A METHODOLOGY TO INCORPORATE HIGHWAY SAFETY INTO TRANSPORTATION PLANNING USING GENERALIZED LINEAR MIXED MODELS

By

Erika Cristina Jaramillo Giraldo
Dissertation submitted in partial fulfillment of the requirements for the degree of:

DOCTOR OF PHILOSOPHY IN CIVIL ENGINEERING
UNIVERSITY OF PUERTO RICO
MAYAGÜEZ
2014

Approved by:

______________________________ ________________
Didier Valdés Díaz, PhD — Date
Chairman, Graduate Committee

______________________________ ________________
Alberto Figueroa Medina, PhD — Date
Member, Graduate Committee

______________________________ ________________
Ivette Cruzado Vélez, PhD
Member, Graduate Committee — Date

______________________________ ________________
Pedro Torres Saavedra, PhD
Member, Graduate Committee — Date

______________________________ ________________
Raúl Macchiavelli, PhD — Date
Member, Graduate Committee

______________________________ ________________
Edgardo Lorenzo, PhD — Date
Representative of Graduate Studies

______________________________ ________________
Ismael Pagán Trinidad, M.S.C.E. — Date
Department Director

ABSTRACT

Highway Safety has been identified as a significant problem worldwide. Crashes have been found as the second cause of death in the world according to the World Health Organization (2010). In fact, road crashes cost billions of dollars per year in the US alone. The US Department of Transportation has established highway safety as one of their main priorities in their Action Plan that mainly consists of establishing countermeasures and engineering strategies for the reduction of crashes. Several efforts are underway but most of the implemented strategies in many states and Puerto Rico have a reactive or a short-term planning approach. Such approaches have generated some improvements to the current system (Lovegrone, 2006). However, a proactive approach is necessary. This approach would require incorporating highway safety aspects in the decision making process from the beginning when planning alternatives are generated and crash data is unknown (de Leur, 2001).

Currently, Safety Performance Functions (SPFs) are considered by many as the main tool in estimating a road's safety and an integral part of decision making. SPFs are mathematical models that are statistically developed to conduct crash data analysis. The models attempt to explain crash occurrence on various road facilities types as a function of the traffic and geometric characteristics of these facilities.

SPFs are not just valuable to the success of the reactive approach to dealing with road safety problems; they are of vital importance to the success of the proactive approach. The primary objective of the proactive approach is to ensure that road safety is an explicit priority in transportation planning policies.

Several crash prediction models have been developed for site or project analysis, but very few of them are for planning purposes. Strategic, mid and short-term planning models with a wide prediction range, due to a wide range of Average Annual Daily Traffic (AADT) measures, were not found in the researched literature. A model of this type could be used to calculate the number of average crashes per type of road while considering conceptual design aspects or design changes. They could also be used in the implementation of safety devices for the whole region. These types of models were developed in this research project.

This research project utilized Generalized Linear Mixed Models (GLMM) so as to use them in the incorporation of highway safety into the strategic planning process. These models can be used to forecast the rate of crashes for different planning and conceptual design scenarios. These GLMM have several advantages, in terms of predicting crash rates, including the incorporation of not only a set of known explanatory variables, but also of random effects present in the system. As a result, the model explains possible temporal correlation and spatial effects in the data. Therefore, these types of models offer great versatility in the modeling of crash rates and its related factors.

The research approach included the filtering of a crash database according to a set of identified explanatory variables, estimation of parameters in a set of candidate GLMM's, evaluation of the estimated models using several statistical tests and goodness of fit methods, and the selection of models that represent a better fit for the phenomena under study.

Once the models were obtained, a methodology for their incorporation into the strategic planning process was developed and reported. Therefore, the deliverables included are GLMM crash prediction models for municipalities and different types of segments on expressways-freeways and arterials (population average and specific subject models), along with a methodology for the incorporation of these models in safety analysis which is a vital part of the strategic planning process.

RESUMEN

La seguridad vial ha sido identificada como un problema significativo a nivel mundial. Se ha encontrado que los choques son la segunda causa de muerte de acuerdo con la organización mundial de la salud (2010). De hecho, los choques cuestan billones de dólares por año en los Estados Unidos. El Departamento de Transportación de Estados Unidos ha establecido la seguridad vial como una de sus prioridades en su plan de acción, el cual consiste principalmente en el establecimiento de medidas de mitigación y estrategias de ingeniería para la reducción de choques. Varios esfuerzos se están llevando a cabo pero la mayoría de las estrategias implementadas en muchos estados y en Puerto Rico tienen un enfoque reactivo de corto plazo. Estos enfoques han generado algunas mejoras al sistema actual (Lovegrone, 2006). Sin embargo, es necesario un enfoque proactivo. Este enfoque requeriría incorporar aspectos de seguridad vial desde el principio del proceso de toma de decisiones cuando se generan las alternativas y los datos de choques son desconocidos (de Leur, 2001).

Actualmente, las Funciones de Desempeño en Seguridad (SPF's por sus siglas en inglés), son consideradas por muchos como la herramienta principal en la estimación de la seguridad de una vía y parte integral de la toma de decisiones. Las funciones SPF son modelos matemáticos que se desarrollan con estadística para conducir análisis de datos de choques. Los modelos intentan explicar los choques que ocurren en varios tipos de instalaciones viales como función de las características del tránsito y la geometría de estas instalaciones.

Las SPF no son solo valiosas para el éxito del enfoque reactivo para tratar con problemas de seguridad vial; son de vital importancia para el éxito del enfoque

proactivo. El objetivo principal del enfoque proactivo es asegurar que la seguridad vial es una prioridad explicita de las políticas de planificación estratégica del transporte.

Varios modelos de predicción de choques han sido desarrollados para análisis de sitios o proyectos, pero pocos han sido desarrollados para propósitos de planificación. Modelos de planificación estratégica, a mediano y corto plazo por tipo de carretera y para un amplio rango de predicción, debido a un rango amplio de medidas de tráfico promedio diario anual (AADT, por sus siglas en inglés), no se encontraron en la literatura. Un modelo de este tipo puede ser usado para calcular el número promedio de choques por tipo de carretera, considerando aspectos de diseño conceptual o cambios en el diseño. También pueden ser usados en la implementación de dispositivos de seguridad para una región completa. Esos tipos de modelos fueron desarrollados en este proyecto.

En este trabajo se ajustaron modelos lineales generalizados mixtos (GLMM) con el propósito de usarlos en la incorporación de la seguridad vial en el proceso de planificación estratégica y la planificación a mediano y corto plazo. Estos modelos pueden ser utilizados para predecir las tasas de choque para diferentes escenarios de planificación y diseño conceptual. Esos modelos GLMM tienen varias ventajas en términos de la predicción de tasas de choques incluyendo la incorporación no sólo de un grupo de variables conocidas, también permiten la inclusión de efectos aleatorios presentes en el sistema. Como resultado, el modelo explica la posible correlación temporal y los efectos espaciales en los datos. Por consiguiente, esos tipos de modelos ofrecen gran versatilidad en la modelación de tasas de choques y sus factores contribuyentes.

El enfoque de esta investigación incluye filtrar los datos de choques de acuerdo a un conjunto de variables explicativas identificadas, la estimación de parámetros en un grupo de modelos GLMM candidatos, la evaluación de los modelos estimados utilizando varias pruebas estadísticas y métodos de bondad de ajuste, y la selección de los modelos que representan mejor el ajuste al fenómeno bajo estudio.

Una vez obtenidos los modelos, se desarrolló una metodología que permite su incorporación dentro del proceso de planificación. Por lo tanto, los aportes de este trabajo incluyen modelos GLMM de predicción por municipios y para diferentes tipos de segmentos para autopistas y arterias (modelos promedio poblacional y sujeto específicos), junto con una metodología para incorporarlos en el análisis de seguridad vial que es una parte vital del proceso de planificación estratégica.

DISCLAIMER

The data used in this research was obtained through multiple agencies in charge of the collection of data in their respective areas. The author does not take responsibility for the accuracy of the data used in the research.

DEDICATION

This work is dedicated to the great living God, whose power has been the engine in my life, to my husband, who is the most important gift in my life, to my father, may he rest in peace, and to my mother for the values she instilled in me and her unconditional support in the course of my lifetime. To my aunt Nora who has been an unconditional help in every moment of my life.

ACKNOWLEDGEMENTS

I would like thank Professor Didier Valdés for his guidance and support, Dr. Cruzado for her work as member of my committee, and a special thanks goes to Professor Raúl Macchiavelli, who allowed me to learn from his statistics courses and thus, easily understand complex concepts. I would also like to extend my gratitude to Professor Pedro Torres, who gave me his technical support in critical moments of my work, to Dr. Alberto Figueroa for his support in the initial phase of this research, to Planner Gabriel Rodriguez for giving me very useful information for my work, and to Andrea Valdés for collaborating with the wording and revision of my work. I thank the members of the road safety office (Engineer Juan Carlos Rivera and Alexis Neváres) and the 'HPMS' office (planner Miguel A. Martínez) for giving me the necessary data for the development of this work. I would also like to thank all the people who in one way or another allowed this work to become a reality.

TABLE OF CONTENTS

ABSTRACT II
RESUMEN V
DISCLAIMER VIII
DEDICATION IX
ACKNOWLEDGEMENTS X
TABLE OF CONTENTS XI
LIST OF TABLES XIV
LIST OF FIGURES XVI
ACRONYMS XVII
CHAPTER 1: INTRODUCTION 1
1.1 INTRODUCTION 1
1.2 JUSTIFICATION 4
1.3 SCOPE 5
1.4 OBJECTIVES 6
1.5 RESEARCH APPROACH 7
1.6 ORGANIZATION 10
CHAPTER 2: BACKGROUND REVIEW 11
2.1 DEFINITION AND GENERAL ASPECTS OF A SAFETY MANAGEMENT SYSTEM 11
2.2 IDENTIFICATION OF AREAS AT HIGH RISK OF CRASHES “HSIP’ AND MICRO MODELS 12
2.3 MACRO LEVEL PREDICTION MODELS FOR A PROACTIVE APPROACH TO SAFETY 15
2.4 PLANNING FOR A PROACTIVE APPROACH TO BE IMPLEMENTED AT THE STATE LEVEL FOR THE SAFETY MANAGEMENT SYSTEM OF SOME STATES 18
2.4.1 Step 1: Incorporating safety into the vision statement *18*
2.4.2 Step 2: Incorporating safety into the set of goals and objectives *20*
2.4.3 Step 3: Incorporating safety into system performance measures *20*
2.4.4 Step 4: Incorporating safety into technical analysis *21*
2.4.5 Step 5: Evaluating alternative projects and strategies *22*
2.4.6 Step 6 Developing the plan and program and Step 7 Monitoring system performance *23*
Step 6: Developing the plan and program *23*
Step 7: Monitoring system performance *23*
2.5 EXAMPLE OF THE PHOENIX SAFETY MANAGEMENT SYSTEM 23
2.6 DISCUSSION 24
CHAPTER 3: TRANSPORTATION PLANNING AND HIGHWAY SAFETY IN PUERTO RICO AND GENERAL STATISTICS 26
3.1 BACKGROUND 26
3.2 THE LOCAL INCORPORATION OF SAFETY INTO THE PLANNING PROCESS AND ITS ASSEMENT 30

3.3 SUMMARY AND ANALYSIS OF CRASH STATISTICS IN THE WESTERN REGION, PER MUNICIPALITY 33

CHAPTER 4: CALIBRATION OF MACRO MODELS FOR HIGHWAY SAFETY PLANNING 38

4.1 DETAILED METHODOLOGY 38
4.2 LITERATURE REVIEW AND IDENTIFICATION OF EXISTING MODELS 40
4.3 PROCEDURES USED IN PUERTO RICO 40
4.4 COLLECTING DATA AND PREPARATION 40
4.5 CALIBRATION AND APPLICATION OF MODELS FOUND IN THE LITERATURE 45
4.5.1 Generation of the expected crash counts in Excel 46
4.5.2 Calculation of correction factors for the base year 46
4.5.3 Prediction of crashes with models found in the literature 48
4.5.4 Comparison of the coefficients of variation of the correction factors to observe if the data fits the model 48
4.5.5 Identification of independent variables used in the western region model of Puerto Rico 51
4.5.6 Development of the Puerto Rican highway crash prediction model and the selection of the best model 52
4.5.6.1 Data collection and analysis of significant variables 52
4.5.6.2 Defining distribution and model structure 53
4.5.6.3 Estimation using the Maximum Likelihood (ML) Method and model selection 53
4.5.6.4 Models for total crashes 53
4.5.6.5 Models for fatal crashes 54
4.5.6.6 Models for injury crashes 55
4.6 RESULTS AND DISCUSSION 56

CHAPTER 5: DEVELOPING MACRO MODELS 61

5.1 INTRODUCTION 61
5.2 DETAILED METHODOLOGY 62
5.2.1 Data collection and analysis of significant variables 62
5.2.2 Defining the model distribution and structure 64
5.2.3 Estimation using the Maximum Likelihood (ML) method 64
5.2.4 Model selection 65
5.2.5 Model integration into the planning process 65
5.3 RESULTS AND DISCUSSION OF APKPY MODELS 66
5.3.1 Model with rate (APKPY) as the dependent variable 67
5.3.2 Estimated parameters for APKPY models and interpretation 67
5.4 RESULTS AND DISCUSSION OF APIPKPY MODELS 69
5.4.1 Comparison of models with double rate (APIPKPY) as the dependent variable 69
5.4.2 Estimated parameters for APIPKPY the models and interpretation 70

CHAPTER 6: DEVELOPMENT OF FREEWAY-EXPRESSWAY AND ARTERIAL MODELS 73

6.1 INTRODUCTION 73
6.2 DETAILED METHODOLOGY 75
6.2.1 Data collection and analysis 78
6.2.2 Defining the distribution and model structure 81
6.2.3 Estimation using the Maximum Likelihood (ML) Method 82
6.2.4 Model selection 83
6.2.5 Work plan and tasks 83
6.3 RESULTS AND DISCUSSION OF FREEWAYS AND ARTERIALS (APKPY) MODELS 85

6.3.1 Comparison of freeway and arterial models with rate (APKPY) as the dependent variable *86*
6.3.1.1 Comparison of significant freeway models 86
6.3.1.2 Comparison of significant arterials models 88
6.3.2 Estimated parameters of APKPY models and interpretation for freeways and arterials *89*
a. Estimated parameters to freeways 89
b. Estimated parameters for arterial models 92

CHAPTER 7: INCORPORATION OF MODELS INTO THE SAFETY MANAGEMENT SYSTEM 98

7.1 INTRODUCTION 98
7.2 DETAILED METHODOLOGY 100
7.2.1 Proactive use or safety planning system and Reactive use or short term planning *103*
7.2.1.1 Population average and specific subject models as tools to proactive technical analysis of alternative strategies for improvement 103
7.2.1.2 Reactive use (HSIP) or short term safety planning 104
7.2.1.3 Specific subject models as tools for identification of hazardous sites and regions respectively 105
7.3 APPLICATION EXAMPLES 107
7.3.1 Example 1 – Proactive approach (municipality model) for planning goals *108*
7.3.2 Example 2- Proactive approach (municipality model) for risk rating *109*
7.3.3 Example 3- Proactive approach (freeway and arterial model) for introducing safety as part of the strategic planning process *113*
7.3.4 Example 4 – Reactive approach (freeway and arterial model) for Risk Rating of Segments *116*

CHAPTER 8: CONCLUSIONS 117

8.1 LOCAL PROCEDURES 117
8.2 CALIBRATION OF ARIZONA MODELS TO WESTERN REGION AND DEVELOPMENT OF WESTERN REGION MODELS 118
8.3 REGIONAL MODELS 119
8.4 FREEWAY AND ARTERIAL MODELS 120
8.5 FRAMEWORK FOR THE INCORPORATING OF SAFETY INTO LONG AND MID RANGE URBAN PLANNING 124
8.6 GENERAL CONCLUSIONS 124

CHAPTER 9: RECOMMENDATIONS 126

9.1 LOCAL PROCEDURES 126
9.2 CALIBRATION OF ARIZONA MODELS FOR THE WESTERN REGION AND THE DEVELOPMENT OF WESTERN REGION MODELS 126
9.3 REGIONAL MODELS 128
9.4 FREEWAY AND ARTERIAL MODELS 129
9.5 FRAMEWORK FOR THE INCORPORATION OF SAFETY INTO LONG AND MID RANGE URBAN PLANNING 129
9.6 GENERAL RECOMMENDATIONS 129
REFERENCES 130

LIST OF TABLES

Table 4.1 Data for the total crashes model for the western region year 2002 (DTPW and HPMS database). *42*

Table 4.2 Data for the total crashes model for the western region year 2006 (DTPW and HPMS database) *43*

Table 4.3 Data for the fatal crashes model for the western region year 2002 (DOTPW and HPMS database) *43*

Table 4.4 Data for the fatal crashes model for the western region year 2006 (DOTPW and HPMS database) *44*

Table 4.5 Total crashes spreadsheet *49*

Table 4.6 Fatal crashes spreadsheet *50*

Table 4.7 Log-normal total crash model prediction analysis. *56*

Table 4.8 Log linear- negative binomial total crash model prediction analysis *57*

Table 5.1 GLM Model comparison of freeways with APKPY as the dependent variable *67*

Table 6.1 Descriptive statistics of significant variables for freeways. *80*

Table 6.2 Descriptive statistics of significant variables for Arterials. *81*

Table 6.3 GLMM Model comparison of freeways with APKPY as the dependent variable *87*

Table 6.4 Model comparison for arterials with APKPY as the dependent variable. *88*

Table 6.5 Estimated parameters for total freeway crash models with APKPY as the dependent variable *89*

Table 6.6 Estimated parameters for freeway crash injury models with APKPY as the dependent variable *90*

Table 6.7 Estimated parameters for fatal freeway crash models with APKPY as the dependent variable *91*

Table 6.8 Estimated parameters for models of total arterial crashes with APKPY as the dependent variable. *92*

Table 6.9 Estimated parameters for arterial injury crash models with APKPY as the dependent variable. *93*

Table 6.10 Estimated parameters for fatal arterial crash models with APKPY as the dependent variable. *94*

Table 7.1 Example 1. .... *109*

Table 7.2 Crash risk of municipalities in the Arecibo region. *110*

Table 7.3 Crash risk of municipalities in the Aguadilla region. *110*

Table 7.4 Crash risk of municipalities in the Humacao region. *111*

Table 7.5 Crash risk of municipalities in the Mayaguez region. *111*

Table 7.6 Crash risk of municipalities in the San Juan region.... *111*

Table 7.7 Crash risk of municipalities in the Fajardo region... *112*

Table 7.8 Crash risk of municipalities in the Ponce region... *112*

Table 7.9 Freeway total crashes SPF without metal barrier (North area). *113*

Table 7.10 Freeway fatal crashes SPF without metal barrier (North area). *114*

Table 7.11 Freeway total crashes SPF without metal barrier (South area)................................... *114*

Table 7.12 Freeway fatal crashes SPF without metal barrier (South area).................................... *114*

Table 7.13 Arterial total crashes SPF without metal barrier (North area)...................................... *114*

Table 7.14 Arterial fatal crashes SPF without metal barrier (North area)...................................... *115*

Table 7.15 Arterial total crashes SPF without metal barrier (South area). *115*

Table 7.16 Arterial fatal crashes SPF without metal barrier (South area). *115*

LIST OF FIGURES

Figure 1.1 The Commonwealth of Puerto Rico- the Area of study. 5

Figure 1.2 General Methodology. 8

Figure 2.1 Incorporating Safety into the Planning Process. 19

Figure 2.2 Phoenix Safety Management System. 24

Figure 3.1 State highway safety program. 28

Figure 3.2 Highway miles and kilometers in the western region year 2006 (HPMS database) 33

Figure 3.3 Population density in the western region year 2002 (Census estimates). 34

Figure 3.4 Population 16-64 in the western region year 2002 (Census estimates) 34

Figure 3.5 Total population in the western region year 2002 (Census estimates). 35

Figure 3.6 Proportion of Interstates in the western region year 2006 (HPMS database) 35

Figure 3.7 Crash rates in the western region (DTPW 2002-2006 database) standardized by length. 36

Figure 3.8 Crash rates in the western region (DTPW 2002-2006 database) standardized by population 36

Figure 4.1 Methodology for the calibration of Arizona Models and the development of Western Puerto Rico Models. 39

Figure 5.1 Methodology for all Puerto Rico municipality models. 63

Figure 6.1 Methodology for Puerto Rico freeway and arterial models. 76

Figure 6.2 Example of SPF of APKPY by AADT for region 1 for freeways 96

Figure 7.1 Detailed methodology to incorporate safety into transportation planning process through SMS (Adapted of Arizona SMS) 100

Figure 7.2 Percentile curves for freeways model 2 and shoulder type 6 106

ACRONYMS

APIPKPY- Accidents per Inhabitant per Kilometer per Year

APMVTPKPY- Accidents per Million of Vehicle Travel per Kilometer per Year

APKPY- Accidents per Kilometer per Year

CARE- Critical Analysis Rate Environment

DTPW- Department of Transportation and Public Works

FARS- Fatality Analysis Reporting System

GLM – Generalized Linear Model

GLMM- Generalized Linear Mixed Models

HSM- Highway Safety Manual

HPMS- Highway Performance Monitoring System

HSIP- Highway Safety Improvement Program

HSP- Highway Safety Plan

ISTEA- Intermodal Surface Transportation Efficiency Act

LRTP- Long Range Transportation Plan

MAP-21- Moving Ahead for Progress in the 21^{st} Century Act

MPO- Metropolitan Planning Organization

NCHRP- National Cooperative Highway Research Program

NHTA-National Highway Traffic Safety Administration

PRTSC- Puerto Rico Traffic Safety Commission

SHSP- Strategic Highway Safety Planning

SAFETY-LU – Safe, Accountable, Flexible and Efficient Transportation Equity Act- A Legacy for Users

SMS- Safety Management System

SPF- Safety Performance Functions

TAZ- Traffic Analysis zones

TIP- Transportation Improvement Program

USC- United Stated Code

VMT- Vehicles Mile Traveled

CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

Around the world the quantity of private vehicles on public roads has experienced a sustained increase since their invention. As a result, the number of collisions have also increased at a rapid pace. Highway safety has been one of the main concerns in urban transportation in the last two decades (Naderan and Shahi 2010) and is of the utmost importance today. To give an example of the scale of impact that this issue has, is the fact that collisions costs the United States approximately $230 billion per year (Cambridge Systematics Inc., 2008).

Highway safety is not only a concern in the United States, it is also an important issue worldwide. The United Nations, private entities, and governments in more than 100 countries have come together in a joint effort to improve highway safety through a project called “A Decade of Action for Road Safety”, which was released on May 11, 2011. The goal of this project is to prevent five million deaths globally by 2020. The overall plan consists of the implementation of activities divided into five pillars: road safety management, safer roads and mobility, safer vehicles, safer road users, and post-crash response.

In the United States, the federal government has raised awareness on the importance of safety, embedding this issue at the core of the transportation law enacted in 2012 called "Moving Ahead for Progress in the 21st Century" (MAP-21). This transportation law establishes measures for the State Departments of Transportation (DOT's) to develop performance targets in conjunction with Metropolitan Planning

Organizations (MPO's) in order to considerably improve the safety of all transportation systems. The law stipulates that each state must make investments that are cost-effective and resource efficient, enabling the achievement of the set objectives.

One of the ways that states are working towards this goal is through the implementation of a Highway Safety Improvement Program (HSIP). These programs aim to reduce highway safety problems by focusing on the identification, diagnosis and remediation of hazardous sites in the short term (planning 3-5 years). These programs typically use crash prediction models at the microscopic level to identify, evaluate, and plan safety measures for each site. These models are designed to predict crashes on new sites, existing sites, or at the project level. These models use traffic volume as the independent variable, and as group characteristics, specific aspects such as the horizontal and vertical geometry of the road, and the operation type of the intersection. The results of these models enable agencies to identify and prioritize hazardous sites, and consequently lead to cost-effective mitigation measures and improvements to the sites. HSIP’s have served their purpose in terms of improving existing conditions. The strategic planning process however, lacks specific tools that could help planners focus on safety issues of the alternatives evaluated at the planning level.

The current planning process in several states and Puerto Rico does take into account the mobility and environmental impacts of each of the alternative being considered; however, that is not the case with regards to safety. In Puerto Rico, a Strategic Highway Safety Plan (SHSP) is utilized during the planning process. This plan was developed by PRHTA and their main objective is that all safety initiatives are fully coordinated and developed based on current Puerto Rico trends and statistical evaluations. Unfortunately, there is no model suggested to perform this analysis. In contrast, highway mobility is analyzed throughout the urban planning process. The typical result of an urban planning process includes long term demand predictions which helps generate alternative solutions or projects that would improve the highway mobility. Similarly, the environmental impacts of all the alternatives being considered are analyzed in detail. A similar approach should be implemented so as to consider estimated highway safety issues (e.g. expected amount of crashes and their severity) during the strategic planning process, in order to understand the safety impacts of each planning alternative and establish safety improvement strategies from day one.

In order to establish a thorough safety approach appropriate prediction tools should be provided for each planning stage. Macro level models are needed for a whole city or region in order to identify existing and future potential safety problems. Typical independent variables used in long range planning macro models include socioeconomic and demographic variables of the areas of interest. The result of this type of model is the prediction of the number and severity of accidents, and not the explanation of their causes.

A review of existing literature did not unearth any evidence of the development of prediction models in Puerto Rico that permit the implementation of a proactive approach during medium and long term planning. The Puerto Rico Transportation Plan for 2040 indicates that safety is one of its main objectives, but it does not describe a methodology to use to integrate this topic long term, nor a manner in which to describe the possible results at a network level. The characteristics of a successful conceptual design or a map identifying road segments with the greatest safety problems are not provided. In other words, there are no projects or specific strategies to mitigate such situations in the long-term planning process. The development of more detailed studies of the zones would make resource investment more cost-effective and the process more efficient for the agencies.

Models by road and segment type are needed in order to identify existing and future potential safety problems and to identify where the regional problems are. The variables used in these models are related to geometric and operational characteristics.

This investigation proposes the development of a methodology that allows for the incorporation of highway safety within planning procedures through the use of models in order to predict average crash rates by region, type of road, and road segment. The developed models will also allow an analysis at the network level. These models will be sensitive to improvements and management strategies of the mobility of the road network. In the models, the temporal correlation and spatial effects of the data will be considered, an aspect which is not considered in prediction models at the segment level found in the Highway Safety Manual (HSM, 2010). A planning process that integrates safety would allow the evaluation of multiple and large project alternatives.

1.2 JUSTIFICATION

Puerto Rico has faced highway safety problems similar to those of the United States. There are various procedures in place for the regional planning of transportation systems that need to be revised in light of the research conducted in this thesis. Therefore, the results of this project have the potential to become an important and necessary part of the regional planning process in the near future. The application of the new models and methodologies will help in the implementation of safety features from the beginning of any transportation project plan and for the future needs of the system.

In terms of worldwide strategic goals safety, mobility, and the environment are taken into consideration so as to support the future benefits viewed in the implementation of the results of this research.

1.3 SCOPE

The scope of the present study is the development of statistical models specific to Puerto Rico, for municipalities, freeway-expressway, and arterial roadways for various crash severity types. These models will determine the relationship between crash frequency, crash rates, and roadway characteristics, and will predict crashes for use in short, mid and long term planning procedures (See Figure 1.1).

Figure 1.1 The Commonwealth of Puerto Rico- the Area of study.

Due to the nature of safety data, which cannot be obtained in a controlled setting, observational studies have been selected as the most practical method to explore the relationships between crashes and their related factors. Controlled laboratory experiments would produce quicker and more accurate results because of the ability to control some factors, but these methods are not typically available in the safety analysis field due to variations in conditions and safety concerns (Hauer, 1997). In order to determine the contribution of a variety of accident prediction variables to the prediction of crashes, and to provide a more proactive model, a multivariate analysis of the independent variables has been developed.

1.4 OBJECTIVES

The main objective of this research work is the development of models that could be incorporated into the short, mid and long-term highway safety planning processes.

The specific objectives that guide the research process are the following:

i. Identify and describe the current procedures used by Puerto Rico address highway safety issues in the mid and long term planning process.

ii. Identify highway safety predictive models used in mid and long range transportation plans in Puerto Rico and in the United States.

iii. Evaluate the effectiveness and possible application of such models using a local test bed area by collecting information on transportation systems, population (census), accidents, and other relevant data.

iv. Develop and calibrate different types of highway safety predictive models. The models will be developed in two phases. The first phase involves the development of planning models that can be applied to the regional level. The second phase will develop planning models using local data by functional classification and road geometry (Safety Performance Functions, SPF).

v. Incorporate the use of highway safety models into a methodology to be applied for the mid and long-term highway safety planning processes in Puerto Rico.

1.5 RESEARCH APPROACH

This section describes the proposed methodology for the development of highway safety prediction models appropriate for strategic planning, developed using Puerto Rico crash data as shown in Figure 1.2. The Western Puerto Rico models were developed by region (Chapter 4) and use crash frequency as the dependent variable. The models developed by region (Chapter 5) uses crashes by region as the dependent variable. Segment length and population represent the offset variable, and road type (primary, secondary, tertiary) represent the independent variables. The models will also take into account the temporal correlation and spatial effects of the different regions and municipalities. This is due to unobserved spatial variables.

The models developed by functional classification (Chapter 6) will use crashes by road type (freeways, arterials) as the dependent variable, and geometric characteristics as the independent variables. The model also takes into account temporal correlation due to repeated measures in time for the same segment. The spatial effects of the different regions and segments are also included in the models.

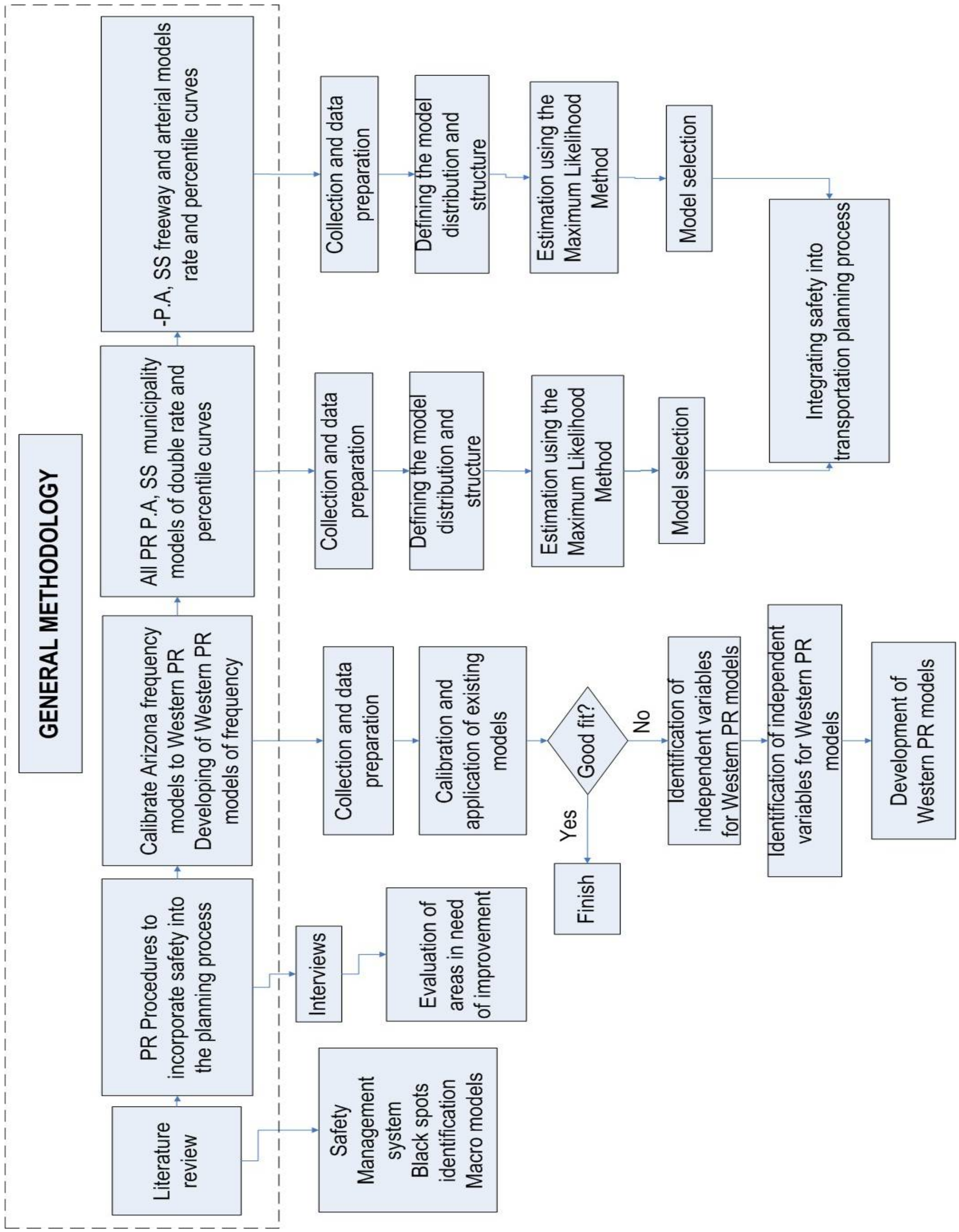


Figure 1.2 General Methodology.

The resulting models will be used to explore the relationship between crash rates and their severities with various variables that may serve as predictors. Possible predictors include traffic intensity, socioeconomic and demographic factors, traffic demand measures, and other geometric characteristics. Examples of geometric characteristics that were taken into consideration include the number of lanes, lane width, slope, curvature, etc. The rest of this section describes the general research approach and the methodology.

For years a wide variety of methods have been implemented in order to deal with the methodological problems associated with accident frequency data. Otherwise, these problems can compromise the validity of the conclusions if not properly handled. Previous investigations have used various approaches for the aggregation and modeling of accidents based on the purpose of the study and the nature of the available data. Some of the approaches are: classical linear regression, Poisson models, negative binomial models, hierarchical Bayes models with spatial effects, and linear models with logarithmic transformations (Anastasopoulos,2009). The present study proposes to apply negative binomial generalized linear mixed models to crash rate data in order to incorporate variables at a segment level, which will allow the data to be temporally and spatially correlated. The models' computations will be done with the Statistical Analysis System, SAS software (SAS Institute, Inc, 2013).

The steps taken in order to estimate the regression parameters in a GLMM are the same regardless of the implemented model. What varies is the probability distribution, the explanatory variables to be used, and the assumptions made. The general procedure used to find the parameter estimates in a GLMM include collection and data preparation, the definition of distribution and the model structure (Negative binomial distribution with random effects of segment and region in this case), estimation of parameters using the maximum likelihood method, model selection, and model integration into the planning process. In Chapters 4, 5 and 6 the corresponding statistical methodology is described in detail.

1.6 ORGANIZATION

This section describes the contents of each of the chapters: Chapter 1 contains an introduction to the highway safety topic, Chapter 2 contains a literature review of highway safety, Chapter 3 describes the elements currently employed in highway safety planning by Puerto Rico and in the United States, Chapter 4 presents the calibration and development of Puerto Rico western region planning models, Chapter 5 presents the development of all the Puerto Rico planning models by region, Chapter 6 describes the development of freeway-expressway and arterial planning models, Chapter 7 presents a framework that may be used to incorporate the developed models into the planning process, Chapter 8 contains the conclusions and Chapter 9 presents the recommendations.

CHAPTER 2: BACKGROUND REVIEW

2.1 DEFINITION AND GENERAL ASPECTS OF A SAFETY MANAGEMENT SYSTEM

This section reviews existing studies related to three main areas of highway safety namely: Safety management systems, high risk crash area identification, and macro level predictions models. The safety management system is presented as a concept first followed by the literature related to the identification of areas at high risk of crashes and, a series of papers that consider a macro level prediction model for a proactive approach to safety.

The first of the three main areas of highway safety, safety management systems (SMS), was a requirement from each state under the "Intermodal Surface Transportation Efficiency Act of 1994" (ISTEA). The Safety Management System is a process by which diverse stakeholders establish strategies to reduce the quantity and severity of crashes (Depue et al., 2003). However, following the approval of the National Highway System Designation Act of 1995, the development and implementation of a highway safety system using the Safety Management System became optional.

The "Safe, Accountable, Flexible and Efficient Transportation Equity Act- A Legacy for Users of 2005" (SAFETEA-LU) created the Highway Safety Improvement Program (HSIP) as a new source of funds in order to reduce traffic accidents and serious fatalities. States submit annual reports that identify the locations that have the highest safety needs and represent more than 5% of crashes in their areas in order to increase public awareness. As part of the MAP-21 Act signed in 2012, 25% of the state funds are invested in HSIP projects.

The Safety Management System is a process that must be carried out by coordinating the various agencies and individuals interested in the common goal of reducing the number of accidents and fatalities at the state and local level. In order for this to be achieved, cooperation between the different stakeholders is paramount. In the NCHRP Report 501 of 2003, the various components and communications between the stakeholders were defined in order to obtain an Integrated Safety Management System (Bahar et al. 2003).

The NCHRP Report 501 of 2003 states that the components of integrated asset management are organizational structure, leadership, mission, vision, resources, integrated process management, and safety tools. The incorporation of the various components mentioned above will lead to improvements in safety requirements and alignment of responsibility for achieving safety with vision, goals, and strategies of support. Using this process, the state agencies can form an integrated organizational structure and can provide the necessary resources for the management and implementation of the integrated system.

2.2 IDENTIFICATION OF AREAS AT HIGH RISK OF CRASHES “HSIP’ AND MICRO MODELS

The localization of high risk sites is an important element of the integrated systems. There are two types of methods used to identify the location of high risk sites or “black spots”. One of them uses basic statistics and rates (frequency, rate, severity, confidence intervals, etc.). The other method is based on crash prediction models or Safety Performance Function (SPF) (Tarko and Kanodia, 2004).

The method of locating black spots in Puerto Rico is based on a weighted average of frequencies, rates, and crash severities similar to the method used in Iowa. Other states use prediction models similar to those presented in the HSM 2010. All of these models are used for the analysis of site or project highway safety at the micro level. Other tools that are used to perform analysis at the project level have been developed, such as in the state of Colorado where prediction models were developed in order to evaluate the impacts of multiple urban freeways design alternatives (Kononov et al., 2007).

Two types of SPF's are used to represent crash frequencies as a function of given variables. The first type of SPF is a Level I, or descriptive analysis model, which determines crash frequencies based only on traffic volumes (AADT). The level II SPF, classified as multivariate models or association models, incorporate a variety of variables other than just traffic volumes. In the level II SPF, variables such as weather conditions, roadway geometries, traffic data, and human factors are used to calculate crash frequencies (SafetyAnalyst, 2002). The additional variables that represent highway geometry, and human factors provide useful information when making future improvements to a given roadway. Taking into account the impacts of changing a given geometry allows engineers to make better safety judgments. The information provided by the additional variables also allows for benefits such as education and enforcement. Law enforcement officers and drivers can adopt safer driving practices by recognizing situations, habits, and other conditions that pose safety concerns.

A significant number of studies have been performed in the development of SPFs by using data analysis and interventions on roadway segments over a specified period of time. SPF's have been the most common approach in studying the factors that affect the likelihood of accident occurrence. Models that look at the crash frequency are abundant in the literature review and encompass a wide variety of modeling approaches including (see Lord and Mannering (2010) for a complete review of this literature): Poisson models (Jovanis and Chang, 1989; Jones et al., 1991; Miaou, 1994; Shankar et al., 1995; Poch and Mannering, 1996; El-Basyouny and Sayed, 2006; Lord 2006; Kim and Washington, 2006; Malyshkina and Mannering, 2010a); Poisson-log normal models (Lord and Miranda-Moreno, 2008); Zero inflated count models (Miaou, 1994; Shankar et al., 1997; Lee and Mannering, 2002; Lord et al. 2005, 2007; Malyshkina and Mannering 2010 b); Conway-Maxwell-Poisson models (Lord et al., 2008; Sellers and Shmueli, 2010); Gamma models (Oh et al., 2006; Daniels et al., 2010); Generalized estimating equation models (Wang and Abdel-Aty, 1996; Lord and Mahlawat, 2009); Generalized additive models (Xie and Zhang, 2008; Li et al., 2011); Random effect models (Shankar et al 1998; Quddus, 2008; Sittikariya and Shankar, 2009; Guo and et al., 2010); Negative binomial models (Ulfarsson and Shankar, 2003; Caliendo et al., 2007); Random parameters count models (Anastasoupoulos and Mannering, 2009; El-Basyouny and Sayed, 2009,2010; and Finite mixture and Markov switching models (Malyshkina et al., 2009; Park and Lord, 2009; Malyshkina and Mannering 2010a; Park et al.; 2010) for various agencies across the United States and Canada.

In 2001, the Federal Highway Administration (FHWA) commissioned the development of software tools for safety management known as the SafetyAnalyst. Even though studies occurred before 2002, the development of this software spurred several explorations concerning the most appropriate method for analyzing highway safety and developing statistical models. Currently, the Safety Analyst program only supports Level l SPF’s, but studies have looked into multivariate models to determine the impact other variables may have on the crash frequency for a given type of roadway.

2.3 MACRO LEVEL PREDICTION MODELS FOR A PROACTIVE APPROACH TO SAFETY

The need for planning models arose from legislation that requires explicit consideration of highway safety at the planning level. Both the "Transportation Equity Act of 1998" (TEA 21), and its reinforcement, Moving Ahead for Progress in the 21th century Act (MAP 21, 2012), require the consideration of safety analyses at all stages of the planning and design processes.

Models that incorporate the analysis of highway safety from the planning stage have been developed for various levels of data aggregation: levels such as traffic analysis zones (TAZ). These models have incorporated several prediction variables in order to determine the number and severity of crashes. Typical explanatory variables include socioeconomic aspects, demographic aspects, and infrastructure characteristics. The following is a summary of relevant work found in the researched literature on this topic.

Laumon Amoros (2003) compared the traffic safety in several French counties by taking different types of roads and socioeconomic characteristics into consideration. The authors found a significant relationship between the county and road type. Noland and Quddus (2004) analyzed data from crashes in England to different types of land use, road characteristic, and demographics. This study suggests that areas with high employment density have a higher accident rate.

Washington et al. (2006) developed nine crash prediction models at the planning level for TAZ 859 in Tucson, Arizona using demographic, socioeconomic and road characteristics as predictors. Noland and Oh (2004) examined the association between crashes in the counties of Illinois, with various road infrastructure networks, some socioeconomic, and demographic variables. They found that the number of lanes was a significant factor for the models. These did not vary significantly when demographic variables were included.

Agüero and Jovanis (2006) investigated the crash risk in Pennsylvania counties using data from fatal and injury crashes with respect to socio-demographic characteristics, weather conditions, transport infrastructure, and the number of trips. Full hierarchical Bayes models were developed with spatial effects, temporal effects, and space-time interaction. The information was then compared with the traditional, negative binomial crash frequency at the county level. The authors concluded that the inclusion of the spatial correlation analysis is more important for crashes on road segments and intersections than for crashes at the county level because the spatial correlation is more pronounced in those cases.

Hadayegui et al. (2006) examined the temporal transferability of accident prediction models for areas using appropriate evaluation measures of predictive performance so as to assess whether the relationship between the dependent and independent variables behave consistently over time. The results show that the models are not transferable in a strict statistical sense. However, measures of transferability indicate that the models provide useful information when transferred to the application context. Hadayeghi et al. (2007) developed 23 crash prediction models at the traffic analysis zones level in the City of Toronto, Canada that are consistent with conventional models commonly used for urban transport planning. The authors used generalized linear regression models with the assumption of a negative binomial error structure in order to explore the relationship between frequency of crashes in a traffic analysis zone and predictive variables such as traffic intensity, demographic, socioeconomic, land use, and traffic demand measures.

Lovegrone and Sayed (2007) investigated the use of predictive models in macro-level driving studies of hazardous locations (traditional use) with 577 urban and rural neighborhoods across Greater Vancouver in British Columbia, Canada. Several areas with a high risk of accidents were identified and prioritized by diagnosis. Two areas were analyzed in detail, revealing several potential improvements to traditional methods. The study used 35 models, 16 for rural areas and 19 for urban areas.

Wier et al. (2009) developed models of vehicle-pedestrian collisions with injuries for 176 census tracts in San Francisco, California, which were spatially disaggregated by counties. Simple regression models were developed to predict changes in vehicle collisions – based on pedestrian traffic volume change. It became clear from the study that traffic volume was the leading cause of pedestrian collision with injuries. Huang et al. (2010) developed a Bayesian spatial model for Florida and concluded that the safety status is worse in areas with low income, low education and high unemployment compared to relatively prosperous areas. Also, counties with high traffic intensity, population density, and a high degree of urbanization are associated with higher crash risk.

Naderan and Shahi (2010) developed crash prediction models based on the generation phase of the travel-demand modeling technique in four steps. Generalized linear models were generated with the assumption of a negative binomial distribution. The developed models predicted crashes in areas of urban traffic. This analysis was based on the predicted number of trips generated by purpose. They can be used to evaluate the effect of the generation of future travel on the frequency of collisions for the use of comprehensive transportation planning studies. The models are an effective way of incorporating safety into the long term transportation plan by helping safety planners develop scenarios of demand management while simultaneously evaluating the perceived effects on the safety of urban areas.

Abdel Aty et al. (2011) developed negative binomial models and investigated the association between crash frequency, various travel productions, and travel attractions combined with the road characteristics TAZ 1349 in the State of Florida.

2.4 PLANNING FOR A PROACTIVE APPROACH TO BE IMPLEMENTED AT THE STATE LEVEL FOR THE SAFETY MANAGEMENT SYSTEM OF SOME STATES

According to the Figure 2.1, Washington et al 2006 (NCHRP 546) suggests that incorporating safety considerations into transportation planning, or daily activities of the agency, consists of the following steps:

2.4.1 Step 1: Incorporating safety into the vision statement

Transportation system planning begins with the creation of a vision. This vision, which is shared by many states, includes “safety” as a desired characteristic of the future travel experience and represents an important “point of departure” for the many planning activities that follow.

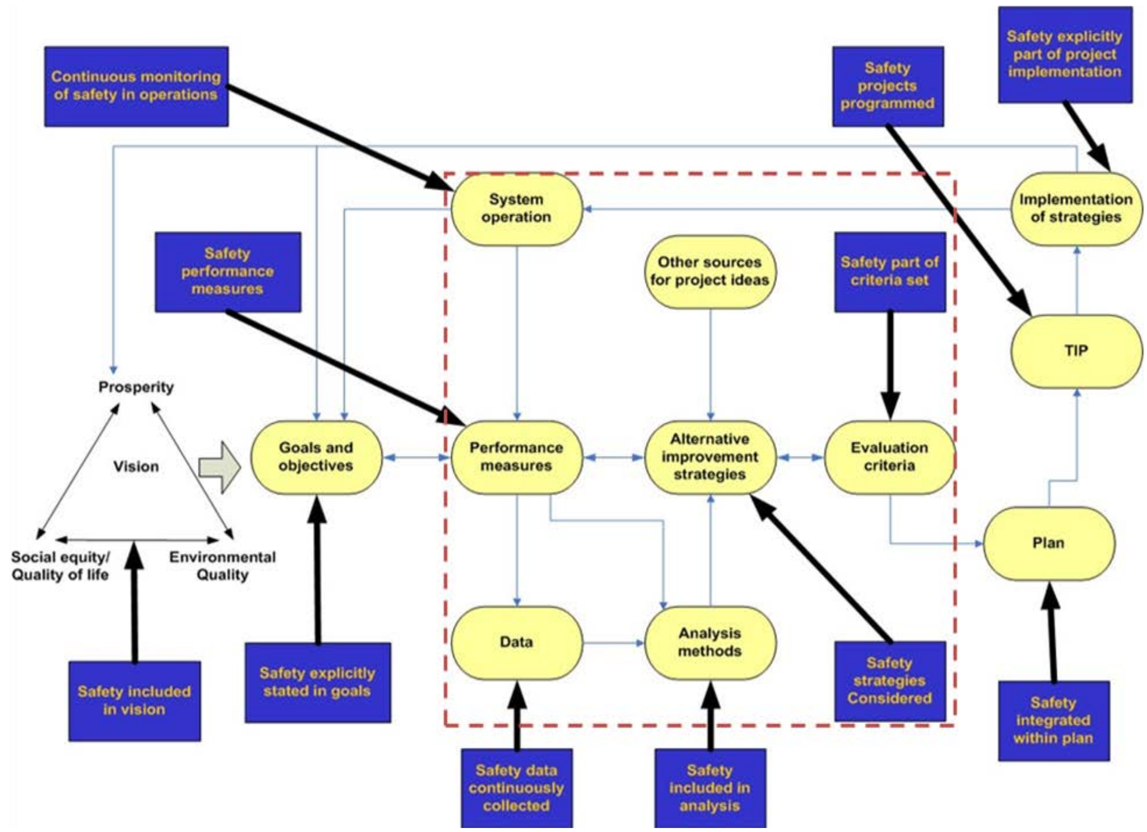


Figure 2.1 Incorporating Safety into the Planning Process. (Source: NCHRP 546).

The following vision statement from the California statewide transportation-planning process illustrates a state's typical vision goal. California's vision is to have: "a safe, sustainable transportation system that is environmentally sound, socially, equitably and economically viable, and developed through collaboration; it provides for the mobility and accessibility of people, goods, services, and information through an integrated, multimodal network." [Caltrans, California Transportation Plan, 2025, Sacramento, CA, March 2004].

2.4.2 Step 2: Incorporating safety into the set of goals and objectives

The goals and objectives for a region are derived from the vision statement. The vision is typically accomplished through well defined goals and objectives. These serve as an assessment of the relative contributions that each possible alternative or strategy and directs the planning process towards its desired outcome.

In metropolitan areas, safety goals and objectives can also be more specific and include targets such as: reducing fatal accidents in the region by 10% over the next three years and reducing fatal and serious injury accidents by drivers aged 16 to 23 by 30%. Specific safety targets such as these may serve as a guidance and motivation to the engineers and planners in order to achieve regional safety goals.

In order to achieve these safety goals, it is necessary to increase funding dedicated to reducing high accident levels in the region. These monies will then be used to undertake safety studies throughout the region, mitigate major accident hot spots at a cost of X million dollars, but with an annual benefit of Y million dollars, support traffic safety education and traffic enforcement efforts, and build an information system that will identify incidents on transportation facilities to continually support these strategic safety investments.

2.4.3 Step 3: Incorporating safety into system performance measures

The evaluation of system performance has traditionally relied on measures of congestion, travel delay, traffic volumes, and the infraestructure's condition, such as pavement and bridge condition. The transportation system's safety performance can be monitored as well.

Performance measures are used to monitor the characteristics of the system operation, and to determine the extent to which desired goals and objectives are being achieved. The use of performance measures is a relatively new trend in transportation-planning, so there is little consistency from one jurisdiction to another on how safety monitoring practices are performed.

The following is an example of the goals laid out in the Texan Comprehensive Safety Plan. This is an example of how performance measures can relate to specific goals. The first goal was to decrease traffic deaths and injuries, reduce vehicular traffic accident rate's per100 million VMT, and traffic accident injury rate's per 100 million VMT. The second goal was to stabilize the increase of the frequency and percentage of all speed related accidents.

Another example of the role of safety performance measures in transportation planning is found in the Minnesota Statewide Transportation Plan. One policy is to, "increase the safety and security of transportation systems and users." There are five specific measures that define what is meant by increased safety: reducing the number of accidents per vehicle-mile traveled, reducing the number of general aviation accidents, reducing the number of accidents between cars and trains at railroad crossings, reducing the total number of roadway fatalities, and reducing the number of general aviation fatalities. The Minnesota DOT analyzed the impacts of different safety policies in achieving safety goals using a trend based projection.

2.4.4 Step 4: Incorporating safety into technical analysis

In this stage, transportation problems are broken down into components that are used to pinpoint the areas in need of improvement. In the technical analysis process there are two aspects that merit special attention. When considering a closer integration of safety into system planning, it is important to focus on safety-related data and their use, and analysis models/tools.

2.4.5 Step 5: Evaluating alternative projects and strategies

The process of reviewing the worth of each of the alternatives presenting this information to decision makers in a comprehensive and useful form is called evaluation. Most safety-related evaluation efforts use one of three methods, 1) simply listing the evaluation criteria to show how the alternatives compare, 2) assigning weights or scores to the evaluation factors, or 3) conducting cost-benefit analysis.

In some situations, funding programs are divided into specific categories. In these cases the effectiveness of each category can be measured with one evaluation criterion, (such as safety, air quality, economic development, etc.), and the selection of the “best” alternative becomes much easier. In comprehensive transportation planning, however, reducing project selection decisions to a single criterion seldom happens.

This evaluation step leads to the incorporation of policies, operations strategies, infrastructure projects, studies, regulations, and education as well as awareness, financing strategies, partnerships and collaborative undertakings into the transportation plan.

2.4.6 Step 6 Developing the plan and program and Step 7 Monitoring system performance

Step 6: Developing the plan and program

It is important that both medium and long-term transportation plans include developed, safety-related projects based on the evaluation performed in the previous step. Examples of such projects are: increasing driver safety awareness, increasing safety belt use and child seats, increasing restraint usage, preventing drowsy and distracted driving, curbing high risk driving behaviors, making sure drivers are fully licensed, competent and insured, reducing impaired driving, creating more effective processes and Safety Management Systems, and developing and encouraging multidisciplinary safety teams.

Step 7: Monitoring system performance

System performance should be monitored to evaluate the effectiveness of the various strategies, programs, and policies that were implemented.

2.5 EXAMPLE OF THE PHOENIX SAFETY MANAGEMENT SYSTEM

Figure 2.2 shows the Safety Management System for Phoenix, Arizona. This system uses a safety goal and several safety performance measures to drive the collection of safety-related data, and the identification of projects and strategies. These projects and strategies reflect the planning, engineering, education, and enforcement aspect of the safety challenge in that metropolitan area. Many states have similar safety management systems, although they are not often tied to the transportation plan.

