C	FLORILAND MALL	LINEBAUGH AVE
31	43	DALE MABRY HWY	F	BUSCH BLVD	LINEBAUGH AVE
32	44	BUSCH BLVD	F	DALE MABRY HWY	HIMES AVE
33	47	SR 39	C	I-4 N FRONTAGE	SAM ALLEN RD
34	49	SR 39	C	I-4	I-4 N FRONTAGE
35	50	PARK RD	B	US 92	I-4
36	51	BUSCH BLVD	F	ARMENIA AVE	N BOULEVARD
37	52	FLORIDA AVE	C	BUSCH BLVD	FLORILAND MALL
38	53	SR 39	C	I-4 FRONTAGE RD	I-4
39	54	BUSCH BLVD	F	FLORIDA AVE	I-275

40	**55**	**BUSCH BLVD**	**E**	I-275	NEBRASKA AVE
41	**56**	**BUSCH BLVD**	**E**	22ND ST	26TH ST
42	**57**	**BUSCH BLVD**	**E**	NEBRASKA AVE	22ND ST
43	**58**	**BUSCH BLVD**	**E**	26TH ST	30TH ST
44	**59**	**BUSCH BLVD**	**D**	30TH ST	MCKINLEY DR
45	**61**	**BUSCH BLVD**	**D**	52ND ST	56TH ST
46	**62**	**FLORIDA AVE**	**C**	YUKON ST	BUSCH BLVD
47	**65**	**FLORIDA AVE**	**C**	WATERS AVE	YUKON ST
48	**66**	**DALE MABRY HWY**	**F**	WATERS AVE	BUSCH BLVD
49	**67**	**FLORIDA AVE**	**C**	BIRD ST	WATERS AVE
50	**69**	**US HWY 92**	**B**	PARK ST	COUNTY LINE RD
51	**70**	**US HWY 92**	**C**	WOODROW WILSON	THONOTOSASSA RD
52	**72**	**PARK RD**	**B**	US 92	I-4
53	**73**	**US HWY 92**	**A**	THONOTOSASSA RD	SR 600
54	**77**	**NEBRASKA AVE**	**D**	BROAD ST	SITKA
55	**82**	**US HWY 92**	**B**	MCINTOSH RD	MOORE'S LAKE RD
56	**89**	**US HWY 301**	**C**	SLIGH AVE	HARNEY RD
57	**91**	**REYNOLDS ST**	**C**	SAMMONDS RD	THONOTOSASSA RD
58	**93**	**M L KING BLVD**	**C**	TURKEY CREEK RD	SAMMONDS RD
59	**95**	**US HWY 92**	**C**	PARSONS AVE	KINGSWAY RD
60	**96**	**NEBRASKA AVE**	**D**	HANNA AVE	SLIGH AVE
61	**98**	**US HWY 92**	**C**	PINE ST	PARSONS AVE
62	**99**	**US HWY 92**	**C**	PEACH ST	PINE ST
63	**100**	**US HWY 92**	**C**	CR 579	PEACH ST
64	**101**	**DALE MABRY HWY**	**F**	CITY LIMITS	LAMBRIGHT ST
65	**102**	**HILLSBOROUGH AVE**	**F**	VETERAN'S EXPWY	VETERAN'S FRONT
66	**103**	**HILLSBOROUGH AVE**	**F**	FRONTAGE RD	VETERAN'S EXPWY
67	**104**	**HILLSBOROUGH AVE**	**F**	HOOVER RD	WESTSHORE BLVD
68	**105**	**HILLSBOROUGH AVE**	**F**	BENJAMIN RD	HOOVER RD
69	**106**	**HILLSBOROUGH AVE**	**F**	VETERAN'S FRONT	BENJAMIN RD
70	**107**	**HILLSBOROUGH AVE**	**D**	DALE MABRY HWY	HIMES AVE
71	**108**	**DALE MABRY HWY**	**F**	HILLSBOROUGH AV	CITY LIMITS
72	**109**	**HILLSBOROUGH AVE**	**D**	HIMES AVE	HABANA AVE
73	**112**	**HILLSBOROUGH AVE**	**D**	NEBRASKA AVE	15TH ST
74	**113**	**HILLSBOROUGH AVE**	**D**	15TH ST	22ND ST
75	**114**	**HILLSBOROUGH AVE**	**B**	40TH ST	50TH ST
76	**115**	**HILLSBOROUGH AVE**	**F**	50TH ST	56TH ST
77	**116**	**HILLSBOROUGH AVE**	**F**	HARNEY RD	SUNCOAST SCHOOL
78	**117**	**HILLSBOROUGH AVE**	**F**	56TH ST	EAST LAKE SQ MA
79	**118**	**HILLSBOROUGH AVE**	**F**	SUNCOAST SCHOOL	ORIENT RD
80	**119**	**SR 39**	**D**	PARK RD EXTENSI	AIRPORT RD
81	**120**	**US HWY 301**	**F**	I-4	SLIGH AVE