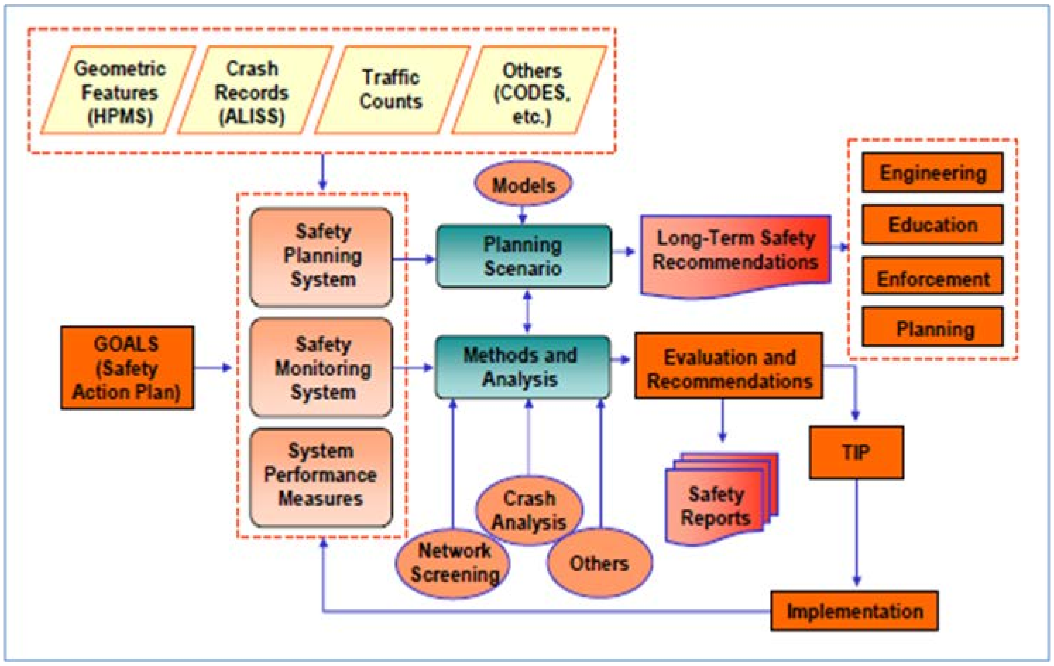


Figure 2.2 Phoenix Safety Management System.
(Source: NCHRP 546)

2.6 DISCUSSION

Several models have been developed and have even been incorporated into the Highway Safety Manual in order to analyze highway safety at the site or project level. However, as indicated by Hadayeghi et al. (2005), highway safety models for the planning process are still under investigation. Additionally, the developed models are mainly Bayesian, binomial or generalized linear models, with few GLMM models. The idea of this research project is to explore Generalized Linear Mixed Models (GLMM), which allow the proper modeling of data sets in cases where the observations are not completely independent. GLMM have been widely used in other fields including economics, agriculture, and forestry. The modeling framework of GLMM is able to handle correlated data by incorporating random variables. Several modeling strategies under the same framework provide great flexibility in order to model the variability of each of the variables involved in these types of models. Several benefits can be obtained by using mixed models. In some cases we obtain more accurate estimates, but in others, the structure of the model is more robust and the inference space is wider (Agresti 2003).

CHAPTER 3: TRANSPORTATION PLANNING AND HIGHWAY SAFETY IN PUERTO RICO AND GENERAL STATISTICS

3.1 BACKGROUND

There are 16,693 miles of roads in Puerto Rico. In 2010 there were 3,102,941 licensed drivers and 3,020,455 registered vehicles. There are approximately 3.8 million citizens distributed over the 78 municipalities of Puerto Rico. This means that there are about 1,000 people per square mile, a ratio higher than any state within the United States. It also ranked among the world's highest ratio. Of the total population, approximately 3.03 million are less than 55 years of age. This shows that Puerto Rico's population is relatively young. Therefore, the tendency of most drivers is to have an active social life. Approximately 261,000 traffic crashes occur in Puerto Rico every year resulting in over 35,000 injuries and approximately 340 fatalities (SHSP, 2014).

Traffic crash data have been reviewed by the Puerto Rico Traffic Safety Commission (PRTSC) staff throughout the years in order to identify problems that are unique to Puerto Rico. The primary and most reliable data source is FARS (Fatality Analysis Reporting System). Also, crash data are obtained from the Accident Information System of the Analysis of Accidents Office within the Department of Transportation and Public Works (DTPW). Data on licensed drivers, registered vehicles, and VMT are also obtained from the DTPW.

In the transportation planning process of the Commonwealth of Puerto Rico, safety appears in the vision statement, and in the set of goals and objectives. However, the road safety audit division is developing a Safety Management System that includes a HSIP related to planning, monitoring and performance measures of safety.

HSIP comes from the SAFETEA-LU source of funds, and is continued by the MAP-21. With these funds, the Road Safety Audit Division in the office of traffic regulations is in the process of developing a HSIP. The HSIP will incorporate planning, evaluation and implementation of safety countermeasures. At the moment, they have a methodology for planning, used to identify black spot priorities similar to the methodology used in Iowa State that consisted of establishing a weighted average of the sites with higher frequencies, rates and crash severities. However, there is no methodology for evaluating design alternatives; a method to later identify the fulfillment of goals and objectives related to infrastructure improvements. The Road Safety Audit Division will apply the Highway Safety Manual's Methodologies and strategies and evaluate its implementation and subsequent monitoring.

Additional federal requirements that are related to safety are incorporated into 23 U.S.C. Section 402, as shown in figure 3.1. This requires the creation of the State Highway Safety Program (SHSP). This program, administered by the National Highway Traffic Safety Administration (NHTSA), requires that the state Governor be responsible for the administration of the State Highway Safety Program through a Governor's Highway Safety office.

The Puerto Rico Highways and Transportation Authority (PRHTA) has the responsibility of overseeing Puerto Rico's network of roads and highways, and to enforce safety. As part of the Moving Ahead for Progress in the 21st Century Act (MAP-21) requirements, the PRHTA shall develop its multi-annual Strategic Highway Safety

Plan (SHSP). This data driven plan shall include the input from other stakeholders that share the same objectives.

The main objective of the SHSP is that all safety initiatives toward reducing fatalities and injuries in highway accidents are fully coordinated and developed based on current Puerto Rico trends and statistical evaluation results. As such, the plan will include emphasis areas as well as strategies to reduce fatalities and injuries, based on Puerto Rico crash data.

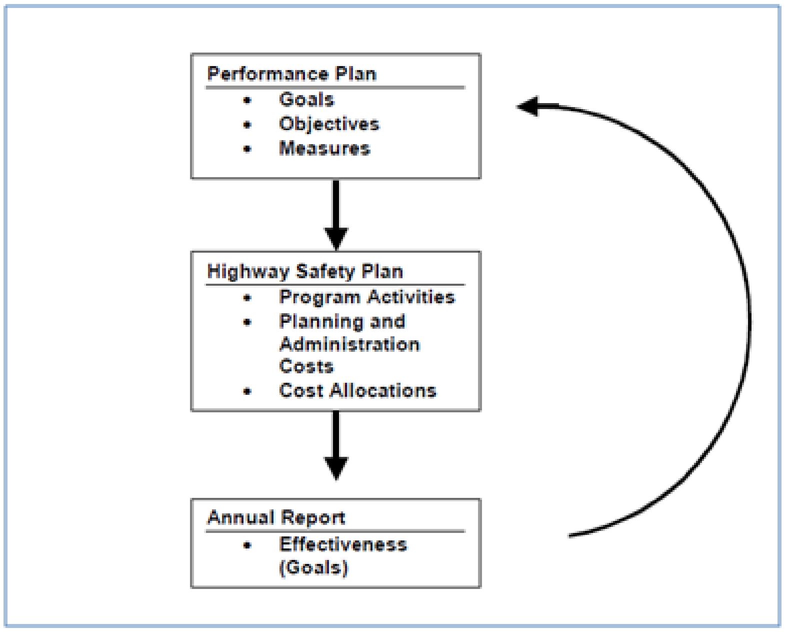


Figure 3.1 State highway safety program.
(Source NCHRP 546)

The majority of initiatives undertaken by the Governor's Highway Safety office are directed towards encouraging the use of the passenger restraint system, minimizing dangers associated with individuals driving under influence of drugs and alcohol, and encouraging safe behavior in school and construction zones. While these activities are associated with the behavioral aspect of transportation usage, it is clear that the substantive safety issues this programs seeks to address are of great interest to transportation planning efforts aimed toward increasing transportation system safety. The relationship between highway safety offices and their safety programs, as well as the planning efforts of transportation agencies, is an area that needs to be strengthened to include strategies to better integrate these processes.

The Puerto Rico Traffic Safety Commission (PRTSC) is in charge of developing the Highway Safety Plan. The Executive Board includes representatives of the following agencies: DOTPW, Department of Health, Puerto Rico Police Department (PRPD), Department of Education, Department of Justice, Courts Administration, and representatives from Public Interest and Youth Representative.The PRTSC lead the development of a comprehensive traffic safety program for Puerto Rico including strategies for Education, Enforcement, Engineering, and Emergency Medical Service to be implemented throughout the 2014-2018 time period.

Several program areas were activated including the following: Traffic records and information systems, emergency medical response, occupant protection, alcohol impaired driving, aggressive driving, young drivers, vulnerable road users, roadway departure, and intersections. Each one of these program areas have performance goals, performance measures, and strategies.

In the traffic engineering program the elimination of hazards on the roadway that may cause or aggravate traffic crashes is one of the engineering strategies that can help improve traffic safety. The engineering component of hazard elimination requires a team of experienced professionals from the Puerto Rico Highway and Transportation

Authority's (PRHTA) Road Safety Audit Division to attend to both the citizen's requests as well as perform a proactive analysis of hazardous road segments and intersections.

The team then provides in-house design for road safety improvements to be bid by the PRHTA. PRTSC provides funds for personnel, vehicle, and equipment as well as for the construction of road improvements through reimbursement. These funds are supplemented by Federal Safety Funds. However, there are some larger projects which cannot be designed in-house and present a challenge to the efforts of improving hazardous conditions on longer corridors or roads with high speeds and/or traffic volumes. Some performance goals for the Traffic Engineering program are: increasing the percent of hazard elimination construction funds liquidated, implementing roadside improvements (Impact Attenuators, NCHRP 350 upgrades), implementing island wide road departure countermeasures, and increasing the amount of projects completed by the Impact Team. The Traffic Engineering program created performance measures in order to achieve these goals

3.2 THE LOCAL INCORPORATION OF SAFETY INTO THE PLANNING PROCESS AND ITS ASSEMENT

In order to assess the planning process, current and former officials of the DTPW were interviewed. The questions asked in this interview were taken from the NHRP 546 report. The questions were:

- Does the vision statement for the planning process (transportation plan) include safety?

• Is there at least one planning goal and at least two objectives related to safety (transportation plan)?

• Are safety-related performance measures a part of the plan being used by the agency?

• Can safety performance measures link to the evaluation criteria that will be used later in the planning process to assess the relative benefits of one project or strategy over others?

• Is safety-related data used in problem identification and for identifying potential solutions?

• Are safety analysis tools used regularly to analyze the potential impacts of prospective strategies and actions?

• Are evaluation criteria used for assessing the relative merits of different strategies and projects including safety-related issues in the area of programming and budget of the agency?

• Do the products of the planning process include at least some actions that focus on transportation safety?

• Is safety one of the main priorities to the extent that a prioritization scheme is used to develop a program of action for an agency?

• Is there a systematic monitoring process that collects data on the safety-related characteristics of the transportation system performance, and feeds this information back into the planning and decision making process?

• Are all of the key safety stakeholders (Governor's Office of Highway Safety, MPO, State Department of Transportation, local Departments of Transportation, Departments of Public Health, Departments of Public Safety, local police agencies, the Department of Education, Federal Highway Administration, Federal Transit Administration, American Automobile Association, etc) involved in the planning process?

The answers for many of the interview questions were negative, indicating that there are still many opportunities for improvement in regards to the integration of safety into the medium and long-term transportation plan. The most important issues found in the interview are as follows:

• In the transportation planning process of the Commonwealth of Puerto Rico, safety appears in the vision statement and in the set of goals and objectives of LRTP.

• There is no Safety Management System in place for planning and monitoring safety.

• The Road Safety Audit Division in the Office of Traffic Regulations is in the process of developing a HSIP with the planning, evaluation and implementation of safety countermeasures.

• Currently the Road Safety Audit Division have a methodology for identifying black spot priorities similar to the methodology applied in state of Iowa.

• Additional federal requirements related to safety are incorporated into 23 U.S.C. Section 402. This code requires the creation of a State Highway Safety Program

SHSP. A SHSP was developed in Puerto Rico for the 2014-2018 time period. However the planning process proposed by Washington et al. 2006 and shown in Chapter 2 Figure 2.1 is incomplete in Puerto Rico procedures. The central procedures in the flowchart, as indicated in the dashed area of Figure 2.1, have not been developed or can be improve in Puerto Rico. The performance measures skips the highlighted procedures and continues to plan and TIP.

3.3 SUMMARY AND ANALYSIS OF CRASH STATISTICS IN THE WESTERN REGION, PER MUNICIPALITY

Figures 3.2 through 3.6 present highways length, population density, population with ages between 16 to 64 in thousand, total population in thousands, and proportion of Interstates of the Western municipalities, respectively.

Figures 3.7 to 3.9 present crash rate plots, due to crash frequencies alone are not comparable. Crash rate allow establishing a hierarchy of risk of crashes in the Western municipalities.

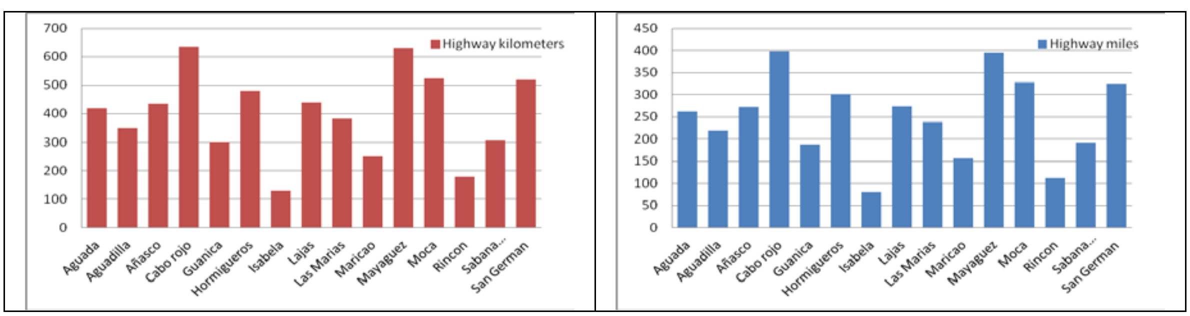


Figure 3.2 Highway miles and kilometers in the western region year 2006 (HPMS database)

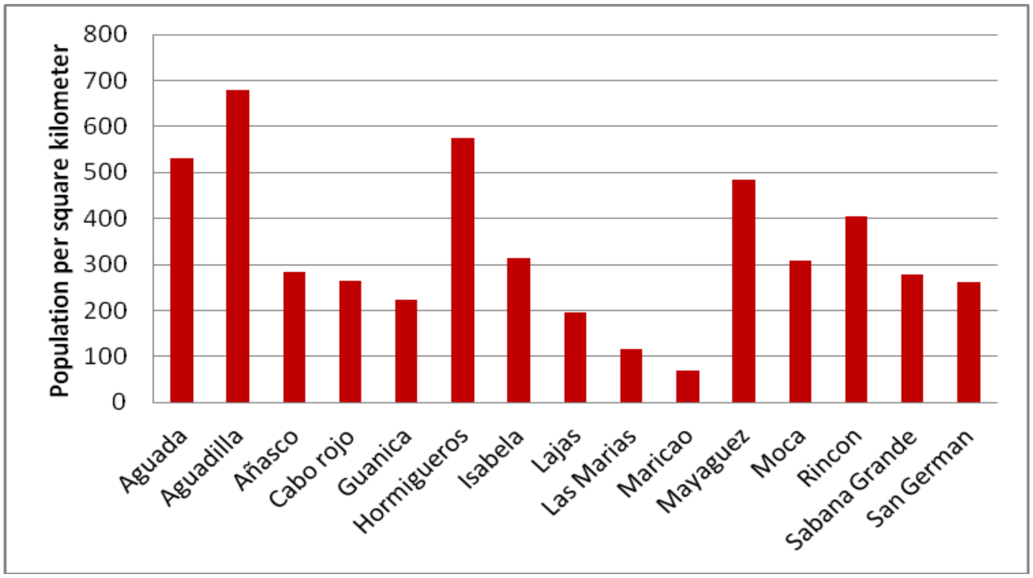


Figure 3.3 Population density in the western region year 2002 (Census estimates)

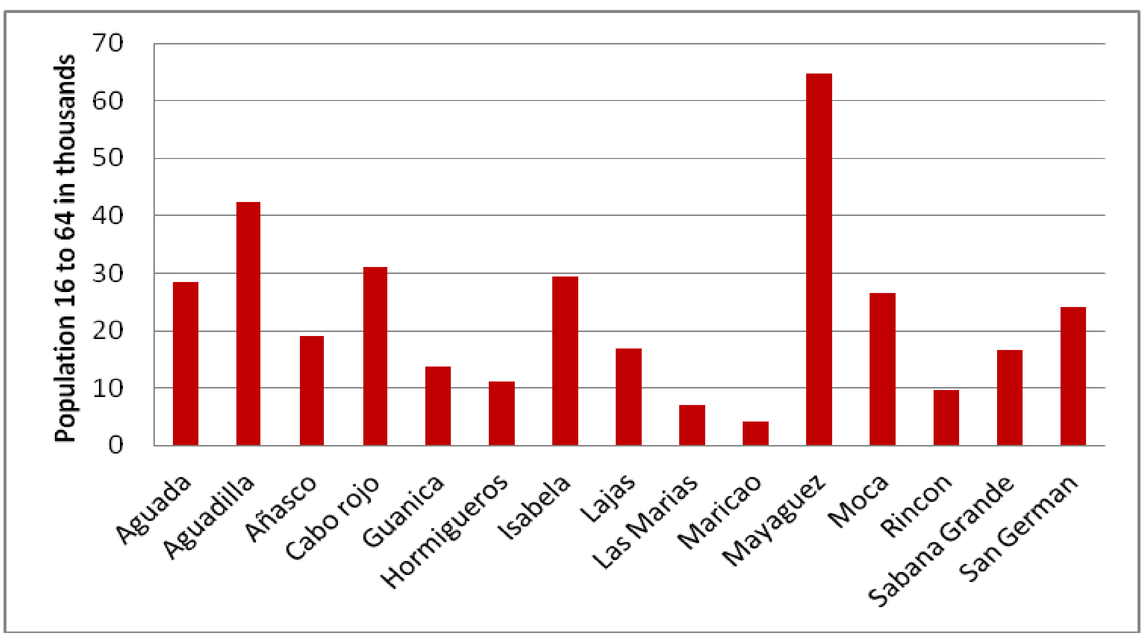


Figure 3.4 Population 16-64 in the western region year 2002 (Census estimates)

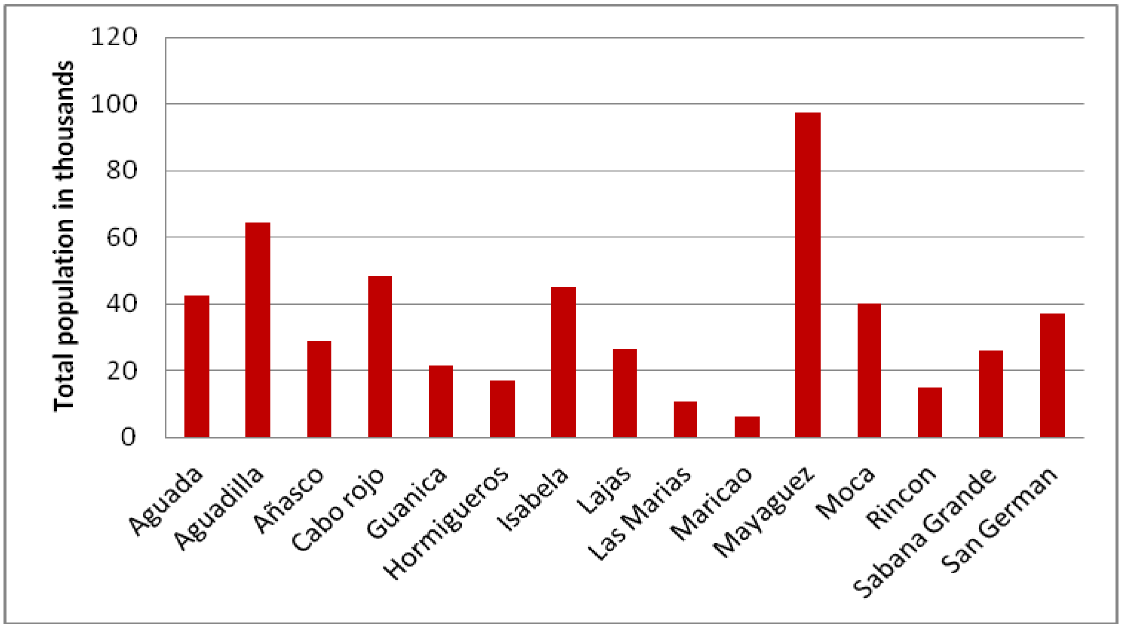


Figure 3.5 Total population in the western region year 2002 (Census estimates)

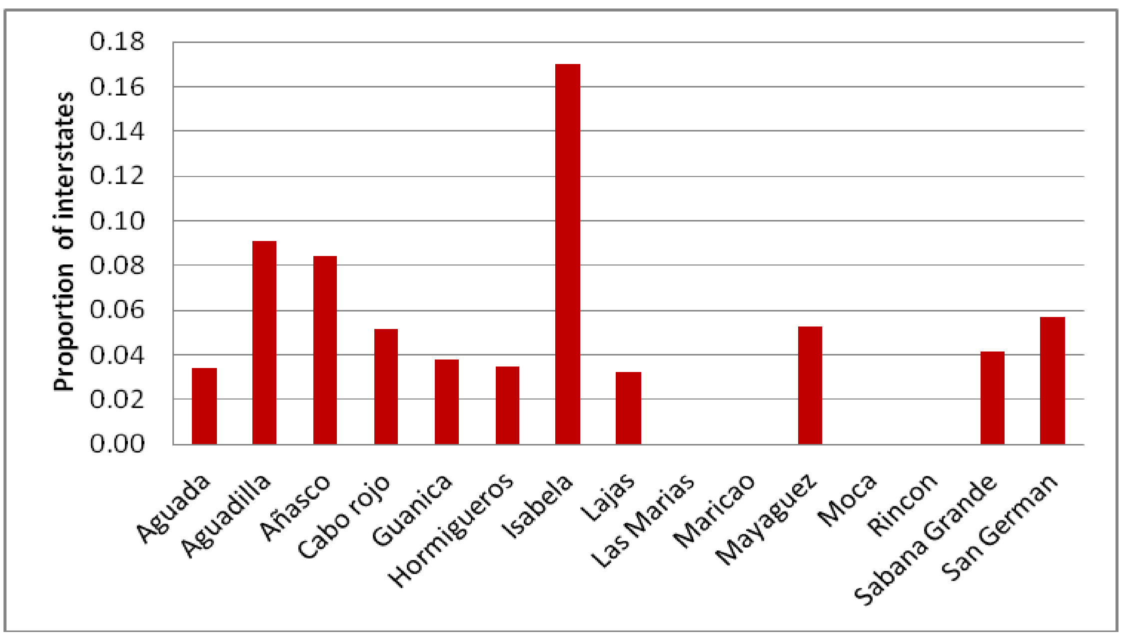


Figure 3.6 Proportion of Interstates in the western region year 2006 (HPMS database)

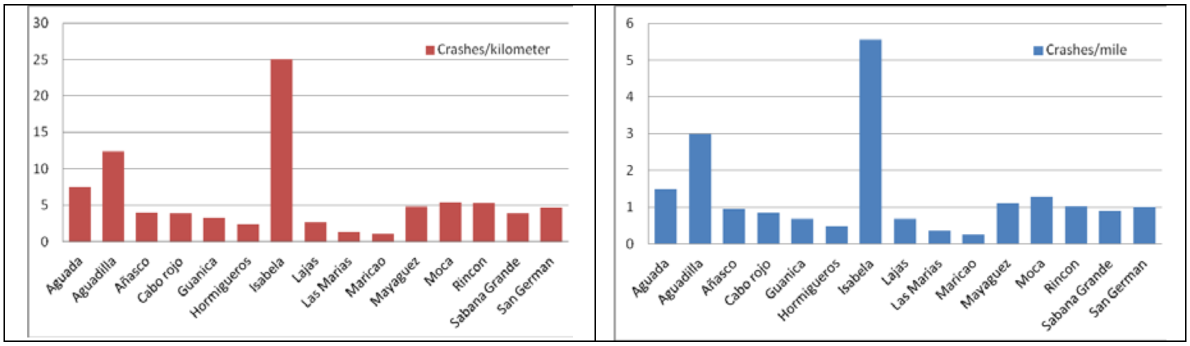


Figure 3.7 Crash rates in the western region (DTPW 2002-2006 database) standardized by length

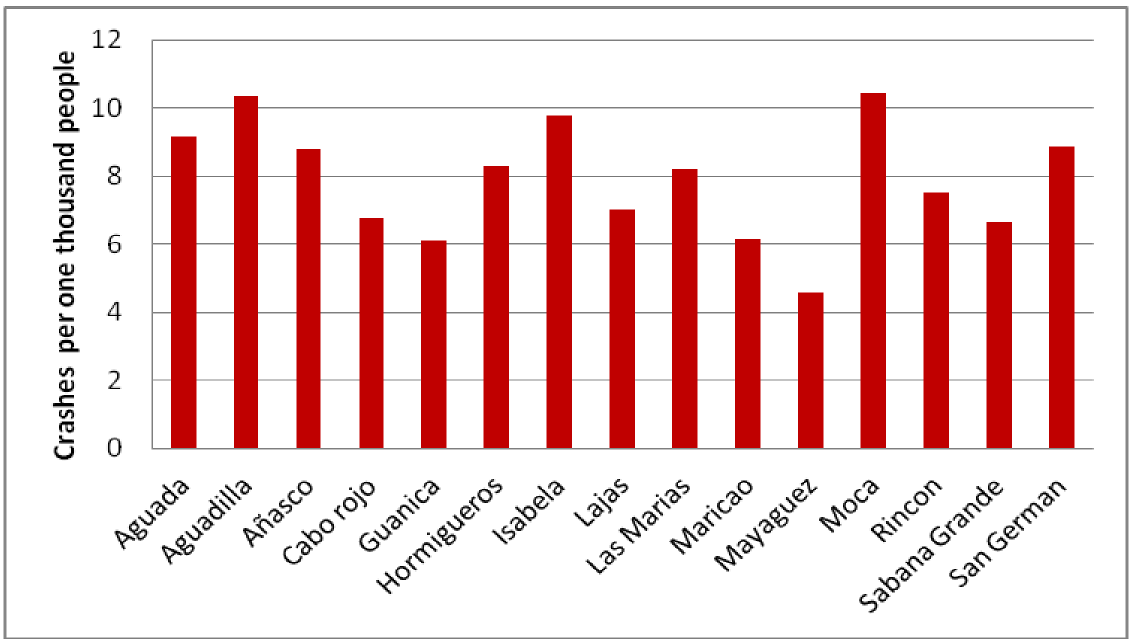


Figure 3.8 Crash rates in the western region (DTPW 2002-2006 database) standardized by population

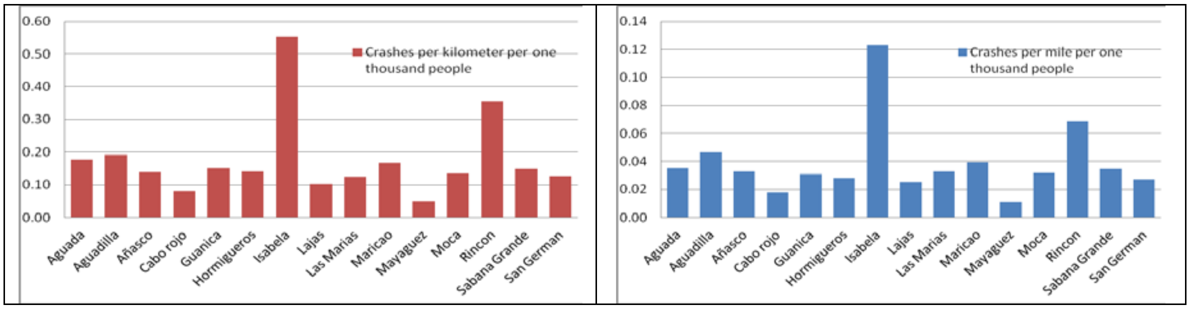


Figure 3.9 Crash double rates in the western region (DTPW 2002-2006 database)

As a summary of the plots, Figure 3.2 shows that Cabo Rojo and Mayagüez are the municipalities with the highest number of kilometers of builded highways. Figure 3.3 shows that Aguadilla and Hormigueros are the municipalities with the greatest population density. Figure 3.4 shows that Mayaguez is the municipality with a higher population of persons within the ages of 16 to 64. Figure 3.5 shows that Mayaguez is the municipality with the most habitants. Figure 3.6 shows that Isabela is the municipality with the highest proportion of interstate highways. Figure 3.7 shows that Isabela is the municipality with the greatest number of crashes per mile. Figure 3.8 shows that Aguadilla and Moca are the municipalities with most crashes per one thousand people. Figure 3.9 shows that Isabela is the municipality with the highest number of crashes per mile per one thousand people followed by Rincon.

CHAPTER 4: CALIBRATION OF MACRO MODELS FOR HIGHWAY SAFETY PLANNING

4.1 DETAILED METHODOLOGY

The methodology of this study consists of six basic steps: a literature review, the identification of existing models and procedures used in Puerto Rico, the collection and preparation of data, the calibration and application of existing models, the identification of independent variables to be used in the mode, and finally, the development of crash prediction models for Puerto Rico. The scheme of the general methodology is shown in Figure 4.1.

This chapter's main goal is to recommend a model for safety planning that can be applied in Puerto Rico, and that can guide the planning department's decision making. The resulting model will be used in the exploration of the relationship between crash frequencies in the municipalities and some variables that may be predictive. These variables include traffic intensity, socioeconomic and demographic factors, types of land use, and traffic demand measures, and other geometric characteristics. The following figure describes each one of the methodology steps in more detail.

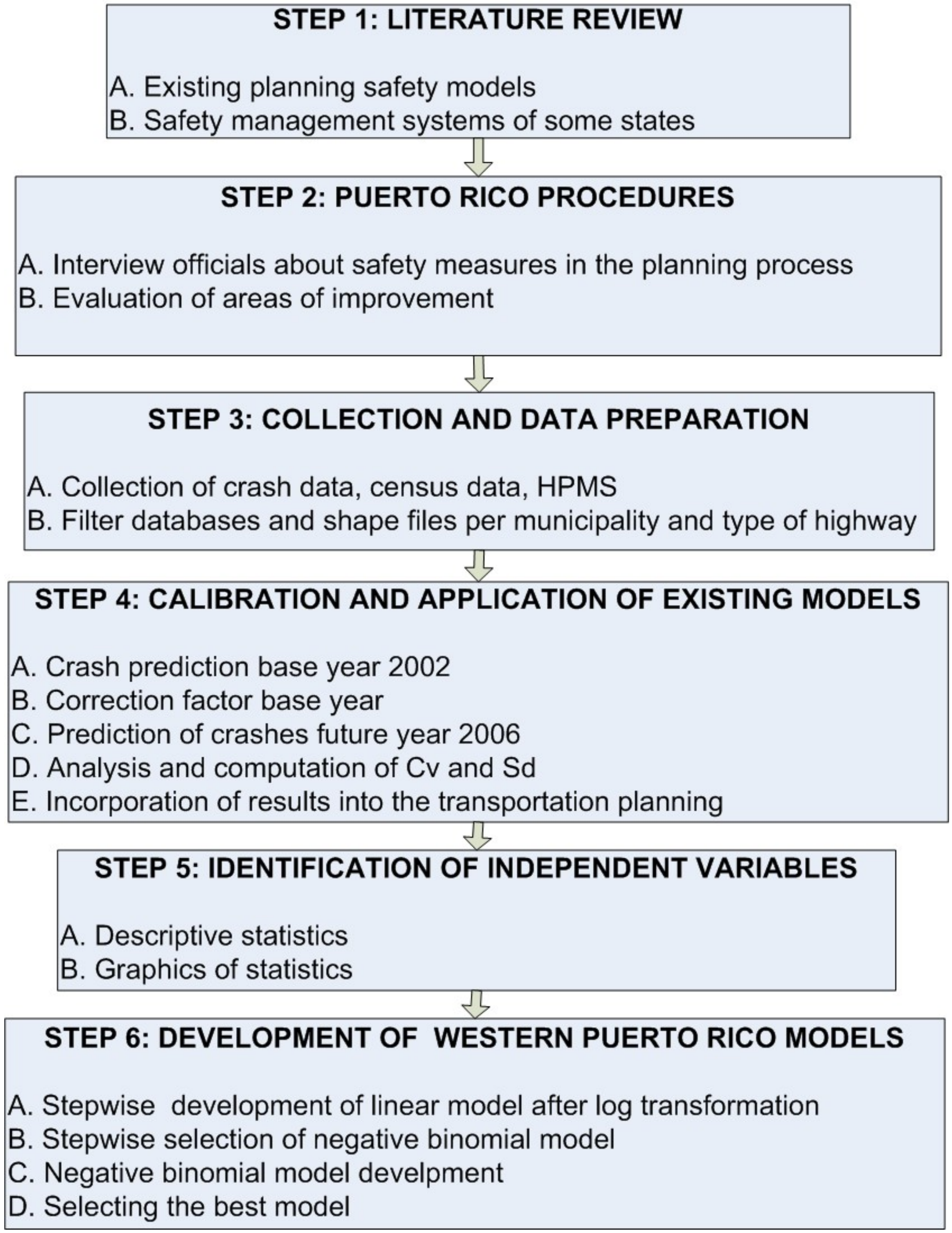


Figure 4.1 Methodology for the calibration of Arizona Models and the development of Western Puerto Rico Models.

4.2 LITERATURE REVIEW AND IDENTIFICATION OF EXISTING MODELS

In this section, the safety management systems of a few states were analyzed where safety was incorporated into the mid and long- term transportation planning process. Additionally, existing models found in the literature were reviewed. This part of the report is included in Chapter 2- Section 2.3.

4.3 PROCEDURES USED IN PUERTO RICO

It was necessary to carry out interviews with transportation agency officials who were responsible for highway safety and with transportation consultants. The Transportation Plan 2040 of Puerto Rico should find ways to include procedures to incorporate safety into mid and long-term transportation plans in Puerto Rico. This was discussed further in Chapter 3.

4.4 COLLECTING DATA AND PREPARATION

The collection of data from the western municipalities (planning level) for the application of the Tucson, Arizona prediction model for safety, requires the cooperation between different traffic agencies in the region. Data are collected at different levels and sharing them between agencies can be difficult. It is necessary to have the support of the ACT (Highways and Transportation Authority) from the initial stage of the data collection process. Data should be collected from the ACT, state and, in some cases, metropolitan planning agencies. The crashes database was obtained through the Division of Safety Audits in the Traffic Regulations Office. The socioeconomic and demographic information was obtained from the American Fact Finder website http://factfinder2.census.gov, and the traffic and geometric characteristics were obtained by using the Office of Highway Performance Monitoring Systems database.

Preparing a dataset by municipality consisted of developing a joint database with crash frequency data, census data, and geometric data, among others. The variables needed to run and calibrate the Arizona models were: number of crashes per municipality, population density, population within the ages 16-64, miles of the roads built, number of intersections per mile, total miles of rural and urban interstate highways as a proportion of total miles, total miles of freeways and expressways as proportion of total miles, population within the ages 0-15, the total number of minorities, total miles of roads in the municipality, total miles of main roads and rural and portion of total miles in the municipality, among others. Some of this data is collected using GIS tools such as dynamic segmentation and spatial joining.

The detailed steps followed to obtain the data were:

i. Filtering crash database

The DTOP crash database in Microsoft Access was filtered by municipality to obtain the total number of crashes for each municipality.

ii. Filtering of the file PRM-EST00INT-AGESEX-5YR

The database of the American Fact Finder web page PRM-EST00INT-AGESEX-5YR was used to obtain the population per municipality and year.

iii. Filtering of miles per municipality

The total road miles per municipality and per road type were obtained using the HPMS GIS shape files and database, and the Arc GIS 9.3 tools were 'selected by': municipality, road type, etc.

The summary of the joint data is presented in tables 4.1 to 4.4.

Table 4.1 Data for the total crashes model for the western region year 2002 (DTPW and HPMS database).

MUNICIPALITY	TOTAL CRASHES	TOT_MILE	POP_PAC	POP16_64
Aguada	2286	262.49	2.145	28490
Aguadilla	3243	219.06	2.751	42410
Añasco	1206	272.07	1.144	18980
Cabo Rojo	1703	397.36	1.070	30969
Guanica	665	187.39	0.907	13755
Hormigueros	820	300.36	2.330	11026
Isabela	2082	80.11	1.273	29414
Lajas	1032	274.53	0.789	16996
Las Marias	370	238.42	0.469	7139
Maricao	174	156.12	0.276	4215
Mayagüez	6282	394.58	1.957	64653
Moca	1979	327.68	1.248	26559
Rincon	621	110.94	1.638	9779
Sabana Grande	862	191.34	1.131	16747
San German	1670	324.97	1.062	23998

In Tables 4.1 and 4.2 the variable POP_PAC represents the Population density (population estimates from U.S. Census SF1) in persons per acre. The variable POP16_64 is the total population of ages between 16 to 64 (from U.S. Census SF1), and TOT_MILE is the total mileage of all functional classes of roads.

Table 4.2 Data for the total crashes model for the western region year 2006 (DTPW and HPMS database)

MUNICIPALITY	TOTAL CRASHES	TOT_MILE	POP_PAC	POP16_64
Aguada	2404	262.49	2.15	28975
Aguadilla	3394	219.06	2.70	41458
Añasco	1329	272.07	1.16	19390
Cabo Rojo	1999	397.36	1.11	31980
Guanica	800	187.39	0.87	13205
Hormigueros	921	300.36	2.37	30067
Isabela	2739	80.11	1.29	16948
Lajas	996	274.53	0.79	16948
Las Marias	333	238.42	0.45	6929
Maricao	212	156.12	0.27	4210
Mayagüez	6282	394.58	1.88	62131
Moca	2096	327.68	1.26	26982
Rincon	727	110.94	1.66	9838
Sabana Grande	940	191.34	1.12	16602
San German	1923	324.97	1.05	23547

Table 4.3 Data for the fatal crashes model for the western region year 2002 (DOTPW and HPMS database)

MUNICIPALITY	FATAL CRASHES	INT_PMI	PNF_0111	PNF_0512	POPMIN
Aguada	7	0.12	0.03	0.00	0.13
Aguadilla	12	0.09	0.09	0.07	0.17
Añasco	5	0.08	0.08	0.02	0.19
Cabo Rojo	5	0.05	0.05	0.00	0.17
Guanica	0	0.1	0.04	0.00	0.2
Hormigueros	4	0.04	0.03	0.00	0.19
Isabela	7	0.35	0.17	0	0.17
Lajas	2	0.1	0.03	0	0.2
Las Marias	0	0.08	0.00	0.00	0.14
Maricao	0	0.08	0.00	0.00	0.11
Mayagüez	9	0.1	0.05	0.00	0.22
Moca	7	0.09	0.00	0.00	0.11
Rincon	1	0.06	0.00	0.00	0.14
Sabana Grande	2	0.07	0.04	0.00	0.15
San German	5	0.11	0.06	0.00	0.17

In Table 4.3, the variable INT_PMI represents the number of intersections per mile (using total mileage in the municipality). The variable PNF_0111 is total mileage of urban and rural interstates as a proportion of the total mileage (federal functional classifications 01 and 11), while PNF_0512 is the total mileage of other freeways and expressways (i.e., not interstate and also not principal arterials) as a portion of the total mileage. POP00_15 represents the total population of ages 0 to 15 (from U.S. Census SF1), and PPOPMIN is the total number of minorities (from U.S. Census SF1) as a proportion of the total population.

Table 4.4 Data for the fatal crashes model for the western region year 2006 (DOTPW and HPMS database)

MUNICIPALITY	FATAL CRASHES	INT_PMI	PNF_0111	PNF_0512	POPMIN
Aguada	6	0.12	0.03	0.00	0.134
Aguadilla	14	0.09	0.09	0.07	0.17
Añasco	3	0.08	0.08	0.02	0.194
Cabo Rojo	4	0.05	0.05	0.00	0.168
Guanica	4	0.10	0.04	0.00	0.201
Hormigueros	1	0.04	0.03	0.00	0.192
Isabela	6	0.35	0.17	0.00	0.166
Lajas	8	0.10	0.03	0.00	0.2
Las Marias	3	0.08	0.00	0.00	0.143
Maricao	0	0.08	0.00	0.00	0.113
Mayagüez	9	0.10	0.05	0.00	0.22
Moca	3	0.09	0.00	0.00	0.105
Rincon	2	0.06	0.00	0.00	0.143
Sabana Grande	6	0.07	0.04	0.00	0.151
San German	1	0.11	0.06	0.00	0.172

4.5 CALIBRATION AND APPLICATION OF MODELS FOUND IN THE LITERATURE

This section investigates the potential of using the models and safety planning characteristics found in the literature review in western Puerto Rico. This model is calibrated with correction factors in order to apply it to the Western area of Puerto Rico and to analyze if the characteristics are suitable.

The models in this section were developed by Washington et al. (2006) for the City of Tucson, Arizona, using the total frequency of crashes and fatal crashes. In these models the dependent variable is transformed and is analyzed using log-linear regression with the ordinary least squares method to calculate estimates of the parameters. In this model, the dependent variable is assumed to be continuous, with a normal distribution and homogeneity of the variance. The mathematical forms of the models are the following:

Total Crash Frequency Model

$$\begin{aligned} Ln(total\_crashes+1) = 5.020 + 0.474\times 10^{-1} \times POP\_PAC + \\ 0.196\times 10^{-3} \times POP16\_64 + \\ 0.151\times 10^{-2} \times TOT\_MILE, \end{aligned} \tag{4-1}$$

Fatal Crash Frequency Model

$$\begin{aligned} Ln(fatal\_crash+1) = 0.652 - 0.924\times 10^{-1} \times INT\_PMI + \\ 1.762\times PNF\_0111 + 1.389\times PNF\_0512 + \\ 0.263\times 10^{-3} \times POP00\_15 + 0.319\times PPOPMIN, \end{aligned} \tag{4-2}$$

The assessment of the adequacy of these models required several steps.

4.5.1 Generation of the expected crash counts in Excel

The model inputs are crash counts by municipality, independent variables for the base year 2002, and the independent variables estimated for four years later, the 2006 scenario.

4.5.2 Calculation of correction factors for the base year

The correction factor is calculated using the average observed frequency of crashes divided by the average estimated frequencies. It is an essential component of the analysis because it corrects the differences between the safety model fitted to a region, and the region or state to be used. The BCF is used to evaluate the goodness of fit of the model.

$$BCF_{unbiased} = \frac{\sum_{i=1}^{N} \frac{O_i}{N}}{\sum_{i=1}^{N} \frac{P_i}{N}} = \frac{\sum_{i=1}^{N} O_i}{\sum_{i=1}^{N} P_i}, \tag{4-3}$$

where O_i is the observed crash frequency for municipality i, P_i is the estimated crash frequency by the model for municipality i, N is the number of municipalities.

To evaluate the goodness of fit, the BCF should be calculated for each municipality using

$$BCF_i = \frac{O_i}{P_i}, \tag{4-4}$$

where O_i is the observed crash frequency for municipality i, P_i is the predicted frequency using the model for municipality i.

The next step is to compute the average BCF through all municipalities using

$$BCF_{average} = \sum_{i=1}^{N} \frac{O_i}{P_i}, \tag{4-5}$$

The standard deviation and coefficient of variation of individual BCF's are computed as follows:

$$SD_{BCF} = \left[\frac{\sum_{i=1}^{N} \left(\frac{O_i}{P_i} - BCF_{average} \right)^2}{N} \right]^{1/2}, \tag{4-6}$$

and.

$$CV_{BCF} = \frac{SD_{BCF}}{BCF_{average}}, \tag{4-7}$$

where SD is the standard deviation, and CV is the coefficient of variation.

The standard deviation is the standard deviation of the population of the correction factors of the municipality, and the coefficient of variation is the standard deviation divided by the average. Those are calculated to compare the goodness of fit of the models.

4.5.3 Prediction of crashes with models found in the literature

The data from the 2002 base year for the region of interest is used to calculate the correction factors described in step 3. The model is then used with the independent variables estimated for the year of interest of (2006) to predict future collisions. Model predictions for all municipalities are multiplied by the correction factor previously computed to obtain the corrected estimate of crashes in the future scenario. These prediction models are thus adjusted for local conditions.

4.5.4 Comparison of the coefficients of variation of the correction factors to observe if the data fits the model

The coefficient of variation, CV is a measure of unexplained variation in the crash model. A CV near zero indicates that the model fits the observed data well. A CV equal to one suggests that the standard deviation is the same as the average. A CV greater than one suggests that there are significant unexplained variations in local models, indicating a lack of fit to the model. Thus, values less than one are preferred.

All steps mentioned above were programmed into an Excel spreadsheet, as summarized in Tables 4.5 and 4.6. The total crash model presents a CV greater than one, which may suggest that there are significant unexplained variations in local models, and indicates a lack of fit to the model.

Table 4.5 Total crashes spreadsheet

BASE YEAR 2002						
MUNICIPALITY	**Oi**	**POP_PAC**	**POP16_64**	**TOT_MILE**	**Pi (Ln Y)**	**BCF**
Aguada	2286	2.14	28490	262.49	11.10	29.01
Aguadilla	3243	2.75	42410	219.06	13.79	301.65
Añasco	1206	1.14	18980	272.07	9.21	8.25
Cabo Rojo	1703	1.07	30969	397.36	11.74	73.74
Guánica	665	0.91	13755	187.39	8.04	4.67
Hormigueros	820	2.33	11026	300.36	7.75	2.82
Isabela	2082	1.27	29414	80.11	10.97	27.81
Lajas	1032	0.79	16996	274.53	8.80	6.45
Las Marías	370	0.47	7139	238.42	6.80	2.43
Maricao	174	0.28	4215	156.12	6.09	2.55
Mayagüez	6282	1.96	64653	394.58	5.05	0.02
Moca	1979	1.25	26559	327.68	10.78	24.27
Rincón	621	1.64	9779	110.94	7.18	2.12
Sabana Grande	862	1.13	16747	191.34	8.64	6.59
San Germán	1670	1.06	23998	324.97	10.26	17.19
Totals	**24995**				**136.21**	
					Unbiased BCF	**183.5**
					Average BCF	**34.0**
					Std.dev.BCF	**73.9**
					CV BCF	**2.2**

PREDICTION YEAR 2006						
Oi	**POP_PAC**	**POP16_64**	**TOT_MILE**	**Pi (Ln Y)**	**BCF**	**ADJUSTED Pi**
2404	2.15	28975	262.49	11.20	183.5	2055
3394	2.70	41458	219.06	13.60	183.5	2496
1329	1.16	19390	272.07	9.29	183.5	1704
1999	1.11	31980	397.36	11.94	183.5	2191
800	0.87	13205	187.39	7.93	183.5	1456
921	2.37	10968	300.36	7.74	183.5	1420
2739	1.29	30067	80.11	11.10	183.5	2036
996	0.79	16948	274.53	8.79	183.5	1614
333	0.45	6929	238.42	6.76	183.5	1240
212	0.27	4210	156.12	6.09	183.5	1118
6282	1.88	62131	394.58	17.88	183.5	3282
2096	1.26	26982	327.68	10.86	183.5	1993
727	1.66	9838	110.94	7.19	183.5	1320
940	1.12	16602	191.34	8.62	183.5	1581
1923	1.05	23547	324.97	10.18	183.5	1867

Table 4.6 Fatal crashes spreadsheet

BASE YEAR 2002								
MUNICIPALITY	**Oi**	**INT_PMI**	**PNF_0111**	**PNF_0512**	**POP00_15**	**POPMIN**	**Pi (Ln Y)**	**BCF**
Aguada	7	0.12	0.03	0.000	9979	0.13	3.37	2.08
Aguadilla	12	0.09	0.07	0.000	14089	0.17	4.53	2.65
Añasco	5	0.08	0.02	0.000	6578	0.19	2.48	2.02
Cabo Rojo	5	0.05	0.05	0.000	10242	0.17	3.48	1.43
Guánica	0	0.10	0.04	0.000	5191	0.20	2.14	0.00
Hormigueros	4	0.04	0.03	0.000	3245	0.19	1.62	2.46
Isabela	7	0.35	0.17	0.000	10161	0.17	3.64	1.92
Lajas	2	0.10	0.03	0.000	5818	0.20	2.29	0.87
Las Marías	0	0.08	0.00	0.000	2627	0.14	1.38	0.00
Maricao	0	0.08	0.00	0.000	1612	0.11	1.10	0.00
Mayagüez	9	0.10	0.05	0.004	18781	0.22	5.75	1.57
Moca	7	0.09	0.00	0.000	9872	0.11	3.27	2.14
Rincón	1	0.06	0.00	0.000	3193	0.14	1.53	0.65
Sabana Grande	2	0.07	0.04	0.000	6011	0.15	2.35	0.85
San Germán	5	0.11	0.06	0.000	7870	0.17	2.87	1.74
Total	**66**						**42**	
							Unbiased BCF	**1.58**
							Average BCF	**1.36**
							Std.dev.BCF	**0.88**
							CV BCF	**0.64**
PREDICTION YEAR 2006								
Oi	**INT_PMI**	**PNF_0111**	**PNF_0512**	**POP00_15**	**POPMIN**	**Pi (Ln Y)**	**BCF**	**ADJ Pi**
6	0.12	0.03	0.00	9020	0.13	3.12	1.58	4.92
14	0.09	0.07	0.00	13055	0.17	4.26	1.58	6.72
3	0.08	0.02	0.00	6167	0.19	2.37	1.58	3.74
4	0.05	0.05	0.00	10102	0.17	3.45	1.58	5.44
4	0.10	0.04	0.00	4569	0.20	1.98	1.58	3.12
1	0.04	0.03	0.00	3161	0.19	1.60	1.58	2.53
6	0.35	0.17	0.00	9613	0.17	3.50	1.58	5.52
8	0.10	0.03	0.00	5329	0.20	2.17	1.58	3.42
3	0.08	0.00	0.00	2304	0.14	1.30	1.58	2.05
0	0.08	0.00	0.00	1450	0.11	1.06	1.58	1.68
1	0.10	0.05	0.00	16795	0.22	5.23	1.58	8.25
3	0.09	0.00	0.00	9309	0.11	3.13	1.58	4.93
2	0.06	0.00	0.00	3037	0.14	1.49	1.58	2.35
6	0.07	0.04	0.00	5600	0.15	2.24	1.58	3.53
1	0.11	0.06	0.00	7248	0.17	2.70	1.58	4.27

The fatal crash model presents a CV less than one. However, it must be noted that the prediction Pi is to be back transformed in order to be expressed in the original scale, and hence the variability will increase. Therefore the Arizona models were not able to explain the phenomena very well. This can be attributed to the variation of geographical and human behavior between states.

4.5.5 Identification of independent variables used in the western region model of Puerto Rico

Due to Arizona model's lack of fit to Puerto Rico's characteristics, a decision was made to find new explanatory variables and develop a western region model with the data of the 15 municipalities that form this region. Road network characteristics, socio-economic and demographics were utilized as possible independent variables. The frequency of crashes by municipality was the dependent variable.

Prior to modeling, all of the variables were examined individually to determine whether the variables were logical. Checks for reasonableness include computing means, medians, modes, maxima, and minima for all the variables in the database. Often times, coding and transcription errors are detected during this process. This helped to eliminate atypical data values on the model on the modeling results.

A descriptive analysis of the different explanatory variables was performed and several graphs were plotted in order to better understand the data (Section 3.3). In addition to the model variables used in the evaluated literature, the number of miles in primary, secondary, tertiary, municipal, and state roads were also considered. This division is more common on the roads of Puerto Rico than the division made according to functional classification adopted in existing models (for more details see Appendix Table 1, chapter 4).

4.5.6 Development of the Puerto Rican highway crash prediction model and the selection of the best model

This section consists of developing an initial model using a set of independent variables generated in section 4.5.5. Stepwise, forward and backward analyses for a 0.15 level of significance were realized. These procedures were applied with the purpose of reducing the group of variables during a preliminary analysis for variable selection. After these analyses were performed, few variables were found to be statistically significant or had a logical interpretation. The initial model was a linear regression with the logarithmic transformation of the dependent count variable, similar to the model done by Washington et al. which utilized the ordinary least squares method.

The steps followed for the development of the Puerto Rican binomial negative macro models were: data collection and analysis of significant variables, the definition of the distribution and model structure, the estimation of the parameters, and model selection (McCullagh et al., 1989). This methodology is described in more detail in chapters 5 and 6 section 2. However a brief description is also provided in this section.

4.5.6.1 Data collection and analysis of significant variables

The stepwise, forward and backward process was repeated several times and a number of candidate models were estimated by used variations of variables and adding, maintaining or dropping variables at a 0.15 level of significance.

4.5.6.2 Defining distribution and model structure

The number of collisions are nonnegative and discrete in nature. Therefore, Poisson and Negative Binomial models were developed. In this case, negative binomial models presented the best fit due to the over dispersion parameter (variances of counts are greater than means).

4.5.6.3 Estimation using the Maximum Likelihood (ML) Method and model selection

The maximum likelihood method was used for parameter estimation. And the Newton Raphson method is used as the optimization technique.

The Negative Binomial models presented the best fit due to the over dispersion parameter (variances are greater than the means). The best of the Negative Binomial models were selected using the Akaike's Information Criterion (AIC).

4.5.6.4 Models for total crashes

The resulting model for total crashes after taking the log transformation is the following log-normal model.

$$Ln\_(crashes+1) = 5.84 + 0.03 \times Population/1000, \tag{4-8}$$

The coefficient of determination, R^2 , is equal to 0.84, indicating that 84% of the variability of the dependent variable is explained by the model predictor.

The resulting log - linear negative binomial model where the mean is not directly related to the variable, if not by a logarithmic link function is the following

$$\begin{aligned} Ln\left\{\hat{E}(crashes)\right\} &= 4.5972 + 0.5098\times10^{-2}\times Highway\_miles \\ &+0.6432\times POP\_PAC\times 9.316\times Interstates, \end{aligned} \tag{4-9}$$

In equation 4.9, highway miles represent the total miles of highways, POP_PAC is the population density (population estimates from U.S. Census SF1) in persons per acre, and INTERSTATES is the proportion of interstate highways.