82	**122**	**M L KING BLVD**	**C**	DOVER RD	FORBES RD
83	**123**	**50TH/56TH ST**	**C**	NET PARK	HILLSBOROUGH AVE
84	**124**	**FLORIDA AVE**	**C**	VIOLET	HILLSBOROUGH AV
85	**126**	**US HWY 301**	**C**	M L KING BLVD	I-4
86	**127**	**HIGHLAND AVE**	**C**	OSBORNE AVE	VIOLET
87	**128**	**FLORIDA AVE**	**C**	OSBORNE AVE	VIOLET
88	**130**	**M L KING BLVD**	**C**	MCINTOSH RD	DOVER RD
89	**132**	**40TH ST**	**D**	OSBORNE AVE	HILLSBOROUGH AV
90	**133**	**M L KING BLVD**	**C**	VALRICO RD	MCINTOSH RD
91	**134**	**50TH/56TH ST**	**C**	HARNEY RD	NET PARK
92	**135**	**50TH/56TH ST**	**C**	LAKE AVE	HARNEY RD
93	**137**	**M L KING BLVD**	**C**	KINGSWAY RD	VALRICO RD
94	**138**	**DALE MABRY HWY**	**F**	M L KING BLVD	HILLSBOROUGH AV
95	**139**	**HIGHLAND AVE**	**C**	M L KING BLVD	OSBORNE AVE
96	**141**	**FLORIDA AVE**	**C**	M L KING BLVD	OSBORNE AVE
97	**148**	**M L KING BLVD**	**C**	15TH ST	22ND ST
98	**149**	**50TH/56TH ST**	**C**	M L KING BLVD	LAKE AVE
99	**150**	**M L KING BLVD**	**C**	50TH ST	I-4
100	**151**	**M L KING BLVD**	**C**	LAKE AVE	50TH ST
101	**153**	**M L KING BLVD**	**C**	I-4	ORIENT RD
102	**154**	**US HWY 301**	**C**	M L KING BLVD	I-4
103	**155**	**M L KING BLVD**	**B**	US HWY 301	RIGA BLVD
104	**156**	**M L KING BLVD**	**F**	HIGHVIEW RD	PINE ST
105	**157**	**M L KING BLVD**	**C**	PARSONS AVE	KINGSWAY RD
106	**158**	**M L KING BLVD**	**B**	RIGA BLVD	FALKENBURG RD
107	**159**	**M L KING BLVD**	**C**	WILLIAMS RD	LAKEWOOD DR
108	**160**	**M L KING BLVD**	**C**	BROADWAY AVE	CR 579
109	**161**	**M L KING BLVD**	**C**	I-75	WILLIAMS RD
110	**162**	**M L KING BLVD**	**B**	FALKENBURG RD	I-75
111	**163**	**SR 39**	**B**	CHARLIE GRIFFIN	ALEXANDER ST
112	**167**	**40TH ST**	**D**	MELBOURNE BLVD	LAKE AVE
113	**168**	**DALE MABRY HWY**	**F**	TAMPA BAY BLVD	M L KING BLVD
114	**170**	**FLORIDA AVE**	**C**	FLORIBRASKA AVE	LAKE AVE S
115	**174**	**40TH ST**	**D**	21ST AVE	MELBOURNE BLVD
116	**175**	**40TH ST**	**D**	19TH AVE	21ST AVE
117	**176**	**MELBURNE BLVD**	**B**	50TH ST	40TH ST
118	**177**	**50TH ST**	**C**	I-4	MELBURNE BLVD
119	**178**	**COURTNEY CAMPBELL PKWY**	**F**	PINELLAS COUNTY	ROCKY POINT DR
120	**180**	**COURTNEY CAMPBELL PKWY**	**F**	ROCKY POINT DR	MEMORIAL HWY
121	**182**	**US HWY 301**	**C**	BROADWAY AVE	M L KING BLVD
122	**184**	**FLORIDA AVE**	**C**	COLUMBUS AVE	FLORIBRASKA AVE
123	**185**	**DALE MABRY HWY**	**F**	COLUMBUS DR	TAMPA BAY BLVD

124	**190**	**40TH ST**	**D**	COLUMBUS DR	19TH AVE
125	**191**	**SR 39**	**B**	TRAPNELL RD	CHARLIE GRIFFIN
126	**192**	**DALE MABRY HWY**	**F**	GOLD TRIANGLE	COLUMBUS DR
127	**196**	**BOY SCOUT BLVD**	**C**	LOIS AVE	COLUMBUS RD
128	**197**	**39TH ST**	**C**	12TH AVE	I-4
129	**204**	**DALE MABRY HWY**	**F**	SPRUCE ST	GOLD TRIANGLE
130	**218**	**SR 60 / ADAMO DR**	**C**	34TH ST	39TH ST
131	**222**	**MEMORIAL HWY**	**E**	BOY SCOUT BLVD	I-275
132	**226**	**CHANNELSIDE DR**	**E**	TWIGGS ST	ADAMO DR
133	**227**	**SR 60 / ADAMO DR**	**F**	CITY LIMITS	78TH ST
134	**228**	**50TH ST**	**C**	ADAMO DR	BROADWAY AVE
135	**229**	**SR 60 / ADAMO DR**	**F**	ORIENT RD	CITY LIMITS
136	**230**	**SR 60 / ADAMO DR**	**F**	MAYDELL DR	ORIENT RD
137	**231**	**SR 60 / ADAMO DR**	**F**	US HWY 41	MAYDELL DR
138	**234**	**US HWY 301**	**C**	WOODBERRY RD	BROADWAY AVE
139	**236**	**KENNEDY BLVD**	**C**	MERIDIAN ST	CHANNELSIDE DR
140	**245**	**US HWY 301**	**C**	ADAMO DR	WOODBERRY RD
141	**252**	**MEMORIAL HWY**	**E**	BOY SCOUT BLVD	I-275
142	**263**	**KENNEDY BLVD**	**C**	HOOVER BLVD	MEMORIAL HWY
143	**264**	**KENNEDY BLVD**	**F**	GARDENIA ST	OCCIDENT ST
144	**265**	**KENNEDY BLVD**	**F**	MEMORIAL HWY	GARDENIA ST
145	**266**	**KENNEDY BLVD**	**F**	OCCIDENT ST	WESTSHORE BLVD
146	**267**	**US HWY 41**	**C**	CITY LIMITS	LEE ROY SELMON
147	**268**	**DALE MABRY HWY**	**F**	KENNEDY BLVD	CYPRESS ST
148	**269**	**KENNEDY BLVD**	**F**	WESTSHORE BLVD	LOIS AVE
149	**270**	**KENNEDY BLVD**	**F**	DALE MABRY HWY	HIMES AVE
150	**271**	**KENNEDY BLVD**	**F**	HIMES AVE	HENDERSON BLVD
151	**272**	**KENNEDY BLVD**	**F**	MACDILL AVE	ARMENIA AVE
152	**273**	**KENNEDY BLVD**	**F**	ARMENIA AVE	HOWARD AVE
153	**274**	**KENNEDY BLVD**	**D**	HOWARD AVE	WILLOW AVE
154	**275**	**KENNEDY BLVD**	**D**	WILLOW AVE	N BOULEVARD
155	**277**	**DALE MABRY HWY**	**F**	ROLAND ST	KENNEDY BLVD
156	**279**	**US HWY 41**	**D**	PALM RIVER RD	CITY LIMITS
157	**280**	**DALE MABRY HWY**	**F**	AZEELE ST	ROLAND ST
158	**281**	**HENDERSON BLVD**	**C**	AZEELE ST	KENNEDY BLVD
159	**282**	**US HWY 301**	**D**	PALM RIVER RD	ADAMO DR
160	**283**	**DALE MABRY HWY**	**F**	SWANN AVE	AZEELE ST
161	**284**	**22ND ST**	**F**	MARITIME BLVD	LINDSEY
162	**285**	**SR 60**	**C**	VALRICO RD	MILLER RD
163	**286**	**SR 60**	**F**	PROVIDENCE RD	LAKEWOOD DR
164	**287**	**SR 60**	**F**	BUILDERS SQUARE	KINGS AVE
165	**288**	**SR 60**	**F**	KINGS AVE	PARSONS AVE