The fit of the model is analyzed using Pearson Chi Square/DF method. This resulted in a value of 1.15, is close to 1, indicating good fit of the model to the data.

4.5.6.5 Models for fatal crashes

The resulting model for fatal crashes after taking the log transformation is log-transformation is the following log-normal model. In this equation, municipal roads represents the miles of municipal roads as a proportion of total miles of roads.

$$\begin{aligned} Ln\_(fatal\_crashes+1) &= 0.01+0.02 \\ &\times Municipality\_roads, \end{aligned} \tag{4-10}$$

The coefficient of determination R^2 , is 0.31, indicating that 31% of the variability of the dependent variable is explained by the model predictor.

4.5.6.6 Models for injury crashes

The resulting model for injury crashes after taking the log transformation, is the following log-normal model

$$\begin{aligned} Ln\_(Injury\_crashes+1) &= 2.54+0.97\times POP\_PAC \\ &+0.03\times Tertiary\_roads, \end{aligned} \tag{4-11}$$

The coefficient of determination R^2 is 0.80, indicating that 80% of the variability of the dependent variable is explained by the model predictors.

Log linear-Negative Binomial Model

$$\begin{aligned} Ln\left\{\hat{E}(crashe\,s)\right\} &= 2.74+0.88\times POP\_PAC+2.57 \\ &\times Interstates+0.003\times Tertiary\_roads, \end{aligned} \tag{4-12}$$

In equation 4.12 POP_PAC represents the population density (population estimates from U.S. Census SF1) in persons per acre, interstate is the proportion of interstate roads, tertiary_ road is the miles of tertiary roads as a proportion of the total road mileage.

The fit of the model was analyzed by using of the Pearson Chi-Square/DF method which resulted in a value of 1.35. The result is near to 1, indicating that the model describes the data adequately (Values below 2 are considered good fit of data to the model).

4.6 RESULTS AND DISCUSSION

The objective of this section is to compare the log normal and log linear negative binomial models in order to determine which one offers a better prediction. The Tables 4.7 and 4.8 present the results for the prediction of crashes in the municipalities of the western region for the year 2006, by lognormal and log linear negative binomial models respectively. First, the models are fit and then the number of crashes are predicted. The best prediction was obtained using negative binomial models because the model yielded more accurate predictions. Table 4.7 presents the log- normal model and Table 4.8 presents the results for the log-linear negative binomial model.

Table 4.7 Log-normal total crash model prediction analysis

	YEAR 2002				YEAR 2006		
MUNICIPALITY	**Oi**	**Population/1000**	**Pi**	**BCF**	**Oi**	**Population/1000**	**Pi**
Aguada	2286	42.46	1533.64	1.49	2404	42.57	1539.89
Aguadilla	3243	64.41	3311.69	0.98	3394	63.13	3166.03
Añasco	1206	28.76	948.69	1.27	1329	29.27	965.84
Cabo Rojo	1703	48.19	1875.07	0.91	1999	50.01	1998.41
Guánica	665	21.54	736.61	0.90	800	20.62	713.35
Hormigueros	820	16.88	625.66	1.31	921	17.19	632.37
Isabela	2082	45.11	1683.49	1.24	2739	45.77	1722.90
Lajas	1032	26.39	873.20	1.18	996	26.28	869.87
Las Marías	370	10.89	507.17	0.73	333	10.46	499.56
Maricao	174	6.46	434.21	0.40	212	6.43	433.68
Mayagüez	6282	97.22	10462.66	0.60	6282	93.60	9214.61
Moca	1979	40.18	1416.02	1.40	2096	40.51	1432.60
Rincón	621	14.98	585.36	1.06	727	15.21	590.10
Sabana Grande	862	25.98	860.68	1.00	940	25.84	856.40
San Germán	1670	37.03	1268.14	1.32	1923	36.52	1245.45
			BCFaverage	**1.05**			

Table 4.8 Log linear- negative binomial total crash model prediction analysis

YEAR 2002						
MUNICIPALITY	**Oi**	**Highway_miles**	**POP_PAC**	**Interstates**	**Pi**	**BCF**
Aguada	2286	262.49	2.14	0.03	2060.56	1.11
Aguadilla	3243	219.06	2.75	0.07	3481.14	0.93
Añasco	1206	272.07	1.14	0.02	1021.48	1.18
Cabo Rojo	1703	397.36	1.07	0.05	2412.59	0.71
Guánica	665	187.39	0.91	0.04	659.36	1.01
Hormigueros	820	300.36	2.33	0.03	2833.73	0.29
Isabela	2082	80.11	1.27	0.17	1649.66	1.26
Lajas	1032	274.53	0.79	0.03	903.27	1.14
Las Marías	370	238.42	0.47	0.00	452.41	0.82
Maricao	174	156.12	0.28	0.00	262.55	0.66
Mayagüez	6282	394.58	1.96	0.05	4260.30	1.47
Moca	1979	327.68	1.25	0.00	1176.93	1.68
Rincón	621	110.94	1.64	0.00	500.99	1.24
Sabana Grande	862	191.34	1.13	0.04	800.11	1.08
San Germán	1670	324.97	1.06	0.06	1749.40	0.95
					BCFaverage	**1.04**

YEAR 2006				
Oi	**Highway_miles**	**POP_PAC**	**Interstates**	**Pi**
2404	262.49	2.15	0.03	2068.35
3394	219.06	2.70	0.07	3360.60
1329	272.07	1.16	0.02	1034.93
1999	397.36	1.11	0.05	2476.04
800	187.39	0.87	0.04	643.22
921	300.36	2.37	0.03	2911.26
2739	80.11	1.29	0.17	1669.55
996	274.53	0.79	0.03	901.38
333	238.42	0.45	0.00	447.04
212	156.12	0.27	0.00	262.30
6282	394.58	1.88	0.05	4065.09
2096	327.68	1.26	0.00	1184.77
727	110.94	1.66	0.00	509.16
940	191.34	1.12	0.04	796.93
1923	324.97	1.05	0.06	1732.86

A summary of the principal insights from the crash models of western region areas is presented here.

• Trial and error is used to derive meaningful and useful models. Knowledge of transportation safety is used to derive a model that is consistent and in agreement with current knowledge of motor vehicle crashes and safety.

• The models developed in this work provide a tool for that can be used by traffic planners. This tool can be used to set targets for meeting safety objectives and performance milestones, and for providing feedback on development and/or growth scenarios

• The focus of macro-level models is prediction. The goal of these models are to inform the analyst what problems can occur in the future.

• The safety analysis of a particular project will be able to predict the expected number of crashes and consider them as an explanatory statement regarding safety, and not as a defining statement. It would represent the injury severity risk expected by changes in the number of intersections, residential development, road mileage, and local population.

• Although valid explanations are provided for the predictor variables in the models, the models are not mainly used to gain insight about variables related to crashes. Instead, the main goal is predicting crash outcomes by municipality. This restriction is not too dissimilar from the restriction placed on travel demand models, whose primary purpose is to predict demand for roadway space of motor vehicles in hypothetical or future scenarios.

• One should simply use the model to forecast possible crash problems that may happen in the future or during a given hypothetical growth scenario. Once these possible outcomes are known, plans can be implemented to remediate these crashes using specific countermeasures, as required by the local conditions.

• The developed models in this chapter have limitations and assumptions. An important assumption of the models is that 'new' safety countermeasures such as rumble strips, crash cushions, etc. are not applied in future scenarios. In other words, the 'average' set of design standards with respect to safety are assumed to exist in the future, Innovative, newly adopted, or progressive safety countermeasure investments are analyzed independently using another model or research study.

• The appropriate use of the developed models fall in the planning, prediction, or forecasting domains, and not for the traffic and safety engineering domains.

• The next step would be to examine design policies and safety investments to ensure the regional safety goals are met.

• It is necessary to develop models that include public transit and others modes of transport as an independent variable in order to capture its effect on safety.

• It is necessary to develop a complete HSIP which monitors and evaluates such as changes in infrastructure, enforcement, education and medical emergencies.

• It is necessary to improve the current system and the collecting data in order to obtain better data quality that would produce better models.

• Transportation planning often focuses on infrastructure-related solutions. A much broader perspective on how the planning process can affect the safety of the transportation system would include recommended policies, processes, studies, and budget priorities.

• It is important that the long-range transportation plan include safety education on topics such as: use safety publicity, bicyclist and pedestrian, work zone, education policy, elderly driver evaluation, mature driver education.

• It is also important to include the following engineering and operation topics in the long-term transportation plan: traffic management safety audits of existing, rehabilitated and new roadways, traffic safety studies, and traffic safety measures in construction zones.

CHAPTER 5: DEVELOPING MACRO MODELS

5.1 INTRODUCTION

Urban transportation planning focuses on mitigating congestion and mobility in a proactive way. Large projects are developed in response to the results from the modeling processes. Then the most favorable alternatives are proposed in the long-term transportation plan recommendations.

Meanwhile, safety has received little attention during the strategic planning process. Highway safety is mostly assessed when the facility has been built, and the problems are evident or, as a recent practice, during the design stage. Incorporating safety in a comprehensive way has surged into the initial planning stage as a strategy to improve highway safety and achieve the zero crashes objective.

Many studies have been developed to predict the crashes since the early planning stages through the use of macro regional models. Some examples of these models are: Laumon Amoros (2003), Noland and Quddus (2004), Washington et al. (2006), Noland and Oh (2004), Agüero and Jovanis (2006), Hadayegui et al. (2006), Hadayegui et al. (2007), Lovegrone and Sayed (2007), Wier et al. (2009), Naderan and Shani (2010), and Abdel Aty et al. (2010) (See details in chapter 2). Macro models are used to predict the aggregated number of global crashes for a region (TAZ, county, municipality, etc) with the objective of comparing different planning alternatives. These models consider the population as an exposition measure and infrastructure, socioeconomic, and socio demographic aspects as independent variables. These models, however, are not sensitive to improvements in the highway system.

In Chapter 4, models for the western region of Puerto Rico were developed. These models had no good fit.

This chapter describes the GLMM models developed for the island, and the infrastructure variables employed for data fit. Section 5.1 presents an introduction to the importance of the topic strategic planning with regional models. Section 5.2 describes the methodology implemented for the development of the models, while Section 5.3 shows the results of regional generalized linear model with length as offset (APKPY- crash rate). Section 5.4 shows the regional models generalized linear mixed models with length and inhabitants as offset (APIPKPY- crash double rate) and region and municipality as random effects.

5.2 DETAILED METHODOLOGY

The steps followed for the development of generalized linear mixed binomial negative macro models are: data collection and analysis of significant variables, the definition of the distribution and model structure, the estimation of the parameters, model selection, and integration into the planning process (McCullagh et al.,1989). Figure 5.1 shows a flowchart with the methodology of this chapter.

5.2.1 Data collection and analysis of significant variables

The data used in this chapter were factors related to the design of highways and the population. The crash data was obtained from the transit regulation office in the safety auditory division of the Highway Transportation Authority from 2002 to 2004, the years in which PDO crashes were available. The population was obtained from American Fact Finder website http://factfinder2.census.gov from 2002 to 2004. Both databases were joined in Excel. The database was filtered by crash severity, including PDO crashes and by municipality.

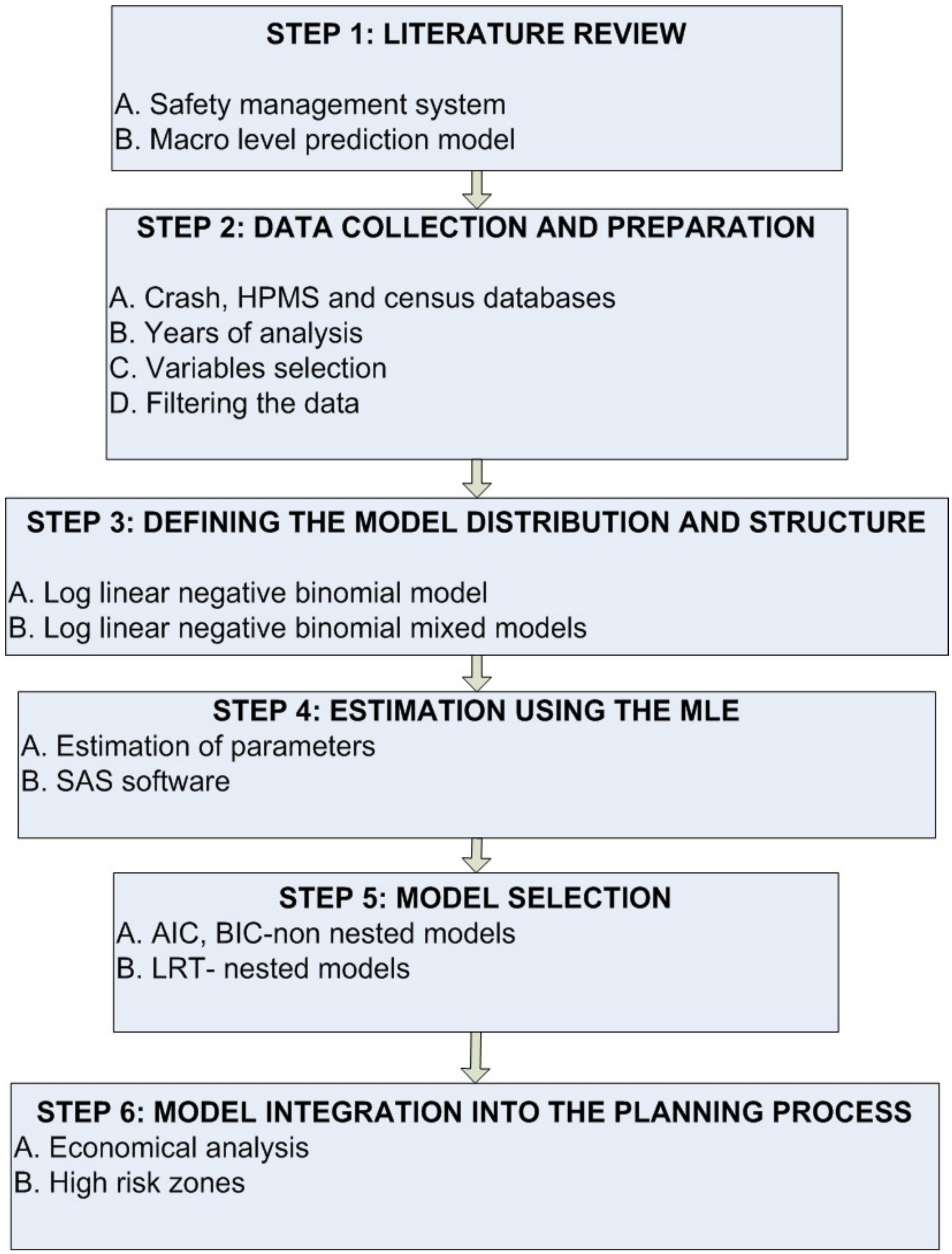


Figure 5.1 Methodology for all Puerto Rico municipality models.

Stepwise analysis was realized with the purpose of finding the significant variables for the models. The significant variables were the proportion of secondary and tertiary roads, and as an offset the number kilometers of highways and the population were used.

5.2.2 Defining the model distribution and structure

The macro region models have been developed considering that crashes are not negative and discrete in nature. As a result, the Poisson and Negative Binomial Models were analyzed. However, due to the dispersion of the data, the Negative Binomial with random effects model, which are called mixed binomial negative models were adjusted.

This section is to establish the conditional distribution of the response vector 'y' (crash counts) given the random effects 'u, representing, in this chapter, the municipality and regional effects. The length of the municipality highways, and length per population were considered as offset variables for modeling crash rates (i.e., number of crashes/length and number of crashes/length*inhabitant).

5.2.3 Estimation using the Maximum Likelihood (ML) method

The estimation of the model parameters was performed using the maximum likelihood method. In order to maximize the approximation of the likelihood function, the Laplace Method was used. The default optimization technique in this procedure was the dual quasi-Newton method. The results are the parameter estimates with their respective standard errors and the information criteria (AIC, BIC) used to compare different models (Litell et al.,2006).

5.2.4 Model selection

Based on the results of each model specification and considering the included explanatory variables, a series of tests and criteria were applied in order to compare the models, verify their goodness of fit, and choose the model that represents the best fit for the data under study. Some of the tests and criteria typically used are: likelihood ratio tests, conditional Pearson Chi-Square tests, AIC and BIC ,(Agresti, 2003).

5.2.5 Model integration into the planning process

A methodological framework was developed with the purpose of integrating the developed models into the strategic planning process activities. These models can be used to obtain a baseline condition (current circumstances) and to determine what would happen, in terms of crash rates, under various planning scenarios. One possible way to analyze the impact of the incorporation of safety devices is to create accident risk maps for each alternative scenario. Then, determine what would happen with the implementation of regional mitigation measurements or safety improvements. Another important impact that can be studied is the economic analysis, considering the direct and indirect costs of accidents in relation to the required strategic investments.

At the end of the planning process, the zones at greater risk of accidents should be included in the transportation plan for the development of more detailed analyses and the implementation of safety projects. Afterwards, those high risk zones should be followed up throughout the planning process to verify whether after the implementation of the countermeasures, there truly was a reduction in the accident rate as a result of the implemented projects.

Furthermore, the results of the model can provide planners with information about the safety data expected in the future. In order to do this, however one must assume that design standards in the future will be similar to those used in the present. Afterwards, planners could prepare new safety plans that would implement new system-wide safety initiatives. These safety initiatives would improve safety in the future and calculate the Crash Modification Factors (CMF's) accordingly. In this way, the planner can estimate the resources required to meet the objectives in regional safety.

Town expansion can affect the population, the number of miles on the roads, intersections density, etc. A point of interest here is that a statewide or regional safety objective predicting an X percent reduction in fatal crashes does not necessarily guarantee the reduction of the current level of total or fatal crashes. This is due to the possibility of population growth and other factors that will most likely increase the number of crashes. Therefore the crash models can become a tool for planning by helping them set targets in order to meet safety objectives and performance milestones.

5.3 RESULTS AND DISCUSSION OF APKPY MODELS

In this section, the models are standardizing by segment length. Models with crash frequency as dependent variable do not allow for a direct comparison between municipalities to be conducted. The models of crash frequency are not standardized by the total number of kilometers and the number of inhabitants. In contrast, crash prediction models standardized by length, allow the comparison of different municipalities that have a similar population so as to identify the most hazardous municipalities. This is done by using logarithm total kilometers as an offset.

In the modeling of municipalities, only the primary road, secondary road, tertiary road and population variables were found to be significant.

5.3.1 Model with rate (APKPY) as the dependent variable

The conditional Pearson/DF criterion was used in order to verify the fit of the model. Table 5.1 presents the goodness of fit and the comparison criteria for total crash prediction models with APKPY as the dependent variable. The response was assumed to have a negative binomial distribution.

Table 5.1 GLM Model comparison of freeways with APKPY as the dependent variable

#	Dependent variable (APKPY severity)	Explanatory Variables (X's)	AIC	BIC	-2log-likelihood	Conditional Pearson	Conditional Pearson/DF
1	Total PDO Injury Fatal	Propredprim,propredsec, propredter Years(2002 to 2004)	1634 1552 1292 688	1645 1563 1304 700	1624 1542 1282 678	387 195 190 229	5.23 2.82 2.75 3.32

APKPY: Accidents by kilometer per year (rate)
***pvalue <0.05**

In Table 5.1 and 5.2 the variables propredprim represents the proportion of primary roads, propredsec represents is the proportion of secondary roads, and propredter represent the proportion of tertiary roads.

5.3.2 Estimated parameters for APKPY models and interpretation

The maximum likelihood method was employed to estimate the parameters of the proposed models. The interpretation of the parameters was done by using the concept of elasticity, which is a general concept used to quantify the response of one variable when another variable changes.

Tables 5.2 presents the results of the goodness of fit using the different degrees of crash severities APKPY as the dependent variable.

Table 5.2 Estimated solutions for all severities model 1

Effect	total crashes Estimate (SE)	PDO crashes Estimate (SE)	Injury crashes Estimate (SE)	fatal crashes Estimate (SE)
Intercept	-41.7947 * (0.7572)	-42.4928 * (0.7666)	-44.1607 * (0.7575)	-50.5197 * (1.2658)
propredprim	18.2144 * (11.4789)	26.5290* (11.26)	22.0113* (10.8849)	23.6483* (14.1222)
propredsec	16.9455* (6.1884)	21.4822* (6.6245)	21.4077* (6.7755)	20.6861* (8.1292)
propredter	10.1253* (3.9344)	10.0348* (3.7355)	10.2377* (3.7373)	23.1328* (4.5024)
Scale	3.5614 (0.4523)	3.5262 (0.4590)	3.5374* (0.4611)	3.3447 (0.4680)

APKPY Accidents by kilometer per year
***p-value<0.05**

The model in Table 5.2 has a specification log-level. A model with specification log-level represents a link function that related the mean with the explanatory variables. The interpretation of β is based on the concept of elasticity, and is expressed as:

$$Log\{E(Y_i)\} = \alpha + \beta x_i, \qquad (5\text{-}1)$$

The value β*100 can be interpreted as the relative change (%) of the expected response by one unit of increase in the explanatory variable. For a detailed derivation see the Appendix, Chapter 5.

5.4 RESULTS AND DISCUSSION OF APIPKPY MODELS

In this section the models are standardized by segment length and population. Some models found in the literature employed the frequency, or crash rate per kilometer, as a dependent variable in the data analysis. However, to perform a direct comparison between different municipalities, obtain a hierarchy for hazardous municipalities, and know how the conceptual designs of the road types affect safety, while keeping the exposure level constant, it is necessary to employ length and population standardization. In the modeling of municipalities, only the primary road, secondary road, tertiary road and population variables were found to be significant as regresors.

5.4.1 Comparison of models with double rate (APIPKPY) as the dependent variable

The conditional Pearson/DF criterion was employed in order to determine the modeling fit and a different criterion was used for comparing the developed models. Table 5.3 presents the three total crash prediction models with APIPKPY as the dependent variable, and the comparison criteria for the models. The selection of the best model was conducted using the Bayesian Information Criterion, which has a penalty for degree of freedom and is useful for non-nested models.

Table 5.3 GLMM comparison of freeways with APIPKPY as the dependent variable.

#	Dependent variable (APIPKPY severity)	Variables (X's)	AIC	BIC	-2loglikelihood	Conditional Pearson	Conditional Pearson/DF
2	Total PDO Injury Fatal	Propredsec, propredter Random (Region) Years(2002 to 2004)	3879 3721 2915 1165	3878 3721 2915 1165	3869 3711 2905 1155	124 125 145 690	0.53 0.55 0.64 3.07
3	**Total** **PDO** **Injury** **Fatal**	**Propredsec,** **propredter** **Random** **(Municipality)** **Years(2002 to 2004)**	**3792** **3664** **2854** **1101**	**3806** **3678** **2868** **1115**	**3780** **3652** **2842** **1089**	**46** **42** **63** **159**	**0.20** **0.19** **0.28** **0.71**
4	Total PDO Injury Fatal*	Propredsec, propredter Random (Region-Municipality) Years(2002 to 2004)	3783 3656 2849 1102	3783 3656 2849 1102	3769 3642 2835 1088	46 42 63 159	0.20 0.19 0.28 0.71

APIPKPY: Accidents by inhabitant by kilometer per year (double rate).

*** Matrix G is not positively defined**

In Table 5.3 the variable propredsec represents the proportion of secondary roads, and propredter is the proportion of tertiary roads.

In order to compare between models 3, 4 and 5, the criteria for non nested models was used. The fit of the models was verified by using the Pearson/DF criterion. According to the BIC and fit, 4 was the best model for all four responses.

5.4.2 Estimated parameters for APIPKPY the models and interpretation

Tables 5.4, 5.5, and 5.6 present the results goodness of the fit of the model for total crashes by inhabitant by kilometer per year as the dependent variable.

Table 5.4 Estimated solutions for all severities model 3 (Random-Region)

Effect	Total crashes Estimate (SE)	PDOI crashes Estimate (SE)	Injury crashes Estimate (SE)	fatal crashes Estimate (SE)
Intercept	-9.9205* (0.2617)	-10.0141* (0.2510)	-11.6668* (0.2382)	-16.4401* (0.3845)
Propredsec	4.1814* (1.7186)	3.8786* (1.6429)	4.1620* (1.5618)	1.9562 (2.2289)
Propredter	4.1525* (1.5324)	3.9187* (1.4677)	2.9438* (1.4110)	8.1348* (2.0921)

APIPKPY: Accidents by inhabitant by kilometer per year (double rate).
***p-value<0.05**

Table 5.5 Estimated solutions for all severities model 4 (Random- Municipality)

Effect	**Total crashes Estimate (SE)**	**PDO crashes Estimate (SE)**	**Injury crashes Estimate (SE)**	**fatal crashes Estimate (SE)**
Intercept	-10.8515* (0.6019)	-10.9148* (0.5894)	-12.5923* (0.4945)	-16.5444* (0.3982)
Propredsec	12.9289* (3.9342)	12.5620* (3.8465)	10.8263* (3.2099)	3.6217 (2.6388)
Propredter	4.6365 (3.8420)	4.2401 (3.7531)	4.5872 (3.1356)	6.3135* (2.4945)

APIPKPY: Accidents by inhabitant by kilometer per year (double rate).
***p-value<0.05**

Table 5.6 Estimated parameters for arterial models for all severities model 5 (Random-Region and Municipality)

Effect	**Total crashes Estimate (SE)**	**PDO crashes Estimate (SE)**	**Injury crashes Estimate (SE)**	**fatal crashes Estimate (SE)**
Intercept	-10.5873* (0.6219)	-10.67981* (0.6111)	-12.3461* (0.5212)	-16.5444* (0.6594)
Propredsec	9.0160* (3.8292)	8.9258* (3.7733)	8.3557* (3.2178)	3.6217** (2.9322)
Propredter	4.5945 (3.6444)	4.2639 (3.5935)	3.9782 (3.0923)	6.3133* (3.8795)

APIPKPY: Accidents by inhabitant by kilometer per year (double rate).
***p-value<0.05**

The interpretation is based on the concept of elasticity for a generalized linear mixed model with specification log-level, expressed as:

$$log\{E(Y_{ij} \mid U_i)\} = \alpha + \beta(x_i) + U_i, \tag{5-2}$$

The value of β*100 can be interpreted as the relative change (%) of the expected response with one unit of increase for the explicative variable, for any given value of the random effect. A detailed derivation is found in the appendix, Chapter 5.

CHAPTER 6: DEVELOPMENT OF FREEWAY-EXPRESSWAY AND ARTERIAL MODELS

6.1 INTRODUCTION

Different micro models have been developed in order to predict the number of crashes in an intersection, highway segment or on the project level with different geometric characteristics. This was done by taking the AADT as an exposure measure, and by using the road geometry and the operational characteristics as independent variables. Some models were developed in AASTHO’s HSM in order to analyze the safety of arterials and two way highways. The models can be currently used in a reactive manner or in the proactive analysis of new highway designs, but should not be used during the early strategic planning stages due to detailed design and operational requirements.

The general goal of this study is to develop tools in order to improve the decision making process related to safety planning. However, unlike the macro models, the micro models are difficult to use during the early stages of the planning process. Macro models are used to obtain global predictions at a regional scale while micro models are used at a smaller scale, focusing on specific sites, project design or existing networks.

This chapter’s main objective consists of the development of generalized linear mixed models that can be used in the strategic planning or long-term planning process. These models are versatile with a wide inference space, which means that they can represent the big picture big picture for any particular functional classifications such as freeways and arterials. Also this means that the forecast can be performed for any segment type of the population of segments and not only for a pre- determined level or type of segment. These models consider spatial effects while using innovative methodologies that improve the efficiency of the estimators, and require a relatively low sample size when compared to GLM. The proposed models will help obtain crash estimates for each segment and highway type. These models are also used to divide crashes in a region by type of highway.

The models developed can be divided into specific subject models and by population average. The specific subject models are those that improve the reliability of the estimation at the site level by using EB correction by means of the random effect of the site and can be used in the short-term planning. The population average models are those that can be used in the general strategic planning process in the mid and long-term.

GLMMs are used to incorporate safety into the strategic planning process. These models allow for the incorporation of random variables that may exist in the system but are not observed. Therefore, they offer versatility in crash rate modeling and its related factors.

The general methodological approach used for this research includes the filtration of the acquired crash data by way of segment analysis and crash severity. The Highway Performance Monitoring System (HPMS) database was filtered in accordance to the segments and the independent variables that were selected to be analyzed. Afterwards, the roads geometry data and the crash databases were incorporated. Next, the model structure was defined; the parameters of a group of candidate GLMMs were defined using the maximum likelihood method. The estimated models were assessed using the significance tests and goodness of fit methods. Then the models that best fitted the studied phenomena were selected. SAS software was used to perform the calculations for the model.

This chapter describes the GLMM models developed for the island and the infrastructure variables employed for data fit. The Section 6.1 presents an introduction to the importance of the topic of strategic planning with freeway and arterial models. Section 6.2 describes the methodology implemented for the development of the models. Section 6.3 shows the results of generalized linear mixed models for freeways and arterials. The models used length as the offset (APKPY-crash rate), and segment type as a random effect.

6.2 DETAILED METHODOLOGY

This section describes the methodology developed with the data available in Puerto Rico that could be used for the development of highway safety prediction models and would be appropriate for strategic planning. The models were developed and classified by road type while considering the possibility of data correlation. These models were used to explore the relationship between crash rates by crash severity, and various variables such as traffic intensity (AADT), and geometric characteristics. Crash severities are defined as crashes that result in a fatality, an injury or property damage only. Traffic intensity refers to the average annual daily traffic .Geometric characteristics describes the number of lanes, their width, slope, curvature, etc. The rest of this section describes the general research approach and the methodological steps followed. Figure 6.1 shows a flowchart with the methodology of this chapter.

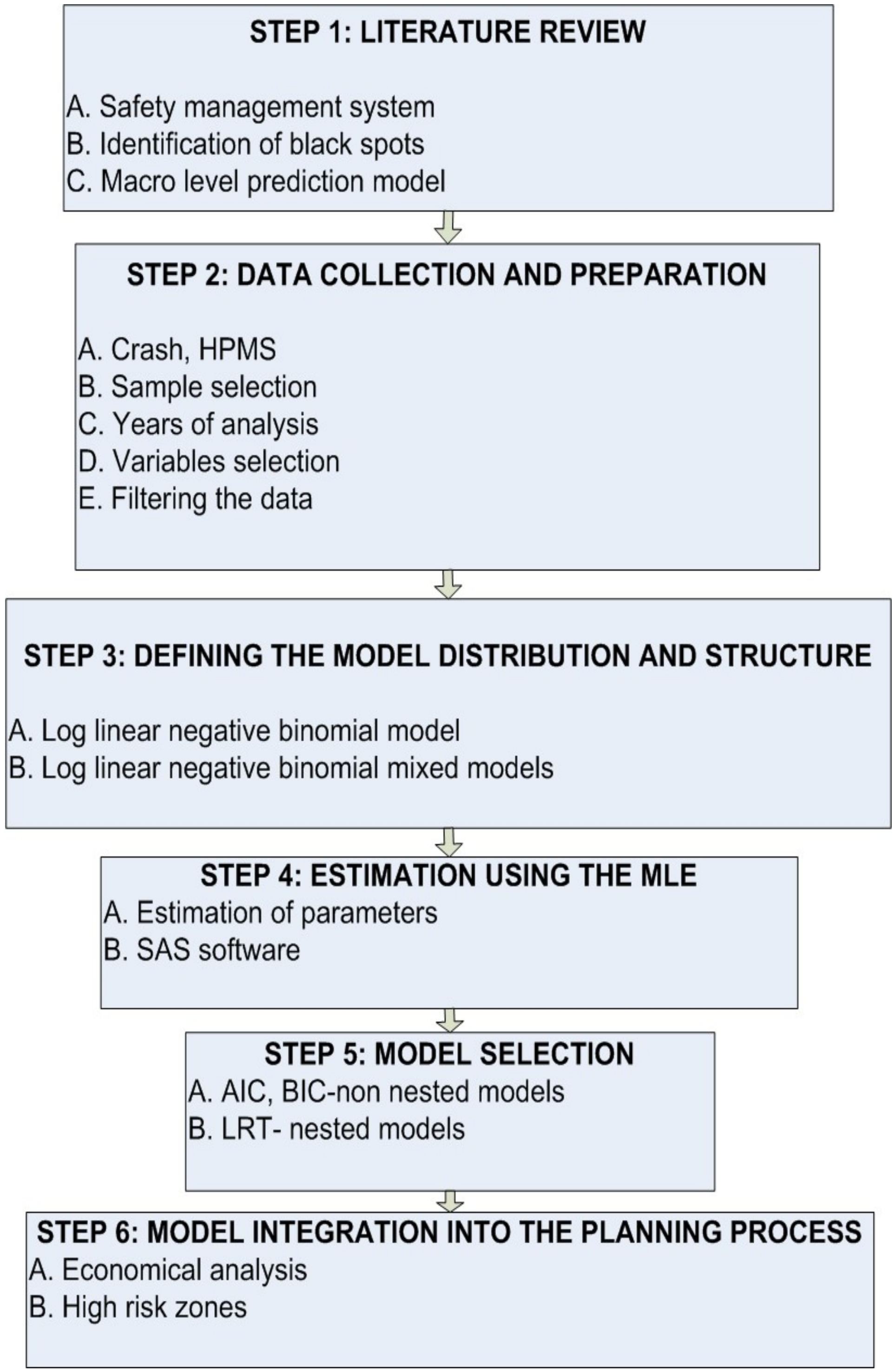


Figure 6.1 Methodology for Puerto Rico freeway and arterial models.

A wide variety of methods have been implemented in the past years. In order to deal with the problems regarding the methodology and the data associated with accident frequency. If not handled correctly, they can compromise the statistical validity of the analysis. Previous studies have used various approaches for the aggregation and modeling of accidents depending on the purpose of the study and the nature of the available data. The most common approaches include lineal regression, Poisson models, the Negative Binomial model, the Full Bayes hierarchical model with spatial effects, and linear models with logarithmic transformations. The present study developed a statistical modeling approach using GLMM with mixed negative binomial distribution. Much like a classical linear regression, a GLMM includes a set of variables that would be used to explain the mean crash rates. However, the GLMM also includes random effect terms in order to incorporate the correlation between observed crash rates in the same segment. The estimation of the parameters in the GLMMs were done using SAS software (SAS V.9.3 Institute Inc.,2013).

Some of the advantages of using a GLMM to model the relationship between crash rates, their severity, and several explanatory variables include: the proper handling of crash rates through the specification of a probability distribution for the number of crashes, and the ability to use the length and AADT of the segments as offsets variables. Different patterns in data variability were modeled by changing the probability distributions for the number of crashes. Additionally, random terms were incorporated into the model in order to account for the correlation between crash rates in the same segment and its spatial effects. The expected crash rates are related to possible predictors by using an appropriate link function that preserves their rate scale (i.e., positive numbers), and the GLMMs allow flexibility when building the model. This is important since GLMM’s embrace a big family of models (for instance, classical linear models become a special case of a GLMM).

For the non-Bayesian models proposed in this work, the estimation of the regression parameters was performed by using the same inferential methods. What varied was the probability distribution for the number of crashes at a given segment and the link function used to relate the number of crashes to potential predictors. The general procedure used to estimate the parameters in a GLMM includes the following (Agresti, 2003).

The steps followed for the development of generalized linear mixed binomial negative models are: data collection and analysis of significant variables, the definition of the distribution and model structure, the estimation of the parameters, model selection, and integration into the planning process (McCullagh et al.,1989).

6.2.1 Data collection and analysis

The data used in this study are associated with a variety of factors related to the freeways and arterials in Puerto Rico, such as the number of crashes and the geometric and operational characteristics of the highways from 2004 to 2009. The crash data was obtained from the Transit Regulation Office in the Safety Auditory Division of the Highway Transportation Authority. The various highway geometric characteristics were obtained from the Highway Performance Monitoring System database. The HPMS office manages a database as part of a federal program that requires an inventory of geometric and operational variables. The inventory contains a random and representative sample of highway segments from the main functional classification system. The geometric and operational variables are taken into consideration for each of the segments in the random sample. A segment is defined as a highway section of varied length that has similar geometric and operational characteristics. The random sample used in this study was provided by the HPMS office.

The analyzed database contained 98 variables pertaining to the horizontal and vertical geometry of the segments, as well as their operational characteristics. The preliminary analysis of the data utilized the 14 variables found to be significant in the literature review. These variables were: type of terrain, grade of curvature, design speed, speed limit, region, type of shoulder, type of median, AADT, width of lane, width of right shoulder, width of left shoulder, proportion of signalized intersections, proportion of stop intersections, other intersections and proportion of ramps. However, only 5 variables were found to be truly significant after further analysis.

The database was filtered by crash severity, and segments were discarded if they were found to have significant changes in the geometry or in the implementation of countermeasures between 2004 and 2009.

Tables 6.1 and 6.2 show the definition of the significant variables in freeways and arterials, respectively. The municipalities within each region are defined in the Appendix, Chapter 6.

Table 6.1 Descriptive statistics of significant variables for freeways

Definition of variables	N(%) Sample size	Mean (SD) APKPY or AADT	[Minimum, Maximum] APKPY or AADT
Region (categorical)			
01-San Juan - (population- 1,070,609)	228 (29.92%)	17.68 (14.01)	[0, 77.78]
02-Arecibo - (population- 628,745)	114 (14.96%)	6.71 (3.68)	[0.71, 21.27]
03-Aguadilla - (population - 380,002)	6 (0.79%)	2.81 (0.32)	[2.37, 3.16]
04-Mayaguez- (population- 366,967)	66 (8.66%)	4.38 (2.53)	[0.74, 11.90]
05-Ponce - (population- 411,596)	108 (14.17%)	7.27 (5.14)	[0, 21.67]
06-Guayama - (population- 311,781)	24 (3.15%)	5.85 (2.32)	[1.95, 13.17]
07-Humacao- (population- 568,651)	216 (28.35%)	9.19 (7.98)	[0, 36.73]
AADT (Covariate)			
Annual Average Daily Traffic	762 (100%)	77880.31(57411.10)	[10400, 290000]
Type of shoulder (Categorical)			
1-None	12 (1.57%)	12.64 (7.29)	[5, 26.88]
2-Surfaced	702 (92.13%)	9.76 (9.93)	[0, 77.78]
3-Stabilized	12 (1.57%)	18.01 (3.78)	[12.5, 25]
4-Combined	36 (4.72%)	21.93 (12.78)	[0.57, 52.31]

In Table 6.1 and 6.2, N (%) represents the sample size of segments (%), the Mean represents the mean of APKPY, SD represent the standard Deviation of APKPY, the Minimum represent the minimum value of APKPY, AADT or signalized intersections by kilometer, respectively. The Maximum represents the maximum value of APKPY, AADT or signalized intersections by kilometer, respectively.

In Table 6.1 the sample size N 762 is, corresponding to a random sample of 146 segments of freeways and 6 years of crash analyses.

Table 6.2 Descriptive statistics of significant variables for Arterials

Definition de variables	N(%) Sample size	Mean (SD) APKPY, AADT or int/kil	[Minimum, maximum] APKPY, AADT or int/kil
Region (categorical)			
01-San Juan- (population- 1,070,609)	155 (25.41%)	8.02 (10.55)	[0, 57.14]
02-Arecibo - (population- 628,745)	72 (11.8%)	2.20 (3.54)	[0, 14.44]
03-Aguadilla - (population - 380,002)	120 (19.67%)	16.72 (14.76)	[0, 84.62]
04-Mayaguez (population- 366,967)	90 (14.75%)	12.05 (6.85)	[0, 40]
05-Ponce (population- 411,596)	53 (8.69%)	0.67 (1.08)	[0, 4.85]
06-Guayama (population- 311,781)	-	-	-
07-Humacao (population- 568,651)	120 (19.67%)	16.37 (14.65)	[0, 76.67]
AADT (Covariate) Annual Average Daily Traffic	595 (100%)	44678.32 (24882.16)	[600, 126300]
Group of signalized intersections (categorical)			
A- 0 to 0.53	298 (48.85%)	7.83 (9.90)	[0, 76.67]
B-0.54 to 1.36	156 (25.57%)	11.37 (8.43)	[0, 40.67]
C- 1.37+	156 (25.57%)	15.29 (17.56)	[0, 84.61]
Signalized intersections by kilometer (continuous)	595 (100%)	0.98 (1.38)	[0, 10]

In Table 6.2 the sample size N is 595, which consists of 107 random arterial segments of arterials and 6 years of crash analyses.

6.2.2 Defining the distribution and model structure

The freeway and arterial models have been developed considering that crashes are not negative and discrete in nature. Therefore, the Poisson and Negative Binomial distribution were analyzed. However, due to the dispersion of the data, Negative Binomial with random effects model, which are called mixed negative binomial models were adjusted.

The purpose of this section is to establish the conditional distribution of the response vector 'y' (crash counts) given the random effects 'u', representing, in this case, the road segments effect. The length and AADT of the segments are considered as offset variables for modeling crash rates (i.e., number of crashes/length and number of crashes/length*AADT). In a classical GLMM, the random terms in ‘u’ are usually assumed to have a normal distribution. A set of potential predictors for crash rates needs to be included in the model’s formulation. This allows us to explain mean crash rates as a function of the predictors. We fit the GLMMs using link functions that are considered natural for the distribution of the response (McCullagh and Nelder, 1989). The logit link function serves as an example when the proportion of crashes are modeled as a Binomial distribution, and the log link function can also be used as an example when the number of crashes are modeled as a Poisson or Negative Binomial distribution and is considered instead.

6.2.3 Estimation using the Maximum Likelihood (ML) Method

The estimation of the parameters in a GLMM is often performed using the ML method based on a Laplace approximation of the marginal distribution of the response variable. The maximization of the likelihood function is achieved using the dual quasi-Newton method. These computational tasks have already been implemented in existing statistical software such as SAS. Other estimation techniques can be considered depending on the complexity of the model. An advantage of using the ML method is that some of the traditional model selection methods are directly applied to our problem. Moreover, the inference process on the parameters in the model is supported by the ML theory (Agresti, 2003).

The Maximum Likelihood Estimation in this work is based on the Laplace Approximation. The idea is to approximate the integral by using a square approximation around the point in which the integrand takes its maximum (Demidenko, 2004).

These days, this method is widely regarded as the most useful one. The approximation can be found implemented in the procedure glimmix of SAS using the Laplace option, which indicates an approximation of first order with a point of quadrature.

6.2.4 Model selection

Based on the results of each model specification, and considering the included variables, a series of tests were done in order to compare candidate models, verify their goodness of fit, and define the model that represents the best fit for the phenomena under study. For instance, likelihood ratio tests were used to compare models that nest with one another. Alternatively, we relied on model selection criteria such as AIC, BIC or HQ to compare models with the same distribution for the variable response and link function.

6.2.5 Work plan and tasks

The following tasks were conducted in this research project:

i. The literature review consisted of identifying existing models for accident forecasting at the planning stage, and identifying methodologies that were to be used for model estimation. Additionally, interviews were conducted at PRHTA and PRDOT to determine the current approach used in Puerto Rico for the inclusion of safety at the planning stage. The procedures used to consider highway safety issues in the mid-term and long term planning processes in Puerto Rico were identified and described.

ii. Data was gathered from several sources such as the Traffic Safety Commission (CST, for its acronym in Spanish), the safety office of the Puerto Rico Highway and Transportation Authority (PRHTA), the Puerto Rico Department of

Transportation and Public Works (PRDOT), and the Census databases. The highway safety predictive models used in the Puerto Rico at mid and long term in the transportation plans were identified.

iii. An accident database was organized and filtered by accident type, functional classification, and geometric and operational characteristics. The database was separated by each group of homogeneous characteristics, while considering the segments reported by the PRHTA for the Highway Performance Monitoring System (HPMS) required by the FHWA.

iv. The list of potential variables that were to be included in the planning models were identified. These variables were selected and tested in the databases, to verify their significance in terms of the number and quality of the data available. At this point, the data was reviewed to ensure the variables were the most representative and appropriate for the model.

v. Typical models that were used in the studied literature to model the highway safety phenomena as part of the planning process were verified. Generalized linear mixed models were used in this project. The model parameters were estimated using traditional inferential methods already implemented in existing statistical analysis software. Models were developed according to their highway functional classification with various segment characteristics.

vi. The significance and importance of each variable included in these models was tested. Goodness of fit tests were also performed in order to make sure that the developed models best represented the modeled phenomena. It is not uncommon that this process the initial models needed some adjustment. During this process the number of variables or the model form were continually adjusted. Therefore, the process was repeated from Task 5 to Task 7, until a good set of models were obtained.

viii. An integrated highway safety strategic planning process was proposed in Chapter 7 considering the methodologies already proposed in the reviewed literature and the unique characteristics of the strategic planning process in Puerto Rico. The developed models were incorporated in the selected methodology and presented to DOT officials for their consideration and to help them comply with the federal requirements by law.

6.3 RESULTS AND DISCUSSION OF FREEWAYS AND ARTERIALS (APKPY) MODELS

In this section the models were standardized by length of segment only. The freeway model included 12 variables, but only region, type of shoulder, number of lanes and AADT were significant. Six planning models were found to consider the number of accidents by kilometer by year and the number of accidents by millions of vehicle- miles traveled by year as dependent variables.

Arterial modeling included 14 variables, but only region, groups of intersections, proportion of signalized intersections, and AADT were found to be significant. Three planning models were found by considering accidents per kilometer per year and accidents per million vehicle-miles traveled by year as dependent variables. Table 6.3 and Table 6.4 presents freeway and arterial models that can be used to incorporate safety into the early stage of planning. The goodness of fit and the information criteria for the comparison of the models are also shown.

6.3.1 Comparison of freeway and arterial models with rate (APKPY) as the dependent variable

A freeway is defined as a major divided highway designed for high-speed travel, having few or no intersections. Also called freeway, limited access highway superhighway, and thruway. Meanwhile an arterial is defined as an arterial road, or arterial thoroughfare, is a high-capacity urban road. The primary function of an arterial road is to deliver traffic from collector roads to freeways or expressways, and between urban centers at the highest level service possible. As such, many arteries are limited-access roads or feature restrictions on private access. Frequency models do not allow for a comparison directly between the sites to be conducted. Therefore, the standardized crash prediction models by length, allow for the comparison of different segments with the same AADT so as to identify the most hazardous sites or characteristics. Information criterion was used for nested models and non nested models when comparing models. The fit of the models were verified by the conditional Pearson/DF criterion.

6.3.1.1 Comparison of significant freeway models

Table 6.3 presents the six predictive models of the total crash rates, the goodness of fit, and the comparison criteria for the freeway models.

The selection of the best models was done with the Bayesian Information Criterion (BIC), which has a penalty related to the degrees of freedom and is used for non nested models and likelihood ratios (LRT) for nested models.

According to the BIC, models 1, 2, 5 and 6 have the best fit. From the planification's point of view, the number of crashes when using the same design or functional classification varies according to the location of the highway segment and its intricate characteristics. Therefore, the likelihood ratio is used between full model 5 and reduced model 6 in order to verify if the region contributes to the prediction of the model's variability. According to the LRT, model 6 is preferred (X^2=10, critical X^2=12.59, df=6). The significance of this is that, the model including only logAADT is preferred to predict crash rate for freeways. Also LRT is used between full model 2 and reduced model 1 in order to verify whether the number of lanes contributes to the prediction of the model's variability. According to the LRT, model 1 is preferred (X^2=6, critical X^2=5.99, df=2).

Table 6.3 GLMM Model comparison of freeways with APKPY as the dependent variable

#	Dependent variable (Severity)	Variables	AIC	BIC	-2log-likelihood	Conditional Pearson	Conditional Pearson/DF
1	**Total Injury Fatal**	**Lanes, type of shoulder, logAADT Random (segment)**	**5066 5065 1022**	**5091 5091 1048**	**5048 5047 1004**	**638 642 850**	**0.84 0.85 1.12**
2	Total Injury Fatal	Type of shoulder, logAADT Random (segment)	5068 5067 1024	5088 5087 1044	5054 5053 1010	633 637 818	0.83 0.84 1.08
3	Total Injury Fatal*	Region, lanes, type shoulder, logAADT Random (segment)	5064 5063 1024	5106 5105 1063	5034 5033 996	643 646 924	0.85 0.85 1.22
4	Total Injury Fatal*	Region, type shoulder, logAADT Random (segment)	5066 5065 1024	5103 5102 1059	5040 5039 1000	637 640 925	0.84 0.84 1.22
5	Total Injury Fatal*	Region, logAADT Random (segment)	5074 5073 1019	5103 5101 1045	5054 5053 1001	635 638 932	0.84 0.84 1.23
6	**Total Injury Fatal**	**LogAADT Random: (segment)**	**5072 5072 1018**	**5084 5083 1030**	**5064 5064 1010**	**633 636 826**	**0.83 0.84 1.09**
7	Total Injury Fatal	LogAADT, Speed limit, Design speed, Region	5047 5045 1021	5098 5096 1046	5011 5009 1003	647 651 811	0.85 0.86 1.07

APKPY: Accidents by kilometer per year (rate)

***Matrix G is not positively defined**

6.3.1.2 Comparison of significant arterials models

Table 6.4 shows the four models used for the prediction of total crashes by kilometer by year, the goodness of fit, and the comparison criteria of the arterial models.

Table 6.4 Model comparison for arterials with APKPY as the dependent variable

#	Dependent variable (Severity)	Variables	AIC	BIC	-2log-likelihood	Conditional Pearson	Conditional Pearson/DF
1	Total Injury Fatal	Region, LogAADT, group of intersections by kilometer Random: segment	3483 3455 716	3512 3484 744	3461 3433 694	575 580 485	0.99 1.00 0.83
2	Total Injury Fatal	Region, log AADT Random: segment	3485 3457 716	3508 3480 740	3467 3439 698	573 578 483	0.98 0.99 0.83
3	**Total Injury Fatal**	**Region, LogAADT, signalized intersections by kilometer Random: segment**	**3481 3453 713**	**3507 3479 739**	**3461 3433 693**	**575 579 454**	**0.99 0.99 0.78**
4	Total Injury Fatal	LogAADT, signalized intersections by Kilometer Random: segment	3527 3500 716	3540 3513 729	3517 3490 706	545 548 456	0.93 0.94 0.78

APKPY: Accidents per kilometer per year (rate)

The best models were found to be 2 and 3 according to the BIC. The number of crashes when using the same design or functional classification varied by highway segment location and its implicit characteristics. Therefore, the likelihood rate is used between models 2 and 3 to check if the region provided additional predictability to the variability of the model. According to the LRT, model 3 is preferred (X^2=6, critical X^2=3.84, df=1). Which means, the model including region, logAADT and signalized intersections by kilometer is preferred to predict arterial APKPY.

6.3.2 Estimated parameters of APKPY models and interpretation for freeways and arterials

The maximum likelihood was employed to estimate the parameters. The interpretation of the parameters is done through the concept of elasticity.

a. Estimated parameters to freeways

Tables 6.5 to 6.7 present the parameter estimates for the six freeway models of total injury and fatal crashes, in which the significant variables are region, number of lanes, type of shoulder and logAADT.