166	**289**	**SR 60**	**C**	KINGSWAY RD	RIDGEWOOD AVE
167	**290**	**SR 60**	**C**	MILLER RD	ST CLOUD AVE
168	**291**	**SR 60**	**F**	LITHIA PINECRES	KINGSWAY RD
169	**292**	**SR 39**	**B**	SR 60	TRAPNELL RD
170	**293**	**HENDERSON BLVD**	**C**	STERLING	SWANN AVE
171	**294**	**DALE MABRY HWY**	**F**	HENDERSON BLVD	SWANN AVE
172	**295**	**DALE MABRY HWY**	**F**	NEPTUNE ST	HENDERSON BLVD
173	**296**	**DALE MABRY HWY**	**F**	ESTRELLA ST	NEPTUNE ST
174	**298**	**DALE MABRY HWY**	**F**	SAN CARLOS ST	ESTRELLA ST
175	**299**	**US HWY 41**	**B**	CAUSEWAY BLVD	PALM RIVER RD
176	**300**	**CAUSEWAY BLVD**	**C**	MARITIME BLVD	50TH ST
177	**301**	**CAUSEWAY BLVD**	**F**	78TH ST	US HWY 301
178	**302**	**CAUSEWAY BLVD**	**F**	MAYDELL DR	78TH ST
179	**303**	**DALE MABRY HWY**	**F**	BAY TO BAY BLVD	SAN CARLOS ST
180	**305**	**US HWY 301**	**C**	FALKENBURG RD	CAUSEWAY BLVD
181	**306**	**DALE MABRY HWY**	**D**	EL PRADO BLVD	BAY TO BAY BLVD
182	**307**	**US HWY 301**	**C**	EVERHART RD	FALKENBURG RD
183	**308**	**DALE MABRY HWY**	**D**	EUCLID AVE	EL PRADO BLVD
184	**310**	**DALE MABRY HWY**	**D**	BAY VISTA AVE	EUCLID AVE
185	**311**	**US HWY 41**	**B**	PORT SUTTON RD	CAUSEWAY BLVD
186	**312**	**US HWY 301**	**C**	I-75	EVERHART RD
187	**313**	**US HWY 41**	**B**	MADISON AVE	PORT SUTTON RD
188	**314**	**DALE MABRY HWY**	**D**	GANDY BLVD	BAY VISTA AVE
189	**315**	**GANDY BLVD**	**F**	LEE ROY SELMON	DALE MABRY HWY
190	**316**	**GANDY BLVD**	**F**	LOIS AVE	LEE ROY SELMON
191	**317**	**GANDY BLVD**	**F**	MANHATTAN AVE	LOIS AVE
192	**318**	**US HWY 301**	**B**	GORNTO LAKE RD	PROGRESS BLVD
193	**319**	**DALE MABRY HWY**	**C**	OKLAHOMA	GANDY BLVD
194	**320**	**US HWY 41**	**B**	RIVERVIEW DR	MADISON AVE
195	**321**	**US HWY 301**	**B**	GIBSONTON DR	BALM RIVERVIEW
196	**322**	**US HWY 41**	**B**	GIBSONTON DR	RIVERVIEW DR
197	**323**	**US HWY 41**	**B**	PALM AVE	GIBSONTON DR
198	**324**	**US HWY 41**	**A**	SYMMES RD	PALM AVE
199	**325**	**US HWY 301**	**C**	SYMMES RD	GIBSONTON DR
200	**326**	**US HWY 41**	**A**	ADAMSVILLE AVE	SYMMES RD
201	**327**	**US HWY 41**	**B**	BIG BEND RD	ADAMSVILLE AVE
202	**328**	**US HWY 301**	**C**	BIG BEND RD	RHODINE RD
203	**330**	**US HWY 301**	**C**	BALM RD	BIG BEND RD
204	**331**	**US HWY 41**	**B**	APOLLO BEACH BL	BIG BEND RD
205	**332**	**US HWY 301**	**C**	19TH AVE	BALM RD
206	**333**	**US HWY 41**	**B**	19TH AVE NE	APOLLO BEACH BL
207	**334**	**US HWY 41**	**B**	RUSKIN WIMAUMA	SHELL POINT RD