Table 6.5 Estimated parameters for total freeway crash models with APKPY as the dependent variable

	M1	M2	M3	M4	M5	M6
Effect	**Estimate (SE)**	Estimate (SE)	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**
Intercept	**-8.0560** (1.1920)**	-7.5270** (0.7565)	-9.3641** (1.2235)	-8.7580** (0.8333)	-9.1392** (0.8581)	-8.7114** (0.8501)
logAADT	**0.9575** (0.09841)**	0.9177** (0.06369)	1.0921** (0.1010)	1.0549** (0.07685)	1.0219** (0.07920)	0.9710** (0.07695)
Shoulder 1	**-0.7754* (0.4154)**	-0.4562 (0.4093)	-0.7570** (0.3813)	-0.4565 (0.3761)	-	-
Shoulder 2	**-0.7652** (0.2129)**	-0.6298** (0.2104)	-0.8484** (0.2021)	-0.7552** (0.2065)	-	-
Shoulder 3	**-0.3394 (0.4081)**	-0.1015 (0.4068)	-0.5991 (0.3884)	-0.4361 (0.3991)	-	-
Shoulder 6	**(Ref)**	(Ref)	(Ref)	(Ref)	-	-
4 Lanes	**0.2190 (0.1930)**	-	0.2565 (0.2298)	-	-	-
6 Lanes	**0.3960** (0.1646)**	-	0.4048* (0.1697)	-	-	-
7 > Lanes	**(Ref)**	-	(Ref)	-	-	-
Region 1	**-**	-	-0.2448 (0.1514)	-0.3370** (0.1399)	-0.1909 (0.1348)	-
Region 2	**-**	-	-0.3877** (0.1384)	-0.3920** (0.1358)	-0.4003** (0.1422)	-
Region 3	**-**	-	-0.5265 (0.4558)	-0.5776 (0.4717)	-0.6093 (0.5012)	-
Region 4	**-**	-	0.0059 (0.1614)	-0.04862 (0.1648)	-0.08181 (0.1743)	-
Region 5	**-**	-	0.1502 (0.1385)	0.1022 (0.1378)	0.07423 (0.1452)	-
Region 6	**-**	-	-0.1221 (0.2317)	-0.1243 (0.2405)	-0.1460 (0.2555)	-
Region 7	**-**	-	(Ref)	(Ref)	(Ref)	-

Ref= Reference level
****p-value<0.05**
***p-value<0.1**

Table 6.6 Estimated parameters for freeway crash injury models with APKPY as the dependent variable

	M1	M2	M3	M4	M5	M6
Effect	**Estimate (SE)**	Estimate (SE)	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**
Intercept	**-8.1702** (1.2003)**	-7.6034** (0.7602)	-9.5078** (1.2282)	-8.8555** (0.8354)	-9.2382** (0.8617)	-8.4901** (0.7267)
logAADT	**0.9664** (0.09910)**	0.9235** (0.0640)	1.1035** (0.1014)	1.0626** (0.0771)	1.0293** (0.0795)	0.9501** (0.0657)
Shoulder 1	**-0.7642* (0.4173)**	-0.4512 (0.4108)	-0.7443* (0.3818)	-0.4528 (0.3763)	-	-
Shoulder 2	**-0.7689** (0.2139)**	-0.6343** (0.2112)	-0.8528** (0.2024)	-0.7611** (0.2067)	-	-
Shoulder 3	**-0.3331 (0.4100)**	-0.09663 (0.4083)	-0.6002 (0.3888)	-0.4329 (0.3995)	-	-
Shoulder 6	**(Ref)**	(Ref)	(Ref)	(Ref)	-	-
4 Lanes	**0.2241 (0.1942)**	-	0.2656 (0.2303)	-	-	-
6 Lanes	**0.3917** (0.1655)**	-	0.4030** (0.1700)	-	-	-
7 > Lanes	**(Ref)**	-	(Ref)	-	-	-
Region 1	**-**	-	-0.2446 (0.1516)	-0.3381** (0.1401)	-0.1912 (0.1353)	-
Region 2	**-**	-	-0.3938** (0.1387)	-0.3959** (0.1360)	-0.4049** (0.1426)	-
Region 3	**-**	-	-0.5231 (0.4565)	-0.5714 (0.4722)	-0.6038 (0.5026)	-
Region 4	**-**	-	0.000566 (0.1618)	-0.05213 (0.1651)	-0.08679 (0.1749)	-
Region 5	**-**	-	0.1610 (0.1387)	0.1156 (0.1379)	0.0866 (0.1456)	-
Region 6	**-**	-	-0.1267 (0.2320)	-0.1263 (0.2407)	-0.1503 (0.2562)	-
Region 7	**-**	-	(Ref)	(Ref)	(Ref)	-

Ref= Reference level
****p-value<0.05**
***p-value<0.1**

Table 6.7 Estimated parameters for fatal freeway crash models with APKPY as the dependent variable

	M1	**M2**	**M3**	**M4**	**M5**	M6
Effect	Estimate (SE)	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**	Estimate (SE)
Intercept	-11.3324** (2.2963)	-9.7660** (1.3681)	-12.4740** (2.3742)	-10.3407** (1.6437)	**-10.3623** (1.5745)**	-9.9148** (1.2324)
logAADT	0.8324** (0.1902)	0.7063** (0.1133)	0.9399** (0.1984)	0.7871* (0.1478)	**0.7697** (0.1426)**	0.7109** (0.1101)
Shoulder 1	-0.9430 (0.8467)	-0.3866 (0.8188)	-0.8577 (0.8232)	-0.3283 (0.7874)	**-**	-
Shoulder 2	-0.4307 (0.3755)	-0.1006 (0.3488)	-0.4312 (0.3610)	-0.1926 (0.3409)	**-**	-
Shoulder 3	-0.6600 (0.6080)	-0.09281 (0.5690)	-0.7462 (0.5754)	-0.4159 (0.5510)	**-**	-
Shoulder 6	(Ref)	(Ref)	(Ref)	(Ref)	**-**	-
4 Lanes	0.5197 (0.3616)		0.7054* (0.4154)	-	**-**	-
6 Lanes	0.6920** (0.2961)	-	0.6868** (0.3146)	-	**-**	-
7 > Lanes	(Ref)	-	(Ref)	-	**-**	-
Region 1	-	-	-0.1427 (0.2552)	-0.3715 (0.2427)	**-0.3063 (0.2168)**	-
Region 2	-	-	-0.2443 (0.2125)	-0.2099 (0.1987)	**-0.1847 (0.1931)**	-
Region 3	-	-	-0.8413 (1.0294)	-0.8983 (1.0293)	**-0.8856 (1.0292)**	-
Region 4	-	-	0.07123 (0.3001)	0.000458 (0.2934)	**0.00754 (0.2930)**	-
Region 5	-	-	-0.6717** (0.2981)	-0.6979** (0.2888)	**-0.6868** (0.2877)**	-
Region 6	-	-	0.1196 (0.2667)	0.09855 (0.2619)	**0.1097 (0.2603)**	-
Region 7	-	-	(Ref)	(Ref)	**(Ref)**	-

Ref= Reference level
****p-value<0.05**
***p-value<0.1**

b. Estimated parameters for arterial models

Tables 6.8 to 6.10 present the parameter estimates for the three arterial models of total injury and fatal crashes, in which the significant variables are the region, group of proportion of signalized intersections, intersections by kilometers and the logAADT.

Table 6.8 Estimated parameters for models of total arterial crashes with APKPY as the dependent variable.

	M1	M2	M3	M4
Effect	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**
Intercept	-8.7525** (1.5669)	-9.3121** (1.5874)	**-9.1803** (1.5500)**	-10.5764** (1.1450)
Region 1	-1.1610** (0.2127)	-1.0857** (0.2158)	**-1.1080** (0.2084)**	-
Region 2	-1.5458** (0.3043)	-1.5785** (0.3123)	**-1.5261** (0.3024)**	-
Region 3	0.1699 (0.2139)	0.2106 (0.2195)	**0.1800 (0.2115)**	-
Region 4	-0.4721** (0.2343)	-0.4224* (0.2407)	**-0.4745** (0.2327)**	-
Region 5	-0.4720 (0.4723)	-0.5463 (0.4796)	**-0.4687 (0.4673)**	-
Region 7	(Ref)	(Ref)	**(Ref)**	-
logAADT	1.0854** (0.1441)	1.1091** (0.1469)	**1.0850** (0.1437)**	1.1657** (0.1095)
Group of proportion for signalized intersections 1	-0.4309** (0.1800)	-	**-**	-
Group of proportion for signalized intersections 2	-0.2928 (0.1946)	-	**-**	-
Group of proportion for signalized intersections 3	(Ref)	-	**-**	-
Signalized intersections by kilometer	-	-	**0.1350** (0.0532)**	0.1541** (0.07003)

Ref= Reference level
*** *p-value<0.05**
***p-value<0.1**

Table 6.9 Estimated parameters for arterial injury crash models with APKPY as the dependent variable.

Effect	M1 Estimate (SE)	M2 Estimate (SE)	M3 Estimate (SE)	M4 Estimate (SE)
Intercept	-9.0637** (1.5754)	-9.6860** (1.5979)	**-9.5555** (1.5638)**	-10.8457** (1.1561)
Region 1	-1.1669** (0.2110)	-1.0934** (0.2143)	**-1.1144** (0.2076)**	-
Region 2	-1.5212** (0.3026)	-1.5482** (0.3108)	**-1.5001** (0.3019)**	-
Region 3	0.1817 (0.2119)	0.2203 (0.2177)	**0.1914 (0.2105)**	-
Region 4	-0.5206** (0.2327)	-0.4713** (0.2392)	**-0.5179** (0.2320)**	-
Region 5	-0.4630 (0.4712)	-0.5160 (0.4788)	**-0.4446 (0.4678)**	-
Region 7	(Ref)	(Ref)	**(Ref)**	-
logAADT	1.1127** (0.1450)	1.1421** (0.1479)	**1.1187** (0.1450)**	1.1895** (0.1105)
Group of proportion for signalized intersections 1	-0.4280** (0.1785)	-	**-**	-
Group of proportion for signalized intersections 2	-0.3045 (0.1930)	-	**-**	-
Group of proportion for signalized intersections 3	(Ref)	-	**-**	-
Signalized intersections by kilometer	-	-	**0.1275** (0.0533)**	0.1441** (0.07026)

Ref= Reference level
*** *p-value<0.05**
***p-value<0.1**

Table 6.10 Estimated parameters for fatal arterial crash models with APKPY as the dependent variable.

	M1	M2	M3	M4
Effect	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**
Intercept	-9.2444** (3.4105)	-9.7507** (3.4340)	**-9.3610** (3.3716)**	-10.2481** (2.0431)
Region 1	-0.9611** (0.3741)	-0.9444** (0.3734)	**-0.9715** (0.3644)**	-
Region 2	-1.0616* (0.5421)	-1.0432* (0.5491)	**-1.0185* (0.5375)**	-
Region 3	-0.5357 (0.3352)	-0.4488 (0.3346)	**-0.4297 (0.3256)**	-
Region 4	0.1385 (0.3594)	0.1858 (0.3602)	**0.1152 (0.3534)**	-
Region 5	-0.1030 (0.9512)	-0.2779 (0.9566)	**-0.2036 (0.9419)**	-
Region 7	(Ref)	(Ref)	**(Ref)**	-
logAADT	0.7429** (0.3132)	0.7802** (0.3173)	**0.7264** (0.3119)**	0.7677** (0.1933)
Group of proportion for signalized intersections 1	-0.3601 (0.3203)	-	**-**	-
Group of proportion for signalized intersections 2	0.2097 (0.3176)	-	**-**	-
Group of proportion for signalized intersections 3	(Ref)	-	**-**	-
Signalized intersections by kilometer	-	-	**0.2058** (0.0880)**	0.2213** (0.0946)

Ref= Reference level
****p-value<0.05**
***p-value<0.1**

The interpretation is based on the concept of elasticity for a model with specification log-log. This specification is related to the logarithmic link function. This log link function as well as the mean is found on the left hand side of the equation 6.1. The explanatory variables, such as the AADT, are located on the right hand side of the equation. The log link function is used to join the mean and the explanatory variables.

$$log\{E(Y_{ij} \mid U_i)\} = \alpha + \beta \log(x_i) + U_i, \tag{6-1}$$

The value of β can be interpreted as the relative change (%) of the expected response by an increase of 1% in the explanatory variable at a given segment U_i (is only valid when β is small). For derivation details see the Appendix, Chapter 6.

In all of the freeway models the coefficient of logAADT is positive, which means that for one given segment (with constant geometry), with each 1% AADT increase, there is an increase of β % in the number of crashes by kilometer by year (APKPY), while maintaining the other elements constant .

The following equation is an example of the model for total crashes for freeway model 5 (M5) in region 1 taken from Table 6.5 is the following and can be observed in Figure 6.2:

$$\hat{E}(Y \mid U_i) = \exp\left(-9.33 + 1.02 \log AADT_{ij} + U_i\right), \tag{6-2}$$

The interpretation of this model is that, for each 1% increase in the proportion, there is an increase of 1.02% in the APKPY, while maintaining the other elements constant (Ui).

Figure 6.2 presents an example of safety performance functions for a typical segment of freeway. The typical segment is the segment with average risk or random effects equal to zero. The solid line is the expected number of crashes and both of the dashed line represent its 95% confidence interval. This means that for a typical segment, the expected number of collisions are expected to be within the city with a 95% confidence level. As an example for an AADT of 300,000 the average APKPY for a 95% confidence interval is between [50-150].

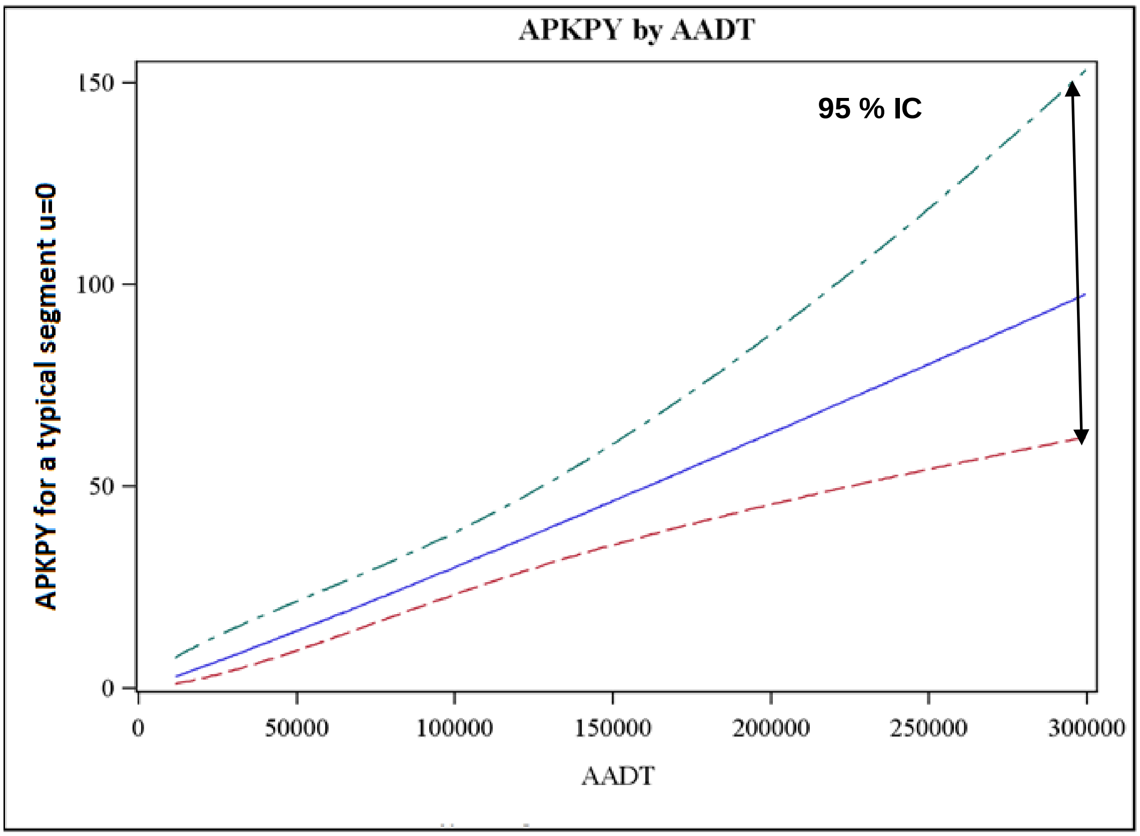


Figure 6.2 Example of SPF of APKPY by AADT for region 1 for freeways

The Figure 6.2 shows that for smaller values of AADT, the APKPY values are less dispersed, but for higher values of AADT the dispersion of APKPY is greater.

Equation 6.3 is an example of the model for total crashes for arterial model 1 (M1) in region 1 and the group of intersections for kilometer number 1 found in Table 6.8,

$$\hat{E}(Y \mid U_i) = \exp\left(-9.18 + 1.08 \log AADT_{ij} + U_i\right), \tag{6-3}$$

The interpretation of this model is that, for each 1% increase in the proportion, there is an increase of 1.08% in the APKPY, while maintaining the other elements constant. The coefficient of AADT is positive, which means that for each change in percentage of the AADT there is a 108 % increase in the APKPY.

CHAPTER 7: INCORPORATION OF MODELS INTO THE SAFETY MANAGEMENT SYSTEM

7.1 INTRODUCTION

The main factors in promotion of the economic productivity of a country are the improvement of mobility and safety. The mobility efforts in place for the planning of urban transport is apparent, however, the safety efforts are underrepresented. This is due to the lack of reliable methodologies for the decision making process in the early planning stage.

Reliable GLMMs enable the establishment of a long, mid and short range plan the last factor can be obtained through the design of projects that consider the established criteria and methods that were developed in this work in the long range transportation plan. These models can be used in order to generate a comprehensive and well coordinated Safety Management System. The Safety Management System (SMS) is a process used to develop strategies that decrease the quantity and severity of crashes. Incorporating safety into the planning process is one way to develop a comprehensive SMS. This plan has seven steps that were described in chapter 2; section 2.4, where step 4, called the technical analysis step, is the most directly related to the use of the developed models into the SMS.

A reference framework is described in this chapter. This chapter explains how regional models and functional classification models can be used in a proactive way; in the decision making process from a micro and macro perspective, in the conceptual design stage considering the hazard and vulnerability of the highway segments, or in the reactive form. The aforementioned information would all be placed into the SMS.

The framework mentioned above is useful in the development of a systematic and comprehensive planning process. The framework identifies the problems and opportunities to improve the transport system for long range planning. It also analyzes the relative affectivity of the different projects or strategies. Later, projects are designed according to the established goals and it is here where statistical models are required for decision making.

In some states, planning models have been developed, but many of the variables employed for the forecast of crashes cannot be manipulated by the planner. In those models, the dependent variable is a frequency, which does not allow direct comparisons between regions. Planners have tried to calibrate such models for other sites, but this has not yet yielded good fits.

The development of a comprehensive safety management system for Puerto Rico requires micro and macro models. The regional macro models are used to identify the current and future hazardous regions. They can also be used to establish long range mitigation plan measures, or to require detailed studies in the most hazardous regions. These models predict crashes, not their causes.

The models developed by functional classification give crash predictions per kilometer by type of road and are used in order to have a more specific idea of the number of regional crashes considering the functional class and their general design and operational characteristics. These features can be controlled in the planning and conceptual design phases.

The joining of the previous models allows for the execution of a network analysis while keeping the sensitivity of the improvements in the conceptual design, location, and mobility management strategies. The spatial effects of the different regions and of the segments are included in the models. The safety planning process is used to analyze different mobility alternatives, while considering the safety effects.

7.2 DETAILED METHODOLOGY

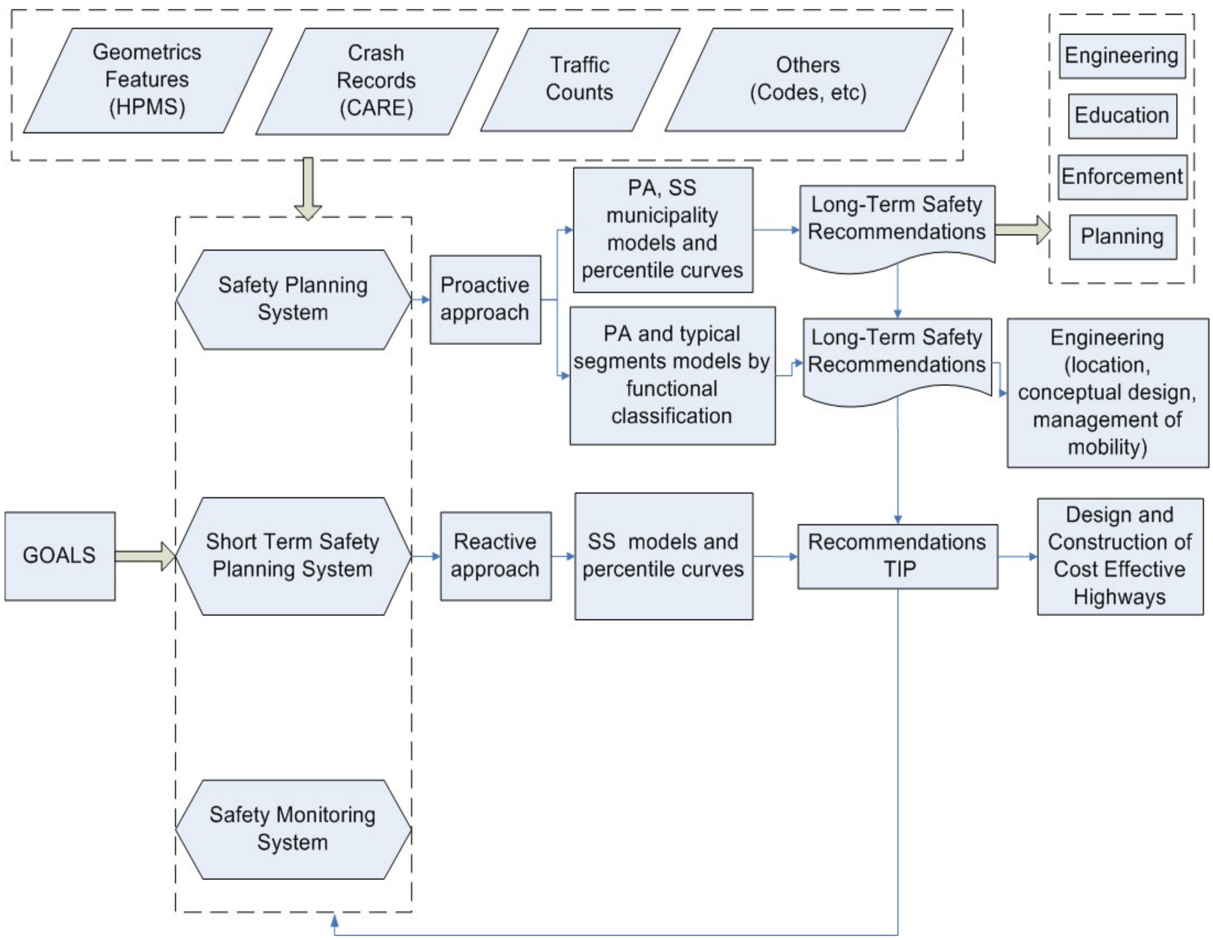


Figure 7.1 Detailed methodology to incorporate safety into transportation planning process through SMS (Adapted of Arizona SMS)

In the conventional seven step planning process, as described in chapter 2, section 2.4 the technical analysis step uses the four step model as a tool. It provides an estimate of future AADT as a result for each one of the highways in a region. With that estimate, the ability of the transportation network is evaluated so as to accommodate future trips. Later, the improvements required by the system from the mobility standpoint are identified, alternatives are established, and then evaluated. As a result, the new transport networks are planned long-term. Safety is evaluated afterwards, specifically in the short term planning phase (3 to 5 years).

The four step model estimates the AADT as well as other predictive variables related to the conceptual design of the project. This serves as the entry data in order to calculate the effect different transportation alternatives on various road types may have. The alternatives are evaluated in any given region by using the developed models. With this estimate, it is possible to evaluate the viability of the projects from the conventional standpoint and from a crash reduction perspective.

In order to conduct an evaluation of the alternatives using multiple criteria, a cost/benefit economical analysis is necessary. This allows an agency to strategically prioritize their projects as part of their long range plan. The development of a new transportation network, and its influence in the reduction of private vehicle voyages, can serves as an example. From these examples, generalized alternatives for the mobility and the conceptual design can be established.

These models can be employed in the development of macro crash modification factors (MCMF), which must consider the establishment of countermeasures at a larger scale. These should have an economic analysis developed and later be recommended in the long term transport plan. It can also be used to better understand the crash at the regional level through functional classification as mentioned in Chapter 5.

Methodologies that assess the hazard, vulnerability, and risk of structural systems in the face of any given event, exists in the structural engineering field. These hazards are associated with the exposure (geological fails in the structural area which can produce an earthquake with a given intensity), and the vulnerability associated with its physical characteristics (size of beams, materials of construction, year of construction, etc). The vulnerability factor is left constant, and it is through different hazard or earthquake scenarios that the risk is evaluated. In other words, the vulnerability of a structural system is evaluated in the event a hazardous situation were to take place. Hazards and structural vulnerability can have a great affect on a structure. These risks are quantified in terms of cost.

A similar fashion, highway risks analysis can be done using a rate model. Risk assessment is evaluated by having the hazard factor remain constant. The hazard (which is associated with the AADT), the functional classification of the highway, and design elements are varied. This affects the vulnerability of the different segments. For example, for arterials, the intensity of intersections per kilometer can be varied in its effect on segment risk. Highways can also be evaluated using the proposed highway risk analysis model. With the developed model, percentile curves can be calculated in order perform a risk assessment or compare the costs of each design alternative from a safety standpoint.

The dependent variable rate model has the advantage of realizing direct comparisons between different segments of built and planned roads. Then, hierarchy can be developed based on the results, as well as develop mitigation measures to address the latent risk foreseen for the sites.

7.2.1 Proactive use or safety planning system and Reactive use or short term planning

The proactive approach consists of taking into consideration the impacts in safety that different strategies of planning or establishment of large projects have on the frequency, rate, and severity of the crashes. It also consists of identifying and recommending mitigation projects at a large scale, improving the conceptual design of the projects, and conducting additional studies pertaining to safety in high risk regions. The reactive uses consist in the risk evaluation of segments with crash history.

7.2.1.1 Population average and specific subject models as tools to proactive technical analysis of alternative strategies for improvement

Step 4 of the planning process consists of model development and of performing an analysis of the available data in order to understand the effects (5 to 10 and 10 to 20 years) that different, large projects will have in the future for safety. It will also analyze the effects depending on different improvement alternatives in the conceptual design and how mitigation measures affect crash frequency, rate or severity.

A prediction is first conducted with the macro regional models, where municipalities with higher rates are found. For those municipalities, the population characteristics, drivers behavior, weather, environment, age of drivers, use of alcohol, and topography are analyzed. Findings or trends found will serve to make recommendations such as fine increases, awareness campaigns, policies, operation strategies, infrastructure projects, studies, regulations, education, awareness, financing strategies, association, and collaboration commitments.

Afterwards, the models by functional classification are used in order to have a greater understanding of what is happening in a particular region, and on each one of its highway types. So, it can tell if a region has a higher crash rate on expressways or arteries, in order to compare its value with the averages of the other regions.

In expressways, the variable to consider in order to rate the models were AADT, region, shoulder type, and the number of lanes. The variables considered in arterial models were AADT, region, group of intersections per kilometers, and intersection per kilometer. These variables take into account the effect of location, some aspects of conceptual design, mobility management, and multimodal aspects. The “region” variable captures the spatial heterogeneity of crashes.

As a result of the functional classification models, analysis recommendations in conceptual design, mobility management, and the multimodality of the transportation network can be established into the step 6 of the planning process, which consists of developing a plan and program with safety in mind (long term plan).

7.2.1.2 Reactive use (HSIP) or short term safety planning

There are six basic components in the reactive process of highway safety management. The six basic components are identification, diagnosis, remediation, economical evaluation, hierarchy of projects, and monitoring of existing highways.

The application of the methods developed from this research can help in identifying the black spot sites, or the sites with the most probability of improvement.

Additionally, applying these methods help agencies to determine if the potential improvements are economically justified, to establish priorities for potential improvements, and to evaluate the effectiveness of improvements after implementation. The development models provide tools for the estimation of average crash frequency in its application, in the screening of the network steps, and the economical evaluation of the HSIP process improvements.

7.2.1.3 Specific subject models as tools for identification of hazardous sites and regions respectively

A methodology at the regional or site level is proposed in this work. It consists of employing the EB method for the prediction of the number of expected crashes in a specific site by means of SS models while considering its random effects.

Through the developed models, crash predictions can be obtained using different options. When put into practice, oftentimes the parameters associated with the random effect are unknown. These parameters are related to the variability of the random effects and the residuals.

Figure 7.2. presents the lineplots of the percentiles for the average total crashes on freeways model 2 and shoulder type 6. Each one of the curves has a different associated random effect. The 50th percentile curve represents the segments with average risk. If we add half variance, a population average model is obtained, and a strategic plan can be developed. On the case of an AADT of 300000 the estimate of APKPY for segments with average risk is between [37.5-75].

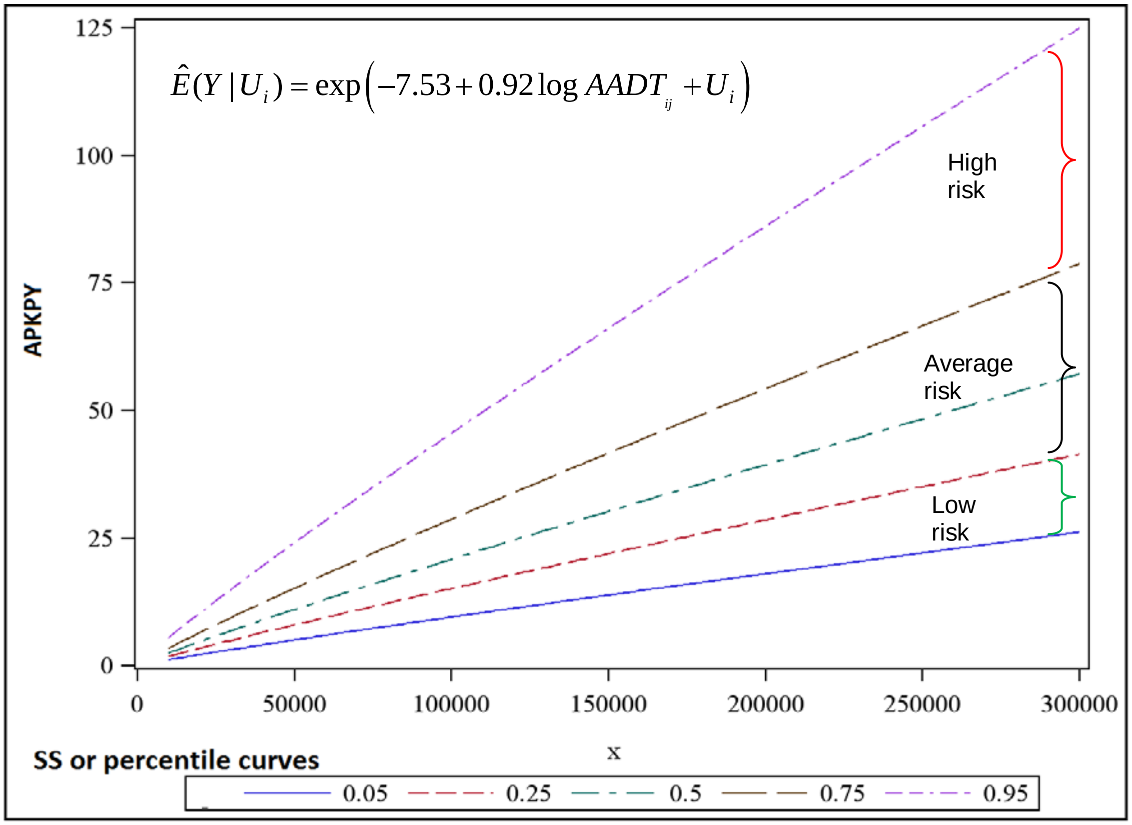


Figure 7.2 Percentile curves for freeways model 2 and shoulder type 6

The most specific details of design are included in the evaluation of design and construction alternatives. Economical evaluations are used to compare the benefits of countermeasures versus their cost. The predictive methods provide procedures for the estimation of crash frequency or average crash rates when the design and operation characteristics of the road are known, this with the purpose of establishment of a short-term plan.

Then of the implementation of the proactive and reactive approach, monitoring the system performance or step 7 of the planning process is conducted. This step consists of taking the collected data and evaluating the performance measures to determine whether there has been a decrease in the quantity and severity of crashes.

Figure 7.2 presents the curves percentiles for the total crashes in the freeway model 2 and shoulder type 6. Each one of the curves has a different associated random effect. The curve with 50^{th} percentile represent the segments with average risk, low percentile represent segments with low risk, and high percentile represent segments with high risk.

7.3 APPLICATION EXAMPLES

The first step during the transportation planning process is the establishment of goals and objectives in order to reduce crashes. However, the objectives must be coherent with the natural growth of the crashes in the long term to adequately monitor compliance with these. The efficient investment of resources in safety can be achieved by risk analysis. This is done by identifying the most hazardous locations, analyzing the characteristics of those locations, and using the available resources to increase the safety of those areas. The municipality risk and rate can be calculated with the models developed in this research, and in accordance with the percentile curves estimates with these models. The proactive approach analysis can be completed using the freeway and arterial models to analyze different alternatives of mobility using the safety approach. In the reactive planning approach, the segment risk can be found with the purpose of establishing the segments that require imminent improvements.

In this section, a series of simple fictitious and real examples are included to exemplify how the models developed can be used in particular situations. The example number one represents the proactive approach for municipality models in the development of strategic planning goals for the reduction of crashes in the long term. The second example is a proactive approach for the risk rating of municipalities. The third example is a proactive approach for analyzing freeway and arterials to calculate the number of crashes depending on different mobility alternatives. The fourth example represents the reactive approach for freeway and arterial models for risk rating, to be used as part of the short term planning process.

7.3.1 Example 1 – Proactive approach (municipality model) for planning goals

The Puerto Rico 2040 strategic plan has as long range objective (20 years) with the goal of a reduction of 10% in the total number of crashes for all municipalities including Mayaguez. The proportion of secondary highways in Mayaguez currently is 0.10, and the population is 100,000 habitants. It is forecasted that in 20 years, the proportion of secondary highways will be 0.12, the length of total highways will increase by 10% (400 to 440 kilometers), and the population will decrease to 93,000. The planning department in the municipality wants to determine how many crashes can be expected in Mayaguez in 20 years and what will be the expected crash reduction after the implementation of safety improvement strategies?. (Please note, the data used in this example are hypothetical only).

According to the results obtained in the municipality models, Mayaguez has an average risk of U=0, as shown in Table 7.5 which was obtained as the output of the models used. The equation to calculate the expected number of crashes is shown in Equation 7.1. The covariates for this equation are the proportion of secondary roads, the proportion of tertiary roads, and the value of U, which is associated to the municipalities as shown in Tables 7.2 through 7.8

$$\hat{E}(Y \mid U_i) = \exp\left(-10.85 + 12.93\,prop\sec + 4.64\,\overline{propter} + U_i\right) \quad (7\text{-}1)$$

Table 7.1 Example 1.

Municipality	Propsec	APKPPPY	Crash frequency
Mayaguez (now)	0.10	0.0001	5673
Mayaguez (after)*	0.12	0.0002	7516
Expected reduction			752
Expected number of crashes**			6764

APKPPPY -Crashes per kilometer per population per year

***without safety measures, **with safety measures.**

The same equation is used to calculate the number of crashes in 20 years without the implemented strategies. This is done in order to monitor the performance of the objectives. Table 7.1 shows the number of crashes expected with and without safety measures. The expected reduction will be 752 crashes and the expected number of crashes after applying the safety measures will be 6764 crashes.

7.3.2 Example 2- Proactive approach (municipality model) for risk rating

Unlike the first example, this example utilizes real crash data obtained from the office of Road Safety in Puerto Rico. The long range plan of the Road Safety Agency should have as one of the objectives to determine the municipalities with the highest risk to make an efficient investment of resources and including additional studies such as Road Safety Audits.

Equation 7.1 allows the calculation of the crash rate of Puerto Rico in 20 years by considering the change in the proportion of secondary roads and tertiary roads in the municipality. The U in the equation represents the risk associated to the municipality being analyzed that is obtained from the outputs calculated from the models. The values of the different percentiles are calculated by multiplying the standard deviation of the model by the representative z-value. Tables 7.2 to 7.9 presents the U values and the characteristics of each municipality by region. The U are indicators of risk. The U values between 95th and 75th percentile [2.31, 0.94], represent high risk and are highlighted in red. The U values between 75th and 25th percentile [0.94,-0.94], represent average risk and are highlighted in yellow. The U values between 25th and 5th [-0.94,-2.31,] represent low risk and are highlighted in green.

Table 7.2 Crash risk of municipalities in the Arecibo region.

Municipality	U	Secondary	Tertiary	Total	Population	km^2
Arecibo	-0.41	10	174	1152	100004	326
Utuado	-0.55	43	149	921	35045	113
Barceloneta	2.00	6	9	189	23407	82
Florida	-0.68	17	21	94	12624	39
Manati	0.46	16	67	476	45591	120
Ciales	-1.23	65	49	436	19667	172
V. Baja	0.53	8	91	520	61923	122
Morovis	0.15	31	66	432	31229	101
V. Alta	0.69	2	61	265	39020	72
Corozal	0.32	21	84	390	37356	110
Dorado	0.78	13	41	262	35794	60
T. Baja	0.54	5	35	371	93399	62
T. Alta	0.46	15	53	290	68345	72
Naranjito	-0.11	31	63	315	30306	72

Table 7.3 Crash risk of municipalities in the Aguadilla region.

Municipality	U	Secondary	Tertiary	Total	Population	km^2
Rincon	1.19	10	32	179	15068	37
Aguada	1.00	9	68	422	42561	78
Aguadilla	0.35	18	54	352	64193	95
Moca	0.81	10	94	527	40326	82
Isabela	0.85	4	88	483	45358	143
San Sebastian	0.08	44	98	645	44083	182
Quebradillas	1.10	3	47	51	25906	23
Camuy	0.87	12	64	468	35635	120
Lares	-0.07	35	110	590	33619	156
Hatillo	1.01	11	65	401	40381	110

Table 7.4 Crash risk of municipalities in the Humacao region.

Municipality	U	Secondary	Tertiary	Total	Population	km^2
Aibonito	0.05893	21	64	296	26621	82
Barranquitas	-0.5058	37	51	288	29798	86
Comerio	-1.0689	37	45	233	20550	72
Aguas Buenas	0.4579	28	40	323	29004	78
Cidra	-0.0951	39	42	337	43509	93
Cayey	-0.1844	18	76	508	48158	130
Salinas	0.7566	0	32	578	31469	180
Guayama	0.5403	0	77	655	45155	169
Arroyo	1.5676	0	23	177	19500	38
Patillas	-0.1195	32	39	367	20109	125

Table 7.5 Crash risk of municipalities in the Mayaguez region.

Municipality	U	Secondary	Tertiary	Total	Population	km^2
Añasco	0.2181	28	68	438	28928	92
Cabo Rojo	0.1733	26	94	639	48717	182
Guanica	0.6936	11	43	297	21344	96
Hormigueros	1.0303	6	26	129	16986	29
Las Marias	-0.4715	51	39	384	10808	121
Mayaguez	0.1418	41	75	635	96490	201
Maricao	-0.4547	26	55	251	6467	169
Lajas	0.1179	25	71	442	26407	158
San German	0.1448	21	104	523	36962	141
Sabana Grande	0.2455	15	57	308	25984	92
Yauco	0.788	0	112	645	45512	179

Table 7.6 Crash risk of municipalities in the San Juan region.

Municipality	U	Secondary	Tertiary	Total	Population	km^2
Catano	2.3369	0	16	93	29823	5
Bayamon	0.8791	4	102	946	221360	115
Guaynabo	0.9333	0	60	502	100495	70
San Juan	-2.039	0	61	1370	426660	124
Trujillo Alto	0.4	5	51	330	76370	55
Carolina	0.2024	9	64	747	184766	117
Loiza	0.551	23	7	158	32145	50
Canovanas	0.7616	15	48	360	45320	73

Table 7.7 Crash risk of municipalities in the Fajardo region.

Municipality	U	Secondary	Tertiary	Total	Population	km^2
Rio Grande	-5.0796	6	102	495	53602	157
Luquillo	-4.2431	0	51	234	20131	68
Fajardo	-4.6483	6	40	314	39944	77
Ceiba	-3.4439	0	32	281	16748	75
Naguabo	0.3109	11	60	439	25018	135
Humacao	0.635	2	60	474	59613	142
Las Piedras	0.6404	15	59	304	36302	35
Juncos	0.7886	13	47	272	38128	68
Gurabo	0.5214	15	62	332	40009	76
San Lorenzo	0.2861	29	79	575	41516	138
Yabucoa	0.4029	8	80	561	39214	142
Maunabo	-3.8085	0	31	188	12714	54
Caguas	0.4594	13	85	775	142927	152

Table 7.8 Crash risk of municipalities in the Ponce region.

Municipality	U	Secondary	Tertiary	Total	Population	km^2
Adjuntas	0.155	22	96	539	19485	173
Guayanilla	0.9069	5	59	399	22866	110
Penuelas	0.7355	16	46	405	26262	117
Ponce	-1.1014	52	122	1283	181971	297
Jayuya	0.0204	28	51	286	17326	102
Villalba	-0.215	36	58	346	27670	96
Juana Diaz	0.262	8	96	605	51236	61
Orocovis	-0.6826	68	63	572	23995	164
Coamo	-0.6185	42	96	555	39020	202
Santa Isabel	0.7299	13	29	380	22457	88

According to Tables 7.2 to 7.8, the municipalities of Rincon, Aguada, Quebradilla, Hatillo, Barceloneta, Cataño, and Hormigueros are the municipalities with the highest risk of crashes. The regional averages of these municipalities in terms of other contributing factor can be compared to the regional averages of other municipalities. Some of the contributing factors include the averages of consumption, of drugs, alcohol, population aging, and topography. Using all these information, recommendations can be made for engineering policies, education measures, medical emergency response management, enforcement strategies, and planning improvements, to effectively use the limited resources in those places with higher risk of crashes.

7.3.3 Example 3- Proactive approach (freeway and arterial model) for introducing safety as part of the strategic planning process

Hypothetical data was used in this example. Suppose that the long range planning process has recognized the need to construct a road between Fajardo and Mayaguez. Different alternatives of mobility have been identified and each one has a different level of safety associated to each alternative, as shown in Tables 7.9 though 7.15. The models M4 for freeways (Table 6.5), and M3 for arterials (Table 6.8) were employed to obtain the expected APKPY for each one of the alternatives. The covariates used in the model M4 are AADT, shoulder type, and regional location. In model M3, the covariates used are regional location, AADT, and signalized intersections by kilometer. These covariates can vary depending on the model chosen.

Alternative 1

Table 7.9 Freeway total crashes SPF without metal barrier (North area).

Mayaguez	Municipalities	AADT	Kilometers	APKPY	APY
Humacao	Fajardo-Rio Grande	55270	14	3	49
San juan	Rio Grande- Bayamon	290000	40	15	601
Arecibo	Bayamon- Arecibo	144000	66	6	421
Aguadillla	Arecibo-Aguadilla	39500	53	1	73
Mayaguez	Aguadilla-Mayaguez	47900	17	3	48
Total			**190**		**1191**

Table 7.10 Freeway fatal crashes SPF without metal barrier (North area).

Region	Municipalities	AADT	Kilometers	APKPY	APY
Humacao	Fajardo-Rio Grande	55270	14	0.18	3
San juan	Rio Grande- Bayamon	290000	40	0.46	18
Arecibo	Bayamon- Arecibo	144000	66	0.31	21
Aguadillla	Arecibo-Aguadilla	39500	53	0.06	3
Mayaguez	Aguadilla-Mayaguez	47900	17	0.16	3
Total			**190**		**47**

Alternative 2

Table 7.11 Freeway total crashes SPF without metal barrier (South area).

Region	Municipalities	AADT	Kilometers	APKPY	APY
Humacao	Fajardo-Yabucoa	55270	31	7	219
Guayama	Yabucoa-Salinas	76700	64	9	561
Ponce	Salinas-Yauco	81700	42	12	495
Mayaguez	Yauco-Mayaguez	47900	45	6	260
Total			**182**		**1535**

Table 7.12 Freeway fatal crashes SPF without metal barrier (South area).

Region	Municipalities	AADT	Kilometers	APKPY	APY
Humacao	Fajardo-Yabucoa	55270	31	0.18	6
Guayama	Yabucoa-Salinas	76700	64	0.26	17
Ponce	Salinas-Yauco	81700	42	0.12	5
Mayaguez	Yauco-Mayaguez	47900	45	0.16	7
Total			**182**		**34**

Alternative 3

Table 7.13 Arterial total crashes SPF without metal barrier (North area).

Region	Municipalities	AADT	Kilometers	APKPY	APY
Humacao	Fajardo-Rio Grande	88100	14	29	403
San juan	Rio Grande- Bayamon	126300	40	14	568
Arecibo	Bayamon- Arecibo	46000	66	3	205
Aguadillla	Arecibo-Aguadilla	66300	53	13	700
Mayaguez	Aguadilla-Mayaguez	88200	17	18	310
Total			**190**		**2186**

Table 7.14 Arterial fatal crashes SPF without metal barrier (North area).

Region	Municipalities	AADT	Kilometers	APKPY	APY
Humacao	Fajardo-Rio Grande	88100	14	0.35	5
San juan	Rio Grande- Bayamon	126300	40	0.21	8
Arecibo	Bayamon- Arecibo	46000	66	0.10	6
Aguadillla	Arecibo-Aguadilla	66300	53	0.23	12
Mayaguez	Aguadilla-Mayaguez	88200	17	0.48	8
Total			**190**		**40**

Alternative 4

Table 7.15 Arterial total crashes SPF without metal barrier (South area).

Region	Municipalities	AADT	Kilometers	APKPY	APY
Humacao	Fajardo-Yabucoa	88100	31	29	893
Guayama	Yabucoa-Salinas	88100	64	29	1844
Ponce	Salinas-Yauco	40000	42	8	323
Mayaguez	Yauco-Mayaguez	88200	45	18	812
Total			**182**		**3872**

Table 7.16 Arterial fatal crashes SPF without metal barrier (South area).

Region	Municipalities	AADT	Kilometers	APKPY	APY
Humacao	Fajardo-Yabucoa	88100	31	0	11
Guayama	Yabucoa-Salinas	88100	64	0	22
Ponce	Salinas-Yauco	40000	42	0	8
Mayaguez	Yauco-Mayaguez	88200	45	0	22
Total			**182**		**63**

This example shows how safety can be analyzed as part of the strategic planning process. For example alternative 1 represents a freeway going through the Northern Region of Puerto Rico, and process and has a total number of 1191 crashes per year, and 47 crash fatalities per year. The second alternative, which represents a freeway going through Southern Region, has a total number of 1535 crashes per year, and 34 crash fatalities per year. The third alternative, which represents an arterial going through the Northern Region, has a total annual number of crashes of 2186 and 40 crash fatalities per year, and alternative 4, arterials going through the Southern Region, has a total number of annual crashes of 3872 and 63 crash fatalities. Each alternative can be analyzed using a cost-benefit analysis in order to rate the safety between the different criteria in the evaluation for the best alternative.

7.3.4 Example 4 – Reactive approach (freeway and arterial model) for Risk Rating of Segments

The following example uses real crash data. The segments with highest risk for crashes can be identified as part of the short term planning process. The segments between the 95th and 75th percentile have high risk. For example a segment with an AADT OF 300,000 and a APKPY between [37.5-75] is considered an average risk segment as can be determined using figure 7.2. Different segments can be evaluated with varying crash rates in order to identify the magnitude of risk of the segments. The covariates will change depending on the model chosen for the segment evaluation. Additional studies can be developed of these segments to identify the problems and possible countermeasures. To identify the appropriate countermeasures for these high risk segments, it is necessary to perform a road safety audit. Afterwards, it is important to monitor the actual reduction in crashes due to the countermeasure chosen. The expected number of crashes before the improvement can then be compared to the observed number of crashes after the improvement.

CHAPTER 8: CONCLUSIONS

8.1 LOCAL PROCEDURES

- The Puerto Rico Department of Transportation and Public Works (PRHTA) and the Puerto Rico Highway and Transportation Authority (PRHTA) have developed the Puerto Rico Strategic Highway Safety Plan (SHSP) 2014-2018, so as to significantly reduce the number of fatalities and serious crash related injuries. The development of this plan is a federal requirement of 23 USC, Section 402, which requires the creation of a state highway safety program.

- During interviews with various experts in the urban planning field of Puerto Rico, the following were found to be true:

- Crash databases exits, but do not have the exact location of the crashes and omit relevant information necessary for crash data analysis.

- Transportation safety appears in the transportation planning process of the Commonweatlh of Puerto Rico, in the LRTP set of goals and objectives, and in their vision statement.

- There is no Safety Management System in place for planning and monitoring safety related to the transport of goods and services on the road.

- The Regulations office has a Road Safety Audit Division which is in the process of developing a HSIP that would include planning, evaluation and implementation of safety countermeasures. The Road Safety Audit Division has a methodology for the identification of black spot priorities and is similar to the one used in the State of Iowa.

8.2 CALIBRATION OF ARIZONA MODELS TO WESTERN REGION AND DEVELOPMENT OF WESTERN REGION MODELS

Due to the lack of fit of Arizona models to Puerto Rico's characteristics, as shown in chapter 4, a decision was made to find new explanatory variables and develop a model for the western region. The model required a set of independent variables using road network characteristics, socio-economic, demographics, and crash history. The conclusions and recommendations are:

- The variables interstate, municipal, tertiary roads and population density were found to be significant.

- The focus of macro- level models is prediction and association. The goal is to inform the analyst what effects a hypothetical planning scenario can have on crash numbers and severities. An example of a hypothetical planning scenario can be the increase in the number of roadway kilometer inside a region. The developed models have limitations and assumptions. An important assumption is that 'new' safety countermeasure investments are to be analyzed independently by some other model or research study.

- The appropriate uses of the developed models are for planning, prediction, or forecasting domains. The inappropriate use of the model would be in the traffic and safety engineering domains.

8.3 REGIONAL MODELS

- A pilot study was conducted in the beginning of this experimentation to verify how some variables would behave and to draw conclusions that would help in the creation of regional models for the whole island. For example, the pilot study revealed that population was not a good prediction variable due to the decrease in population in some municipalities.

- The pilot study results also found that the frequency models do not allow for a direct comparison between municipalities in determining which one is more dangerous. The number of crashes per municipality is only one factor among many. Therefore, information related to the number of kilometers in primary, secondary and tertiary roads was investigated to use them as explanatory variables. Information related to the population and the length of roads was used to "offset" the models, or, to standardize variables. The developed prediction models were used to obtain a base condition (current circumstances), and to determine what would happen under various planning scenarios in terms of crash rates

- When the planning process has reached its final stage, the zones at greater risk of accidents should be included within the transportation plan for the development of more detailed analysis and for the implementation of safety projects. Afterwards, the high risk zones should be analyzed as part of the planning process, to verify whether there is truly a decrease in the accident rate, and see if it was a result of the implemented measures, or if there is another possible explanation.

- The use of these models could provide planners with information concerning future road safety. This data could be achieved while assuming similar design characteristics that to those used now. As a result, planners could prepare new safety plans that would implement new system-wide safety initiatives in order to improve safety in the future and calculate the Crash Modification Factors (CMF's) accordingly. In this way, planners could estimate the amount of resources needed to meet the regional safety objectives.

- Regional expansion can affect the population growth, the number of the miles in the roads, and intersection density. A statewide or regional safety objective may present an X% reduction of total and fatal crashes as a goal, but this does not necessarily mean that there will be a crash reduction because population growth and other factors can still cause an increase in the number of crashes. The fatal crash model simply provides planners with a tool for setting targets and for meeting objectives and performance milestones.

8.4 FREEWAY AND ARTERIAL MODELS

- In the APKPY models for expressways, AADT, number of lanes, shoulder type and region were found to be significant variables; while in the arterial APKPY models, the variables that were found to be significant were the AADT, intersections by kilometer, and segment location or region. The AADT, shoulder type and lanes were the most important elements in predicting crashes on expressways. Meanwhile, AADT, region and signalized intersection by kilometer were found to be the most important in crash predictions on arterials.

- In the expressway model, as the AADT variable increased, the number of crashes per kilometer also increased. The expressway segments with a paved shoulder had a lower total crash rate. Arterial segments located in Arecibo have a lower total crash rate, while those located in Aguadilla had a higher total crash rate, keeping constant with the AADT independent variable. These results concurred with the statistical analysis by region shown in chapter 3.

- The values of β found by the maximum likelihood method describe the typical values of the relationship between the response and covariates, while keeping the effect of a subject (segment) constant. These models are called Specific Subject 'SS' or conditional models and can be employed in crash predictions for improving existing, specific sites, or for the creation of new infrastructures with specific characteristics by means of the EB method.

- In the long range planning level, the typical predictions can be changed into marginal means or Population Average, PA in a simple way for predictions or dependent variables. This characteristic allows for versatility and multiple applications for developed models.

- Alongside total crash models, others were developed in order to separate different crash severities. These models allow an economic analysis in order to evaluate the different alternatives planned for long, mid, and short term projects. Unlike the models developed for arterials in the HSM, these models require a low sample size, little information. In addition the use of a model for functional classification is needed in order to conduct an estimation at an aggregated or regional level.

- A factor is considered arbitrary if the levels studied can be treated as a random sample of a population of levels for the factor; this would mean that an associated probability exits. The random selection of the sample of the levels is done with the purpose of treating this as a representation of the population of effects for which the inference is be done. In this study the factor is the segment, and the levels are the different characteristics or their treatments.

- In the analysis of random segment crashes, these being representative of the entire island, the segments were incorporated as a random factor with the purpose of inferring the result of the segments for the whole island (different degrees of inference in the consideration of different experimental unit orders). When the modeling did not include these effects, the inference could only be done for modeled segment levels.

- The conditional model objectives are: estimating the typical values of the parameters (random effects = 0, in the center of its distribution) of f, studying how the parameters vary in the population, and studying if this variation is related to the characteristics of the individuals (covariates, etc). These objectives were obtained through inference around parameters β and D.

- The components of β described both typical values and the relationship between the response and the covariates; while maintaining the effect of the subject constant.

- As in many regression models, the main objective is to identify a parsimonious functional form so as to describe the relationship between the observations and the covariates. Characterizing the subject can be of interest (predicting the random effects), in order to predict the places according to their improvement potential in crash reduction.

- The Generalized Linear Mixed Models are of the Specific Subject (SS) type: the values relate the observations with a covariate for a given subject, not for the Population

Average. On the other hand, the parameters of a PA represent the relationship between the marginal average, $E(Y \mid u)$ and covariates.

- Both strategies can use non linear models like the ones in this work, through the SS approach is more common because it is richer. It is possible to study the marginal relation (PA) in a simpler way through a SS model, but it is impossible to study SS relationship from a PA model.

- When a prediction for an effect population (levels) for a given factor (segment) is needed in the modeling process, the segments at the time of the study should be treated as random variables.

- The functional classification planning models allow for the consideration of the possible consequences in the case of different infrastructure alternatives in regards to safety. This then enables the mitigation alternatives to be considered from an early stage. The models can be used to find the crash rate reduction as a result of an increase of public transportation.

- The planning models developed in this work can be used to develop macro crash modification factors at the network level. This can be used in conceptual design and to make recommendations for any long term transportation plan.

- Traffic growth can affect the number of required highways. Therefore, when considering crash reduction objectives at a regional level, it is necessary to take normal population growth into account when faced with an increase in the number of road crashes. If this is done, crash reduction objectives can be wholly achieved.

- In addition to the mentioned applications, the models could also be used using a more traditional or reactive approach. Existing infrastructures can be improved when this Empirical Bayes EB method is used.

8.5 FRAMEWORK FOR THE INCORPORATING OF SAFETY INTO LONG AND MID RANGE URBAN PLANNING

- This work found that the developed models are versatile. This means, that they can be used to assess the consequences of safety on different projects, to realize a risk analysis of the existing highways, or to implement on a new infrastructure and to conduct a comprehensive SMS.

8.6 GENERAL CONCLUSIONS

The main contributions of this research are the following:

- The development of SPF models for municipalities, freeways and arterials while taking into account the spatial effects and temporal correlation in collision data and including random effects.

- The development of improved SPF models through the inclusion of random effects, obtaining both conservative and realistic models. This model focuses on explaining part of the extra-variation by improving the mean function.

- The development of different tools and a framework for their incorporation into the strategic planning process.

- The development of arterial SPF models has been restricted to intersections or road segments modeling. This practice ignores the relationship between intersection safety levels on a regional road segment level or vice versa. The incorporation of spatial heterogeneity into the models developed in this thesis could be used to assess the safety of a network composed of both intersections and road segments.

- Preliminary analyses concerning this topic has provided some valuable insights into the relationship between collisions on intersections and road segments in a network.

- The GLMM network models with spatial heterogeneity were compared to GLM models, and the independent models, in terms of goodness-of-fit, inference, and the identification of high risk zones.

- Typically, the *n* segments under consideration in this study belong to mutually exclusive *K* regions. In such cases, an additional component of variability can be included in the model so as to allow for the possibility that different regions have different collision risks due to traffic, geometric and environmental conditions that vary among them.

- Therefore, in order to benefit from the advantages of using the random parameters approach, and avoid over-fitting, it is recommended to cluster the road entities into homogeneous groups (e.g., districts, municipalities, zones, etc.) and fitting a different regression curve for each group rather than for each individual site.

- The models allow for global predictions to be developed at a regional level, while keeping the same sensitivity to changes in geometry, location and safety features for segment safety.

CHAPTER 9: RECOMMENDATIONS

9.1 LOCAL PROCEDURES

- It is necessary to improve the current system and the data collection format in order to obtain a better data quality that would, in turn, produce better models.

- The development of average population models, marginal models, subject specific models or conditional models is necessary in order to fully complete the safety monitoring processes in Puerto Rico and be able to establish an all encompassing safety system.

9.2 CALIBRATION OF ARIZONA MODELS FOR THE WESTERN REGION AND THE DEVELOPMENT OF WESTERN REGION MODELS

- The modeling processes consists of developing the trial and error process and of knowing the contributing factor for crashes.

- The safety analysis of any project will provide expected crashes as a result. This should be considered as an explanatory statement regarding safety and not a defining statement about safety. It could represent the amount of risk expected by changes in the number of intersections, residential development, road mileage, and local population.

- Although many explanations are provided for models predictive variables, models are not used for explaining. They are used instead for the prediction of crash outcomes and their severity per municipality in hypothetical situations. This restriction is not to dissimilar from the restriction placed on travel demand models', whose primary purpose is to predict demand for roadway space of motor vehicles in hypothetical or future scenarios.

- The municipalities at higher risk are found by using this prediction. Afterwards, comparisons between the socio-economic and demographic data of other municipalities can be conducted so as to acquire a complete understanding of the vehicular accidents and of safety measure tendencies. It is also necessary to conduct additional studies and audits for the safety in major risk zones after these areas are identified.

- Crash data is modeled so as to predict any future safety problems that may surge in a hypothetical population growth scenario. This type of models is important in order to plan for specific countermeasures to avoid those crashes typical of regional conditions.

- The next step would be to examine design policies and safety investments so as to meet the regional safety goals.

- It is necessary to develop models that include public transportation and non motorized means as independent variables in order to capture the effect of these important aspects into highway safety.

- It is also necessary to develop a complete HSIP to monitor and evaluate the improvements of safety as a result of in infrastructure, enforcement, education, and medical emergencies through SPF. This complete program cannot be found in Puerto Rico at the moment.

- Transportation planning often focuses solely on infrastructure related solutions. A much broader perspective on how the planning process can affect the safety of the transportation system should include recommended policies, processes, studies, and budget priorities.

- The long range transportation plan must include topics such as safety education programs for motor vehicles , safety awareness, cyclist and pedestrians, work zones, education policies, elderly driver evaluations and mature driver education. It should also include engineering and operation topics such as traffic management safety audits of existing, rehabilitated and new roadways, traffic safety studies, and traffic safety measures in construction zones.

- Municipalities at higher risk can be analyzed in a detailed way through the comparison of average variables of other municipalities, that way providing a better insight of the high occurrence of crashes.

9.3 REGIONAL MODELS

- One possible way to analyze the impact of the incorporation of safety devices is to create accident risk maps for each alternative or scenario, and to determine what would happen with the implementation of mitigation measurements or safety improvements. Another important impact that can be studied is the economic analysis. This analysis considers the direct and indirect costs of accidents versus the required strategic investments.

9.4 FREEWAY AND ARTERIAL MODELS

- The developed models can be used to conduct economic studies while considering both the direct and indirect cost of accidents and the improvements to safety measures.

9.5 FRAMEWORK FOR THE INCORPORATION OF SAFETY INTO LONG AND MID RANGE URBAN PLANNING

- The incorporation of the models into a SMS can assist in monitoring the results of the safety improvements done on the different roads.

9.6 GENERAL RECOMMENDATIONS

- The resulting models can provide the planners with information about future safety data. Assuming of course, that the design standards will be similar to those used in the present. However, the planners can prepare new safety plans to put into practice, or suggest new initiatives at a network level. This will help to improve safety in the future and calculate the corresponding MCMF. This way, the scheduler can have a better idea of resources needed to achieve the regional safety objectives for the future.

REFERENCES

Abdel-Aty, M., Siddiqui, C., Huang, H., and Wang, X. (2011). "Integrating trip and roadway characteristics to manage safety in traffic analysis zones", *Transportation Research Record*, No. 2213, pp. 20-28.

Anastasopoulos, P.Ch., Mannering, F. L. (2009). "A note on modeling vehicle-accident frequencies with random parameter count models", *Accident Analysis and Prevention* Vol. 41, No.1, pp. 153-159.

Anastasopoulos, P.Ch., Mannering, F.L, Shankar, V.N., Haddock, J.E. (2012). " A study of factor affecting highway accident rates using the random-parameters tobit model", *Accident Analysis and Prevention* Vol. 45, pp. 628-633.

Agresti, A. (2003). *Categorical Data Analysis*, 2nd, Wiley Series in Probability and Statistics, Boca-Raton, FL.

Amoros, E., and Laumon, L. M. (2003). "Comparison of road crashes incidence and severity between some french counties", *Accident Analysis and Prevention*, Vol. 35, pp. 537-547.

Agüero-Valverde, J., and Jovanis, P.P. (2006). "Spatial analysis of fatal and injury crashes in Pennsylvania", *Accident Analysis and Prevention*, Vol. 38, pp. 618–625.

Bahar, G., Masliah, M., Mollett, C. (2003). "Integrated Safety Management Process", NCHRP REPORT 501, Transportation Research Board of the National Academies, Washington.

Caliendo, C., Guida, M., Parisi, A., (2007). "A crash- prediction model for multilane roads, *Accident Analysis and Prevention,* Vol. 39, No.4, pp. 657-670.

Cambridge Systematics Inc., (2008). “Crashes vs. Congestion: What’s the Cost to Society?”, Presented to the Automobile Association of America, March.

Christensen, R. (1997). *Log-linear models and logistic regression*. Springer Texts in Statistics, 2nd ed., New York: Springer-Verlag.

Daniels, S., Brijs, T., Nuyts, E., Wets, G. (2010). “Explaining variation in safety performance of roundabouts, *Accident Analysis and Prevention,* Vol. 42, No. 2, pp. 393-402.

Demidenko, E. (2004). *Mixed Models,* Theory and Applications. Wiley.

Depue, L. (2003). “Safety Management Systems A Synthesis of Highway Practice”, NCHRP SYNTHESIS 322, Transportation Research Board Of the National Academies, Washington.

De Guevara, F. L., Washington, S. P., and Oh, J. (2004). “Forecasting crashes at the planning Level: simultaneous negative binomial crash model applied in Tuckson Arizona”, *Transportation Research Record*, No. 1987, pp. 191-199

de Leur, P. and Sayed, T. (2001). “Using claims prediction model for road safety evaluation”, *Canadian Journal Civil Engineering*, Vol. 28, No. 5, pp. 804–812.

El- Basyouny, K., Sayed, T. (2006). “Comparison of two negative binomial regression techniques in developing accident prediction models”, *Transportation Research Record* , No. 1950, pp. 9-16.

El- Basyouny, K., Sayed, T. (2009). “Accident prediction models with random corridor parameters”, *Accident Analysis and Prevention,* Vol. 41, pp. 1118-1123.

El- Basyouny, K., Sayed, T. (2010). “Accident prediction models with random corridor parameters”, *Accident Analysis and Prevention*, Vol. 42, pp. 131-139.

El-Basyouny, K. and Sayed, T. (2009). “Collision prediction models using multivariate Poisson-lognormal regression”, *Accident Analysis and Prevention*, Vol. 41, pp. 820-828.

El Basyouny, K. and Sayed, T. (2009.). “Accident prediction models with random corridor parameters, *Accident Analysis and Prevention*, Vol. 41, pp. 1118-1123.

El-Basyouny, K. and Sayed, T. (2009). “Urban arterial accident prediction models with spatial effects. *Transportation Research Record*, No. 2102, pp. 27-33.

El-Basyouny, K. and Sayed, T. (2010). “A method to account for outliers in the development of accident prediction models”. *Accident Analysis and Prevention*, Vol. 42 , pp.1266–1272.

El-Basyouny, K. and Sayed, T. (2010). “A full Bayes approach to before-after safety evaluation with matched comparisons: a case study of stop-sign in-fill program”, *Transportation Research Record*, No. 2148, pp. 1–8.

El-Basyouny, K. and Sayed, T. (2011). “A multivariate intervention model with random parameters among matched pairs for before-after safety evaluation, *Accident Analysis and Prevention*, Vol. 43, pp. 87–94.

Guo, F., Wang. X., Abdel-Aty, M. (2010). “Modeling signalized intersection safety with corridor spatial correlations. *Accident Analysis and Prevention,* Vol. 42, No. 1, pp. 84-92.

Hadayeghi, A., Shalaby, A.S., and Persaud, B.N. (2007). “Safety prediction models proactive tool for safety evaluation in urban transportation planning applications”, *Transportation Research Record*, No. 2019, pp.225-236.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., Cheung, C. (2006). “Temporal transferability and updating of zonal level accident prediction models”, *Accident Analysis and Prevention*, Vol. 38, pp. 579–589.

Hauer, E., (1997). “Observational Before and After Studies In Road Safety”, Pergamon, Oxford, UK.

Hilbe, J.M. (2007). “Negative Binomial Regression”, Cambridge University Press.

Highway Safety Manual. American Association of State and Highway Transportation Officials (AASHTO), Washington, D.C., 2010.

Highway Performance Monitoring System Manual. Federal Highway Administration, Washington D.C., 2014.

Huang, H., Abdel- Aty, M.A., Darwiche, A. L. (2010). “County- level crash risk analysis in Florida: Bayesian spatial modeling”, *Transportation Research Record*, No. 2148, pp. 27-37.

Jones, B., Janssen, L., Mannering, F.L. (1991). “Analysis of frequency and duration of freeway accidents in Seattle, *Accident Analysis and Prevention,* Vol. 23, No. 2, pp. 239-255.

Jovanis , P., Chang, H. (1989). “Disaggregate model of highway accident occurrence using survival theory”, *Accident Analysis and Prevention,* Vol. 21, No. 5, pp. 445-458.

Kim, D., Washignton, S.P. (2006). “The significance of endogeneity problems in crash models: an examination of left turn lanes in intersection crash models”, *Accident Analysis and Prevention,* Vol. 38, No. 6, pp. 1094-1100.

Kononov, J., Allery, B., Znamenacek, Z. (2003). “Safety planning study of urban freeways”, *Transportation Research Record,* No. 2019, pp 146-155.

Lee, J., Mannering, F.L. (2002). “Impact of roadside features on the frequency and severity of run-Off-roadway accidents: an empirical analysis”, *Accident Analysis and Prevention*, Vol. 34, No. 2, pp. 149-161.

Li, X., Lord, D., Zhang, Y. (2011). “Development of accident modification factors for rural frontage road segments in Texas using results from generalized additive models”, *Journal of Transportation Engineering*, Vol. 137, No. 1, pp. 74-83.

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., Schabenberger, O. (2006), SAS System for Mixed Models, 2nd.ed. SAS Institute Inc.

Lord, D., Washington, S. P., Ivan, J.N. (2005). “Poisson, Poisson-Gamma and Zero inflated regression models of vehicle crashes: a Bayesian perspective”, *Safety Science*, Vol. 46, No. 5, pp. 751-770.

Lord, D., (2006). “Modeling motor vehicles crashes using Poisson-Gamma models: examining the effects of low sample mean values and small mean values and small sample size on the estimation of the fixed dispersion parameter”, *Accident Analysis and Prevention*, Vol. 38, No. 4, pp. 751- 766.

Lord, D. Washington, S. P., Ivan, J.N. (2007). “Further notes on the application of zero inflated models in highway safety”, *Accident Analysis and Prevention*, Vol. 39, No. 1, pp. 53-57.

Lord, D., Guikema, S., Geedipally, S.R. (2008). “Application of conway maxwell- Poisson generalized linear model for analyzing motor vehicle crashes”, *Accident Analysis and Prevention,* Vol. 40, No. 3, pp. 1123-1134.

Lord, D., Mannering, F. (2010). “The Statistical Analysis of Crash-Frequency data: a review of methodological alternatives”, *Transportation Research Part A* 44, 291-305.

Lord, D., Miranda-Moreno, L.F. (2008). “Effects of low sample mean values and small sample size on the estimation of fixed dispersion parameter of Poisson Gamma models for modeling motor vehicle crashes: a Bayesian perspective”, *Safety Science*, Vol. 46, No. 5, pp. 751-770.

Lovegrove, G.R. (2006). “Community-based, macro-level collision prediction models”. Ph.D. Dissertation. Department of Civil Engineering, University of British Columbia, Vancouver, B.C., Canada.

Lovegrone, G.R., and Sayed, T. (2007). “Macrolevel collision predicition models to enhance traditional reactive road safety improvement programs”, *Transportation Research Record*, No. 2019, pp.65-73.

Malyshkina, N., Mannering, F.L., Tarko, A.P. (2009). “Markov switching negative binomial models: an application to vehicle accident frequencies”, *Accident Analysis and Prevention*, Vol. 41, No. 2, pp. 217-226.

Malyshkina, N., Mannering, F.L. (2010a). “Empirical assessment of the impact of highway design exceptions on the frequency and severity of vehicle accidents”, *Accident Analysis and Prevention,* Vol. 42, No. 1, pp. 131-139.

Malyshkina, N., Mannering, F.L., (2010b). “Zero state Markov switching count- data models: an empirical assessment”, *Accident Analysis and Prevention,* Vol. 42, No. 1, pp. 122-130.

Miaou, S.-P., Lum, H. (1993). “Modeling vehicle accidents and highway geometric design relationships”, *Accident Analysis and Prevention*, Vol. 25, No. 6, pp. 689-709.

Miaou, S.-P., (1994). “The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions”, *Accident Analysis and Prevention,* Vol. 40, No. 1, pp. 260-266.

McCullagh, P.; Nelder, J. (1989). *Generalized Linear Models*, 2nd ed., Chapman and Hall, London.

Miller, J.S., Garber, N.J., and Kamatu, J.N. (2011). "Resource guide for enhancing the incorporation of safety into the regional planning process", *Transportation Research Record*, No. 2244, pp. 50-60.

Naderan, A., and Shani, J. (2010). "Crash generation models forecasting crashes in urban areas", *Transportation Research Record*, No. 2148, pp.1-10.

Noland, R. B., and Oh, L. (2004). "The effect of infrastructure and demographic change on traffic- related fatalities and crashes: a case study of Illinois county- level data", *Accident Analysis and Prevention*, Vol. 36, pp. 525-532.

Noland, R.B., and Quddus, M.A.A. (2004). "Spatially disaggregate analysis of road casualties in England", *Accident Analysis and Prevention*, Vol. 36, pp. 973-984.

Naderan, A., and Shani, J. (2010). "Aggregate crash prediction moldels: Introducing crash generation concept", *Accident Analysis and Prevention*, Vol. 42, pp.339–346.

Oh, J., Washington, S.P., Nam, D. (2006). "Accident prediction model for railway-highway interfaces", *Accident Analysis and Prevention*, Vol. 38, No. 2, pp. 346-356.

Park, B.-J., Lord, D. (2009). "Application of finite mixture models for vehicle crash data analysis", *Accident Analysis and Prevention*, Vol. 41, No. 4, pp. 683-691.

Park, B.-J., Lord, D., Hart, J.D. (2010). "Bias properties of Bayesian statistics in finite mixture of negative regression models for crash analysis', *Accident Analysis and Prevention*, Vol. 42, No. 2, pp. 741-749.

Poch, M., Mannering, F.L. (1996). "Negative binomial analysis of intersection-accident frequencies", *Journal of Transportation Engineering*, Vol. 122, No. 2, pp. 105-113.

Quddus, M.A. (2008). “Time series count data models: an empirical application to traffic accidents”, *Accident Analysis and Prevention*, Vol. 40, pp. 1732-1741.

SafetyAnalyst: Software Tools for Safety Management of Specific Highway Sites, Task , SafetyAnalyst for Module 1 – Network Screening, FHWA, 2002.

Strategic Highway Safety Plan. Puerto Rico Safety Commission. San Juan, P.R, 2014.

SAS Institute Inc. (2013). SAS 9.3., Help and Documentation, Cary, NC: SAS Institute Inc.

Sellers, K.F., Shmueli, G. (2010). “A flexible regression model for count data”, *Annals of Applied Statistics*, Vol. 4, No. 2, pp. 943-961.

Shankar, V., Mannering, F.L., Barfield, W. (1995). “Effect of roadway geometrics and environmental factors on rural accident frequencies”, *Accident Analysis and Prevention,* Vol. 27, No. 3, pp. 371-389.

Shankar, V., Milton, J., Mannering, F.L. (1997). “Modeling accident frequencies as zero altered probability processes: an empirical inquiry”, *Accident Analysis and Prevention*, Vol. 29, No. 6, pp. 829-837.

Shankar, V., Albin R., Milton, J., Mannering, F.L. (1998). “Evaluating median cross-over likelihoods with clustered accident counts: an empirical inquiry using random effects negative binomial model”, *Transportation Research Record*, No. 1635, pp. 44-48.

Sittikariya, S., Shankar, V. (2009). “Modeling heterogeneity: traffic accidents”, VDM-Verlag, 80pp.

Tarko, A. P., and M. Kanodia. (2004). “Hazard elimination program - manual on improving safety of Indiana road intersections and sections”, Joint Transportation

Research Program, Indiana Department of Transportation and Purdue University, West Lafayette, Indiana.

Wang, X., Abdel- Aty, M. (1996). “Temporal and spatial analyses of rear- end crashes at signalized intersections”, *Accident Analysis and Prevention,* Vol. 38, No. 6, pp. 1137-1150.

Lord, D., Mahlawat, M., M. (2009). “Examining the application of aggregated and disaggregated Poisson-Gamma models subject to low sample mean bias. *Transportation Research Record*, No. 2136, pp. 1-10.

Ulfarsson, G., Shankar, V. (2003). “An accident count model based on multi- year cross sectional roadway data with serial correlation”, *Transportation Research Record*, No. 1840, pp. 193-197.

Wier, M., Weintraub, J., Humphreys, E. H., Seto, E., and Bhatia, R. (2009). “An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning”, *Accident Analysis and Prevention*, Vol. 41, pp. 137–145.

Xie, Y., Zhang, Y. (2008). “Crash frequency analysis with generalized additive models”. *Transportation Research Record*, No. 2061, pp. 39-45.

APPENDICES

TABLE OF CONTENTS

APPENDICES 139
1. APPENDIX CHAPTER 4 143
1.1 LOG-NORMAL TOTAL CRASHES MODEL 145
1.2 LOG LINEAR NEGATIVE BINOMIAL TOTAL CRASH MODEL 150
1.3 LOG NORMAL FATAL CRASHES MODEL 153
1.4 LOG NORMAL INJURY CRASHES MODEL 158
1.5 LOG LINEAR NEGATIVE BINOMAIL INJURY CRASHES MODEL 163
2. APPENDIX CHAPTER 5 166
2.1 CHAPTER 5-MODEL 1 167
2.1.1 TOTAL CRASHES-M1 (GLM) 168
2.1.2 PDO-M1 171
2.1.3 INJURY-M1 174
2.1.4 FATAL-M1 177
2.2 CHAPTER 5-MODEL 2 180
2.2.1 TOTAL CRASHES-M2 181
2.2.2 PDO-M2 185
2.2.3 INJURY-M2 189
2.2.4 FATAL-M2 193
2.3 CHAPTER 5-MODEL 3 197
2.3.1 TOTAL CRASHES-M3 198
2.3.2 PDO-M3 201
2.3.3 INJURY-M3 204
2.3.4 FATAL-M3 207
2.4 CHAPTER 5-MODEL 4 210
2.4.1TOTAL CRASHES-M4 211
2.4.2 PDO-M4 214
2.4.3 INJURY-M4 217
2.4.4 FATAL-M4 220
3. APPENDIX CHAPTER 6 223
3.1 CHAPTER 6 FREEWAY MODEL 1 240
3.1.1 TOTAL CRASHES-M1 241

3.1.2 INJURY-M1 245
3.1.3 FATAL-M1 249
3.2 CHAPTER 6 FREEWAY MODEL 2 253
3.2.1 TOTAL CRASHSE-M2 254
3.2.2 INJURY –M2 257
3.2.3 FATAL-M2 260
3.3 CHAPTER 6 FREEWAY MODEL 3 263
3.3.1 TOTAL CRASHES-M3 264
3.3.2 INJURY-M3 268
3.3.3 FATAL-M3 272
3.4 CHAPTER 6 FREEWAY MODEL 4 276
3.4.1 TOTAL CRASHES-M4 277
3.4.2 INJURY-M4 281
3.4.3 FATAL-M4 285
3.5 CHAPTER 6 FREEWAY MODEL 5 289
3.5.1 TOTAL CRASHES-M5 290
3.5.2 INJURY-M5 294
3.5.3 FATAL-M5 298
3.6 CHAPTER 6 FREEWAY MODEL 6 302
3.6.1 TOTAL CRASHES-M6 303
3.6.2 INJURY-M6 306
3.6.3 FATAL-M6 309
3.7 CHAPTER 6 ARTERIAL MODEL 1 312
3.7.1 TOTAL CRASHES-M1 313
3.7.2 INJURY-M1 317
3.7.3 FATAL-M1 321
3.8 CHAPTER 6 ARTERAL MODEL 2 325
3.8.1 TOTAL CRASHES-M2 326
3.8.2 INJURY-M2 330
3.8.3 FATAL-M2 334
3.9 CHAPTER 6 ARTERIAL MODEL 3 338
3.9.1 TOTAL CRASHES-M3 339

3.9.2 INJURY-M3 343
3.9.3 FATAL-M3 347
3.10 CHAPTER 6 ARTERIAL MODEL 4 352
3.10.1 TOTAL CRASHES-M4 353
3.10.2 INJURY-M4 357
3.10.3 FATAL-M4 360
4. SCRIPT SAS FREEWAY AND ARTERIAL MODELS 363

1. APPENDIX CHAPTER 4

Table. Data used for Puerto Rico western region models (2002 year)

Municipality	Highway_miles	Population_thousand	Acres	POP_PAC	Intersections	Int_mile	Intestates	Freeways	Pop_young_thousand	POP_drive_thousand	Total_crashes	Fatal_crashes	Injury_crashes	Primary	Secundary	Tertiary	Municipal	Millas_acre
Aguada	262.49	42.46	19795	2.14	32	0.12	0.03	0.00	9.98	28.49	2286	7	361	5.55	5.67	42.33	208.94	0.013
Aguadilla	219.06	64.41	23417	2.75	20	0.09	0.09	0.07	14.09	42.41	3243	12	592	15.64	10.94	33.49	158.82	0.009
Añasco	272.07	28.76	25132	1.14	23	0.08	0.08	0.02	6.58	18.98	1206	5	197	3.78	17.53	42.40	208.35	0.011
Cabo_rojo	397.36	48.19	45023	1.07	21	0.05	0.05	0.00	10.24	30.97	1703	5	262	16.15	0.01	58.11	323.09	0.009
Guanica	187.39	21.54	23750	0.91	18	0.10	0.04	0.00	5.19	13.76	665	0	117	4.43	7.00	27.03	146.06	0.008
Hormigueros	300.36	16.88	7244	2.33	11	0.04	0.03	0.00	3.25	11.03	820	4	130	5.69	3.74	16.45	54.23	0.041
Isabela	80.11	45.11	35430	1.27	28	0.35	0.17	0.00	10.16	29.41	2082	7	397	6.92	2.74	54.50	236.20	0.002
Lajas	274.53	26.39	33463	0.79	27	0.10	0.03	0.00	5.82	17.00	1032	2	122	0.00	15.33	44.00	215.21	0.008
Las_Marias	238.42	10.89	23206	0.47	18	0.08	0.00	0.00	2.63	7.14	370	0	77	0.00	31.55	24.29	182.57	0.010
Maricao	156.12	6.46	23442	0.28	12	0.08	0.00	0.00	1.61	4.22	174	0	23	0.00	16.36	33.94	105.82	0.007
Mayaguez	394.58	97.22	49683	1.96	40	0.10	0.05	0.00	18.78	64.65	6282	9	727	21.71	25.59	46.89	300.39	0.008
Moca	327.68	40.18	32185	1.25	30	0.09	0.00	0.00	9.87	26.56	1979	7	317	0.00	6.48	58.39	262.80	0.010
Rincon	110.94	14.98	9145	1.64	7	0.06	0.00	0.00	3.19	9.78	621	1	130	0.00	6.13	19.89	84.92	0.012
Sabana_Grande	191.34	25.98	22969	1.13	14	0.07	0.04	0.00	6.01	16.75	862	2	155	4.88	9.58	35.58	141.30	0.008
San_German	324.97	37.03	34886	1.06	35	0.11	0.06	0.00	7.87	24.00	1670	5	227	7.64	12.78	64.67	239.89	0.009

1.1 LOG-NORMAL TOTAL CRASHES MODEL

4.5.6.1 Models for total crashes

Log- normal model Eq(4.8 and 4.9)

The REG Procedure
Model: MODEL1
Dependent Variable: Total_crashes

Number of Observations Read	16
Number of Observations Used	15
Number of Observations with Missing Values	1

Analysis of Variance

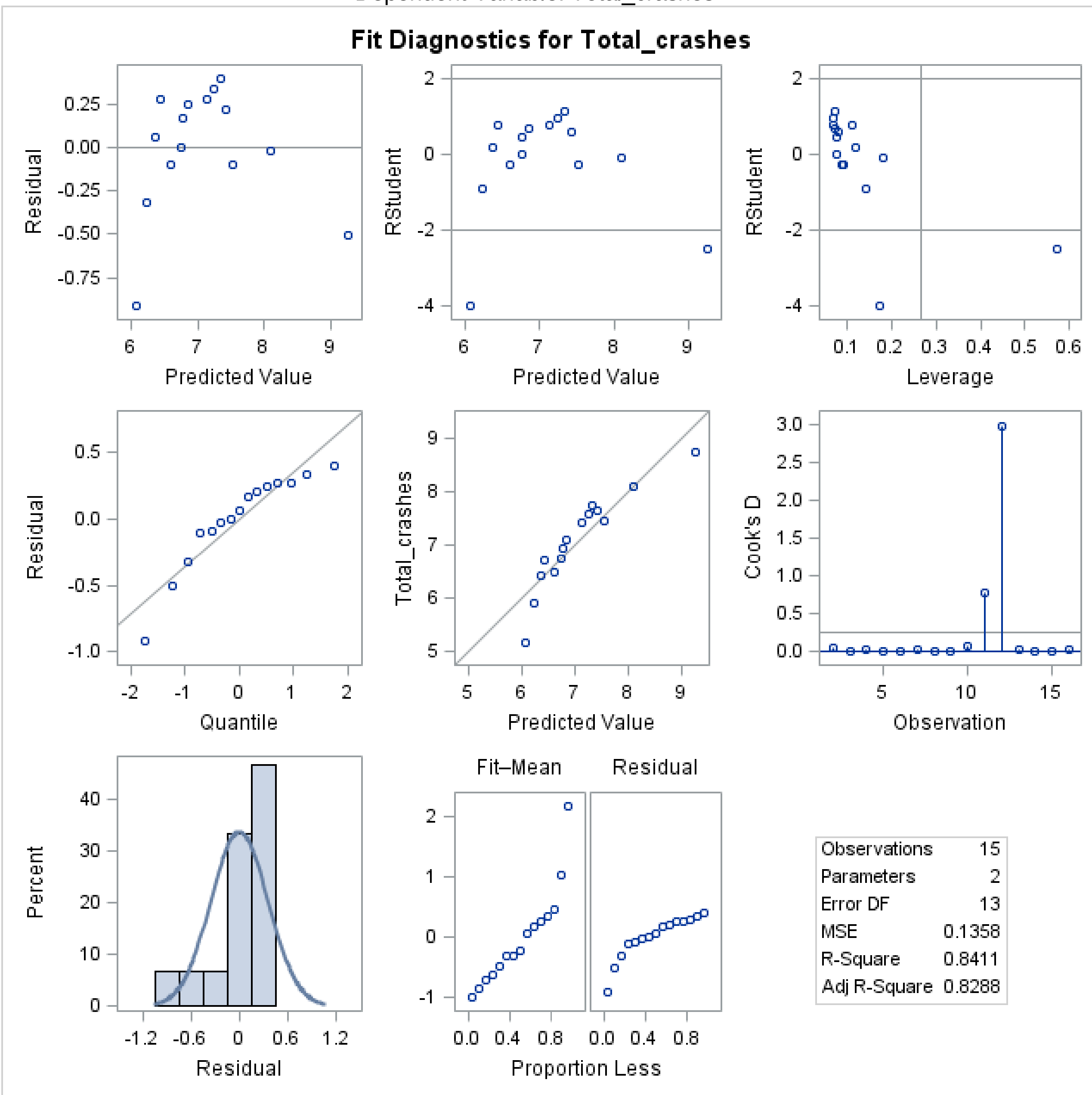
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	9.33887	9.33887	68.79	<.0001
Error	13	1.76483	0.13576		
Corrected Total	14	11.10369			

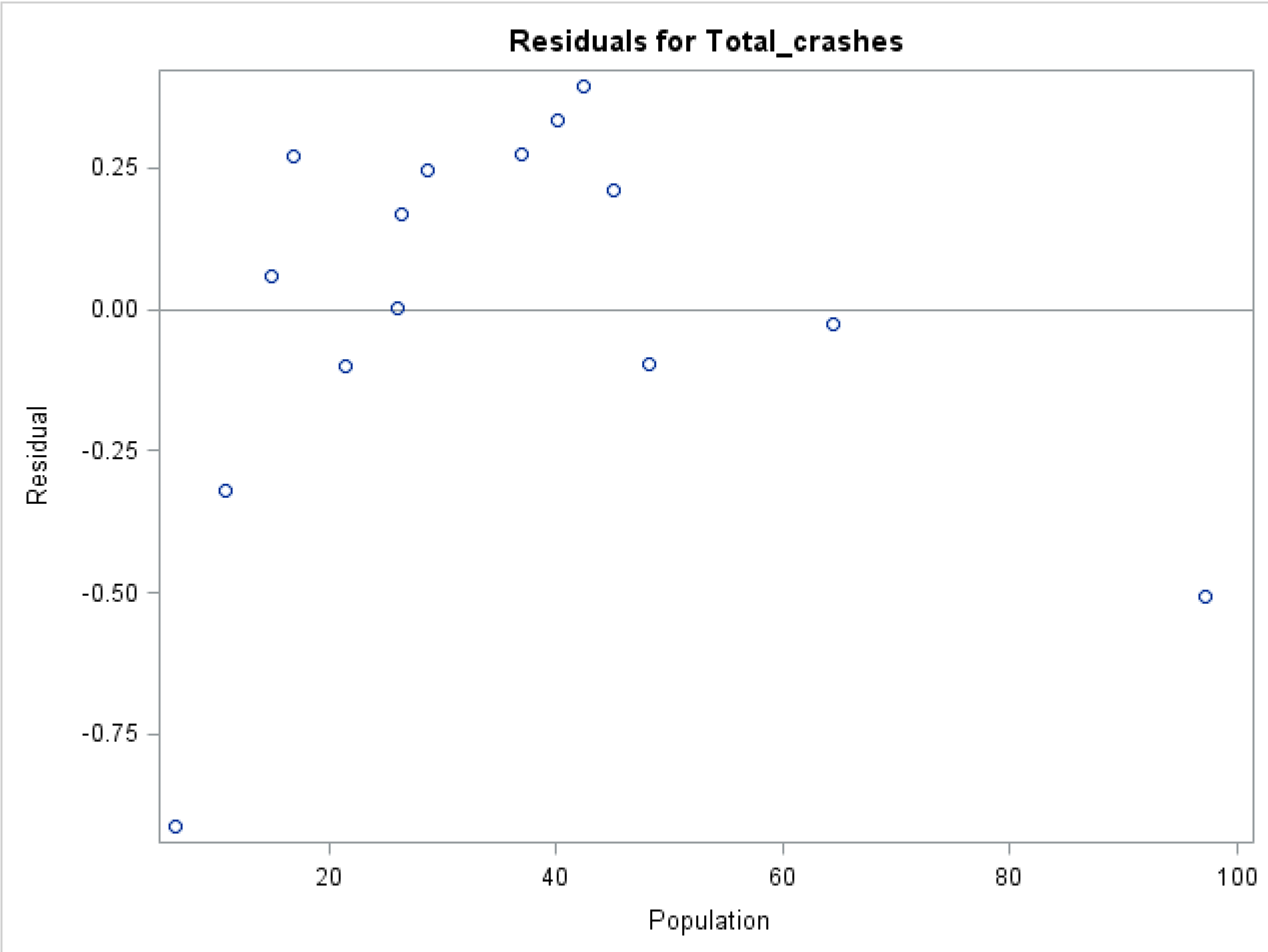
Root MSE	0.36845	R-Square	0.8411
Dependent Mean	7.07733	Adj R-Sq	0.8288
Coeff Var	5.20606		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	5.84693	0.17623	33.18	<.0001
Population	1	0.03506	0.00423	8.29	<.0001

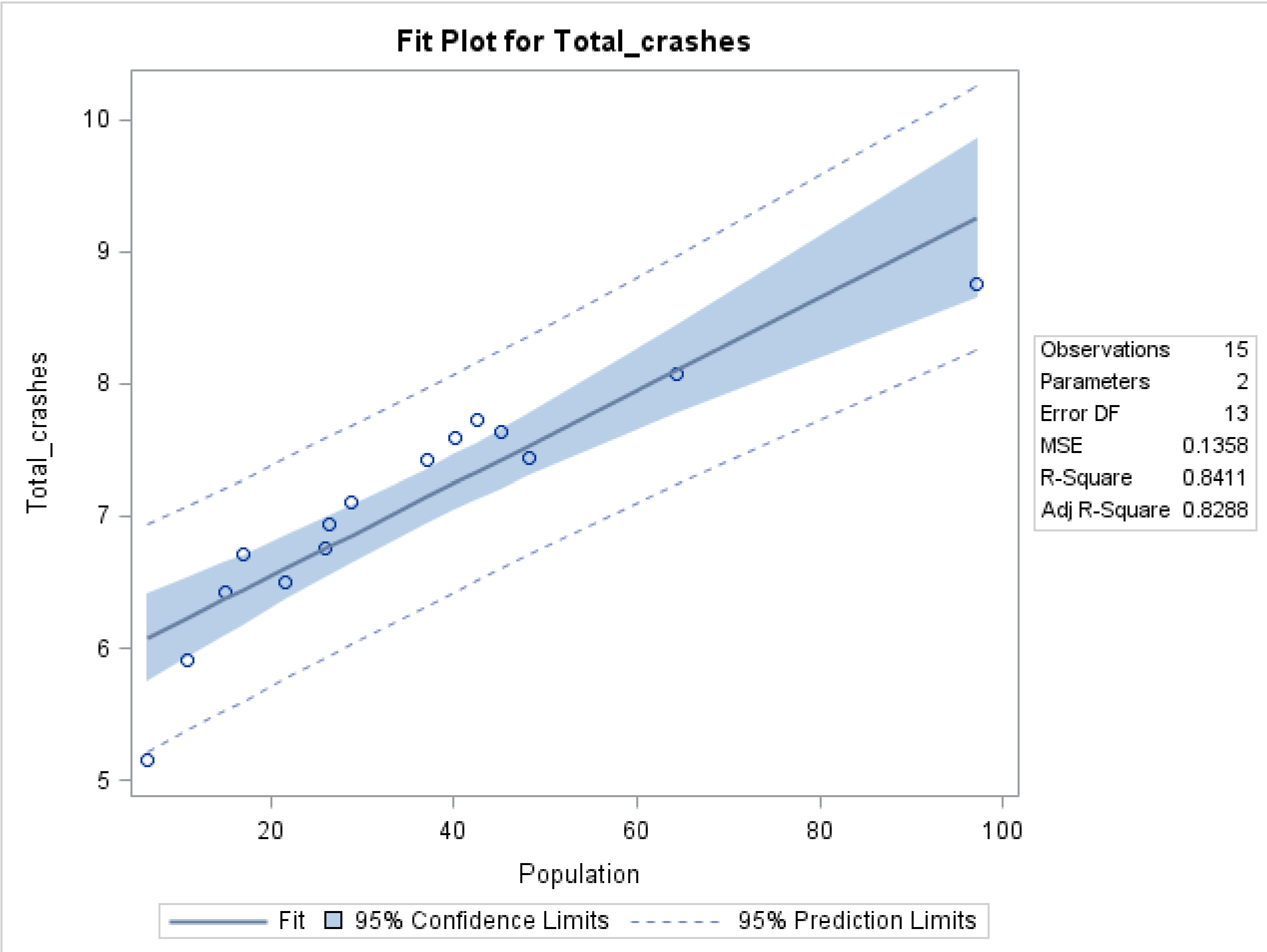
The REG Procedure
Model: MODEL1
Dependent Variable: Total_crashes





Residuals for Total_crashes
Residual
0.25
0.00
-0.25
-0.50
-0.75
20
40
60
80
100
Population

The GLIMMIX Procedure



1.2 LOG LINEAR NEGATIVE BINOMIAL TOTAL CRASH MODEL

Log linear- negative binomial model Eq. (4.10 and 4.11)

The GLIMMIX Procedure

Model 1 Information	
Data Set	WORK.EISENHOWERTOTCRASH
Response Variable	Total_crashes
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Variance Matrix	Diagonal
Estimation Technique	Maximum Likelihood
Degrees of Freedom Method	Residual

Number of Observations Read	15
Number of Observations Used	15

Dimensions	
Covariance Parameters	1
Columns in X	4
Columns in Z	0
Subjects (Blocks in V)	1
Max Obs per Subject	15

Optimization Information	
Optimization Technique	Newton-Raphson
Parameters in Optimization	5
Lower Boundaries	1
Upper Boundaries	0
Fixed Effects	Not Profiled

Iteration History					
Iteration	Restarts	Evaluations	Objective Function	Change	Max Gradient
0	0	4	124.26343235	.	6.25745
1	0	22	113.17329085	11.09014150	4.734797
2	0	3	113.12241859	0.05087226	2.439712
3	0	3	113.085962	0.03645659	0.545013
4	0	3	113.06165824	0.02430376	0.221281
5	0	3	113.05619386	0.00546438	0.034789
6	0	3	113.05602263	0.00017123	0.003138
7	0	3	113.05602093	0.00000170	0.000154
8	0	3	113.05602092	0.00000000	4.014E-6

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	226.11
AIC (smaller is better)	236.11
AICC (smaller is better)	242.78
BIC (smaller is better)	239.65
CAIC (smaller is better)	244.65
HQIC (smaller is better)	236.07
Pearson Chi-Square	12.62
Pearson Chi-Square / DF	1.15

Parameter Estimates					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	4.5972	0.3515	11	13.08	<.0001
Highway_miles	0.005098	0.001079	11	4.73	0.0006
POP_PAC	0.6432	0.1682	11	3.82	0.0028
Intestates	9.3160	2.2690	11	4.11	0.0017

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Highway_miles	1	11	22.33	0.0006
POP_PAC	1	11	14.63	0.0028
Intestates	1	11	16.86	0.0017

1.3 LOG NORMAL FATAL CRASHES MODEL

4.5.6.2 Models for fatal crashes

Log-normal model Eq (4.12 and 4.13)

The REG Procedure
Model: MODEL1
Dependent Variable: Fatal_crashes

Number of Observations Read	16
Number of Observations Used	15
Number of Observations with Missing Values	1

Stepwise Selection: Step 1

Variable Municipal Entered: R-Square = 0.3068 and C(p) = 2.0000

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.00858	0.00858	5.75	0.0322
Error	13	0.01939	0.00149		
Corrected Total	14	0.02797			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	0.01113	0.01717	0.00062592	0.42	0.5284
Municipal	0.02775	0.01157	0.00858	5.75	0.0322

Bounds on condition number: 1, 1

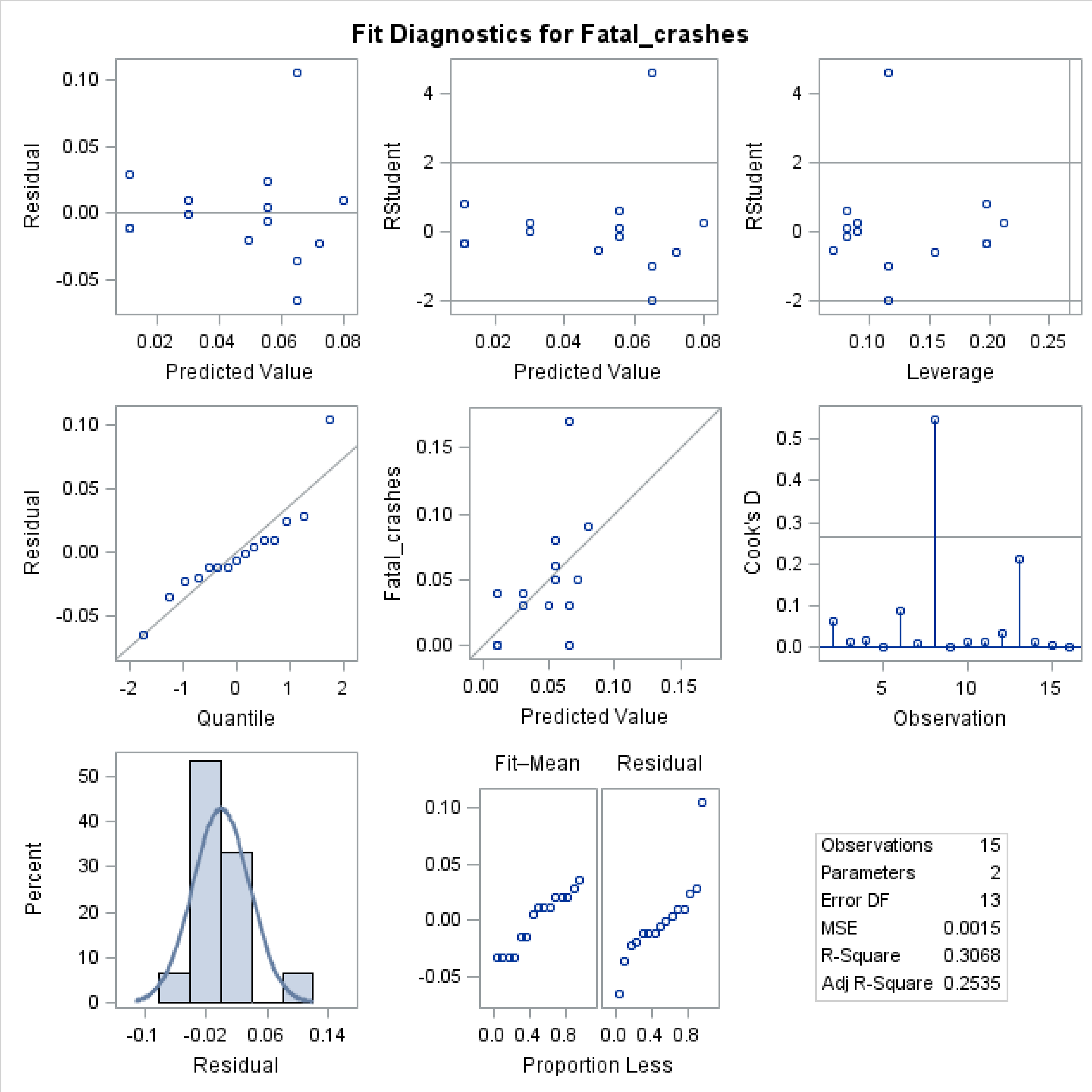
The GLIMMIX Procedure

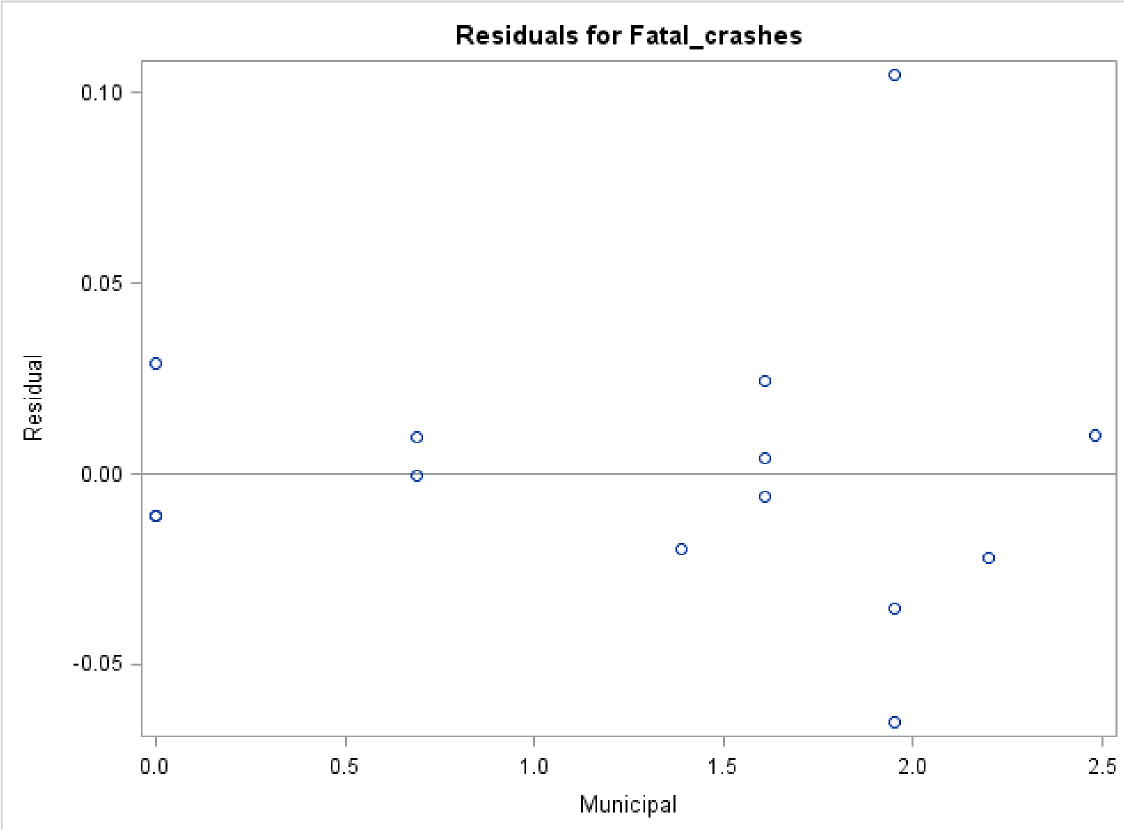
All variables left in the model are significant at the 0.1500 level.

All variables have been entered into the model.

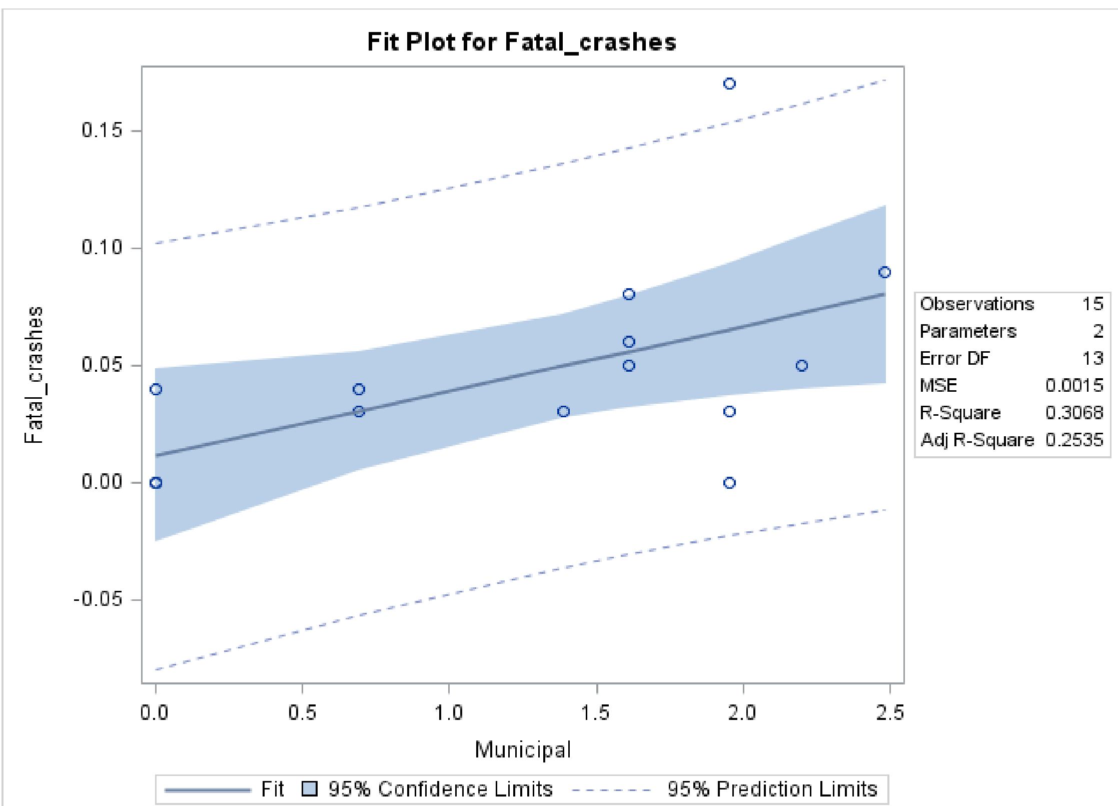
Summary of Stepwise Selection								
Step	Variable Entered	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	Municipal		1	0.3068	0.3068	2.0000	5.75	0.0322

The REG Procedure
Model: MODEL1
Dependent Variable: Fatal_crashes


Fit Diagnostics for Fatal_crashes
Residual
Predicted Value
RStudent
Predicted Value
RStudent
Leverage
Residual
Quantile
Fatal_crashes
Predicted Value
Cook's D
Observation
Percent
Residual
Fit–Mean
Residual
Proportion Less
Observations 15
Parameters 2
Error DF 13
MSE 0.0015
R-Square 0.3068
Adj R-Square 0.2535



Residuals for Fatal_crashes
Residual
0.10
0.05
0.00
-0.05
0.0
0.5
1.0
1.5
2.0
2.5
Municipal



Fit Plot for Fatal_crashes
Fatal_crashes
0.15
0.10
0.05
0.00
-0.05
0.0
0.5
1.0
1.5
2.0
2.5
Municipal
Observations 15
Parameters 2
Error DF 13
MSE 0.0015
R-Square 0.3068
Adj R-Square 0.2535
Fit
95% Confidence Limits
95% Prediction Limits

1.4 LOG NORMAL INJURY CRASHES MODEL

4.5.6.3 Models for injury crashes
Log-normal model Eq (4.14 and 4.15)
The REG Procedure
Model: MODEL1
Dependent Variable: Injury_crashes

Number of Observations Read	15
Number of Observations Used	15

Stepwise Selection: Step 1

Variable POP_PAC Entered: R-Square = 0.4697 and C(p) = 22.0122

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	4.92280	4.92280	11.51	0.0048
Error	13	5.55791	0.42753		
Corrected Total	14	10.48071			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	4.09449	0.37928	49.82443	116.54	<.0001
POP_PAC	0.85621	0.25232	4.92280	11.51	0.0048

Bounds on condition number: 1, 1

Stepwise Selection: Step 2

Variable Tertiary Entered: R-Square = 0.8072 and C(p) = 3.0000

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	8.46040	4.23020	25.13	<.0001
Error	12	2.02031	0.16836		
Corrected Total	14	10.48071			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	2.54768	0.41294	6.40853	38.06	<.0001
POP_PAC	0.97375	0.16040	6.20439	36.85	<.0001
Tertiary	0.03460	0.00755	3.53760	21.01	0.0006

Bounds on condition number: 1.0262, 4.1049

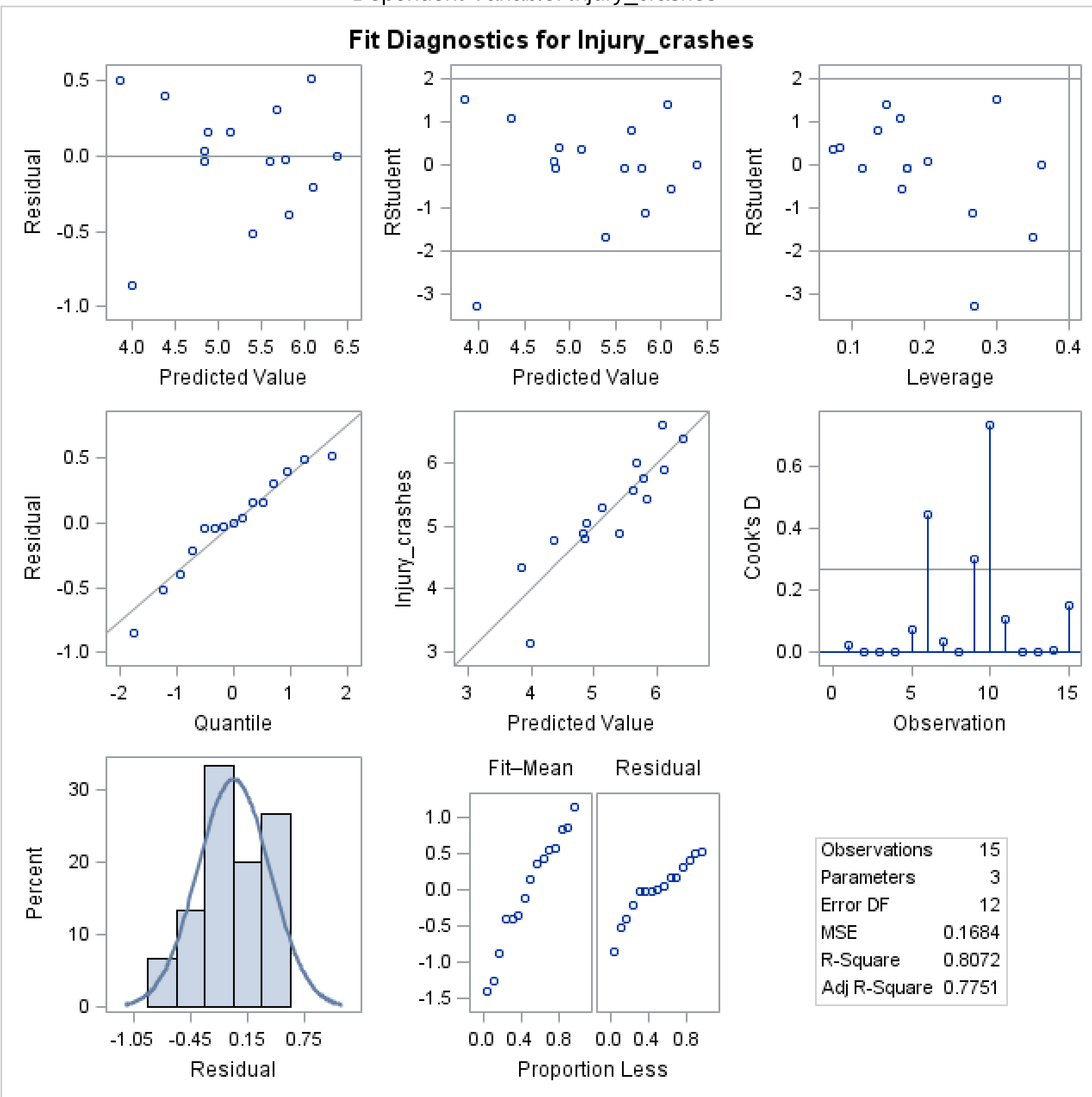
All variables left in the model are significant at the 0.1500 level.