208	**336**	**SR 674**	**D**	I-75	CYPRESS LAKES B
209	**337**	**SR 674**	**B**	US HWY 301	CR 579
210	**338**	**US HWY 301**	**C**	RUSKIN WIMAUMA	19TH AVE
211	**339**	**SR 674**	**B**	US HWY 301	CR 579
212	**340**	**SR 674**	**B**	US HWY 301	CR 579
213	**341**	**SR 674**	**B**	CR 579	CARLTON LAKE RD
214	**342**	**US HWY 41**	**B**	14TH AVE	RUSKIN WIMAUMA
215	**343**	**US HWY 41**	**B**	7TH ST SW	14TH AVE
216	**344**	**SR 674**	**B**	CR 579	CARLTON LAKE RD
217	**345**	**SR 674**	**B**	CARLTON LAKE RD	CR 39
218	**346**	**US HWY 41**	**A**	COCKROACH BAY R	GULF CITY RD
219	**347**	**US HWY 301**	**A**	MANATEE COUNTY	RUSKIN WIMAUMA
220	**357**	**US HWY 301**	**C**	HARNEY RD	WILLIAMS ROAD
221	**358**	**39TH ST**	**C**	7TH AVE	12TH AVE
222	**361**	**US HWY 41**	**A**	MANATEE COUNTY	COCKROACH BAY R
223	**362**	**US HWY 41**	**A**	GULF CITY RD	7TH ST SW
224	**363**	**SR 674**	**B**	15TH ST	30TH ST
225	**364**	**SR 674**	**B**	30TH ST	I-75
226	**365**	**SR 674**	**B**	2ND ST	15TH ST
227	**366**	**SR 674**	**B**	US HWY 41	2ND ST
228	**367**	**SR 674**	**D**	VALLEY FORGE BL	TRINITY POINT B
229	**368**	**SR 674**	**D**	TRINITY POINT B	N PEBBLE BEACH
230	**369**	**SR 674**	**D**	N PEBBLE BEACH	US HWY 301
231	**370**	**SR 674**	**D**	CYPRESS LAKES B	VALLEY FORGE BL
232	**371**	**US HWY 41**	**B**	SHELL POINT RD	19TH AVE NE
233	**372**	**US HWY 301**	**D**	RHODINE RD	SYMMES RD
234	**373**	**SR 674**	**A**	CR 39	POLK COUNTY LIN
235	**374**	**SR 60**	**C**	ST CLOUD AVE	DOVER RD
236	**375**	**SR 60**	**C**	RIDGEWOOD AVE	VALRICO RD
237	**376**	**SR 60**	**A**	CR 39	SMITH-RYALS RD
238	**377**	**SR 60**	**A**	SMITH-RYALS RD	POLK COUNTY LIN
239	**378**	**US HWY 301**	**B**	BALM RIVERVIEW	RIVERVIEW DR
240	**379**	**SR 60**	**F**	PARSONS AVE	LITHIA PINECRES
241	**380**	**SR 60**	**F**	HILLTOP RD	PAULS DR
242	**381**	**SR 60**	**F**	LAKEWOOD DR	HILLTOP RD
243	**382**	**SR 60**	**F**	PAULS DR	BUILDERS SQUARE
244	**383**	**SR 60**	**F**	GORNTO LAKE RD	PROVIDENCE RD
245	**384**	**SR 60**	**B**	DOVER RD	TURKEY CREEK RD
246	**385**	**BAKER ST**	**C**	SR 39	REYNOLDS ST
247	**386**	**US HWY 92**	**C**	REYNOLDS ST	PARK ST
248	**387**	**SR 39**	**D**	ALEXANDER ST	PARK RD EXTENSI
249	**388**	**BAKER ST**	**B**	SR 600	N WHEELER ST

250	**389**	**US HWY 301**	**C**	MCINTOSH RD	PASCO COUNTY LI
251	**390**	**US HWY 301**	**B**	STACY RD	MCINTOSH RD
252	**391**	**US HWY 92**	**C**	MOORE'S LAKE RD	FORBES RD
253	**392**	**M L KING BLVD**	**C**	FORBES RD	TURKEY CREEK RD
254	**393**	**M L KING BLVD**	**C**	LAKEWOOD DR	BROADWAY AVE
255	**394**	**M L KING BLVD**	**F**	PINE ST	PARSONS AVE
256	**395**	**US HWY 92**	**C**	FALKENBURG RD	WILLIAMS RD
257	**398**	**M L KING BLVD**	**C**	ORIENT RD	US HWY 301
258	**399**	**US HWY 301**	**D**	LEE ROY SELMON	PALM RIVER RD
259	**400**	**US HWY 301**	**C**	CAUSEWAY BLVD	LEE ROY SELMON
260	**401**	**US HWY 301**	**B**	PROGRESS BLVD	I-75
261	**403**	**CAUSEWAY BLVD**	**F**	50TH ST	MAYDELL DR
262	**404**	**M L KING BLVD**	**F**	CR 579	HIGHVIEW RD
263	**405**	**US HWY 92**	**C**	WILLIAMS RD	CR 579
264	**406**	**US HWY 41**	**B**	LUTZ LAKE FERN	COUNTY LINE RD
265	**407**	**US HWY 41**	**B**	4TH AVE SE	LUTZ LAKE FERN
266	**408**	**US HWY 41**	**B**	CRYSTAL LAKE RD	SUNSET LANE
267	**409**	**US HWY 41**	**B**	DEBUEL RD	CRYSTAL LAKE RD
268	**410**	**US HWY 41**	**B**	CRENSHAW LAKE R	DEBUEL RD
269	**411**	**US HWY 41**	**B**	SUNSET LANE	4TH AVE SE
270	**412**	**DALE MABRY HWY**	**C**	CHEVAL BLVD	GERACI RD
271	**413**	**DALE MABRY HWY**	**C**	VETERAN'S EXPWY	CHEVAL BLVD
272	**414**	**FLORIDA AVE**	**B**	BEARSS AVE	LAKE MAGDALENE
273	**415**	**FLORIDA AVE**	**B**	LAKE MAGDALENE	FLORIDA NEBRASK
274	**416**	**NEBRASKA AVE**	**B**	HAYES RD	FLORIDA NEBRASK
275	**417**	**NEBRASKA AVE**	**B**	BEARSS AVE	HAYES RD
276	**418**	**NEBRASKA AVE**	**B**	SKIPPER RD	BEARSS AVE
277	**419**	**DALE MABRY HWY**	**D**	EHRLICH RD	NORTHDALE BLVD
278	**420**	**FLETCHER AVE**	**F**	FLORIDA AVE	I-275
279	**421**	**FLETCHER AVE**	**F**	I-275	NEBRASKA AVE
280	**422**	**BEARSS AVE**	**F**	I-275	NEBRASKA AVE
281	**423**	**NEBRASKA AVE**	**D**	131ST AVE	FLETCHER AVE
282	**424**	**FOWLER AVE**	**F**	30TH STREET	MCKINLEY BLVD
283	**425**	**NEBRASKA AVE**	**D**	109TH AVE	FOWLER AVE
284	**426**	**FOWLER AVE**	**D**	NEBRASKA AVE	15TH ST
285	**427**	**FLORIDA AVE**	**C**	124TH ST	FLETCHER AVE
286	**430**	**FOWLER AVE**	**F**	FLORIDA AVE	I-275
287	**431**	**NEBRASKA AVE**	**D**	LINEBAUGH AVE	BOUGAINVILLEA A
288	**432**	**BUSCH BLVD**	**D**	MCKINLEY DR	46TH ST
289	**433**	**NEBRASKA AVE**	**D**	BOUGAINVILLEA A	109TH AVE
290	**434**	**NEBRASKA AVE**	**D**	BUSCH BLVD	LINEBAUGH AVE
291	**437**	**NEBRASKA AVE**	**D**	YUKON ST	BUSCH BLVD

292	**438**	**NEBRASKA AVE**	**D**	SLIGH AVE	BROAD ST
293	**439**	**FLORIDA AVE**	**C**	BOUGAINVILLEA A	FOWLER AVE
294	**440**	**BUSCH BLVD**	**D**	46TH ST	50TH ST
295	**441**	**FLORIDA AVE**	**C**	SLIGH AVE	BROAD ST
296	**442**	**FLORIDA AVE**	**C**	BROAD ST	BIRD ST
297	**443**	**NEBRASKA AVE**	**D**	WATERS AVE	YUKON ST
298	**444**	**BUSCH BLVD**	**F**	N BOULEVARD	FLORIDA AVE
299	**445**	**NEBRASKA AVE**	**D**	BIRD ST	WATERS AVE
300	**446**	**FLORIDA AVE**	**C**	HANNA AVE	SLIGH AVE
301	**447**	**BUSCH BLVD**	**F**	TWIN LAKES BLVD	ORANGE GROVE DR
302	**448**	**BUSCH BLVD**	**F**	HIMES AVE	TWIN LAKES BLVD
303	**449**	**BUSCH BLVD**	**F**	ORANGE GROVE DR	ARMENIA AVE
304	**452**	**HILLSBOROUGH AVE**	**D**	22ND ST	30TH ST
305	**453**	**HILLSBOROUGH AVE**	**D**	34TH ST	40TH ST
306	**456**	**FLORIDA AVE**	**C**	HILLSBOROUGH AV	HANNA AVE
307	**458**	**NEBRASKA AVE**	**D**	HILLSBOROUGH AV	HANNA AVE
308	**459**	**HILLSBOROUGH AVE**	**D**	30TH ST	34TH ST
309	**461**	**BUSCH BLVD**	**D**	50TH ST	52ND ST
310	**462**	**50TH/56TH ST**	**C**	MELBURNE BLVD	M L KING BLVD
311	**463**	**50TH ST**	**C**	BROADWAY AVE	COLUMBUS DR
312	**464**	**SR 60 / ADAMO DR**	**C**	39TH ST	US HWY 41
313	**465**	**M L KING BLVD**	**C**	40TH ST	LAKE AVE
314	**466**	**M L KING BLVD**	**C**	34TH ST	40TH ST
315	**467**	**M L KING BLVD**	**C**	30TH ST	34TH ST
316	**468**	**M L KING BLVD**	**C**	29TH ST	30TH ST
317	**469**	**M L KING BLVD**	**C**	22ND ST	29TH ST
318	**470**	**56TH ST**	**F**	PURITAN RD	RIVERHILLS DR
319	**471**	**56TH ST**	**F**	SLIGH AVE	PURITAN RD
320	**472**	**56TH ST**	**F**	HANNA RD	SLIGH AVE
321	**473**	**56TH ST**	**F**	HILLSBOROUGH AV	HANNA RD
322	**474**	**40TH ST**	**D**	CHELSEA ST	OSBORNE AVE
323	**475**	**40TH ST**	**D**	M L KING BLVD	CHELSEA ST
324	**476**	**40TH ST**	**D**	LAKE AVE	M L KING BLVD
325	**477**	**M L KING BLVD**	**C**	NEBRASKA AVE	15TH ST
326	**487**	**HILLSBOROUGH AVE**	**F**	EAST LAKE SQ MA	HARNEY RD
327	**488**	**M L KING BLVD**	**C**	50TH ST	I-4
328	**490**	**SR 60 / ADAMO DR**	**F**	US HWY 301	BRANDON CROSSIN
329	**503**	**SR 60 / ADAMO DR**	**C**	22ND ST	34TH ST
330	**504**	**SR 60 / ADAMO DR**	**E**	19TH ST	21ST ST
331	**508**	**22ND ST**	**F**	LINDSEY	DURHAM
332	**509**	**US HWY 41**	**C**	LEE ROY SELMON	ADAMO DR
333	**513**	**CHANNELSIDE DR**	**E**	KENNEDY BLVD	TWIGGS ST