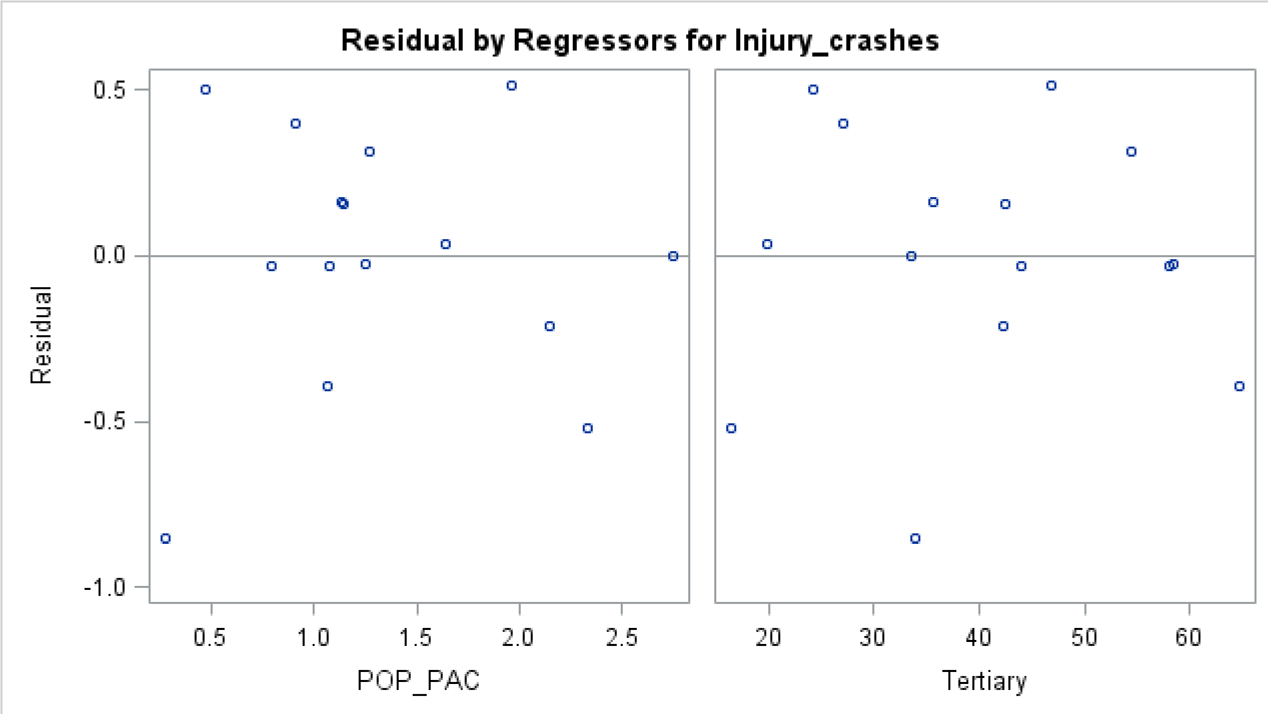
All variables have been entered into the model.

Summary of Stepwise Selection								
Step	Variable Entered	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	POP_PAC		1	0.4697	0.4697	22.0122	11.51	0.0048
2	Tertiary		2	0.3375	0.8072	3.0000	21.01	0.0006

The GLIMMIX Procedure

The REG Procedure
Model: MODEL1
Dependent Variable: Injury_crashes




Residual by Regressors for Injury_crashes
Residual
0.5
0.0
-0.5
-1.0
0.5
1.0
1.5
2.0
2.5
POP_PAC
20
30
40
50
60
Tertiary

1.5 LOG LINEAR NEGATIVE BINOMAIL INJURY CRASHES MODEL

Log-linear negative binomial model Eq(4.16 and 4.17)

The GLIMMIX Procedure

Model 1 Information	
Data Set	WORK.NEGBININJURYCRASH
Response Variable	Injury_crashes
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Variance Matrix	Diagonal
Estimation Technique	Maximum Likelihood
Degrees of Freedom Method	Residual

Number of Observations Read	15
Number of Observations Used	15

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	166.83
AIC (smaller is better)	176.83
AICC (smaller is better)	183.49
BIC (smaller is better)	180.37
CAIC (smaller is better)	185.37
HQIC (smaller is better)	176.79
Pearson Chi-Square	14.80
Pearson Chi-Square / DF	1.35

The GLIMMIX Procedure

Parameter Estimates					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	2.7984	0.3259	11	8.59	<.0001
POP_PAC	0.8831	0.1350	11	6.54	<.0001
Intestates	2.5705	2.1313	11	1.21	0.2531
Tertiary	0.02992	0.006555	11	4.57	0.0008
Scale	0.09430	0.03775	.	.	.

2. APPENDIX CHAPTER 5

2.1 CHAPTER 5 MODEL 1

2.1.1 TOTAL CRASHES-M1 (GLM)

Model Information	
Data Set	LB.MUNICIPALITIESGLIMMIX
Response Variable	Choques_totbase
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Diagonal
Estimation Technique	Maximum Likelihood
Degrees of Freedom Method	Residual

Number of Observations Read	234
Number of Observations Used	233

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	4283.57
AIC (smaller is better)	4293.57
AICC (smaller is better)	4293.84
BIC (smaller is better)	4310.83

Fit Statistics	
CAIC (smaller is better)	4315.83
HQIC (smaller is better)	4300.53
Pearson Chi-Square	1346.45
Pearson Chi-Square / DF	5.88

Parameter Estimates					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	-43.1324	0.4459	229	-96.73	<.0001
propredprim	21.8998	6.7773	229	3.23	0.0014
propredsec	18.7888	3.7237	229	5.05	<.0001
propredter	10.1845	2.2938	229	4.44	<.0001
Scale	3.7840	0.2781	.	.	.

2.1.2 PDO-M1

Model Information	
Data Set	LB.MUNICIPALITIESGLIMMIX
Response Variable	PDO
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Diagonal
Estimation Technique	Maximum Likelihood
Degrees of Freedom Method	Residual

Number of Observations Read	234
Number of Observations Used	226

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	4138.47
AIC (smaller is better)	4148.47
AICC (smaller is better)	4148.74
BIC (smaller is better)	4165.57
CAIC (smaller is better)	4170.57
HQIC (smaller is better)	4155.37
Pearson Chi-Square	1600.63
Pearson Chi-Square / DF	7.21

Parameter Estimates					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	-43.3656	0.4500	222	-96.37	<.0001
propredprim	23.3480	6.7340	222	3.47	0.0006
propredsec	19.5358	3.7795	222	5.17	<.0001
propredter	10.0331	2.2633	222	4.43	<.0001
Scale	3.7157	0.2772	.	.	.

2.1.3 INJURY-M1

Model Information	
Data Set	LB.MUNICIPALITIESGLIMMIX
Response Variable	Choques_herbase
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Diagonal
Estimation Technique	Maximum Likelihood
Degrees of Freedom Method	Residual

Number of Observations Read	234
Number of Observations Used	226

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3334.60
AIC (smaller is better)	3344.60
AICC (smaller is better)	3344.87
BIC (smaller is better)	3361.70
CAIC (smaller is better)	3366.70
HQIC (smaller is better)	3351.50
Pearson Chi-Square	529.04
Pearson Chi-Square / DF	2.38

Parameter Estimates					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	-45.1718	0.4473	222	-100.99	<.0001
propredprim	20.9664	6.5561	222	3.20	0.0016
propredsec	20.4108	3.9102	222	5.22	<.0001
propredter	10.3070	2.2690	222	4.54	<.0001
Scale	3.7772	0.2871	.	.	.

2.1.4 FATAL-M1

Model Information	
Data Set	LB.MUNICIPALITIESGLIMMIX
Response Variable	Choques_fatalbase
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Diagonal
Estimation Technique	Maximum Likelihood
Degrees of Freedom Method	Residual

Number of Observations Read	234
Number of Observations Used	225

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	1454.44
AIC (smaller is better)	1464.44
AICC (smaller is better)	1464.71
BIC (smaller is better)	1481.52
CAIC (smaller is better)	1486.52
HQIC (smaller is better)	1471.33
Pearson Chi-Square	1644.74
Pearson Chi-Square / DF	7.44

Parameter Estimates					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-53.1217	0.8957	221	-59.30	<.0001
propredprim	28.4822	8.9936	221	3.17	0.0018
propredsec	24.0546	5.5625	221	4.32	<.0001
propredter	28.4056	3.2733	221	8.68	<.0001
Scale	3.6307	0.3362	.	.	.

2.2 CHAPTER 5 MODEL 2

2.2.1 TOTAL CRASHES-M2

Model 2 Information	
Data Set	LB.MUNICIPALITIESGLIMMIX2
Response Variable	Choques_totbase
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Class Level Information		
Class	Levels	Values
County	7	1 2 3 4 5 6 7

Number of Observations Read	234
Number of Observations Used	233

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3868.60
AIC (smaller is better)	3878.60
AICC (smaller is better)	3878.86
BIC (smaller is better)	3878.33
CAIC (smaller is better)	3883.33
HQIC (smaller is better)	3875.26

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totbase | r. effects)	3857.46
Pearson Chi-Square	123.77
Pearson Chi-Square / DF	0.53

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
County	0.05331	0.04548
Scale	1.0671	0.09132

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-9.9205	0.2617	6	-37.91	<.0001
propredsec	4.1814	1.7186	224	2.43	0.0158
propredter	4.1525	1.5324	224	2.71	0.0073

2.2.2 PDO-M2

The GLIMMIX Procedure

Model 2 Information	
Data Set	LB.MUNICIPALITIESGLIMMIX2
Response Variable	PDO
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Class Level Information		
Class	Levels	Values
County	7	1 2 3 4 5 6 7

Number of Observations Read	234
Number of Observations Used	226

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3711.04
AIC (smaller is better)	3721.04
AICC (smaller is better)	3721.31
BIC (smaller is better)	3720.77
CAIC (smaller is better)	3725.77
HQIC (smaller is better)	3717.70

Fit Statistics for Conditional Distribution	
-2 log L(PDO | r. effects)	3701.21
Pearson Chi-Square	125.17
Pearson Chi-Square / DF	0.55

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
County	0.03995	0.03813
Scale	0.9693	0.08483

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-10.0141	0.2510	6	-39.89	<.0001
propredsec	3.8786	1.6429	217	2.36	0.0191
propredter	3.9187	1.4677	217	2.67	0.0082

2.2.3 INJURY-M2

Model 2 Information	
Data Set	LB.MUNICIPALITIESGLIMMIX2
Response Variable	Choques_herbase
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Class Level Information		
Class	Levels	Values
County	7	1 2 3 4 5 6 7

Number of Observations Read	234
Number of Observations Used	226

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	2904.78
AIC (smaller is better)	2914.78
AICC (smaller is better)	2915.05
BIC (smaller is better)	2914.51
CAIC (smaller is better)	2919.51
HQIC (smaller is better)	2911.44

Fit Statistics for Conditional Distribution	
-2 log L(Choques_herbase | r. effects)	2895.30
Pearson Chi-Square	144.68
Pearson Chi-Square / DF	0.64

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
County	0.03379	0.03310
Scale	0.8644	0.08030

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-11.6668	0.2382	6	-48.97	<.0001
propredsec	4.1620	1.5618	217	2.66	0.0083
propredter	2.9438	1.4110	217	2.09	0.0381

2.2.4 FATAL-M2

The GLIMMIX Procedure

Model 2 Information	
Data Set	LB.MUNICIPALITIESGLIMMIX2
Response Variable	Choques_fatalbase
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Class Level Information		
Class	Levels	Values
County	7	1 2 3 4 5 6 7

Number of Observations Read	234
Number of Observations Used	225

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	1155.03
AIC (smaller is better)	1165.03
AICC (smaller is better)	1165.30
BIC (smaller is better)	1164.76
CAIC (smaller is better)	1169.76
HQIC (smaller is better)	1161.69

Fit Statistics for Conditional Distribution	
-2 log L(Choques_fatalbase | r. effects)	1136.75
Pearson Chi-Square	689.84
Pearson Chi-Square / DF	3.07

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
County	0.2319	0.1520
Scale	1.0364	0.1194

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-16.4401	0.3845	6	-42.76	<.0001
propredsec	1.9562	2.2289	216	0.88	0.3811
propredter	8.1348	2.0921	216	3.89	0.0001

2.3 CHAPTER 5 MODEL 3

2.3.1 TOTAL CRASHES-M3

Model 3 Information	
Data Set	LB.MUNICIPALITIESGLIMMIX2
Response Variable	Choques_totbase
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	234
Number of Observations Used	233

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3779.73
AIC (smaller is better)	3791.73
AICC (smaller is better)	3792.10
BIC (smaller is better)	3805.87
CAIC (smaller is better)	3811.87

HQIC (smaller is better)	3797.39

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totbase | r. effects)	3480.10
Pearson Chi-Square	45.66
Pearson Chi-Square / DF	0.20

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error
MUNICIPALITIES	2.0044	0.3562
Scale	0.3229	0.03770

Solutions for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-11.3782	0.7785	74	-14.62	<.0001
propredprim	7.4670	7.0497	155	1.06	0.2912
propredsec	15.6079	4.6609	155	3.35	0.0010
propredter	5.8372	3.9824	155	1.47	0.1447

2.3.2 PDO-M3

Model 3 Information	
Data Set	LB.MUNICIPALITIESGLIMMIX2
Response Variable	PDO
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	234
Number of Observations Used	226

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3651.64
AIC (smaller is better)	3663.64
AICC (smaller is better)	3664.03
BIC (smaller is better)	3677.78
CAIC (smaller is better)	3683.78
HQIC (smaller is better)	3669.30

Fit Statistics for Conditional Distribution	
-2 log L(PDO | r. effects)	3359.85
Pearson Chi-Square	41.92
Pearson Chi-Square / DF	0.19

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
MUNICIPALITIES	1.8997	0.3525
Scale	0.3249	0.03895

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-11.3862	0.7659	74	-14.87	<.0001
propredprim	6.6360	6.9174	148	0.96	0.3390
propredsec	14.9447	4.5676	148	3.27	0.0013
propredter	5.3276	3.9010	148	1.37	0.1741

2.3.3 INJURY-M3

The GLIMMIX Procedure

Model 3 Information	
Data Set	LB.MUNICIPALITIESGLIMMIX2
Response Variable	Choques_herbase
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	234
Number of Observations Used	226

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	2842.12
AIC (smaller is better)	2854.12
AICC (smaller is better)	2854.51
BIC (smaller is better)	2868.26

Fit Statistics	
CAIC (smaller is better)	2874.26
HQIC (smaller is better)	2859.78

Fit Statistics for Conditional Distribution	
-2 log L(Choques_herbase | r. effects)	2575.13
Pearson Chi-Square	62.29
Pearson Chi-Square / DF	0.28

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error
MUNICIPALITIES	1.2867	0.2593
Scale	0.2846	0.03742

Solutions for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-13.0035	0.6578	74	-19.77	<.0001
propredprim	5.5926	5.8772	148	0.95	0.3429
propredsec	12.8725	3.8688	148	3.33	0.0011
propredter	5.5726	3.2959	148	1.69	0.0930

2.3.4 FATAL-M3

Model 3 Information	
Data Set	LB.MUNICIPALITIESGLIMMIX2
Response Variable	Choques_fatalbase
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	234
Number of Observations Used	225

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	1088.52
AIC (smaller is better)	1100.52
AICC (smaller is better)	1100.91
BIC (smaller is better)	1114.66

The GLIMMIX Procedure

Fit Statistics	
CAIC (smaller is better)	1120.66
HQIC (smaller is better)	1106.18

Fit Statistics for Conditional Distribution	
-2 log L(Choques_fatalbase | r. effects)	960.15
Pearson Chi-Square	160.89
Pearson Chi-Square / DF	0.72

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error
MUNICIPALITIES	0.5285	0.1293
Scale	0.4335	0.07828

Solutions for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-17.0821	0.5511	74	-31.00	<.0001
propredprim	6.5320	4.6657	147	1.40	0.1636
propredsec	6.2013	3.1942	147	1.94	0.0541
propredter	7.7692	2.6771	147	2.90	0.0043

2.4 CHAPTER 5 MODEL 4

2.4.1TOTAL CRASHES-M4

Model 4 Information	
Data Set	LB.MUNICIPALITIESGLIMMIX2
Response Variable	Choques_totbase
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	234
Number of Observations Used	233

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3769.29
AIC (smaller is better)	3783.29
AICC (smaller is better)	3783.78
BIC (smaller is better)	3782.91
CAIC (smaller is better)	3789.91
HQIC (smaller is better)	3778.61

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totbase | r. effects)	3480.15
Pearson Chi-Square	46.36
Pearson Chi-Square / DF	0.20

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
County	0.4390	0.3068
MUNICIPALITIES	1.5407	0.2866
Scale	0.3225	0.03759

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-11.2077	0.7605	6	-14.74	<.0001
propredprim	8.9775	6.3935	155	1.40	0.1623
propredsec	12.3830	4.4793	155	2.76	0.0064
propredter	5.9142	3.7201	155	1.59	0.1139

2.4.2 PDO-M4

Model Information	
Data Set	LB.MUNICIPALITIESGLIMMIX2
Response Variable	PDO
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	234
Number of Observations Used	226

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3642.05
AIC (smaller is better)	3656.05
AICC (smaller is better)	3656.57
BIC (smaller is better)	3655.67
CAIC (smaller is better)	3662.67
HQIC (smaller is better)	3651.37

Fit Statistics for Conditional Distribution	
-2 log L(PDO | r. effects)	3359.20
Pearson Chi-Square	42.25
Pearson Chi-Square / DF	0.19

Covariance Parameter Estimates		
Cov Parm	**Estimate**	**Standard Error**
County	0.4005	0.2855
MUNICIPALITIES	1.4920	0.2864
Scale	0.3236	0.03857

Solutions for Fixed Effects					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	-11.2509	0.7531	6	-14.94	<.0001
propredprim	8.1757	6.3464	148	1.29	0.1997
propredsec	11.9937	4.4282	148	2.71	0.0076
propredter	5.5039	3.6816	148	1.50	0.1370

2.4.3 INJURY-M4

Model 4 Information	
Data Set	LB.MUNICIPALITIESGLIMMIX2
Response Variable	Choques_herbase
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	234
Number of Observations Used	226

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	2834.88
AIC (smaller is better)	2848.88
AICC (smaller is better)	2849.40
BIC (smaller is better)	2848.50

Fit Statistics	
CAIC (smaller is better)	2855.50
HQIC (smaller is better)	2844.20

Fit Statistics for Conditional Distribution	
-2 log L(Choques_herbase | r. effects)	2573.14
Pearson Chi-Square	62.09
Pearson Chi-Square / DF	0.27

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error
County	0.2472	0.1885
MUNICIPALITIES	1.0707	0.2201
Scale	0.2833	0.03701

Solutions for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-12.8896	0.6606	6	-19.51	<.0001
propredprim	7.5537	5.5872	148	1.35	0.1785
propredsec	11.1856	3.8214	148	2.93	0.0040
propredter	5.2176	3.2072	148	1.63	0.1059

2.4.4 FATAL-M4

Model 4 Information	
Data Set	LB.MUNICIPALITIESGLIMMIX2
Response Variable	Choques_fatalbase
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	logtotkilPOPESTIMATEBASE
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	234
Number of Observations Used	225

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	1088.36
AIC (smaller is better)	1102.36
AICC (smaller is better)	1102.87
BIC (smaller is better)	1101.98

Fit Statistics	
CAIC (smaller is better)	1108.98
HQIC (smaller is better)	1097.68

Fit Statistics for Conditional Distribution	
-2 log L(Choques_fatalbase | r. effects)	959.93
Pearson Chi-Square	160.92
Pearson Chi-Square / DF	0.72

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error
County	0.02078	0.05872
MUNICIPALITIES	0.5120	0.1326
Scale	0.4336	0.07829

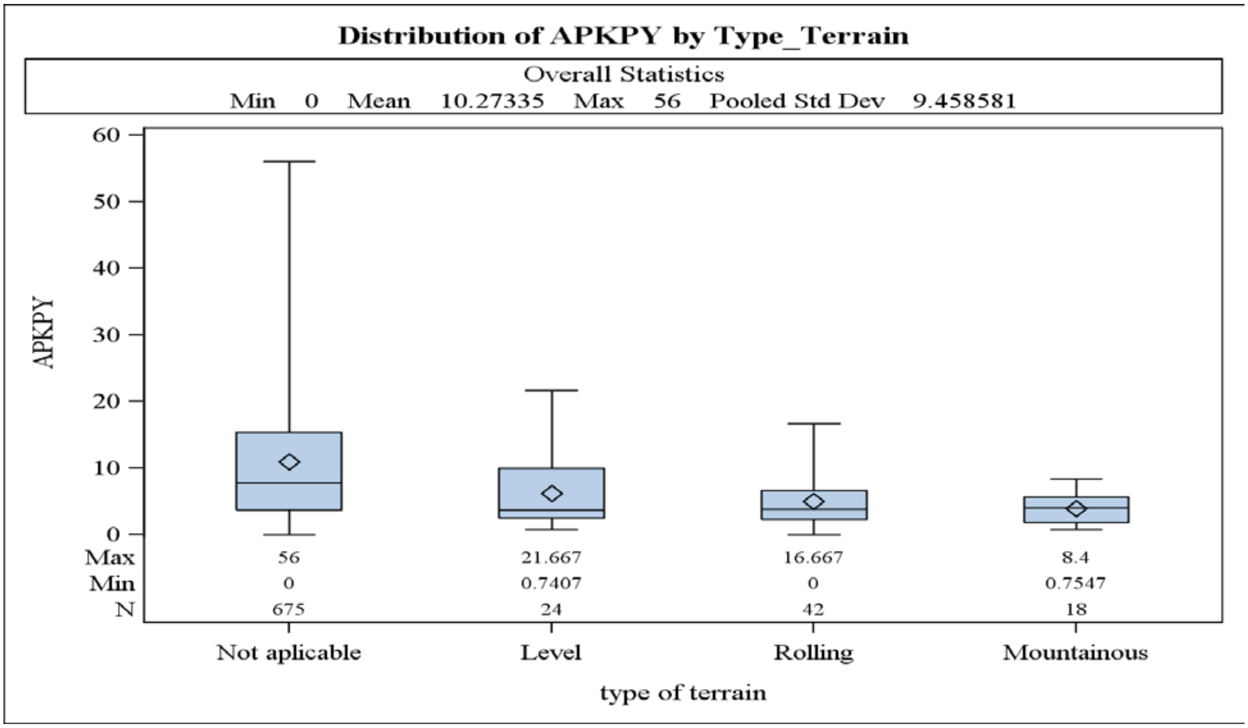
Solutions for Fixed Effects

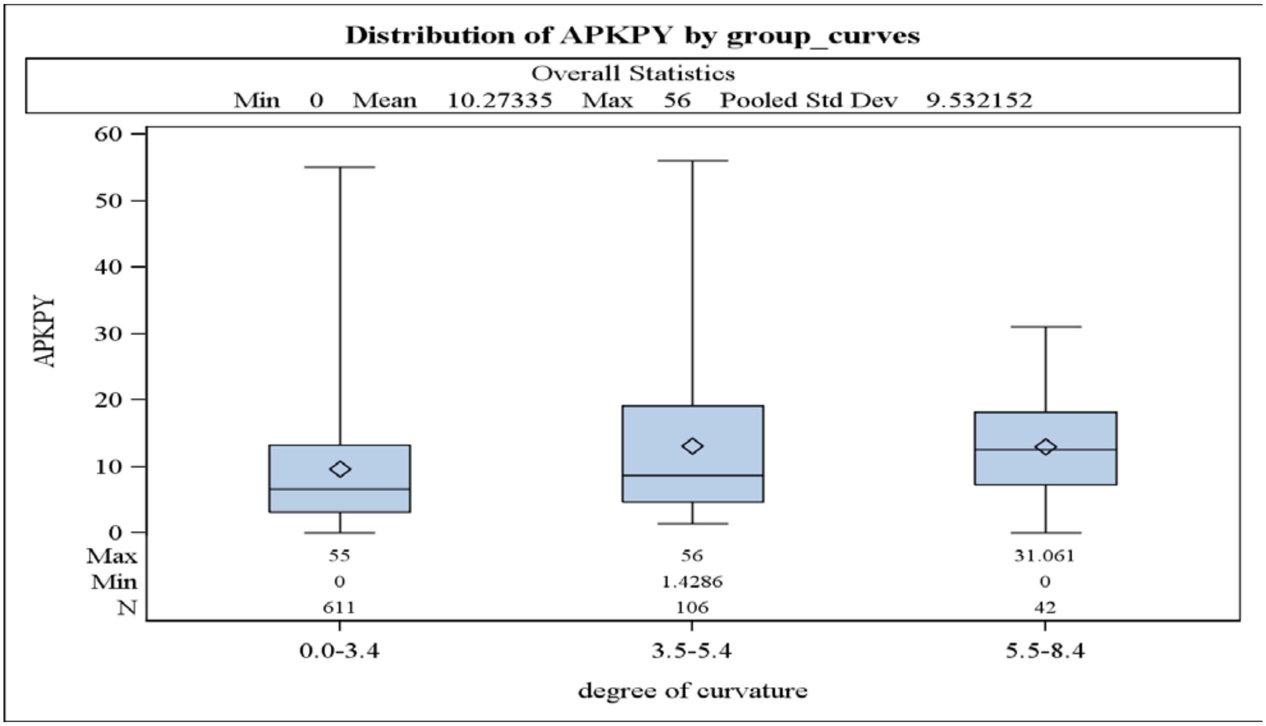
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-17.1187	0.5640	6	-30.35	<.0001
propredprim	7.1941	4.9122	147	1.46	0.1452
propredsec	6.5493	3.3372	147	1.96	0.0516
propredter	7.7829	2.7038	147	2.88	0.0046

3. APPENDIX CHAPTER 6

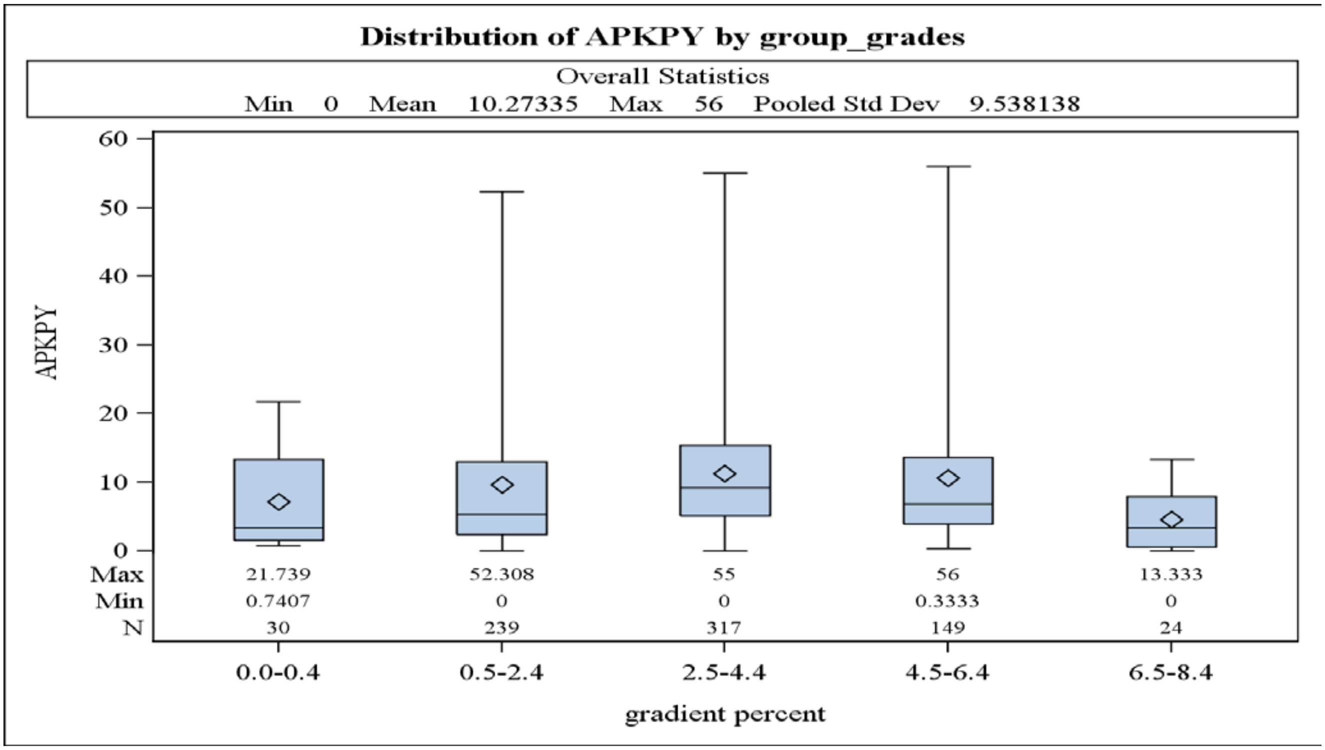
APPENDIX CHAPTER 6

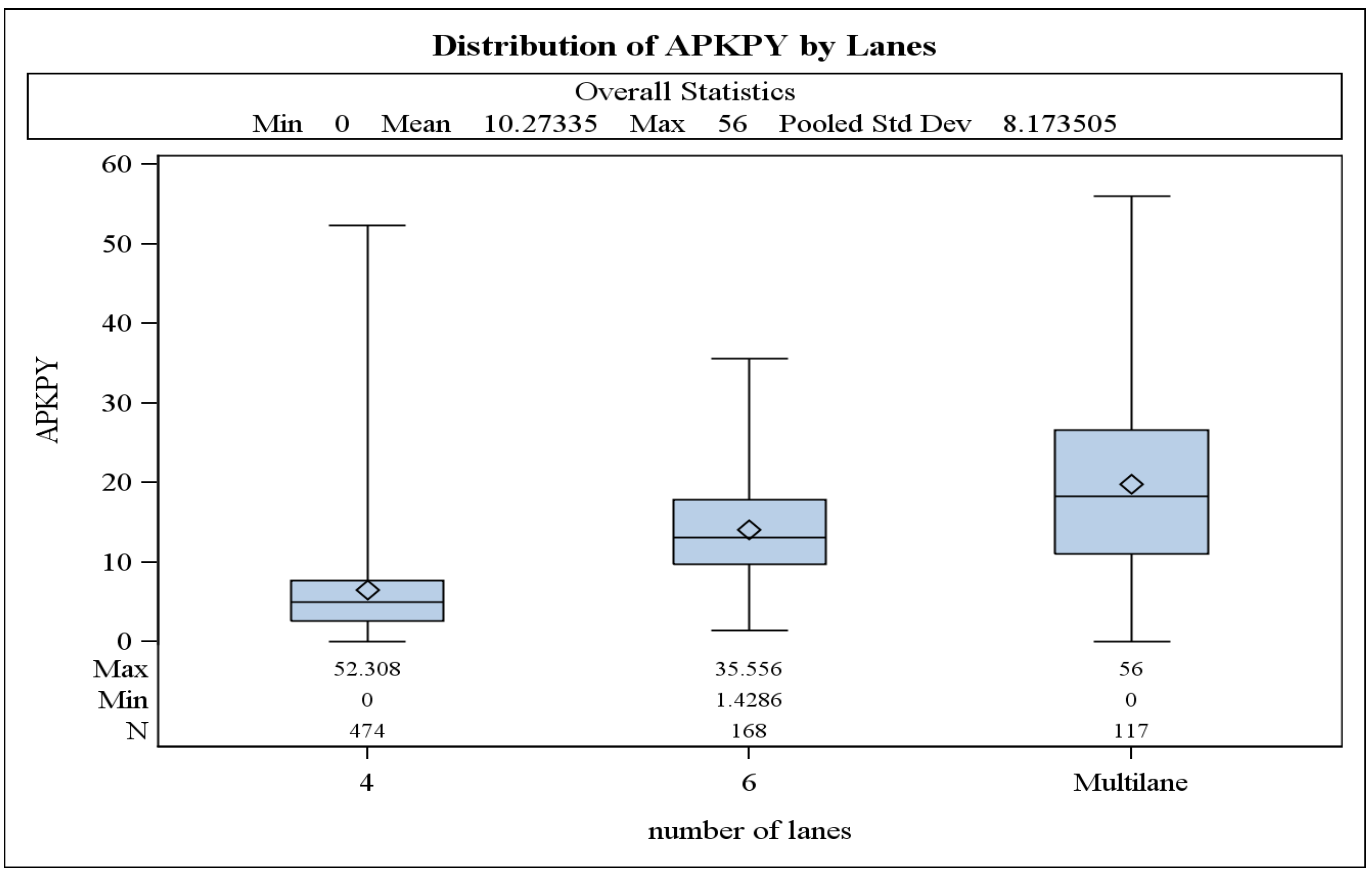
DESCRIPTIVE STATISTICS FREEWAYS

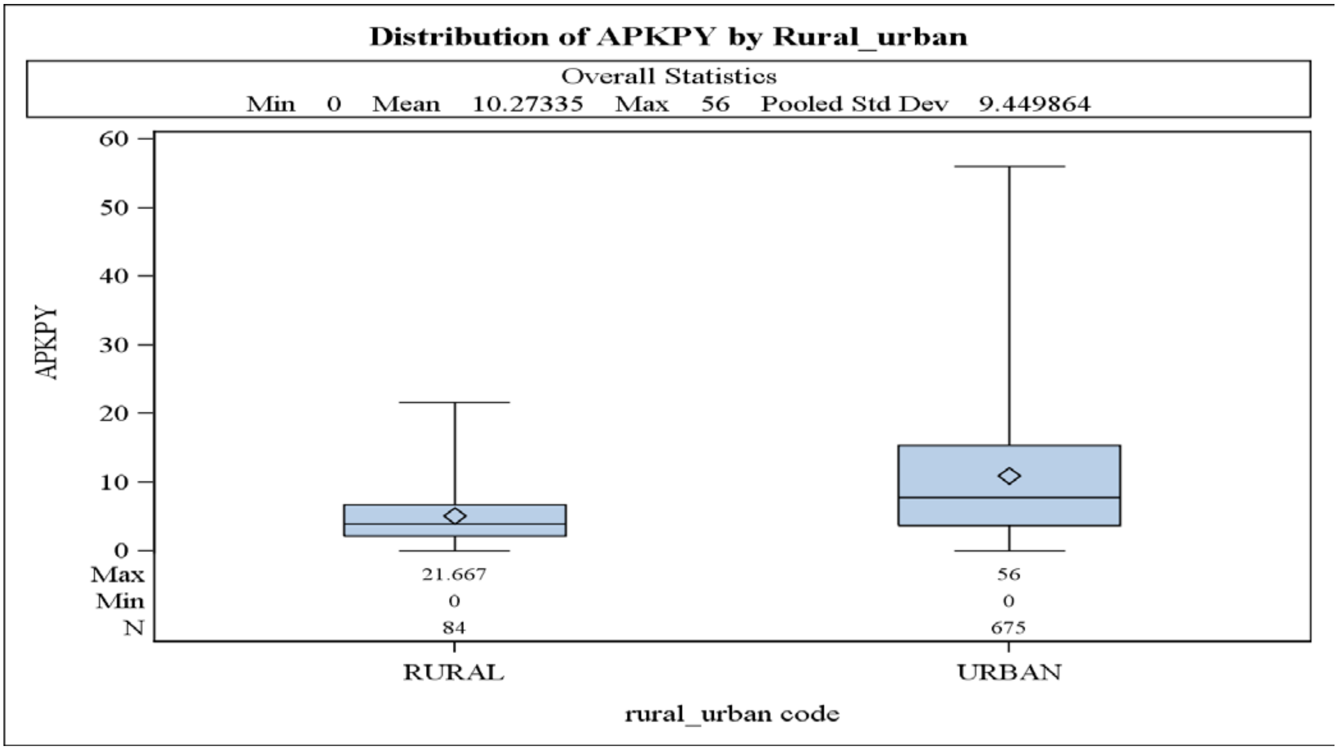




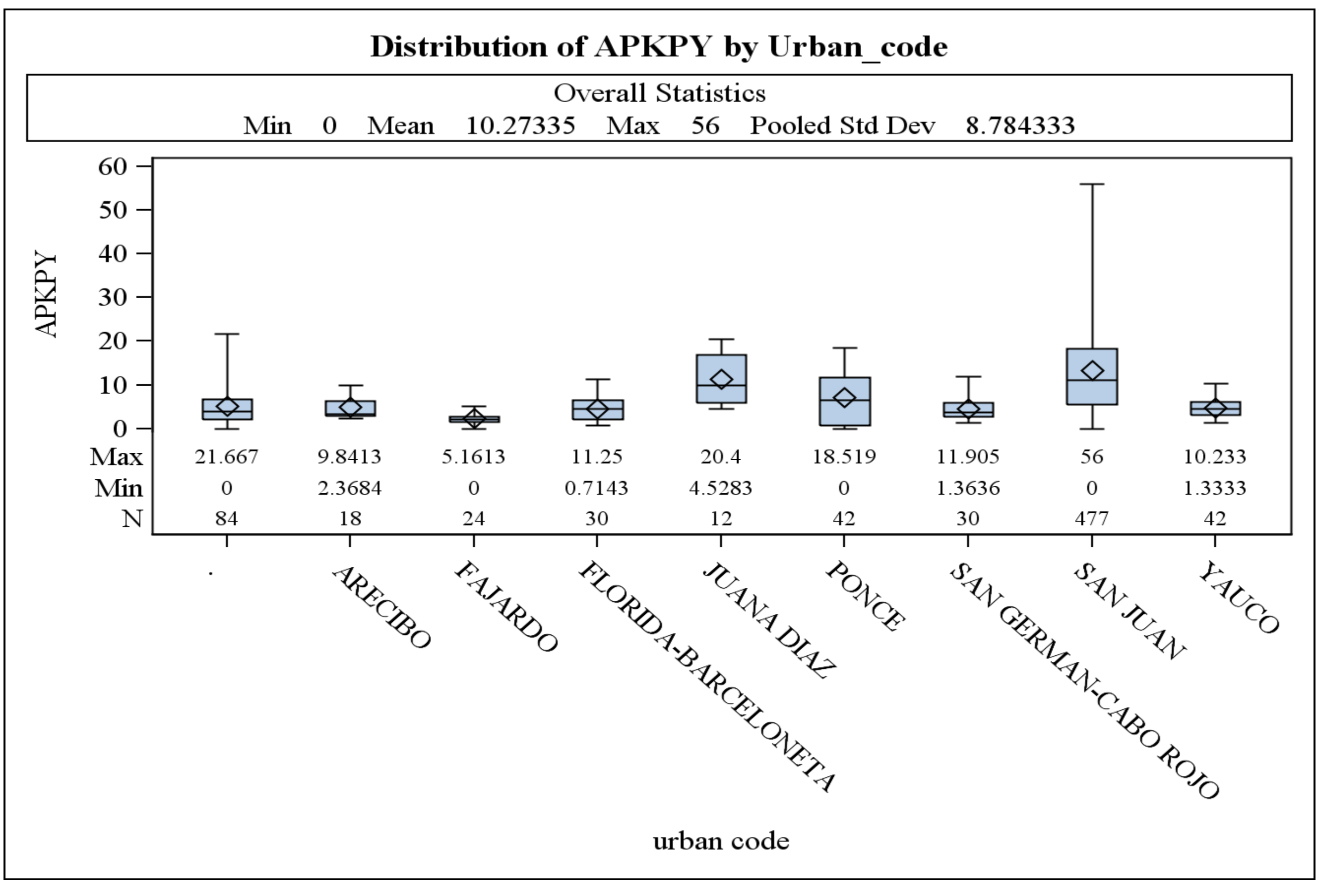
The GLIMMIX Procedure



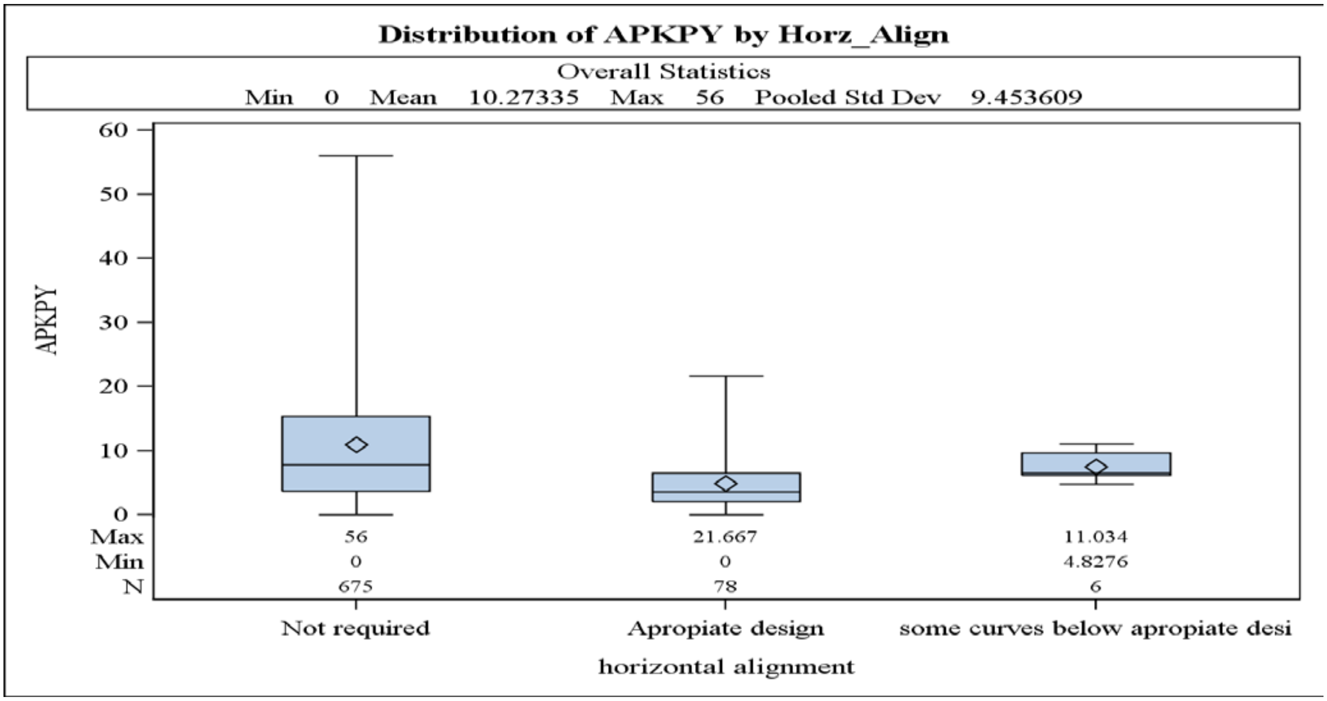


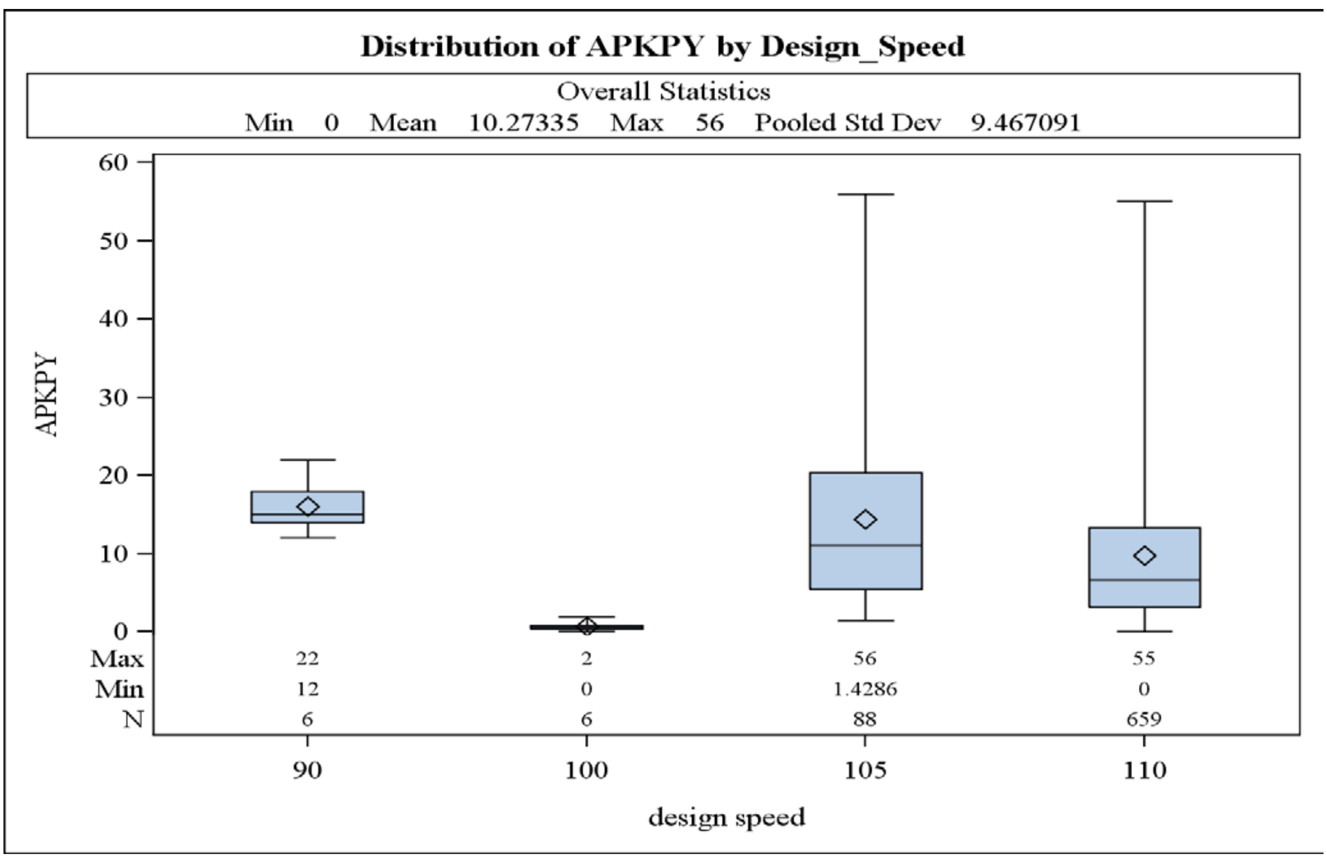


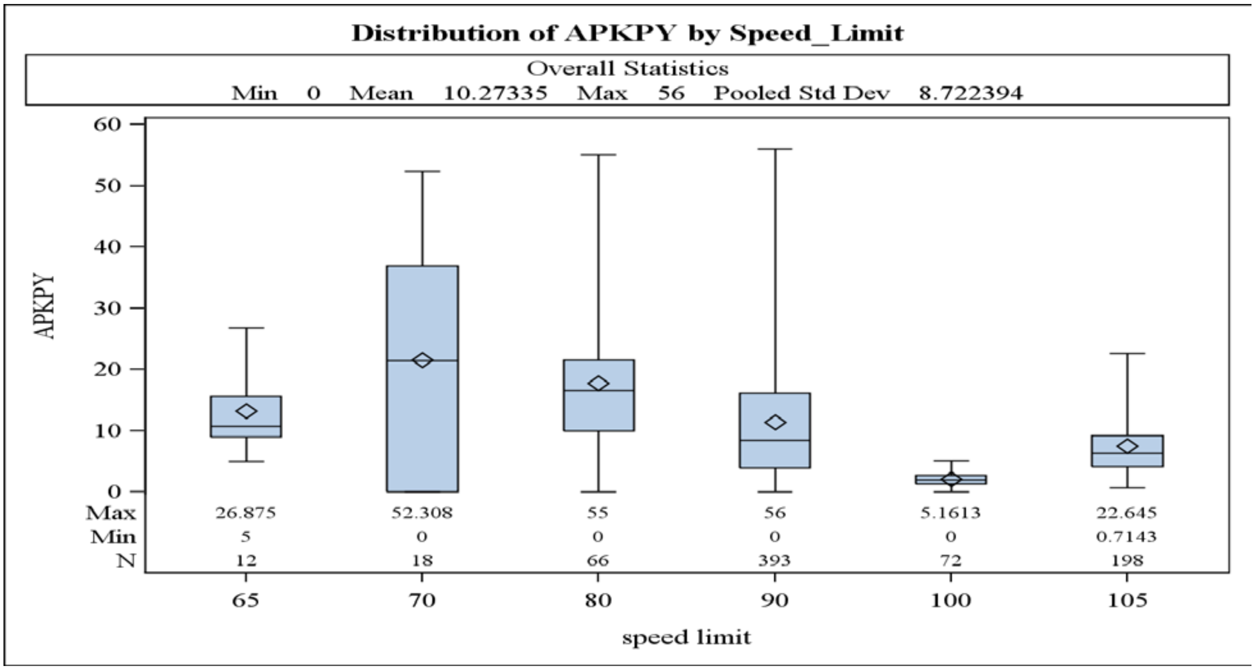
Distribution of APKPY by Rural_urban
Overall Statistics
Min 0 Mean 10.27335 Max 56 Pooled Std Dev 9.449864
APKPY
Max 21.667 56
Min 0 0
N 84 675
RURAL
URBAN
rural_urban code

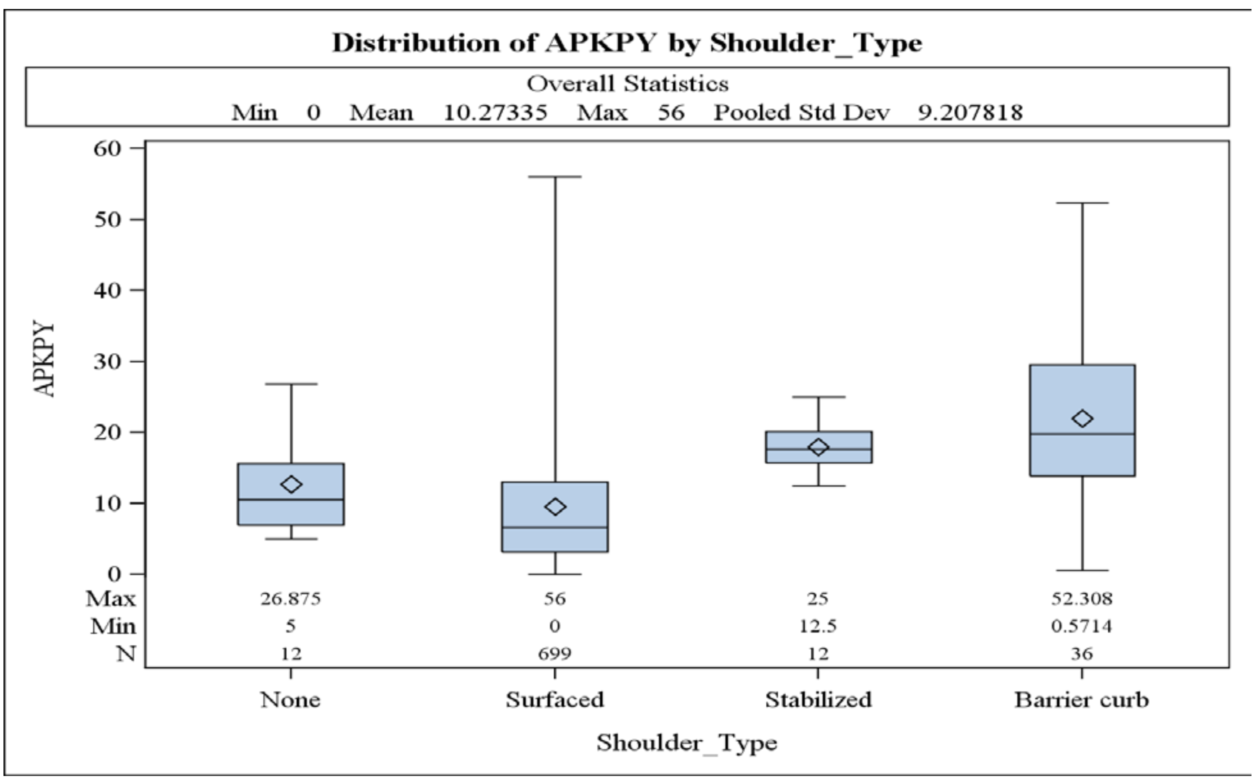


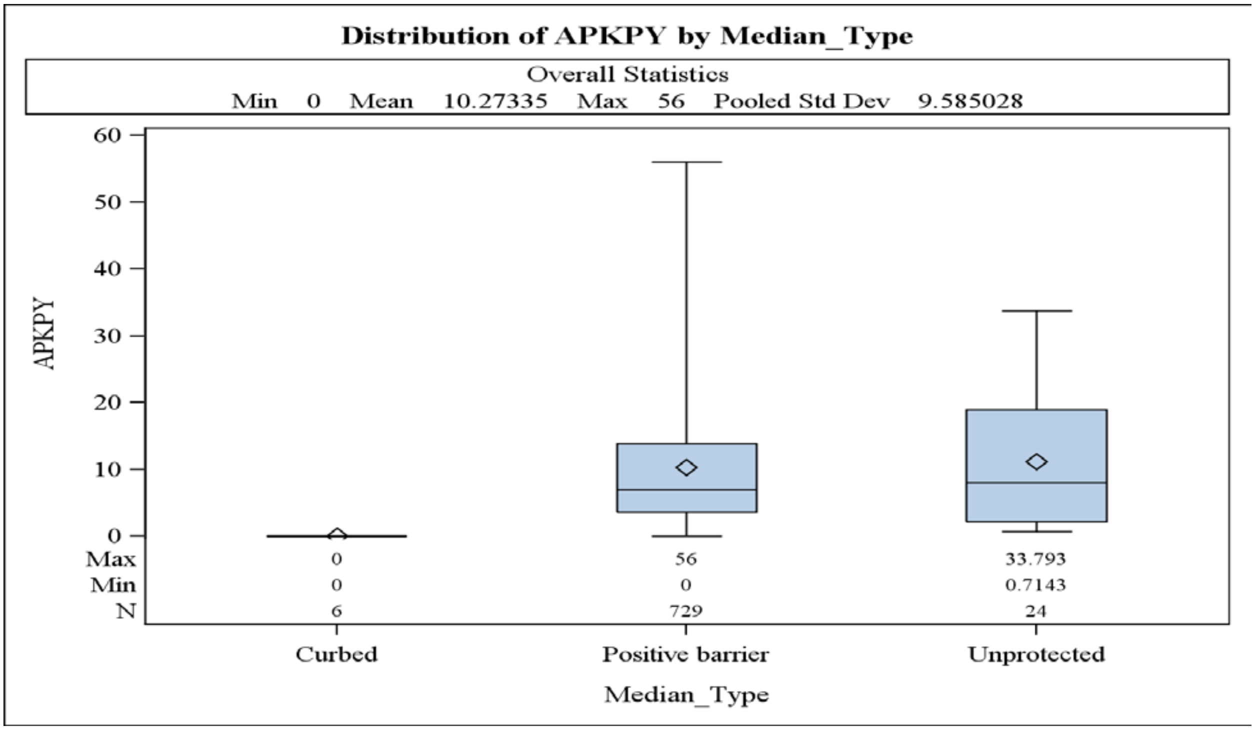
Distribution of APKPY by Urban_code
Overall Statistics
Min 0 Mean 10.27335 Max 56 Pooled Std Dev 8.784333
APKPY
Max 21.667 9.8413 5.1613 11.25 20.4 18.519 11.905 56 10.233
Min 0 2.3684 0 0.7143 4.5283 0 1.3636 0 1.3333
N 84 18 24 30 12 42 30 477 42
.
ARECIBO
FAJARDO
FLORIDA-BARCELONETA
JUANA DIAZ
PONCE
SAN GERMAN-CABO ROJO
SAN JUAN
YAUCO
urban code