334	**514**	**SR 60 / ADAMO DR**	**E**	CHANNELSIDE DR	19TH ST
335	**531**	**FOWLER AVE**	**F**	50TH ST	52ND ST
336	**532**	**FOWLER AVE**	**E**	56TH STREET	RAINTREE BLVD
337	**533**	**FOWLER AVE**	**E**	RAINTREE BLVD	HOYT AVE
338	**534**	**FOWLER AVE**	**E**	HOYT AVE	GILLETTE AVE
339	**535**	**FOWLER AVE**	**E**	GILLETTE AVE	RIVERHILLS BLVD
340	**536**	**FOWLER AVE**	**E**	RIVERHILLS BLVD	I-75
341	**537**	**50TH ST**	**C**	COLUMBUS DR	I-4
342	**538**	**40TH ST**	**D**	I-4	COLUMBUS DR
343	**540**	**FLORIDA AVE**	**C**	LAKE AVE N	M L KING BLVD
344	**541**	**FLORIDA AVE**	**C**	LAKE AVE S	LAKE AVE N
345	**542**	**NEBRASKA AVE**	**D**	SITKA	BIRD ST
346	**548**	**KENNEDY BLVD**	**F**	HENDERSON BLVD	MACDILL AVE
347	**549**	**DALE MABRY HWY**	**F**	LAMBRIGHT ST	WATERS AVE
348	**552**	**DALE MABRY HWY**	**F**	I-275	SPRUCE ST
349	**553**	**DALE MABRY HWY**	**F**	CYPRESS ST	I-275
350	**554**	**HENDERSON BLVD**	**C**	SWANN AVE	AZEELE ST
351	**555**	**HENDERSON BLVD**	**C**	DALE MABRY HWY	STERLING
352	**557**	**DALE MABRY HWY**	**C**	INTERBAY BLVD	OKLAHOMA
353	**558**	**GANDY BLVD**	**F**	WESTSHORE BLVD	MANHATTAN AVE
354	**563**	**BOY SCOUT BLVD**	**C**	TRASK ST	LOIS AVE
355	**564**	**BOY SCOUT BLVD**	**C**	MEMORIAL HWY	WESTSHORE BLVD
356	**565**	**BOY SCOUT BLVD**	**C**	WESTSHORE BLVD	TRASK
357	**567**	**HILLSBOROUGH AVE**	**F**	LOIS AVE	DALE MABRY HWY
358	**568**	**HILLSBOROUGH AVE**	**F**	WESTSHORE BLVD	LOIS AVE
359	**571**	**HILLSBOROUGH AVE**	**B**	OLD MEMORIAL HW	COUNTRYWAY BLVD
360	**572**	**HILLSBOROUGH AVE**	**E**	WEBB RD	KELLY RD
361	**573**	**HILLSBOROUGH AVE**	**E**	MEMORIAL HWY	WEBB RD
362	**574**	**HILLSBOROUGH AVE**	**E**	SAWYER RD	GEORGE RD
363	**575**	**HILLSBOROUGH AVE**	**E**	HANLEY RD	SAWYER RD
364	**576**	**HILLSBOROUGH AVE**	**E**	KELLY RD	HANLEY RD
365	**577**	**HILLSBOROUGH AVE**	**E**	GEORGE RD	FRONTAGE RD
366	**579**	**HILLSBOROUGH AVE**	**B**	PINELLAS COUNTY	OLD MEMORIAL HW
367	**582**	**GANDY BRIDGE**	**B**	HILLSBOROUGH CO	WESTSHORE BLVD
368	**583**	**DALE MABRY HWY**	**C**	MACDILL AFB	INTERBAY BLVD
369	**585**	**FOWLER AVE**	**F**	52ND ST	56TH ST
370	**586**	**FOWLER AVE**	**E**	RIVERHILLS BLVD	I-75
371	**587**	**US HWY 92**	**C**	KINGSWAY RD	MCINTOSH RD
372	**591**	**US HWY 92**	**C**	FORBES RD	TURKEY CREEK RD
373	**592**	**US HWY 92**	**C**	TURKEY CREEK RD	WALTER DR
374	**593**	**US HWY 92**	**C**	WALTER DR	WOODROW WILSON
375	**595**	**REYNOLDS/SR 600**	**C**	US 92	ALEXANDER ST

376	**599**	**SR 60 / ADAMO DR**	**F**	78TH ST	US HWY 301
377	**601**	**SR 60 / ADAMO DR**	**F**	BRANDON CROSSIN	FALKENBURG RD
378	**602**	**SR 60**	**F**	GRAND REGENCY B	MEMORIAL GARDEN
379	**603**	**SR 60**	**F**	I-75	GRAND REGENCY B
380	**604**	**SR 60 / ADAMO DR**	**E**	FALKENBURG RD	I-75
381	**605**	**SR 60**	**F**	MEMORIAL GARDEN	GORNTO LAKE RD
382	**606**	**KENNEDY BLVD**	**F**	LOIS AVE	CHURCH ST
383	**607**	**KENNEDY BLVD**	**F**	CHURCH ST	DALE MABRY HWY
384	**611**	**39TH ST**	**C**	ADAMO DR	4TH AVE
385	**612**	**39TH ST**	**C**	4TH AVE	7TH AVE
386	**613**	**SR 60**	**B**	MUD LAKE RD	CR 39
387	**614**	**SR 60**	**B**	TURKEY CREEK RD	MUD LAKE RD
388	**617**	**TAMPA ST**	**C**	LAKE AVE	M L KING BLVD
389	**621**	**US HWY 301**	**B**	RIVERVIEW RD	PROVIDENCE CONN
390	**622**	**US HWY 301**	**B**	PROVIDENCE CONN	GORNTO LAKE RD
391	**623**	**HILLSBOROUGH AVE**	**C**	COUNTRYWAY BLVD	MONTAGUE ST
392	**624**	**HILLSBOROUGH AVE**	**C**	MONTAGUE ST	MEMORIAL HWY
393	**627**	**US HWY 301**	**A**	CR 579	KNIGHTS GRIFFEN
394	**628**	**US HWY 301**	**A**	KNIGHTS GRIFFEN	STACY RD
395	**632**	**SR 39**	**C**	ALEXANDER ST EXT	KNIGHTS-GRIFFIN
396	**633**	**SR 39**	**C**	ALEXANDER ST EXT	KNIGHTS-GRIFFIN
397	**639**	**SR 39**	**B**	KNIGHTS-GRIFFIN	PASCO COUNTY LI
398	**640**	**SR 39**	**B**	KNIGHTS-GRIFFIN	PASCO COUNTY LI
399	**641**	**COURTNEY CAMPBELL PKWY**	**D**	MCMULLEN BOOTH	PINELLAS COUNTY

APPENDIX E:
SAMPLE ARTERIAL DATA COLLECTION SHEET
(SOURCE: QUALITY/LEVEL OF SERVICE HANDBOOK)

Arterial Data Collection Sheet

1. Arterial Name:										
2. Direction:										
3. Study Period:										
4. Area Type:										
5. Class:										
6. Segment	1	2	3	4	5	6	7	8	9	10
7. From:										
8. To:										
9. Length:										
10. Lanes:										
11. Posted Speed:										
12. Left Turn Lanes:										
13. AADT:										
14. K*										
15. D*										
16. g/C:										
17. % Sidewalk:										
18. Paved Shoulder/ Bicycle Lane/Outside Lane Width:										
19. Buses/Hour:										
20. Other**										

* Determine at a facility level, not at segment level.

** Generally, defaults are recommended for other input values in the Q/LOS Handbook. Collection of data (e.g., percent turns from exclusive turn lanes at selected intersections) may be appropriate for an individual facility.

LIST OF REFERENCES

1. Abdel-Aty, M.A., H. Salkapuram, C. Lee and P.A. Brady. “A Simplistic, Practical Approach To Identify Traffic Crash Profiles at Signalized Intersections”. ITE Journal, Vol. 76, No. 4, 28–33, 2006.
2. Abdel-Aty, M.A. and J. Keller. “Exploring the overall and specific crash severity levels at signalized intersections”. Accident Analysis and Prevention, Vol. 37, No. 3, 417–425, 2004.
3. Abdel-Aty, M.A. and A.E. Radwan. “Modeling Traffic Accident Occurrence and Involvement”. Accident Analysis and Prevention, Vol. 32, No. 5, 633–642, 2000.
4. Abdel-Aty, M.A. and X. Wang. “Crash Estimation at Signalized Intersections along Corridors: Analyzing Spatial Effect and Identifying Significant Factors”. Transportation Research Record, No. 1953, 98–111, 2006.
5. American Association of State Highway and Transportation Officials. A Policy on Geometric Design of Highways and Streets. 5th ed. Washington, D.C.: AASHTO, 2005.
6. Aven, Paula. “CDOT weighs Concrete vs. Asphalt”. Denver Business Journal. < http://www.bizjournals.com/denver/stories/2000/05/08/story6.html>. Accessed on November 24, 2008.
7. Bhesania, R. “Using Accident Statistics and Characteristics to Improve Safety”. ITE Journal, Vol. 61, No. 3, 37-41, 1991.
8. Das, A., A. Pande, M.A. Abdel-Aty and J.B. Santos. “Urban Arterial Crash Characteristics Related with Proximity to Intersections and Injury Severity”. Presented at the 87th Annual Meeting of the Transportation Research Board, Washington, DC, 2008.

9. ESRI. Official Website. <http://www.esri.com/company/about/history.html>. Accessed June 17, 2008.
10. Fitzpatrick, K., W. Schneider IV and J. Carvell. “Using the Rural Two-Lane Highway Draft Prototype Chapter”. Transportation Research Record, No. 1950, 44–54, 2006.
11. Florida Department of Transportation. Quality/Level of Service Handbook. <http://www.dot.state.fl.us/planning/systems/sm/los/pdfs/QLOS2002.pdf>. Accessed July 5, 2008.
12. Florida Department of Highway Safety and Motor Vehicles. “Traffic Crash Statistics Report: A Compilation of Motor Vehicle Crash Data From the Florida Crash Records Database”. Tallahassee: FDHSMV, 2007.
13. Fontaine, M.D. and S.W. Read. “Development and Evaluation of Virginia’s Highway Safety Corridor Program”. FHWA/VTRC 06-R30. Virginia Transportation Research Council, Charlottesville, Virginia, 2006.
14. Frantzeskakis, J.M. and D.I. Iordanis. “Volume-to-Capacity Ratio and Traffic Accidents on Interurban Four-Lane Highways in Greece”. Transportation Research Record, No. 1112, 29–38, 1987.
15. Garber, N.J. and S. Subramanyan. “Incorporating Crash Risk in Selecting Congestion-Mitigation Strategies – Hampton Roads Area (Virginia) Case Study”. Transportation Research Record, No. 1746, 1–5, 2001.
16. Google Inc. Google Earth [Computer Software]. <http://earth.google.com/>. Accessed June 17, 2008.
17. Green, E.R. and K.R. Agent. “Evaluation of High Traffic Crash Corridors”. KTC-02-08/SPR231-01-1F. Kentucky Transportation Center, Lexington, Kentucky, 2002.

18. Green, P.E. and D. Blower. "Potential Effectiveness of Signal Optimization for Various Corridors in Michigan". UMTRI-2007-5-1. Transportation Research Institute, Ann Arbor, Michigan, 2007.

19. Ha, T. and W.D. Berg. "Development of Safety-Based Level-of-Service Criteria for Isolated Signalized Intersections". Transportation Research Record, No. 1484, 98–104, 1995.

20. Hall, J.W. and M. Polanco de Hurtado. "Effect of Intersection Congestion on Accident Rates". Transportation Research Record, No. 1376, 71–77, 1992.

21. Handy, S. "Regional transportation planning in the US: An examination of changes in technical aspects of the planning process in response to changing goals". Transport Policy, Vol. 15, No. 2, 113–126, 2008.

22. Hardin, J.W. and J.M. Hilbe. Generalized Estimating Equations. Boca Raton: Chapman & Hall/CRC, 2003.

23. Hauer, E. Observational Before-After Studies in Road Safety. Oxford: Pergamon Press, 2002.

24. Hauer, E., J.C.N. Ng and J. Lovell. "Estimation of Safety at Signalized Intersections". Transportation Research Record, No. 1185, 48–61, 1988.

25. Indiana Department of Transportation. "US31 Kokomo Corridor Project, Howard and Tipton Counties, Indiana". FHWA-IN-EIS-05-01-FS. Indiana Department of Transportation, Indianapolis, Indiana, 2003.

26. Jernigan, J.D. "Comparative Case Studies of Corridor Safety Improvement Efforts". Final Report. FHWA/VTRC 00-R17. Virginia Transportation Research Council, Charlottesville, Virginia, 1999.

27. Kononov, J. and B.K. Allery. “Explicit Consideration of Safety in Transportation Planning and Project Scoping”. Transportation Research Record, No. 1897, 116–125, 2004.

28. Kononov, J. and B.K. Allery. “Level of Service of Safety: Conceptual Blueprint and Analytical Framework”. Transportation Research Record, No. 1840, 57–66, 2003.

29. Kramer, J. “Review of MPO Long Range Transportation Plans and Regional MPO Planning Activities and Products”. Center for Urban Transportation Research, Tampa, Florida, 2005.

30. Ladrón de Guevara, F., S.P. Washington and J. Oh. “Forecasting Crashes at the Planning Level: Simultaneous Negative Binomial Crash Model Applied in Tucson, Arizona”. Transportation Research Record, No. 1897, 191–199, 2004.

31. Lee, S. and W.D. Berg. “Development of Safety-Based Level-of-Service Parameters for Two-Way Stop-Controlled Intersections”. Transportation Research Record, No. 1635, 127–132, 1998.