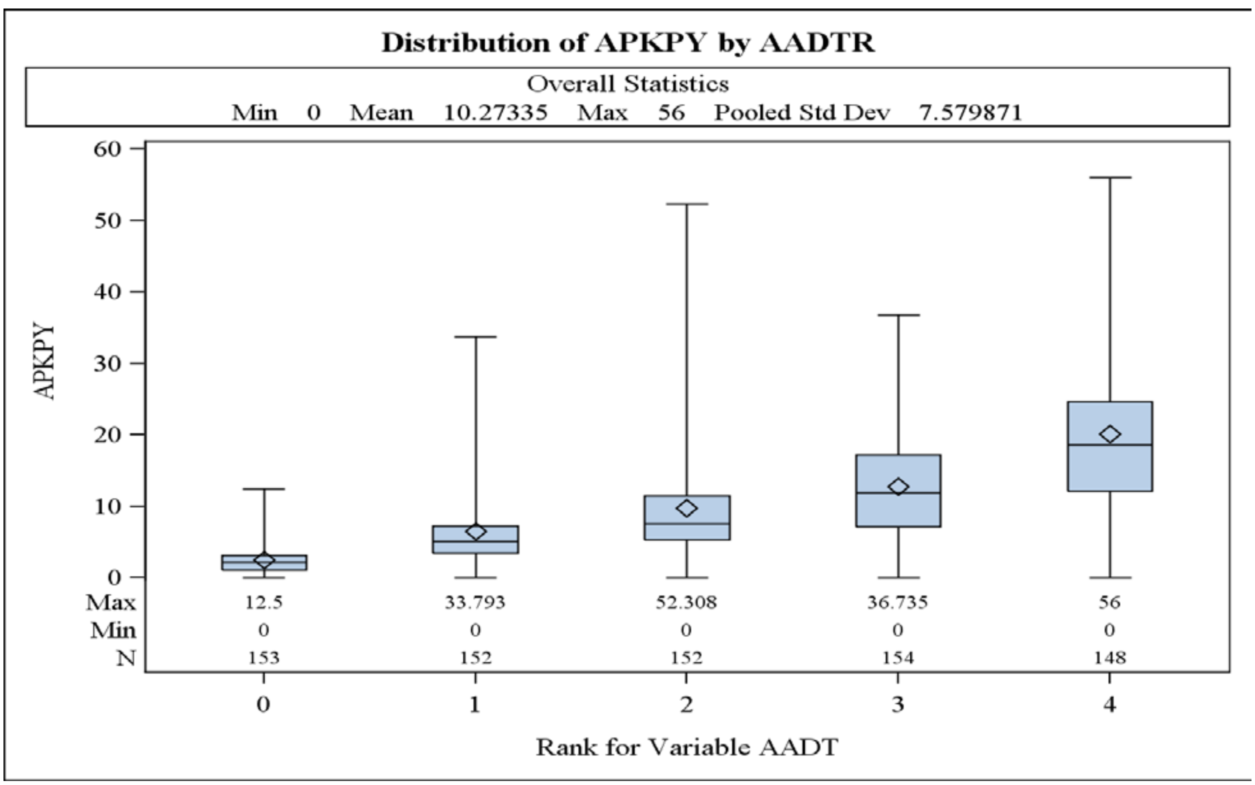

Distribution of APKPY by Horz_Align
Overall Statistics
Min 0 Mean 10.27335 Max 56 Pooled Std Dev 9.453609
APKPY
60
50
40
30
20
10
0
Max 56 21.667 11.034
Min 0 0 4.8276
N 675 78 6
Not required
Apropiate design
some curves below apropiate desi
horizontal alignment

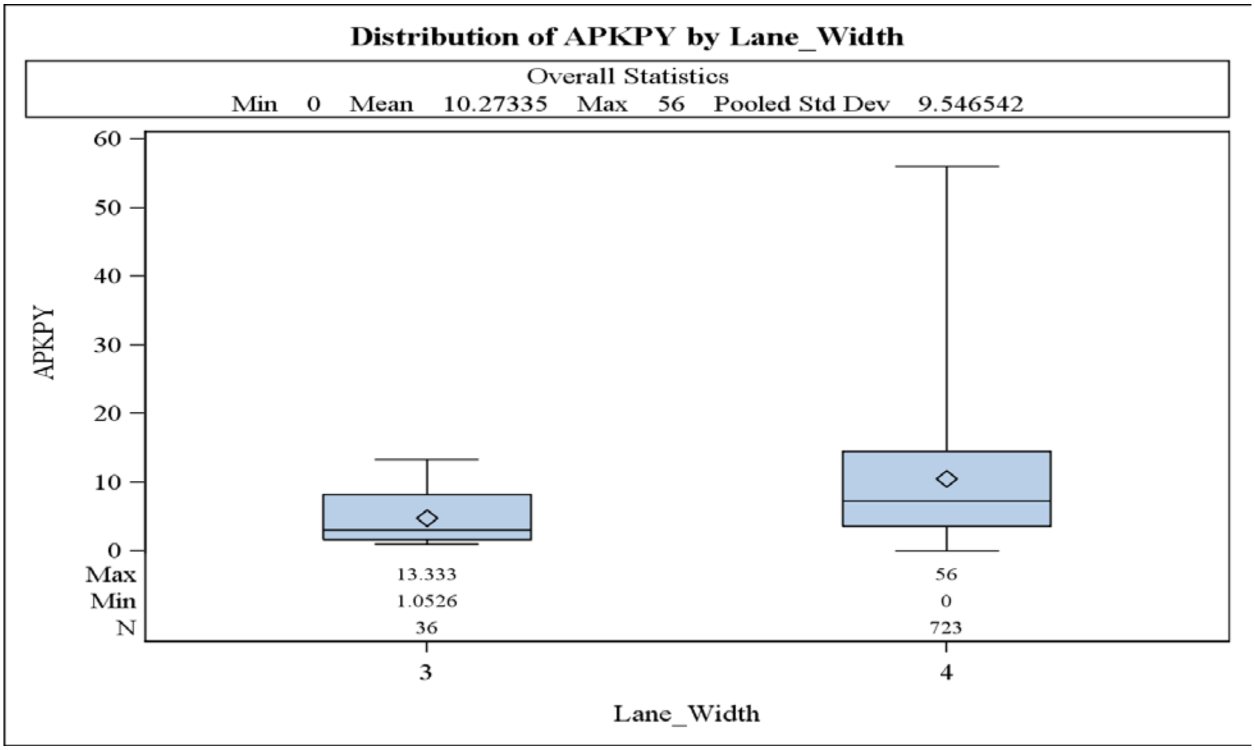

Distribution of APKPY by Design_Speed
Overall Statistics
Min 0 Mean 10.27335 Max 56 Pooled Std Dev 9.467091
APKPY
60
50
40
30
20
10
0
Max 22 2 56 55
Min 12 0 1.4286 0
N 6 6 88 659
90
100
105
110
design speed

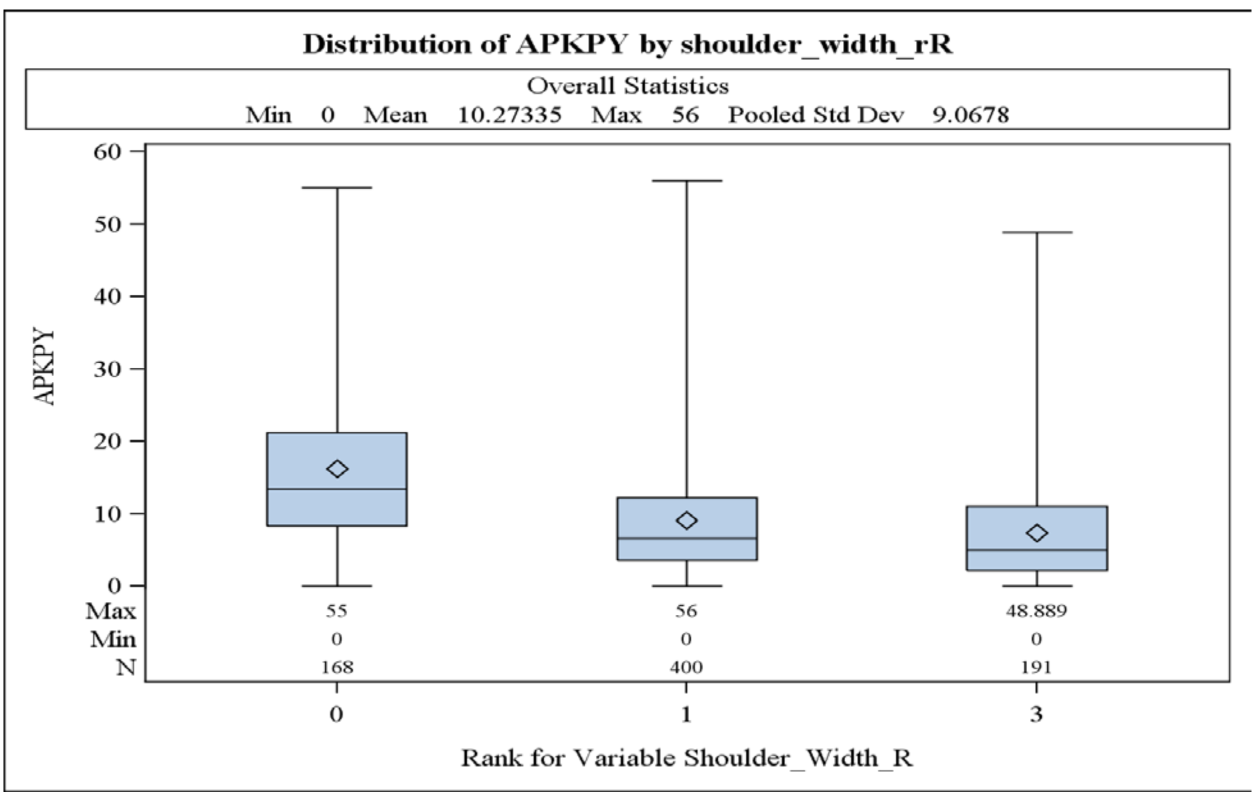

Distribution of APKPY by Speed_Limit
Overall Statistics
Min 0 Mean 10.27335 Max 56 Pooled Std Dev 8.722394
APKPY
0
10
20
30
40
50
60
Max 26.875 52.308 55 56 5.1613 22.645
Min 5 0 0 0 0 0.7143
N 12 18 66 393 72 198
65
70
80
90
100
105
speed limit

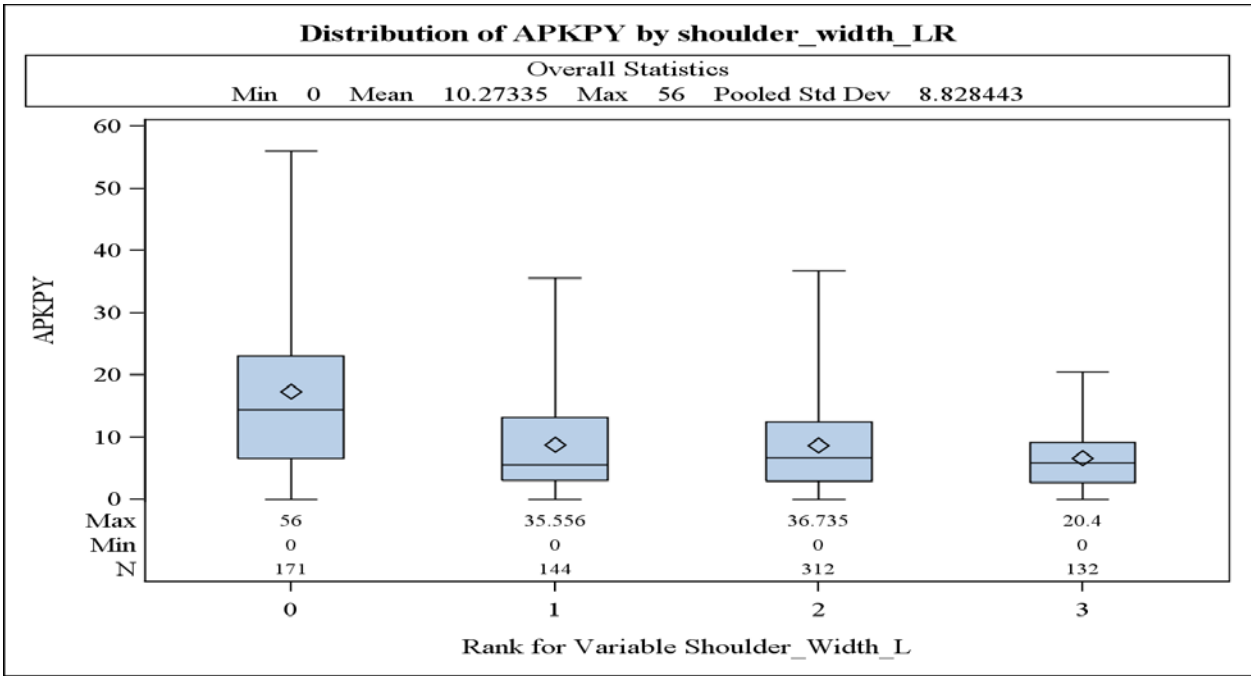

Distribution of APKPY by Shoulder_Type
Overall Statistics
Min 0 Mean 10.27335 Max 56 Pooled Std Dev 9.207818
APKPY
0
10
20
30
40
50
60
Max 26.875 56 25 52.308
Min 5 0 12.5 0.5714
N 12 699 12 36
None
Surfaced
Stabilized
Barrier curb
Shoulder_Type


Distribution of APKPY by Median_Type
Overall Statistics
Min 0 Mean 10.27335 Max 56 Pooled Std Dev 9.585028
APKPY
60
50
40
30
20
10
0
Max 0 56 33.793
Min 0 0 0.7143
N 6 729 24
Curbed
Positive barrier
Unprotected
Median_Type

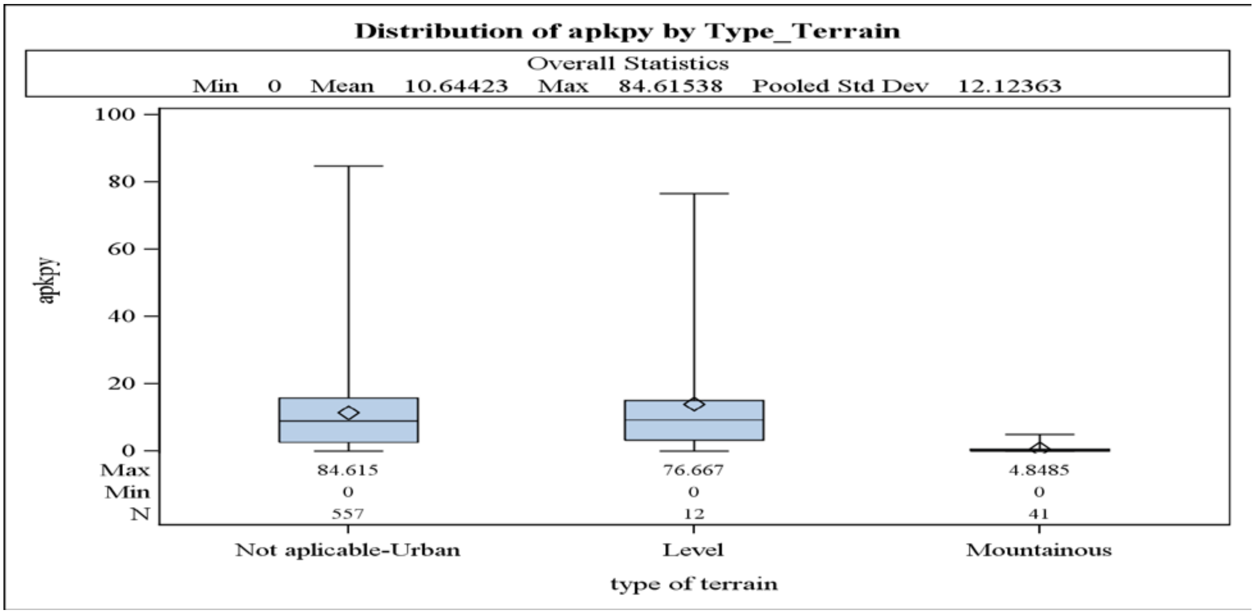

Distribution of APKPY by AADTR
Overall Statistics
Min 0 Mean 10.27335 Max 56 Pooled Std Dev 7.579871
APKPY
60
50
40
30
20
10
0
Max 12.5 33.793 52.308 36.735 56
Min 0 0 0 0 0
N 153 152 152 154 148
0
1
2
3
4
Rank for Variable AADT

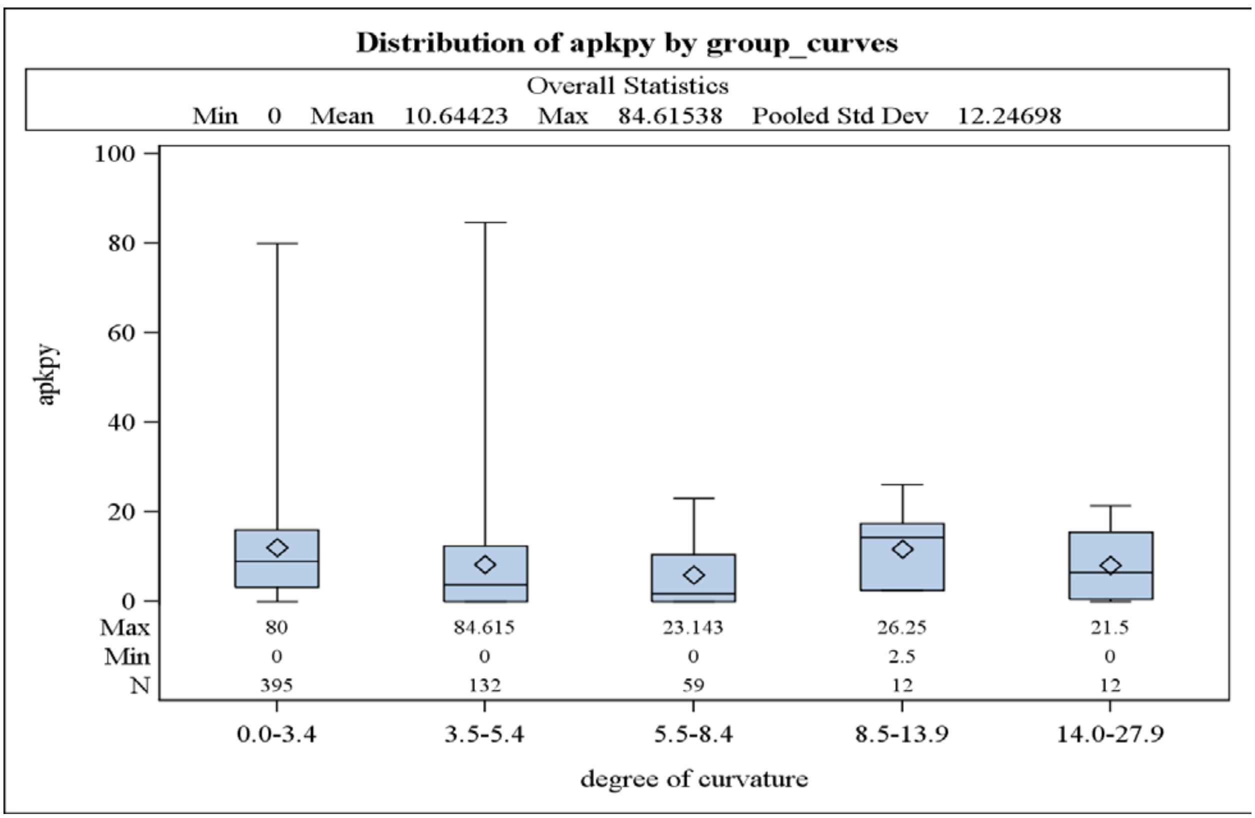

Distribution of APKPY by Lane_Width
Overall Statistics
Min 0 Mean 10.27335 Max 56 Pooled Std Dev 9.546542
APKPY
60
50
40
30
20
10
0
Max 13.333 56
Min 1.0526 0
N 36 723
3
4
Lane_Width

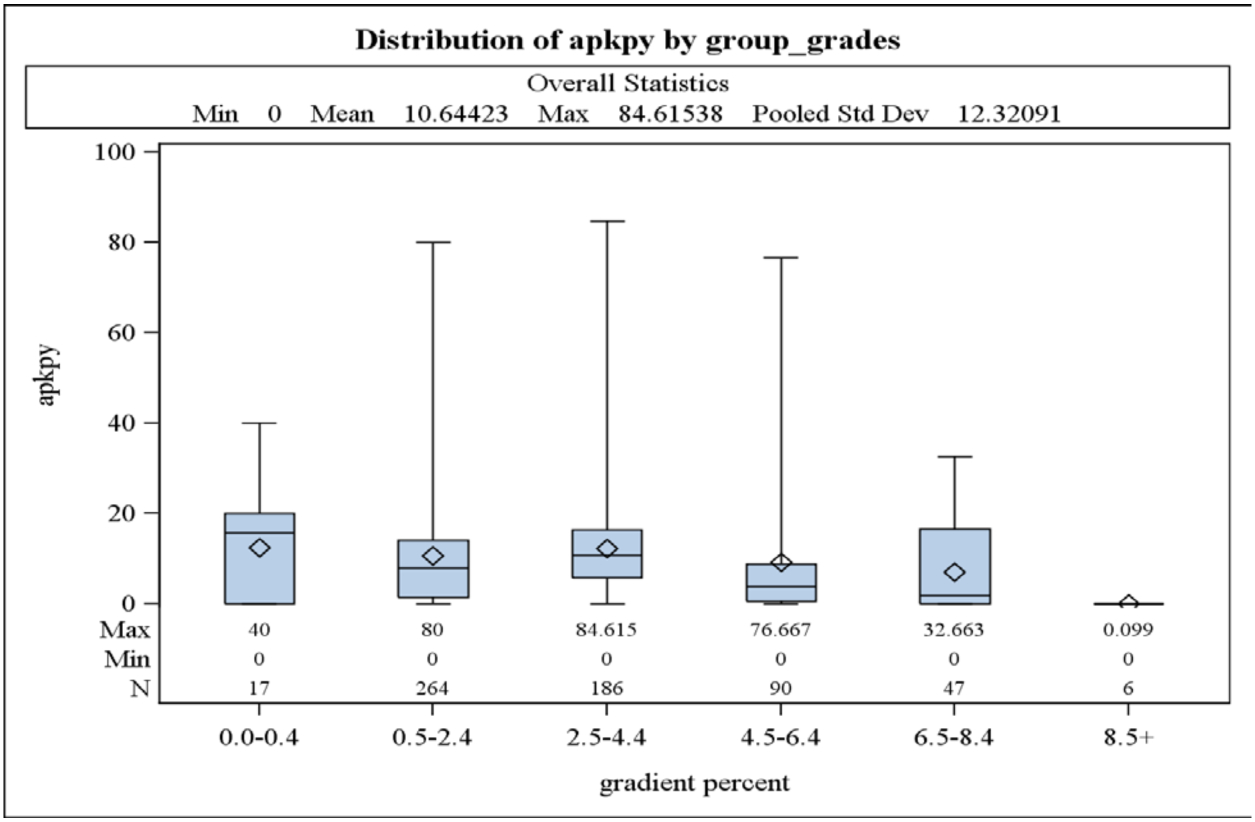

Distribution of APKPY by shoulder_width_rR
Overall Statistics
Min 0 Mean 10.27335 Max 56 Pooled Std Dev 9.0678
APKPY
60
50
40
30
20
10
0
Max 55 56 48.889
Min 0 0 0
N 168 400 191
0
1
3
Rank for Variable Shoulder_Width_R


Distribution of APKPY by shoulder_width_LR
Overall Statistics
Min 0 Mean 10.27335 Max 56 Pooled Std Dev 8.828443
APKPY
Max 56 35.556 36.735 20.4
Min 0 0 0 0
N 171 144 312 132
0 1 2 3
Rank for Variable Shoulder_Width_L

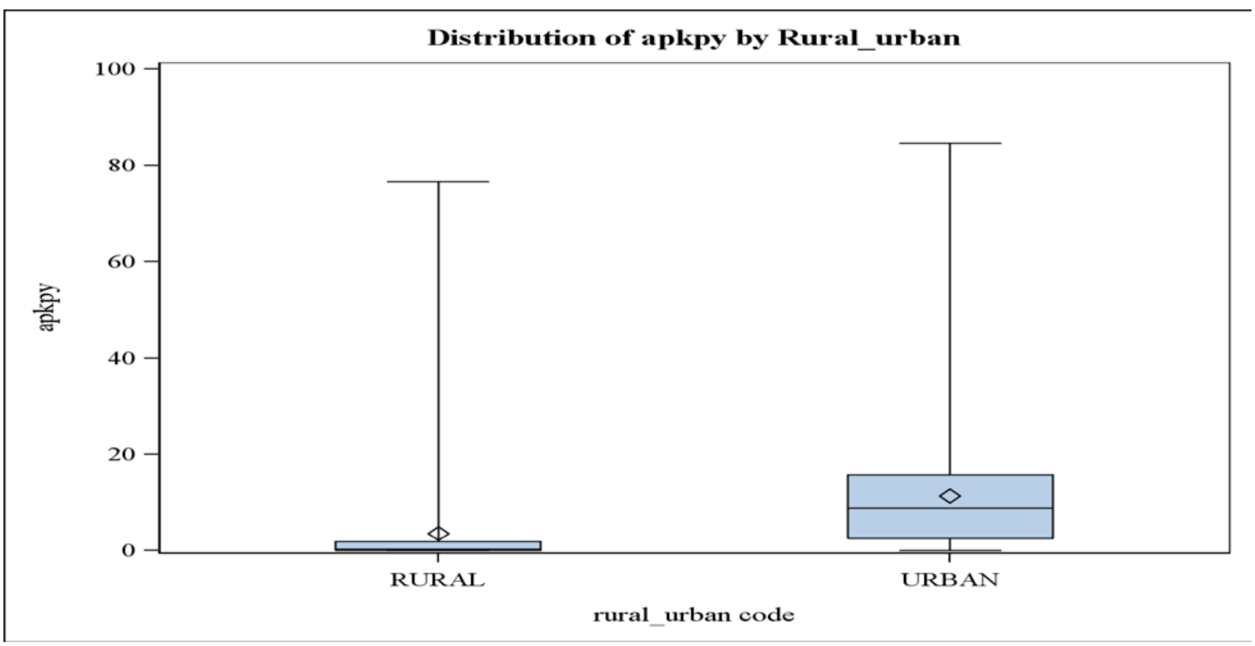
DESCRIPTIVE STATISTICS ARTERIALS



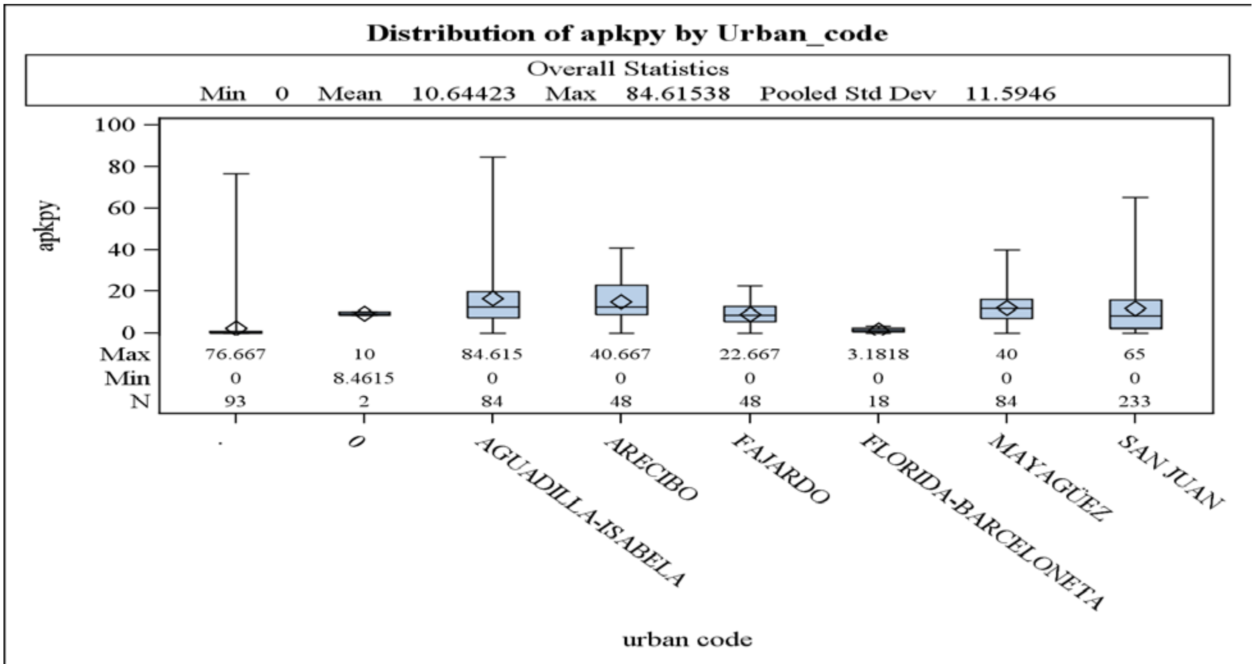


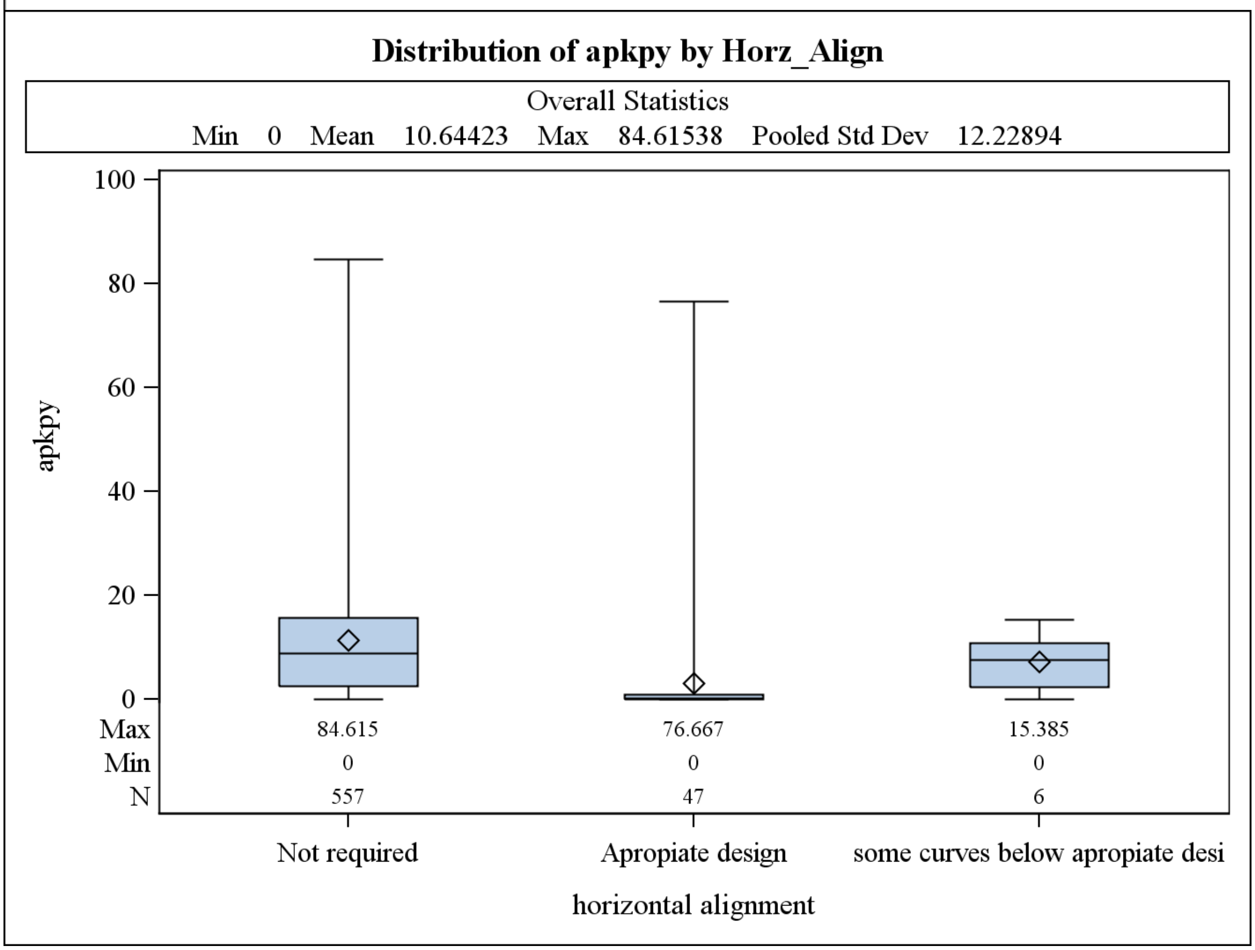


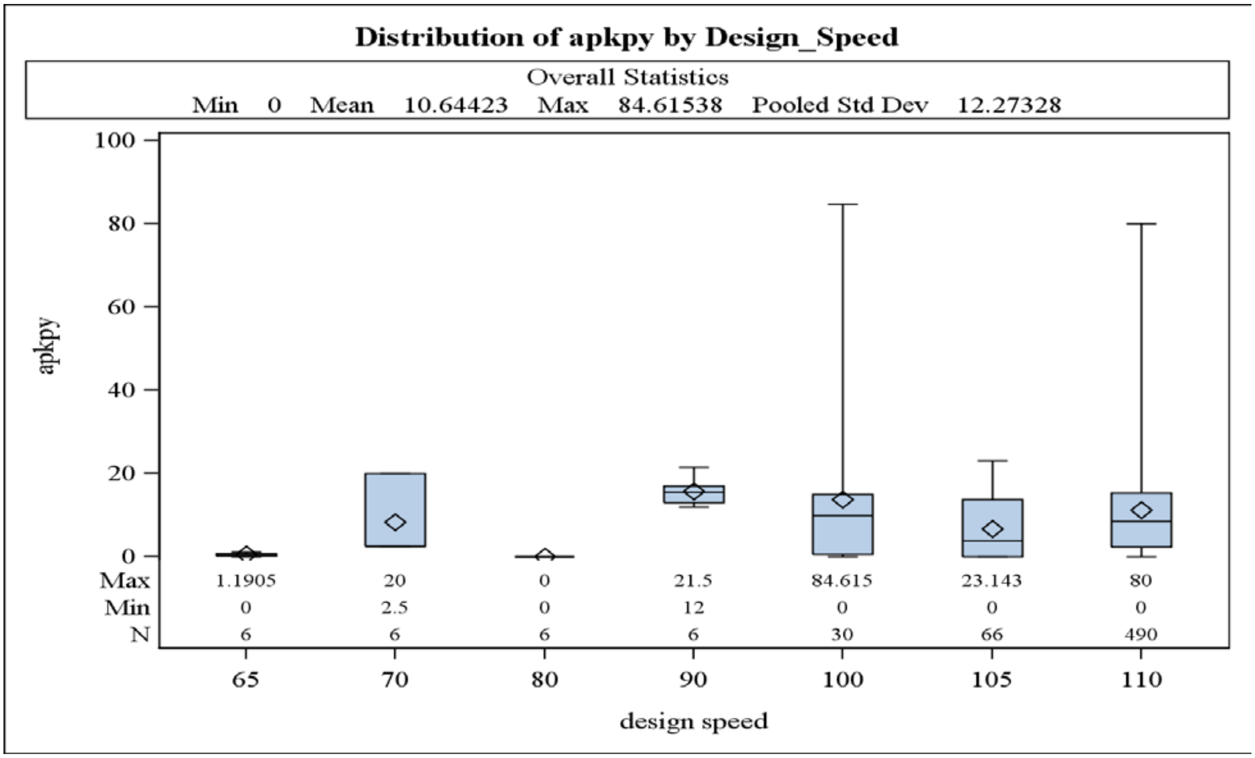
Distribution of apkpy by group_grades
Overall Statistics
Min 0 Mean 10.64423 Max 84.61538 Pooled Std Dev 12.32091
apkpy
100
80
60
40
20
0
Max 40 80 84.615 76.667 32.663 0.099
Min 0 0 0 0 0 0
N 17 264 186 90 47 6
0.0-0.4
0.5-2.4
2.5-4.4
4.5-6.4
6.5-8.4
8.5+
gradient percent

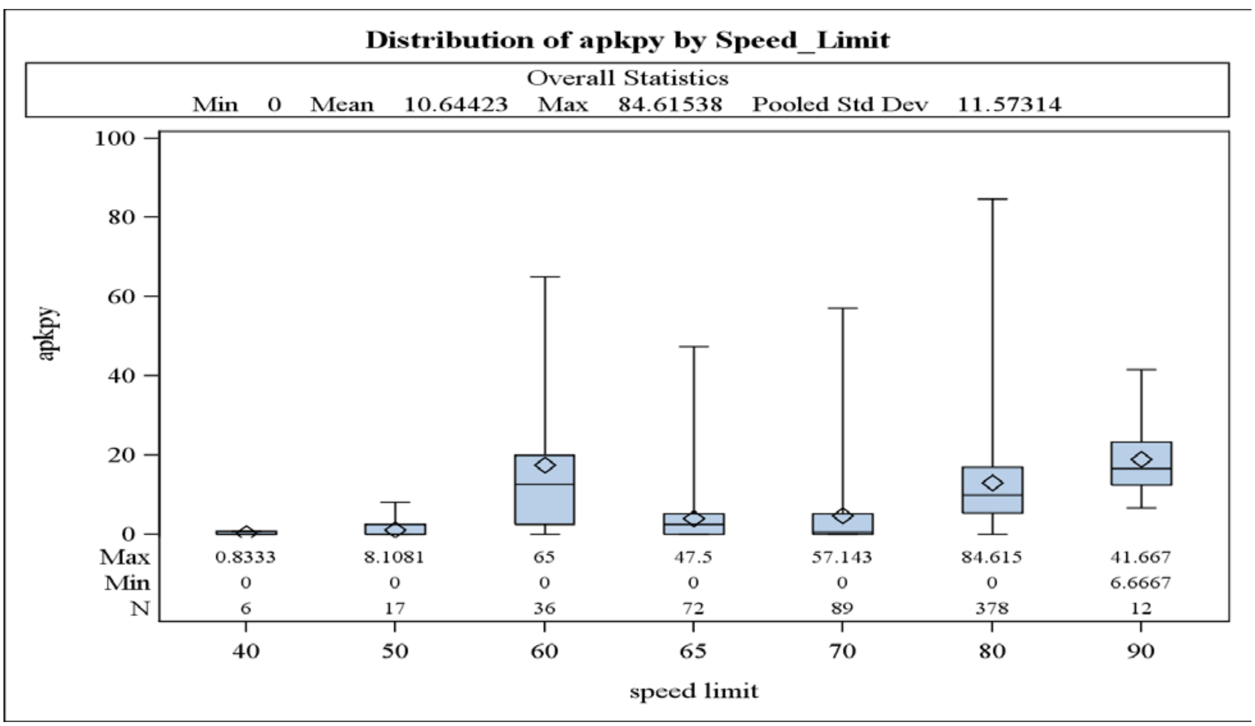


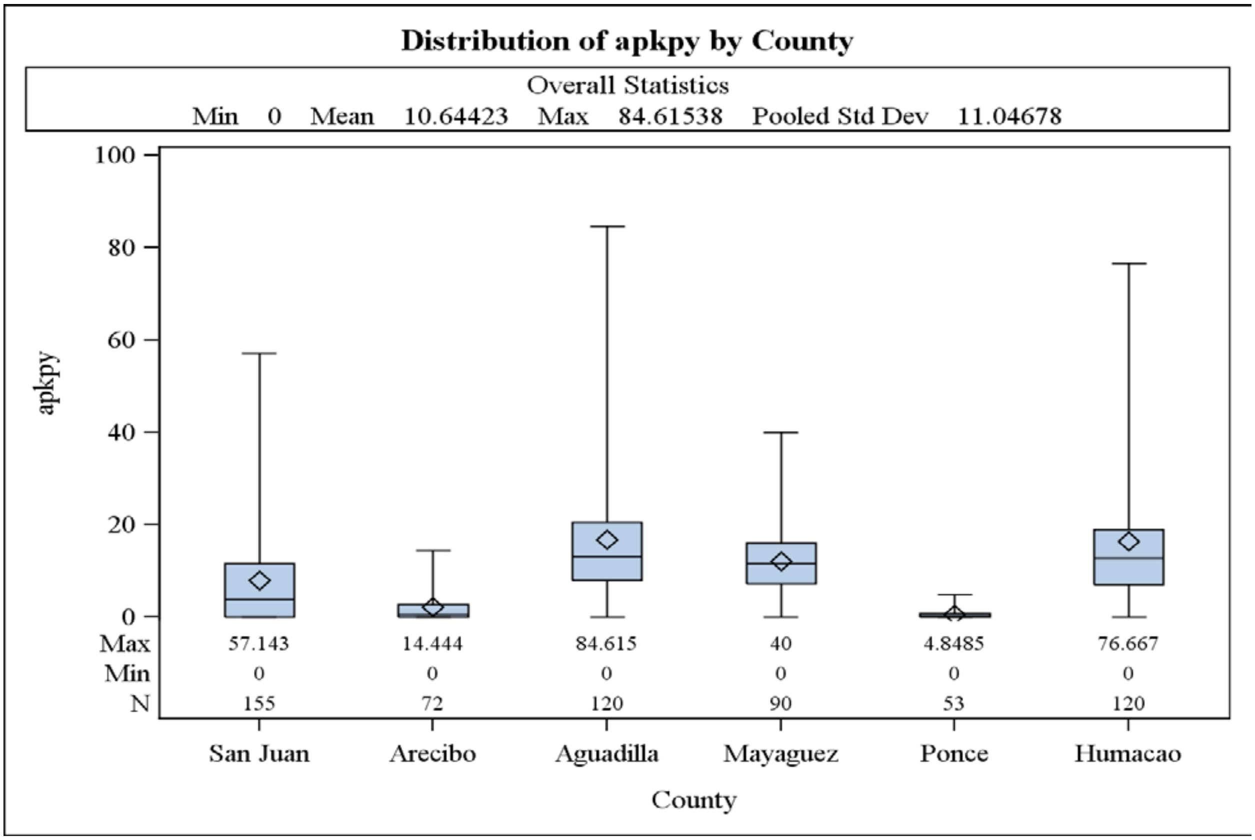
Distribution of apkpy by Rural_urban
apkpy
100
80
60
40
20
0
RURAL
URBAN
rural_urban code


Distribution of apkpy by Urban_code
Overall Statistics
Min 0 Mean 10.64423 Max 84.61538 Pooled Std Dev 11.5946
apkpy
100
80
60
40
20
0
Max 76.667 10 84.615 40.667 22.667 3.1818 40 65
Min 0 8.4615 0 0 0 0 0 0
N 93 2 84 48 48 18 84 233
.
0
AGUADILLA-ISABELA
ARECIBO
FAJARDO
FLORIDA-BARCELONETA
MAYAGÜEZ
SAN JUAN
urban code

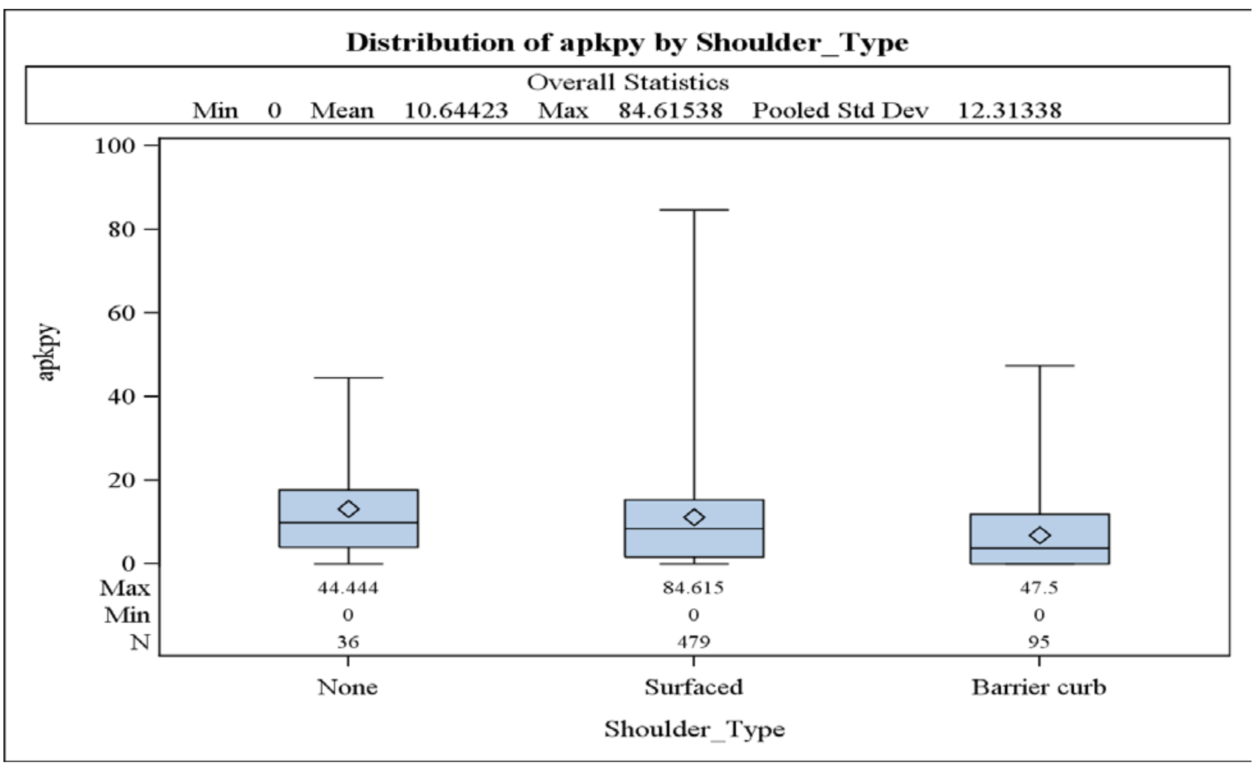

Distribution of apkpy by Horz_Align
Overall Statistics
Min 0 Mean 10.64423 Max 84.61538 Pooled Std Dev 12.22894
apkpy
100
80
60
40
20
0
Max 84.615 76.667 15.385
Min 0 0 0
N 557 47 6
Not required
Apropiate design
some curves below apropiate desi
horizontal alignment


Distribution of apkpy by Design_Speed
Overall Statistics
Min 0 Mean 10.64423 Max 84.61538 Pooled Std Dev 12.27328
apkpy
100
80
60
40
20
0
Max 1.1905 20 0 21.5 84.615 23.143 80
Min 0 2.5 0 12 0 0 0
N 6 6 6 6 30 66 490
65 70 80 90 100 105 110
design speed

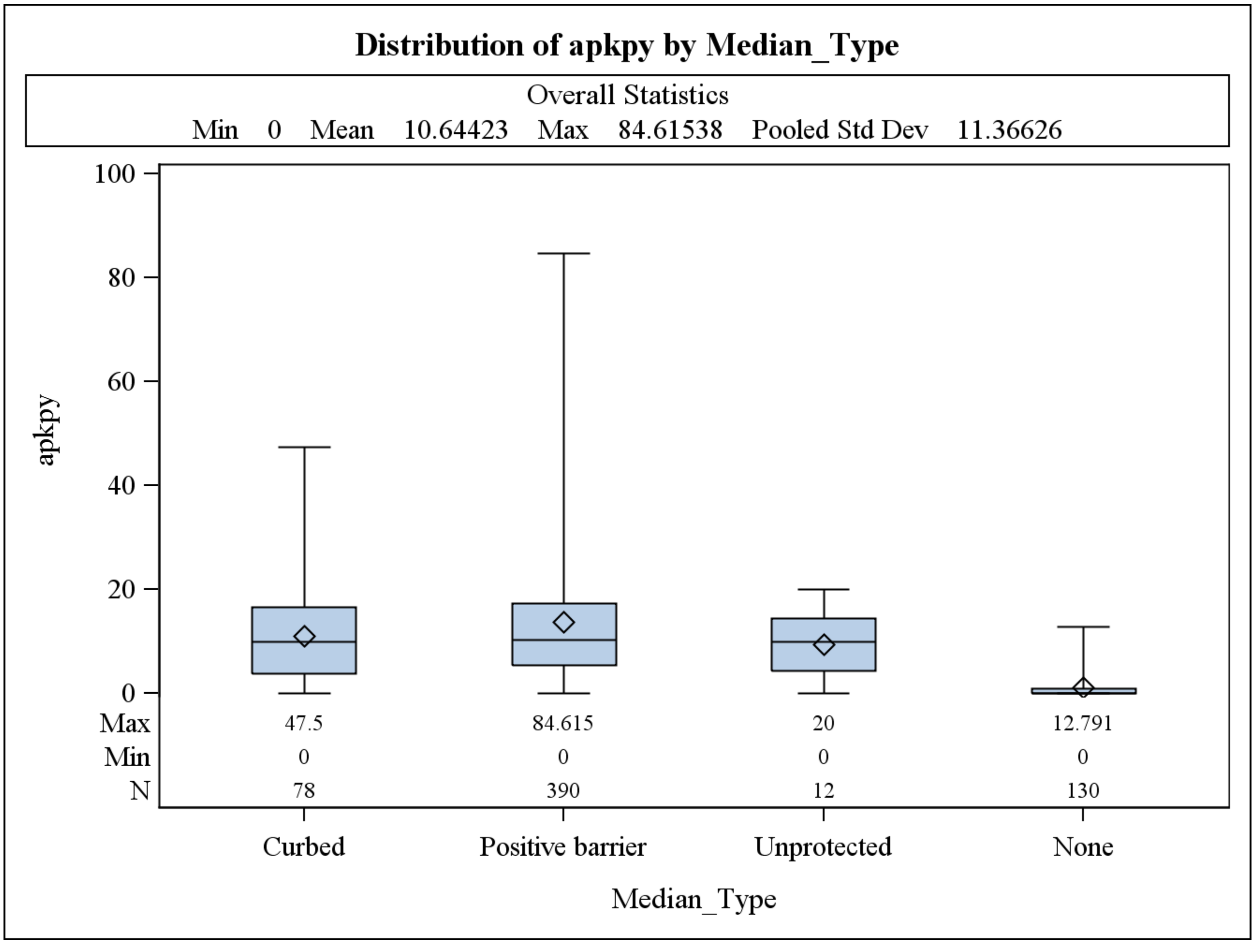

Distribution of apkpy by Speed_Limit
Overall Statistics
Min 0 Mean 10.64423 Max 84.61538 Pooled Std Dev 11.57314
apkpy
100
80
60
40
20
0
Max 0.8333 8.1081 65 47.5 57.143 84.615 41.667
Min 0 0 0 0 0 0 6.6667
N 6 17 36 72 89 378 12
40 50 60 65 70 80 90
speed limit