32. Liang, K.Y. and S.L. Zeger. “Longitudinal Data Analysis Using Generalized Linear Models”. Biometrika, Vol. 73, No. 1, 13–22, 1986.

33. Lin, D.Y., L.J. Wei and Z. Ying. “Model-Checking Techniques Based on Cumulative Residuals”. Biometrics, Vol. 58, No. 1, 1–12, 2002.

34. Lord, D., A. Manar and A. Vizioli. “Modeling crash-flow-density and crash-flow-V/C ratio relationships for rural and urban freeway segments”. Accident Analysis and Prevention, Vol. 37, No. 1, 185–199, 2005.

35. Lord, D. and B. Persaud. “Accident Prediction Models With and Without Trend: Application of the Generalized Estimating Equations (GEE) Procedure”. Transportation Research Record, No. 1717, 102–108, 2000.

36. Lord, D., S.P. Washington and J.N. Ivan. “Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory”. Accident Analysis and Prevention, Vol. 37, No. 1, 35–46, 2005.

37. Lu, J., F. Pan, Q. Xiang and G. Zhang. “Level of Safety Service for Safety Performance Evaluation of Highway Intersections”. In 87th Annual Meeting of the Transportation Research Board Compendium of Papers, CD-ROM, Washington, D.C., 2008.

38. McTrans. McTrans January Newsletter, Volume 8. Gainesville: University of Florida, 2008.

39. National Highway Traffic Safety Administration. “Traffic Safety Facts 2005: A compilation of motor vehicle crash data from the Fatality Analysis Reporting System and the General Estimate System”. Washington, D.C.: NHTSA, 2006.

40. Ogden, K.W., S. Newstead, P. Ryan and S. Gantzer. “Factors affecting crashes at signalised intersections”. Report No. 62. Accident Research Centre, Victoria, Australia, 1994.

41. Persaud, B.N. and L. Dzbik. “Accident Prediction Models for Freeways”. Transportation Research Record, No. 1401, 55–60, 1993.

42. Persaud, B.N. and T. Nguyen. “Disaggregate Safety Performance Models for Signalized Intersections on Ontario Provincial Roads”. Transportation Research Record, No. 1635, 113–120, 1998 (a).

43. Persaud, B.N. and T. Nguyen. “Safety Considerations in Capacity Analysis”. Presented at the Third International Symposium on Highway Capacity, Copenhagen, Denmark, 817–831, 1998 (b).

44. Plazak, D.J. and R.R. Souleyrette. “Process to Identify High Priority Corridors for Access Management Near Large Urban Areas in Iowa Using Spatial Data”. Proceedings of the 2003 Mid-Continent Transportation Research Symposium, Ames, IA, 2003.

45. Rees, J. “Corridor Management: Identifying Corridors with Access Problems and Applying Access Management Treatments, A U.S. 20 Study”. Center for Transportation Research and Education, Ames, Iowa, 2003.

46. SAS Institute Inc. SAS OnlineDoc 9.1.3. < http://support.sas.com/onlinedoc/913/docMainpage.jsp>. Accessed September 19, 2008.

47. Souleyrette, R.R., M.M. O’Bryen, T. McDonald, H. Preston and R. Storm. “Effectiveness of All-Red Clearance Interval of Intersection Crashes”. MN/RC-2004-26. Center for Transportation Research and Education, Ames, Iowa, 2004.

48. Stamatiadis, N. and G. Puccini. “Fatal Crash Rates in the Southeastern United States: Why Are They Higher?” Transportation Research Record, No. 1665, 118–124, 1999.

49. Sun, X., Y. Li, D. Magri and H.H. Shirazi. “Application of Highway Safety Manual Draft Chapter: Louisiana Experience”. Transportation Research Record, No. 1950, 55–64, 2006.

50. Transportation Research Board. Highway Capacity Manual. 4th ed. Washington, D.C.: National Research Council, 2000.

51. University of California Los Angeles. SAS Annotated Output: Negative Binomial Regression. < http://www.ats.ucla.edu/stat/sas/output/sas_negbin_output.htm>. Accessed November 3, 2008.

52. United States Department of Transportation - Federal Highway Administration. SAFETEA-LU Official Website. <http://www.fhwa.dot.gov/safetealu/index.htm>. Accessed May 5, 2008.

53. Wang, X. “Safety Analyses at Signalized Intersections Considering Spatial, Temporal and Site Correlation”. Ph.D. Dissertation. Orlando: University of Central Florida, 2006.

54. Wang, X. and M.A. Abdel-Aty. "Right-Angle Crash Occurrence at Signalized Intersections". Transportation Research Record, No. 2019, 156–168, 2007.

55. Wang, X. and M.A. Abdel-Aty. "Temporal and Spatial Analyses of Rear-end Crashes at Signalized Intersections". Accident Analysis and Prevention, Vol. 38, No. 6, 1137–1150, 2006.

56. Wang, X., M.A. Abdel-Aty, A.M. Almonte and A.L. Darwiche. "Incorporating Traffic Operation Measures in Safety Analysis at Signalized Intersections". Forthcoming presentation at the 88th Annual Meeting of the Transportation Research Board, Washington, DC, 2009.

57. Zegeer, J.D., M. Vandehey, M. Blogg, K. Nguyen and M. Ereti. NCHRP Report 599: Default Values for Highway Capacity and Level of Service Analyses. Transportation Research Board of the National Academies, Washington, D.C., 2008.

58. Zhang, L. and P.D. Prevedouros. "Signalized Intersection LOS that accounts for Safety Risk". In 82nd Annual Meeting of the Transportation Research Board Compendium of Papers, CD-ROM, Washington, D.C., 2003.

59. Zhou, M. and V.P. Sisiopiku. "Relationship Between Volume-to-Capacity Ratios and Accident Rates". Transportation Research Record, No. 1581, 47–52, 1997.

60. Ziegler, A., C. Kastner and M. Blettner. "The Generalised Estimating Equations: An Annotated Bibliography". Biometrical Journal, Vol. 40, No. 2, 115–139, 1998.

61. Zorn, C.J.W. "Generalized Estimating Equation Models for Correlated Data: A Review with Applications". American Journal of Political Science, Vol. 45, No. 2, 470–490, 2001.