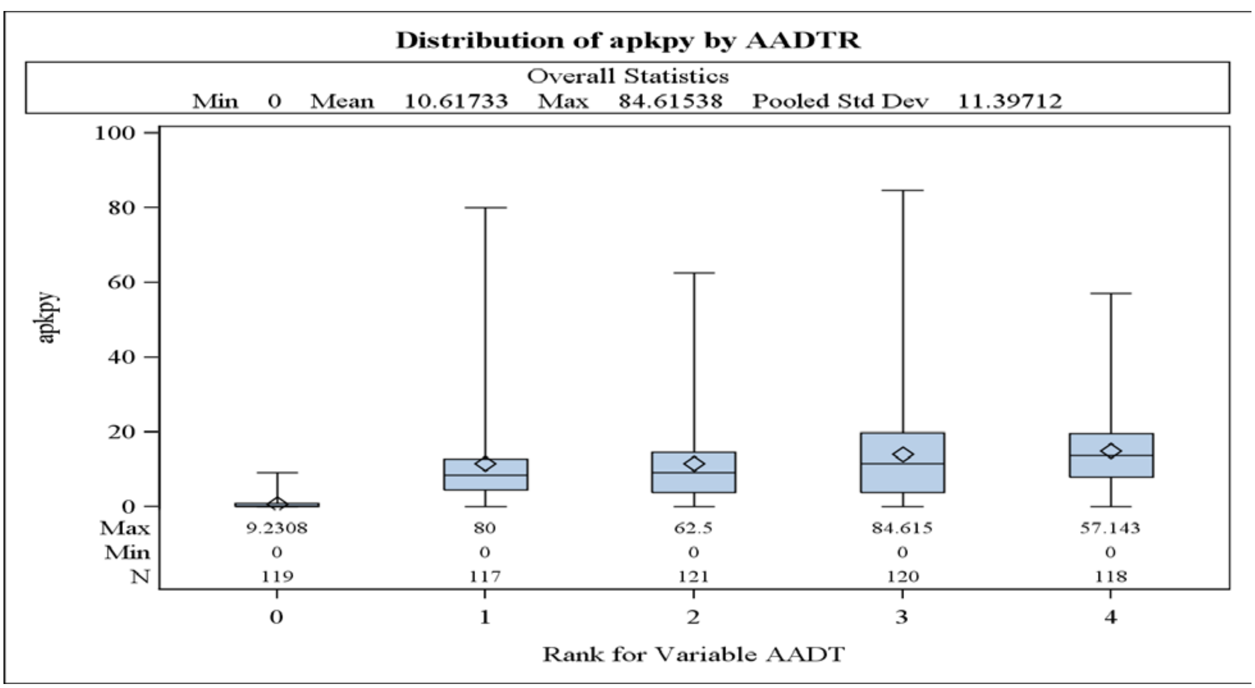


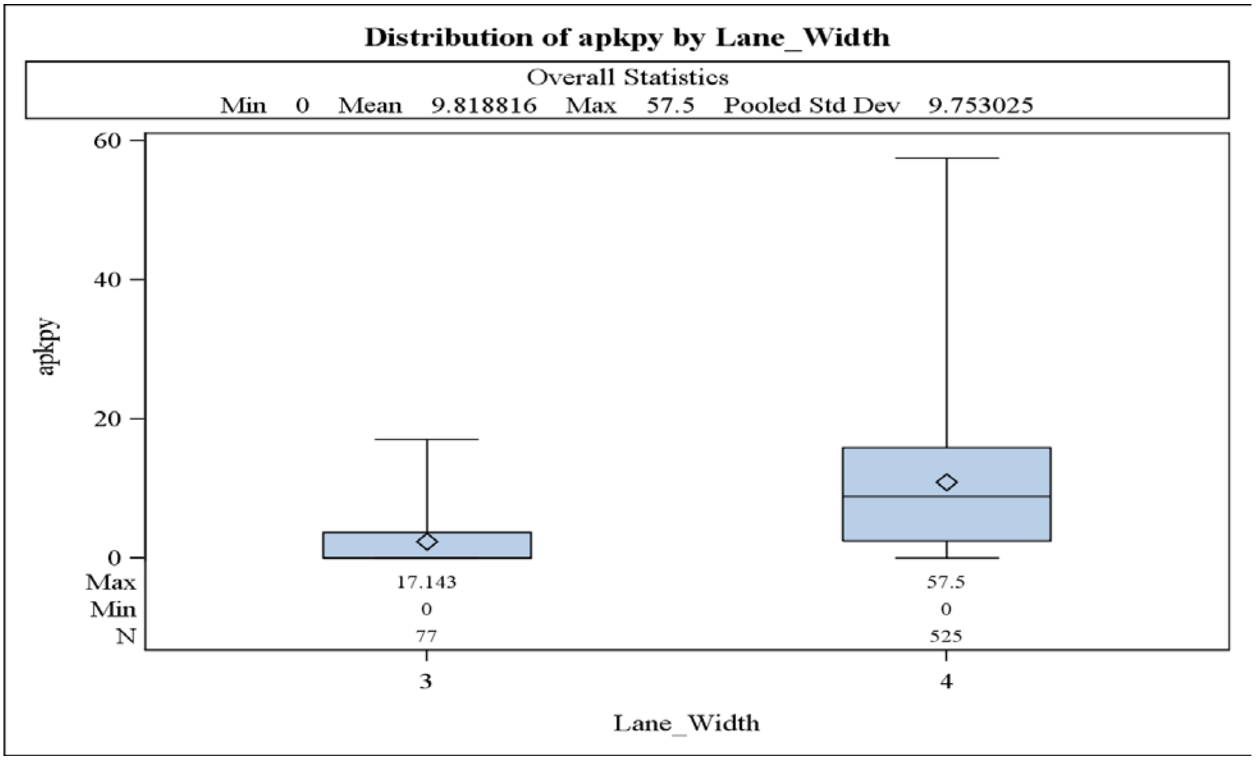
Distribution of apkpy by County
Overall Statistics
Min 0 Mean 10.64423 Max 84.61538 Pooled Std Dev 11.04678
apkpy
100
80
60
40
20
0
Max 57.143 14.444 84.615 40 4.8485 76.667
Min 0 0 0 0 0 0
N 155 72 120 90 53 120
San Juan
Arecibo
Aguadilla
Mayaguez
Ponce
Humacao
County

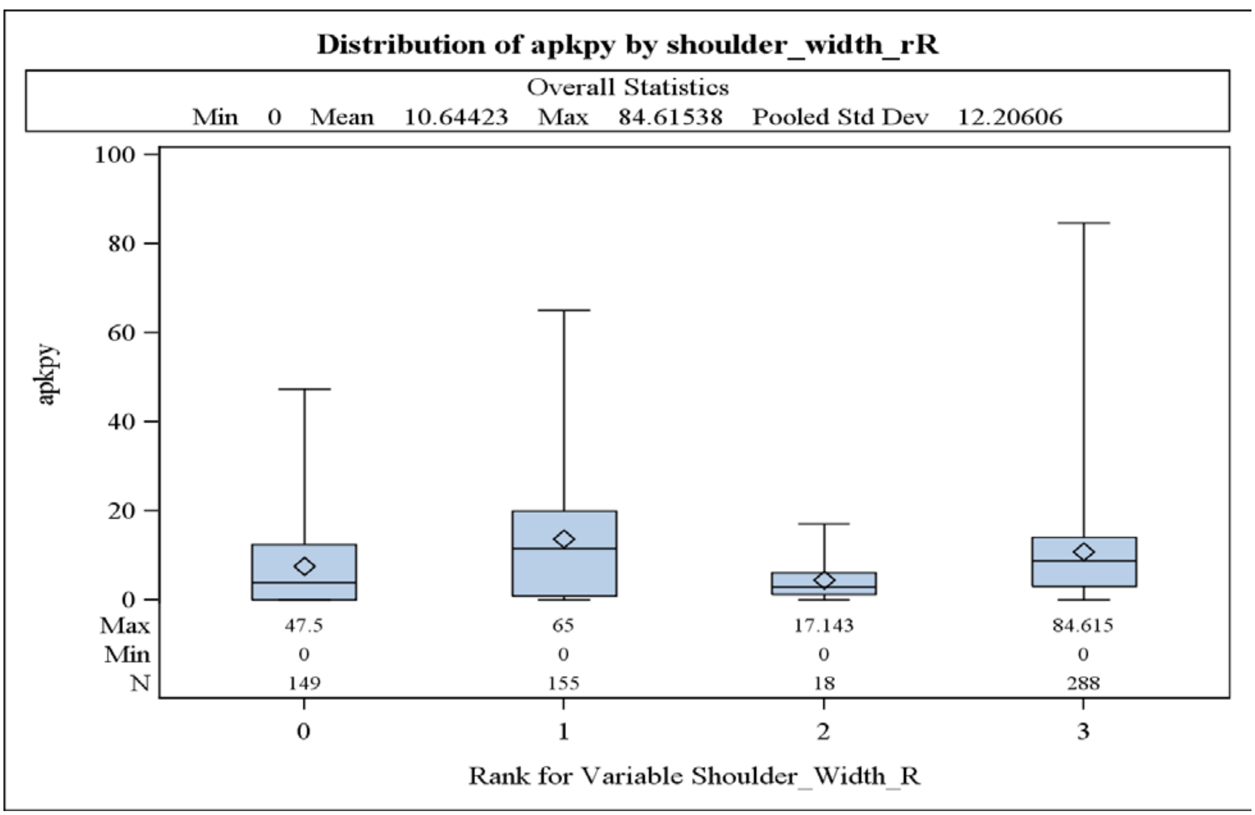


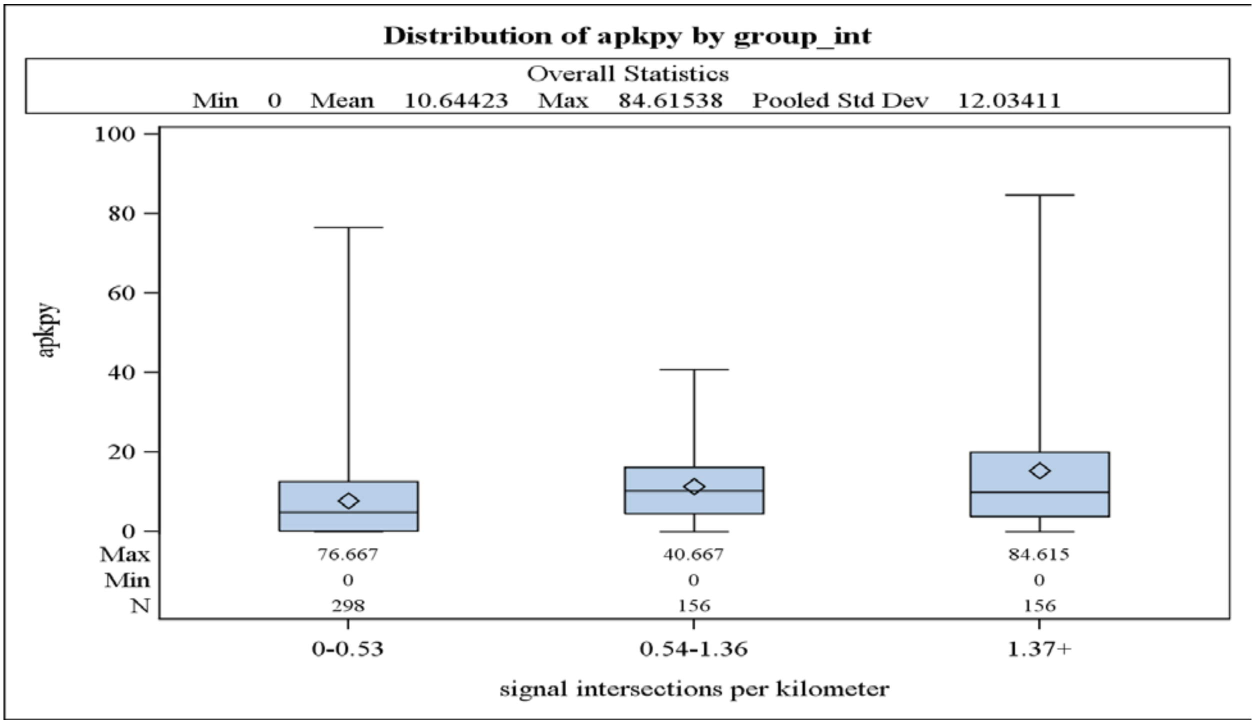
Distribution of apkpy by Shoulder_Type
Overall Statistics
Min 0 Mean 10.64423 Max 84.61538 Pooled Std Dev 12.31338
apkpy
100
80
60
40
20
0
Max 44.444 84.615 47.5
Min 0 0 0
N 36 479 95
None
Surfaced
Barrier curb
Shoulder_Type

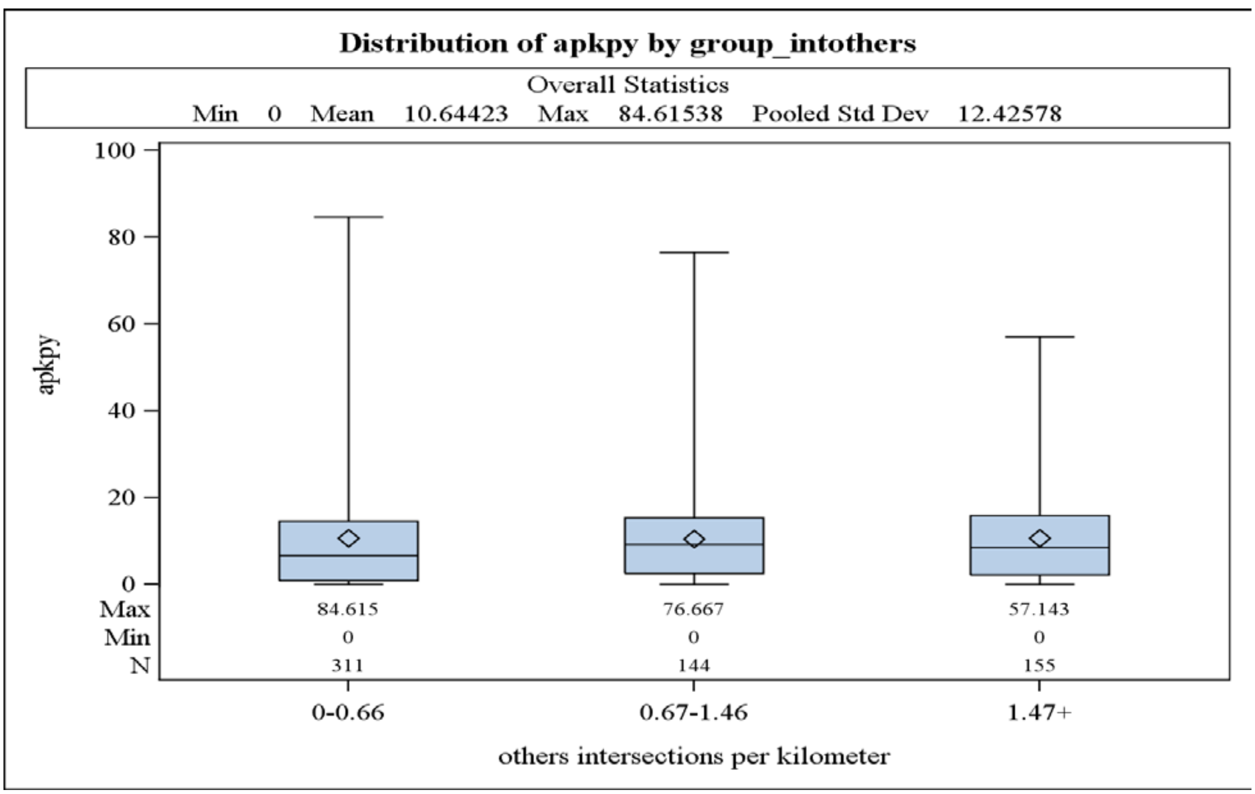

Distribution of apkpy by Median_Type
Overall Statistics
Min 0 Mean 10.64423 Max 84.61538 Pooled Std Dev 11.36626
apkpy
100
80
60
40
20
0
Max 47.5 84.615 20 12.791
Min 0 0 0 0
N 78 390 12 130
Curbed
Positive barrier
Unprotected
None
Median_Type


Distribution of apkpy by AADTR
Overall Statistics
Min 0 Mean 10.61733 Max 84.61538 Pooled Std Dev 11.39712
apkpy
100
80
60
40
20
0
Max 9.2308 80 62.5 84.615 57.143
Min 0 0 0 0 0
N 119 117 121 120 118
0
1
2
3
4
Rank for Variable AADT


Distribution of apkpy by Lane_Width
Overall Statistics
Min 0 Mean 9.818816 Max 57.5 Pooled Std Dev 9.753025
apkpy
60
40
20
0
Max 17.143 57.5
Min 0 0
N 77 525
3
4
Lane_Width


Distribution of apkpy by shoulder_width_rR
Overall Statistics
Min 0 Mean 10.64423 Max 84.61538 Pooled Std Dev 12.20606
apkpy
100
80
60
40
20
0
Max 47.5 65 17.143 84.615
Min 0 0 0 0
N 149 155 18 288
0
1
2
3
Rank for Variable Shoulder_Width_R


Distribution of apkpy by group_int
Overall Statistics
Min 0 Mean 10.64423 Max 84.61538 Pooled Std Dev 12.03411
apkpy
100
80
60
40
20
0
Max 76.667 40.667 84.615
Min 0 0 0
N 298 156 156
0-0.53
0.54-1.36
1.37+
signal intersections per kilometer


Distribution of apkpy by group_intothers
Overall Statistics
Min 0 Mean 10.64423 Max 84.61538 Pooled Std Dev 12.42578
apkpy
100
80
60
40
20
0
Max 84.615 76.667 57.143
Min 0 0 0
N 311 144 155
0-0.66
0.67-1.46
1.47+
others intersections per kilometer

3.1 CHAPTER 6 FREEWAY MODEL 1

3.1.1 TOTAL CRASHES-M1

Statistical expressway models

In These tables the región 1 is formed by 8 municipalities (Bayamon, Catano, Guaynabo, San Juan, Carolina, Loiza, Canovanas, Trujillo Alto). The región 2 is formed by 14 municipalities (Arecibo, Barceloneta, Manati, Vega Baja, Vega Alta, Dorado, Toa Baja, Toa Alta, Naranjito, Corozal, Morovis, Ciales, Utuado, Florida). The region 3 is formed by 10 Municipios (Aguadilla, Isabela, Quebradillas, Camuy, Hatillo, Lares, San Sebastian, Moca, Aguada, Rincón). The región 4 is formed by 11 municipios (Mayaguez, Anasco, las Marias, Maricao, Yauco, Guanica, Lajas, Cabo Rojo, Hormigueros, San German, Sabana Grande). The region 5 is formed by 10 municipios (Ponce, Juana Diaz, Santa Isabel, Coamo, Orocovis, Villalba, Jayuya, Adjuntas, Guayanilla, Penuelas). The región 6 is formed by 10 municipios (Guayama, Arroyo, Patillas, Cayey, Cidra, Aguas Buenas, Comerio, Barranquitas, Aibonito, Salinas). The región 7 is formed by 13 municipios (Humacao, Naguabo, Ceiba, Fajardo, Luquillo, rio grande, las piedras, juncos, Gurabo, Caguas, San Lorenzo, Yabucoa, Murabo).

Model 1 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	Choques_totales
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	5047.82
AIC (smaller is better)	5065.82
AICC (smaller is better)	5066.06
BIC (smaller is better)	5091.42
CAIC (smaller is better)	5100.42
HQIC (smaller is better)	5076.22

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totales | r. effects)	4663.70
Pearson Chi-Square	638.19
Pearson Chi-Square / DF	0.84

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.2100	0.03320
Scale	0.07806	0.008393

Solutions for Fixed Effects							
Effect	number of lanes	Shoulder_Type	Estimate	Standard Error	DF	t Val ue	Pr > |t|
Intercept			-8.0560	1.1920	121	-6.76	<.0001
logAADT			0.9575	0.09841	631	9.73	<.0001
Lanes	4		0.2190	0.1930	631	1.13	0.2570
Lanes	6		0.3960	0.1646	631	2.41	0.0165
Lanes	7		0	.	.	.	.
Shoulder_Type		1	-0.7754	0.4154	631	-1.87	0.0624
Shoulder_Type		2	-0.7652	0.2129	631	-3.59	0.0004
Shoulder_Type		3	-0.3394	0.4081	631	-0.83	0.4059
Shoulder_Type		6	0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	94.66	<.0001
Lanes	2	631	3.26	0.0389
Shoulder_Type	3	631	4.81	0.0026

3.1.2 INJURY-M1

The GLIMMIX Procedure

Model 1 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	No_fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	5047.31
AIC (smaller is better)	5065.31
AICC (smaller is better)	5065.55
BIC (smaller is better)	5090.91
CAIC (smaller is better)	5099.91
HQIC (smaller is better)	5075.71

Fit Statistics for Conditional Distribution	
-2 log L(No_fatal | r. effects)	4666.07
Pearson Chi-Square	641.59
Pearson Chi-Square / DF	0.85

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.2116	0.03345
Scale	0.08137	0.008754

Solutions for Fixed Effects							
Effect	number of lanes	Shoulder_Type	Estimate	Standard Error	DF	t Valu e	Pr > |t|
Intercept			-8.1702	1.2003	121	-6.81	<.0001
logAADT			0.9664	0.09910	631	9.75	<.0001
Lanes	4		0.2241	0.1942	631	1.15	0.2490
Lanes	6		0.3917	0.1655	631	2.37	0.0183
Lanes	7		0	.	.	.	.
Shoulder_Type		1	-0.7642	0.4173	631	-1.83	0.0676
Shoulder_Type		2	-0.7689	0.2139	631	-3.59	0.0004
Shoulder_Type		3	-0.3331	0.4100	631	-0.81	0.4169
Shoulder_Type		6	0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	95.09	<.0001
Lanes	2	631	3.12	0.0450
Shoulder_Type	3	631	4.82	0.0025

3.1.3 FATAL-M1

Model 1 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	Fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	1004.01
AIC (smaller is better)	1022.01
AICC (smaller is better)	1022.25
BIC (smaller is better)	1047.61
CAIC (smaller is better)	1056.61
HQIC (smaller is better)	1032.41

Fit Statistics for Conditional Distribution	
-2 log L(Fatal | r. effects)	983.84
Pearson Chi-Square	849.79
Pearson Chi-Square / DF	1.12

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.05081	0.08891
Scale	0.2384	0.1551

Solutions for Fixed Effects							
Effect	number of lanes	Shoulder_Type	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			-11.3324	2.2963	121	-4.94	<.0001
logAADT			0.8324	0.1902	631	4.38	<.0001
Lanes	4		0.5197	0.3616	631	1.44	0.1512
Lanes	6		0.6920	0.2961	631	2.34	0.0197
Lanes	7		0	.	.	.	.
Shoulder_Type		1	-0.9430	0.8467	631	-1.11	0.2658
Shoulder_Type		2	-0.4307	0.3755	631	-1.15	0.2518
Shoulder_Type		3	-0.6600	0.6080	631	-1.09	0.2781
Shoulder_Type		6	0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	19.15	<.0001
Lanes	2	631	2.90	0.0557
Shoulder_Type	3	631	0.65	0.5819

3.2 CHAPTER 6 FREEWAY MODEL 2

3.2.1 TOTAL CRASHSE-M2

Model 2 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	Choques_totales
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	5054.12
AIC (smaller is better)	5068.12
AICC (smaller is better)	5068.27
BIC (smaller is better)	5088.03
CAIC (smaller is better)	5095.03
HQIC (smaller is better)	5076.21

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totales | r. effects)	4660.53
Pearson Chi-Square	633.29
Pearson Chi-Square / DF	0.83

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.2261	0.03477
Scale	0.07766	0.008326

Solutions for Fixed Effects						
Effect	Shoulder_Type	Estimate	Standard Error	DF	t Valu e	Pr > |t|
Intercept		-7.5270	0.7565	123	-9.95	<.0001
logAADT		0.9177	0.06369	631	14.41	<.0001
Shoulder_Type	1	-0.4562	0.4093	631	-1.11	0.2655
Shoulder_Type	2	-0.6298	0.2104	631	-2.99	0.0029
Shoulder_Type	3	-0.1015	0.4068	631	-0.25	0.8030
Shoulder_Type	6	0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	207.64	<.0001
Shoulder_Type	3	631	3.64	0.0126

3.2.2 INJURY –M2

The GLIMMIX Procedure

Model Information	
Data Set	LB.AUTOPISTAS2
Response Variable	No_fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	5053.33
AIC (smaller is better)	5067.33
AICC (smaller is better)	5067.48
BIC (smaller is better)	5087.24
CAIC (smaller is better)	5094.24
HQIC (smaller is better)	5075.42

The GLIMMIX Procedure

Fit Statistics for Conditional Distribution	
-2 log L(No_fatal | r. effects)	4662.80
Pearson Chi-Square	636.63
Pearson Chi-Square / DF	0.84

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.2274	0.03498
Scale	0.08092	0.008682

Solutions for Fixed Effects						
Effect	Shoulder_Type	Estimate	Standard Error	DF	t Valu e	Pr > |t|
Intercept		-7.6034	0.7602	123	- 10.00	<.0001
logAADT		0.9235	0.06400	631	14.43	<.0001
Shoulder_Type	1	-0.4512	0.4108	631	-1.10	0.2725
Shoulder_Type	2	-0.6343	0.2112	631	-3.00	0.0028
Shoulder_Type	3	-0.09663	0.4083	631	-0.24	0.8130
Shoulder_Type	6	0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	208.19	<.0001
Shoulder_Type	3	631	3.69	0.0119

3.2.3 FATAL-M2

Model 2 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	Fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	1010.07
AIC (smaller is better)	1024.07
AICC (smaller is better)	1024.22
BIC (smaller is better)	1043.98
CAIC (smaller is better)	1050.98
HQIC (smaller is better)	1032.16

Fit Statistics for Conditional Distribution	
-2 log L(Fatal | r. effects)	985.88
Pearson Chi-Square	817.80
Pearson Chi-Square / DF	1.08

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.06164	0.08891
Scale	0.2372	0.1559

Solutions for Fixed Effects						
Effect	Shoulder_Type	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		-9.7660	1.3681	123	-7.14	<.0001
logAADT		0.7063	0.1133	631	6.23	<.0001
Shoulder_Type	1	-0.3866	0.8188	631	-0.47	0.6370
Shoulder_Type	2	-0.1006	0.3488	631	-0.29	0.7731
Shoulder_Type	3	-0.09281	0.5690	631	-0.16	0.8705
Shoulder_Type	6	0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	38.86	<.0001
Shoulder_Type	3	631	0.08	0.9718

3.3 CHAPTER 6 FREEWAY MODEL 3

3.3.1 TOTAL CRASHES-M3

The GLIMMIX Procedure

Model 3 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	Choques_totales
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	5033.55
AIC (smaller is better)	5063.55
AICC (smaller is better)	5064.20
BIC (smaller is better)	5106.21
CAIC (smaller is better)	5121.21
HQIC (smaller is better)	5080.88

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totales | r. effects)	4675.86
Pearson Chi-Square	643.07
Pearson Chi-Square / DF	0.85

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.1711	0.02825
Scale	0.07953	0.008523

Solutions for Fixed Effects								
Effect	County	number of lanes	Shoulder_Type	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept				-9.3641	1.2235	115	-7.65	<.0001
logAADT				1.0921	0.1010	631	10.82	<.0001
Shoulder_Type			1	-0.7570	0.3813	631	-1.99	0.0476
Shoulder_Type			2	-0.8484	0.2021	631	-4.20	<.0001
Shoulder_Type			3	-0.5991	0.3884	631	-1.54	0.1234
Shoulder_Type			6	0	.	.	.	.
Lanes		4		0.2565	0.2298	631	1.12	0.2647
Lanes		6		0.4048	0.1697	631	2.39	0.0173
Lanes		7		0	.	.	.	.
County	1			-0.2448	0.1514	631	-1.62	0.1063
County	2			-0.3877	0.1384	631	-2.80	0.0053
County	3			-0.5265	0.4558	631	-1.16	0.2484
County	4			0.005895	0.1614	631	0.04	0.9709
County	5			0.1502	0.1385	631	1.08	0.2784
County	6			-0.1221	0.2317	631	-0.53	0.5983
County	7			0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	117.02	<.0001
Shoulder_Type	3	631	6.13	0.0004
Lanes	2	631	3.47	0.0317
County	6	631	2.66	0.0150

3.3.2 INJURY-M3

Model 3 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	No_fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	5032.53
AIC (smaller is better)	5062.53
AICC (smaller is better)	5063.18
BIC (smaller is better)	5105.19
CAIC (smaller is better)	5120.19
HQIC (smaller is better)	5079.86

Fit Statistics for Conditional Distribution	
-2 log L(No_fatal | r. effects)	4678.60
Pearson Chi-Square	646.46
Pearson Chi-Square / DF	0.85

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.1710	0.02822
Scale	0.08297	0.008898

Solutions for Fixed Effects								
Effect	County	number of lanes	Shoulder_Type	Estimate	Standard Error	DF	t Valu e	Pr > |t|
Intercept				-9.5078	1.2282	115	-7.74	<.0001
logAADT				1.1035	0.1014	631	10.89	<.0001
Shoulder_Type			1	-0.7443	0.3818	631	-1.95	0.0517
Shoulder_Type			2	-0.8528	0.2024	631	-4.21	<.0001
Shoulder_Type			3	-0.6002	0.3888	631	-1.54	0.1232
Shoulder_Type			6	0	.	.	.	.
Lanes		4		0.2656	0.2303	631	1.15	0.2492
Lanes		6		0.4030	0.1700	631	2.37	0.0181
Lanes		7		0	.	.	.	.
County	1			-0.2446	0.1516	631	-1.61	0.1073
County	2			-0.3938	0.1387	631	-2.84	0.0047
County	3			-0.5231	0.4565	631	-1.15	0.2523
County	4			0.000566	0.1618	631	0.00	0.9972
County	5			0.1610	0.1387	631	1.16	0.2461
County	6			-0.1267	0.2320	631	-0.55	0.5851
County	7			0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	118.50	<.0001
Shoulder_Type	3	631	6.18	0.0004
Lanes	2	631	3.35	0.0356
County	6	631	2.76	0.0117

3.3.3 FATAL-M3

Model 3 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	Fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Estimated G matrix is not positive definite.

Fit Statistics	
-2 Log Likelihood	995.66
AIC (smaller is better)	1023.66
AICC (smaller is better)	1024.23
BIC (smaller is better)	1063.48
CAIC (smaller is better)	1077.48
HQIC (smaller is better)	1039.84

Fit Statistics for Conditional Distribution	
-2 log L(Fatal | r. effects)	995.66
Pearson Chi-Square	924.25
Pearson Chi-Square / DF	1.22

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	5.49E-19	.
Scale	0.2244	0.1431

Solutions for Fixed Effects								
Effect	County	number of lanes	Shoulder_Type	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept				-12.4740	2.3742	115	-5.25	<.0001
logAADT				0.9399	0.1984	631	4.74	<.0001
Shoulder_Type			1	-0.8577	0.8232	631	-1.04	0.2979
Shoulder_Type			2	-0.4312	0.3610	631	-1.19	0.2327
Shoulder_Type			3	-0.7462	0.5754	631	-1.30	0.1951
Shoulder_Type			6	0	.	.	.	.
Lanes		4		0.7054	0.4154	631	1.70	0.0900
Lanes		6		0.6868	0.3146	631	2.18	0.0294
Lanes		7		0	.	.	.	.
County	1			-0.1427	0.2552	631	-0.56	0.5761
County	2			-0.2443	0.2125	631	-1.15	0.2508
County	3			-0.8413	1.0294	631	-0.82	0.4141
County	4			0.07123	0.3001	631	0.24	0.8125
County	5			-0.6717	0.2981	631	-2.25	0.0246
County	6			0.1196	0.2667	631	0.45	0.6541
County	7			0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	22.43	<.0001
Shoulder_Type	3	631	0.75	0.5223
Lanes	2	631	2.40	0.0916
County	6	631	1.34	0.2373

3.4 CHAPTER 6 FREEWAY MODEL 4

3.4.1 TOTAL CRASHES-M4

Model 4 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	Choques_totales
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	5040.28
AIC (smaller is better)	5066.28
AICC (smaller is better)	5066.76
BIC (smaller is better)	5103.25
CAIC (smaller is better)	5116.25
HQIC (smaller is better)	5081.30

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totales | r. effects)	4671.60
Pearson Chi-Square	636.99
Pearson Chi-Square / DF	0.84

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.1868	0.02952
Scale	0.07915	0.008459

Solutions for Fixed Effects							
Effect	County	Shoulder_Type	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			-8.7580	0.8333	117	-10.51	<.0001
logAADT			1.0549	0.07685	631	13.73	<.0001
Shoulder_Type		1	-0.4565	0.3761	631	-1.21	0.2252
Shoulder_Type		2	-0.7552	0.2065	631	-3.66	0.0003
Shoulder_Type		3	-0.4361	0.3991	631	-1.09	0.2750
Shoulder_Type		6	0	.	.	.	.
County	1		-0.3370	0.1399	631	-2.41	0.0163
County	2		-0.3920	0.1358	631	-2.89	0.0040
County	3		-0.5776	0.4717	631	-1.22	0.2212
County	4		-0.04862	0.1648	631	-0.30	0.7681
County	5		0.1022	0.1378	631	0.74	0.4587
County	6		-0.1243	0.2405	631	-0.52	0.6055
County	7		0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	188.42	<.0001
Shoulder_Type	3	631	4.88	0.0023
County	6	631	2.53	0.0197

3.4.2 INJURY-M4

Model 4 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	No_fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	5039.03
AIC (smaller is better)	5065.03
AICC (smaller is better)	5065.51
BIC (smaller is better)	5102.00
CAIC (smaller is better)	5115.00
HQIC (smaller is better)	5080.05

Fit Statistics for Conditional Distribution	
-2 log L(No_fatal | r. effects)	4674.11
Pearson Chi-Square	640.49
Pearson Chi-Square / DF	0.84

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.1864	0.02946
Scale	0.08249	0.008821

Solutions for Fixed Effects							
Effect	County	Shoulder_Type	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			-8.8555	0.8354	117	-10.60	<.0001
logAADT			1.0626	0.07706	631	13.79	<.0001
Shoulder_Type		1	-0.4528	0.3763	631	-1.20	0.2294
Shoulder_Type		2	-0.7611	0.2067	631	-3.68	0.0003
Shoulder_Type		3	-0.4329	0.3995	631	-1.08	0.2789
Shoulder_Type		6	0	.	.	.	.
County	1		-0.3381	0.1401	631	-2.41	0.0161
County	2		-0.3959	0.1360	631	-2.91	0.0037
County	3		-0.5714	0.4722	631	-1.21	0.2267
County	4		-0.05213	0.1651	631	-0.32	0.7523
County	5		0.1156	0.1379	631	0.84	0.4022
County	6		-0.1263	0.2407	631	-0.52	0.5998
County	7		0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	190.16	<.0001
Shoulder_Type	3	631	4.96	0.0021
County	6	631	2.63	0.0159

3.4.3 FATAL-M4

Model 4 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	Fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Estimated G matrix is not positive definite.

Fit Statistics	
-2 Log Likelihood	1000.45
AIC (smaller is better)	1024.45
AICC (smaller is better)	1024.87
BIC (smaller is better)	1058.58
CAIC (smaller is better)	1070.58
HQIC (smaller is better)	1038.31

Fit Statistics for Conditional Distribution	
-2 log L(Fatal | r. effects)	1000.45
Pearson Chi-Square	925.29
Pearson Chi-Square / DF	1.22

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	1.22E-21	.
Scale	0.2297	0.1442

Solutions for Fixed Effects							
Effect	County	Shoulder_Type	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			-10.3407	1.6437	117	-6.29	<.0001
logAADT			0.7871	0.1478	631	5.32	<.0001
Shoulder_Type		1	-0.3283	0.7874	631	-0.42	0.6768
Shoulder_Type		2	-0.1926	0.3409	631	-0.56	0.5723
Shoulder_Type		3	-0.4159	0.5510	631	-0.75	0.4506
Shoulder_Type		6	0	.	.	.	.
County	1		-0.3715	0.2427	631	-1.53	0.1263
County	2		-0.2099	0.1987	631	-1.06	0.2912
County	3		-0.8983	1.0293	631	-0.87	0.3832
County	4		0.000458	0.2934	631	0.00	0.9988
County	5		-0.6979	0.2888	631	-2.42	0.0159
County	6		0.09855	0.2619	631	0.38	0.7068
County	7		0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	28.35	<.0001
Shoulder_Type	3	631	0.21	0.8860
County	6	631	1.56	0.1564

3.5 CHAPTER 6 FREEWAY MODEL 5

3.5.1 TOTAL CRASHES-M5

Model 5 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	Choques_totales
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	5054.06
AIC (smaller is better)	5074.06
AICC (smaller is better)	5074.36
BIC (smaller is better)	5102.51
CAIC (smaller is better)	5112.51
HQIC (smaller is better)	5085.62

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totales | r. effects)	4667.51
Pearson Chi-Square	635.14
Pearson Chi-Square / DF	0.84

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.2151	0.03344
Scale	0.07875	0.008424

Solutions for Fixed Effects						
Effect	County	Estimate	Standard Error	DF	t Valu e	Pr > |t|
Intercept		-9.1392	0.8581	120	-10.65	<.0001
logAADT		1.0219	0.07920	631	12.90	<.0001
County	1	-0.1909	0.1348	631	-1.42	0.1573
County	2	-0.4003	0.1422	631	-2.81	0.0050
County	3	-0.6093	0.5012	631	-1.22	0.2246
County	4	-0.08181	0.1743	631	-0.47	0.6390
County	5	0.07423	0.1452	631	0.51	0.6095
County	6	-0.1460	0.2555	631	-0.57	0.5678
County	7	0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	166.51	<.0001
County	6	631	1.91	0.0765

3.5.2 INJURY-M5

Model 5 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	No_fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	5053.01
AIC (smaller is better)	5073.01
AICC (smaller is better)	5073.31
BIC (smaller is better)	5101.46
CAIC (smaller is better)	5111.46
HQIC (smaller is better)	5084.57

Fit Statistics for Conditional Distribution	
-2 log L(No_fatal | r. effects)	4669.91
Pearson Chi-Square	638.46
Pearson Chi-Square / DF	0.84

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.2156	0.03354
Scale	0.08206	0.008785

Solutions for Fixed Effects						
Effect	County	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		-9.2382	0.8617	120	-10.72	<.0001
logAADT		1.0293	0.07952	631	12.94	<.0001
County	1	-0.1912	0.1353	631	-1.41	0.1580
County	2	-0.4049	0.1426	631	-2.84	0.0047
County	3	-0.6038	0.5026	631	-1.20	0.2301
County	4	-0.08679	0.1749	631	-0.50	0.6199
County	5	0.08656	0.1456	631	0.59	0.5525
County	6	-0.1503	0.2562	631	-0.59	0.5577
County	7	0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	167.53	<.0001
County	6	631	1.98	0.0658

3.5.3 FATAL-M5

The GLIMMIX Procedure

Model 5 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	Fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Estimated G matrix is not positive definite.

Fit Statistics	
-2 Log Likelihood	1001.06
AIC (smaller is better)	1019.06
AICC (smaller is better)	1019.30
BIC (smaller is better)	1044.66
CAIC (smaller is better)	1053.66
HQIC (smaller is better)	1029.46

Fit Statistics for Conditional Distribution	
-2 log L(Fatal | r. effects)	1001.06
Pearson Chi-Square	932.07
Pearson Chi-Square / DF	1.23

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0	.
Scale	0.2308	0.1444

Solutions for Fixed Effects						
Effect	County	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		-10.3623	1.5745	120	-6.58	<.0001
logAADT		0.7697	0.1426	631	5.40	<.0001
County	1	-0.3063	0.2168	631	-1.41	0.1581
County	2	-0.1847	0.1931	631	-0.96	0.3393
County	3	-0.8856	1.0292	631	-0.86	0.3898
County	4	0.007543	0.2930	631	0.03	0.9795
County	5	-0.6868	0.2877	631	-2.39	0.0173
County	6	0.1097	0.2603	631	0.42	0.6736
County	7	0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
logAADT	1	631	29.14	<.0001
County	6	631	1.49	0.1774

3.6 CHAPTER 6 FREEWAY MODEL 6

3.6.1 TOTAL CRASHES-M6

Model 6 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	Choques_totales
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

The GLIMMIX Procedure

Fit Statistics	
-2 Log Likelihood	5064.05
AIC (smaller is better)	5074.05
AICC (smaller is better)	5074.12
BIC (smaller is better)	5088.27
CAIC (smaller is better)	5093.27
HQIC (smaller is better)	5079.82

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totales | r. effects)	4662.13
Pearson Chi-Square	633.29
Pearson Chi-Square / DF	0.83

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.2366	0.03825
County	0.008574	0.01426
Scale	0.07819	0.008377

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-8.7114	0.8501	0	-10.25	.
logAADT	0.9710	0.07695	631	12.62	<.0001

3.6.2 INJURY-M6

Model 6 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	No_fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

The GLIMMIX Procedure

Fit Statistics	
-2 Log Likelihood	5064.00
AIC (smaller is better)	5072.00
AICC (smaller is better)	5072.05
BIC (smaller is better)	5083.38
CAIC (smaller is better)	5087.38
HQIC (smaller is better)	5076.62

Fit Statistics for Conditional Distribution	
-2 log L(No_fatal | r. effects)	4662.01
Pearson Chi-Square	636.17
Pearson Chi-Square / DF	0.84

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.2500	0.03794
Scale	0.08118	0.008700

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-8.4901	0.7267	126	-11.68	<.0001
logAADT	0.9501	0.06571	631	14.46	<.0001

3.6.3 FATAL-M6

Model 6 Information	
Data Set	LB.AUTOPISTAS2
Response Variable	Fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	759
Number of Observations Used	759

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	1010.32
AIC (smaller is better)	1018.32
AICC (smaller is better)	1018.37
BIC (smaller is better)	1029.69
CAIC (smaller is better)	1033.69
HQIC (smaller is better)	1022.94

Fit Statistics for Conditional Distribution	
-2 log L(Fatal | r. effects)	986.97
Pearson Chi-Square	826.29
Pearson Chi-Square / DF	1.09

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error
Section_ID	0.05922	0.08814
Scale	0.2376	0.1560

Solutions for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-9.9148	1.2324	126	-8.05	<.0001
logAADT	0.7109	0.1101	631	6.46	<.0001

3.7 CHAPTER 6 ARTERIAL MODEL 1

3.7.1 TOTAL CRASHES-M1

Model 4 Information	
Data Set	LB.ARTERIAS2GLMSELECT
Response Variable	Choques_totales
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	599
Number of Observations Used	583

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3461.27
AIC (smaller is better)	3483.27
AICC (smaller is better)	3483.73
BIC (smaller is better)	3511.93
CAIC (smaller is better)	3522.93
HQIC (smaller is better)	3494.87

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totales | r. effects)	3186.13
Pearson Chi-Square	575.36
Pearson Chi-Square / DF	0.99

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.4008	0.07428
Scale	0.1827	0.02229

Solutions for Fixed Effects							
Effect	County	signal intersections per kilometer	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			-8.7525	1.5669	92	-5.59	<.0001
County	1		-1.1610	0.2127	482	-5.46	<.0001
County	2		-1.5458	0.3043	482	-5.08	<.0001
County	3		0.1699	0.2139	482	0.79	0.4274
County	4		-0.4721	0.2343	482	-2.01	0.0445
County	5		-0.4720	0.4723	482	-1.00	0.3181
County	7		0	.	.	.	.
logAADT			1.0854	0.1441	482	7.53	<.0001
group_int		1	-0.4309	0.1800	482	-2.39	0.0170
group_int		2	-0.2928	0.1946	482	-1.50	0.1331
group_int		3	0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
County	5	482	14.51	<.0001
logAADT	1	482	56.70	<.0001
group_int	2	482	2.88	0.0569

3.7.2 INJURY-M1

Model 1 Information	
Data Set	LB.ARTERIAS2GLMSELECT
Response Variable	No_fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	599
Number of Observations Used	583

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3433.31
AIC (smaller is better)	3455.31
AICC (smaller is better)	3455.78
BIC (smaller is better)	3483.97
CAIC (smaller is better)	3494.97
HQIC (smaller is better)	3466.91

Fit Statistics for Conditional Distribution	
-2 log L(No_fatal | r. effects)	3161.46
Pearson Chi-Square	580.17
Pearson Chi-Square / DF	1.00

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.3923	0.07279
Scale	0.1830	0.02239

Solutions for Fixed Effects							
Effect	County	signal intersections per kilometer	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			-9.0637	1.5754	92	-5.75	<.0001
County	1		-1.1669	0.2110	482	-5.53	<.0001
County	2		-1.5212	0.3026	482	-5.03	<.0001
County	3		0.1817	0.2119	482	0.86	0.3918
County	4		-0.5206	0.2327	482	-2.24	0.0257
County	5		-0.4630	0.4712	482	-0.98	0.3263
County	7		0	.	.	.	.
logAADT			1.1127	0.1450	482	7.68	<.0001
group_int		1	-0.4280	0.1785	482	-2.40	0.0169
group_int		2	-0.3045	0.1930	482	-1.58	0.1153
group_int		3	0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
County	5	482	14.78	<.0001
logAADT	1	482	58.91	<.0001
group_int	2	482	2.91	0.0554

3.7.3 FATAL-M1

Model 1 Information	
Data Set	LB.ARTERIAS2GLMSELECT
Response Variable	Fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	599
Number of Observations Used	583

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	693.82
AIC (smaller is better)	715.82
AICC (smaller is better)	716.28
BIC (smaller is better)	744.48
CAIC (smaller is better)	755.48
HQIC (smaller is better)	727.42

Fit Statistics for Conditional Distribution	
-2 log L(Fatal | r. effects)	615.56
Pearson Chi-Square	485.04
Pearson Chi-Square / DF	0.83

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.4518	0.1909
Scale	0.04662	0.1838

Solutions for Fixed Effects							
Effect	County	signal intersections per kilometer	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept			-9.2444	3.4105	92	-2.71	0.0080
County	1		-0.9611	0.3741	482	-2.57	0.0105
County	2		-1.0616	0.5421	482	-1.96	0.0508
County	3		-0.5357	0.3352	482	-1.60	0.1106
County	4		0.1385	0.3594	482	0.39	0.7001
County	5		-0.1030	0.9512	482	-0.11	0.9138
County	7		0	.	.	.	.
logAADT			0.7429	0.3132	482	2.37	0.0181
group_int		1	-0.3601	0.3203	482	-1.12	0.2614
group_int		2	0.2097	0.3176	482	0.66	0.5094
group_int		3	0	.	.	.	.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
County	5	482	2.72	0.0196
logAADT	1	482	5.62	0.0181
group_int	2	482	2.09	0.1244

3.8 CHAPTER 6 ARTERAL MODEL 2

3.8.1 TOTAL CRASHES-M2

Model 2 Information	
Data Set	LB.ARTERIAS2GLMSELECT
Response Variable	Choques_totales
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	599
Number of Observations Used	583

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3466.87
AIC (smaller is better)	3484.87
AICC (smaller is better)	3485.18
BIC (smaller is better)	3508.32
CAIC (smaller is better)	3517.32
HQIC (smaller is better)	3494.36

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totales | r. effects)	3185.14
Pearson Chi-Square	573.19
Pearson Chi-Square / DF	0.98

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.4301	0.07873
Scale	0.1828	0.02230

Solutions for Fixed Effects						
Effect	County	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		-9.3121	1.5874	94	-5.87	<.0001
County	1	-1.0857	0.2158	482	-5.03	<.0001
County	2	-1.5785	0.3123	482	-5.05	<.0001
County	3	0.2106	0.2195	482	0.96	0.3378
County	4	-0.4224	0.2407	482	-1.75	0.0800
County	5	-0.5463	0.4796	482	-1.14	0.2552

Solutions for Fixed Effects						
Effect	County	Estimate	Standard Error	DF	t Value	Pr > |t|
County	7	0	.	.	.	.
logAADT		1.1091	0.1469	482	7.55	<.0001

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
County	5	482	13.65	<.0001
logAADT	1	482	57.01	<.0001

3.8.2 INJURY-M2

Model 2 Information	
Data Set	LB.ARTERIAS2GLMSELECT
Response Variable	No_fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	599
Number of Observations Used	583

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3438.86
AIC (smaller is better)	3456.86
AICC (smaller is better)	3457.17
BIC (smaller is better)	3480.30
CAIC (smaller is better)	3489.30
HQIC (smaller is better)	3466.35

Fit Statistics for Conditional Distribution	
-2 log L(No_fatal | r. effects)	3160.37
Pearson Chi-Square	578.10
Pearson Chi-Square / DF	0.99

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.4219	0.07738
Scale	0.1832	0.02240

Solutions for Fixed Effects						
Effect	County	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		-9.6860	1.5979	94	-6.06	<.0001
County	1	-1.0934	0.2143	482	-5.10	<.0001
County	2	-1.5482	0.3108	482	-4.98	<.0001
County	3	0.2203	0.2177	482	1.01	0.3121
County	4	-0.4713	0.2392	482	-1.97	0.0494
County	5	-0.5160	0.4788	482	-1.08	0.2817
County	7	0	.	.	.	.
logAADT		1.1421	0.1479	482	7.72	<.0001

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
County	5	482	13.85	<.0001
logAADT	1	482	59.63	<.0001

3.8.3 FATAL-M2

Model 3- Fatal crashes (Table 6.4 and 6.10)

Model 2 Information	
Data Set	LB.ARTERIAS2GLMSELECT
Response Variable	Fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	599
Number of Observations Used	583

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	698.14
AIC (smaller is better)	716.14
AICC (smaller is better)	716.45
BIC (smaller is better)	739.58
CAIC (smaller is better)	748.58
HQIC (smaller is better)	725.62

Fit Statistics for Conditional Distribution	
-2 log L(Fatal | r. effects)	616.37
Pearson Chi-Square	483.32
Pearson Chi-Square / DF	0.83

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.4790	0.1952
Scale	0.04547	0.1835

Solutions for Fixed Effects						
Effect	County	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		-9.7507	3.4340	94	-2.84	0.0055
County	1	-0.9444	0.3734	482	-2.53	0.0117
County	2	-1.0432	0.5491	482	-1.90	0.0580
County	3	-0.4488	0.3346	482	-1.34	0.1804
County	4	0.1858	0.3602	482	0.52	0.6061
County	5	-0.2779	0.9566	482	-0.29	0.7715
County	7	0	.	.	.	.
logAADT		0.7802	0.3173	482	2.46	0.0143

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
County	5	482	2.63	0.0234
logAADT	1	482	6.05	0.0143

3.9 CHAPTER 6 ARTERIAL MODEL 3

3.9.1 TOTAL CRASHES-M3

Model 3 Information	
Data Set	LB.ARTERIAS2GLMSELECT
Response Variable	Choques_totales
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	599
Number of Observations Used	583

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3460.67
AIC (smaller is better)	3480.67
AICC (smaller is better)	3481.05
BIC (smaller is better)	3506.72
CAIC (smaller is better)	3516.72
HQIC (smaller is better)	3491.21

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totales | r. effects)	3186.80
Pearson Chi-Square	574.60
Pearson Chi-Square / DF	0.99

Covariance Parameter Estimates		
Cov Parm	**Estimate**	**Standard Error**
Section_ID	0.3948	0.07358
Scale	0.1827	0.02229

Solutions for Fixed Effects						
Effect	**County**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept		-9.1803	1.5500	93	-5.92	<.0001
logAADT		1.0850	0.1437	482	7.55	<.0001
County	1	-1.1080	0.2084	482	-5.32	<.0001
County	2	-1.5261	0.3024	482	-5.05	<.0001
County	3	0.1800	0.2115	482	0.85	0.3953
County	4	-0.4745	0.2327	482	-2.04	0.0420
County	5	-0.4687	0.4673	482	-1.00	0.3163
County	7	0	.	.	.	.
Int_signal		0.1350	0.05316	482	2.54	0.0114

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
logAADT	1	482	57.04	<.0001
County	5	482	14.20	<.0001
Int_signal	1	482	6.45	0.0114

3.9.2 INJURY-M3

The GLIMMIX Procedure

Model 3 Information	
Data Set	LB.ARTERIAS2GLMSELECT
Response Variable	No_fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	599
Number of Observations Used	583

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3433.35
AIC (smaller is better)	3453.35
AICC (smaller is better)	3453.73
BIC (smaller is better)	3479.40
CAIC (smaller is better)	3489.40
HQIC (smaller is better)	3463.89

Fit Statistics for Conditional Distribution	
-2 log L(No_fatal | r. effects)	3161.97
Pearson Chi-Square	578.50
Pearson Chi-Square / DF	0.99

Covariance Parameter Estimates		
Cov Parm	**Estimate**	**Standard Error**
Section_ID	0.3903	0.07288
Scale	0.1830	0.02240

Solutions for Fixed Effects						
Effect	**County**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept		-9.5555	1.5638	93	-6.11	<.0001
logAADT		1.1187	0.1450	482	7.72	<.0001
County	1	-1.1144	0.2076	482	-5.37	<.0001
County	2	-1.5001	0.3019	482	-4.97	<.0001
County	3	0.1914	0.2105	482	0.91	0.3636
County	4	-0.5179	0.2320	482	-2.23	0.0260
County	5	-0.4446	0.4678	482	-0.95	0.3423
County	7	0	.	.	.	.
Int_signal		0.1275	0.05330	482	2.39	0.0171

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
logAADT	1	482	59.56	<.0001
County	5	482	14.36	<.0001
Int_signal	1	482	5.72	0.0171

3.9.3 FATAL-M3

Model 3 Information	
Data Set	LB.ARTERIAS2GLMSELECT
Response Variable	Fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	599
Number of Observations Used	583

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	693.32
AIC (smaller is better)	713.32
AICC (smaller is better)	713.71
BIC (smaller is better)	739.37
CAIC (smaller is better)	749.37
HQIC (smaller is better)	723.87

Fit Statistics for Conditional Distribution	
-2 log L(Fatal | r. effects)	615.97
Pearson Chi-Square	454.03
Pearson Chi-Square / DF	0.78

Covariance Parameter Estimates		
Cov Parm	**Estimate**	**Standard Error**
Section_ID	0.4348	0.1794
Scale	0.04580	0.1837

Solutions for Fixed Effects						
Effect	**County**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept		-9.3610	3.3716	93	-2.78	0.0066
logAADT		0.7264	0.3119	482	2.33	0.0203
County	1	-0.9715	0.3644	482	-2.67	0.0079
County	2	-1.0185	0.5375	482	-1.89	0.0587
County	3	-0.4297	0.3256	482	-1.32	0.1875
County	4	0.1152	0.3534	482	0.33	0.7445
County	5	-0.2036	0.9419	482	-0.22	0.8290
County	7	0	.	.	.	.
Int_signal		0.2058	0.08797	482	2.34	0.0197

Type III Tests of Fixed Effects				
Effect	**Num DF**	**Den DF**	**F Value**	**Pr > F**
logAADT	1	482	5.42	0.0203
County	5	482	2.62	0.0235
Int_signal	1	482	5.47	0.0197

3.10 CHAPTER 6 ARTERIAL MODEL 4

3.10.1 TOTAL CRASHES-M4

Model 4- Total crashes (Table 6.4 and 6.8)

Model 4 Information	
Data Set	LB.ARTERIAS2GLMSELECT
Response Variable	Choques_totales
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	599
Number of Observations Used	583

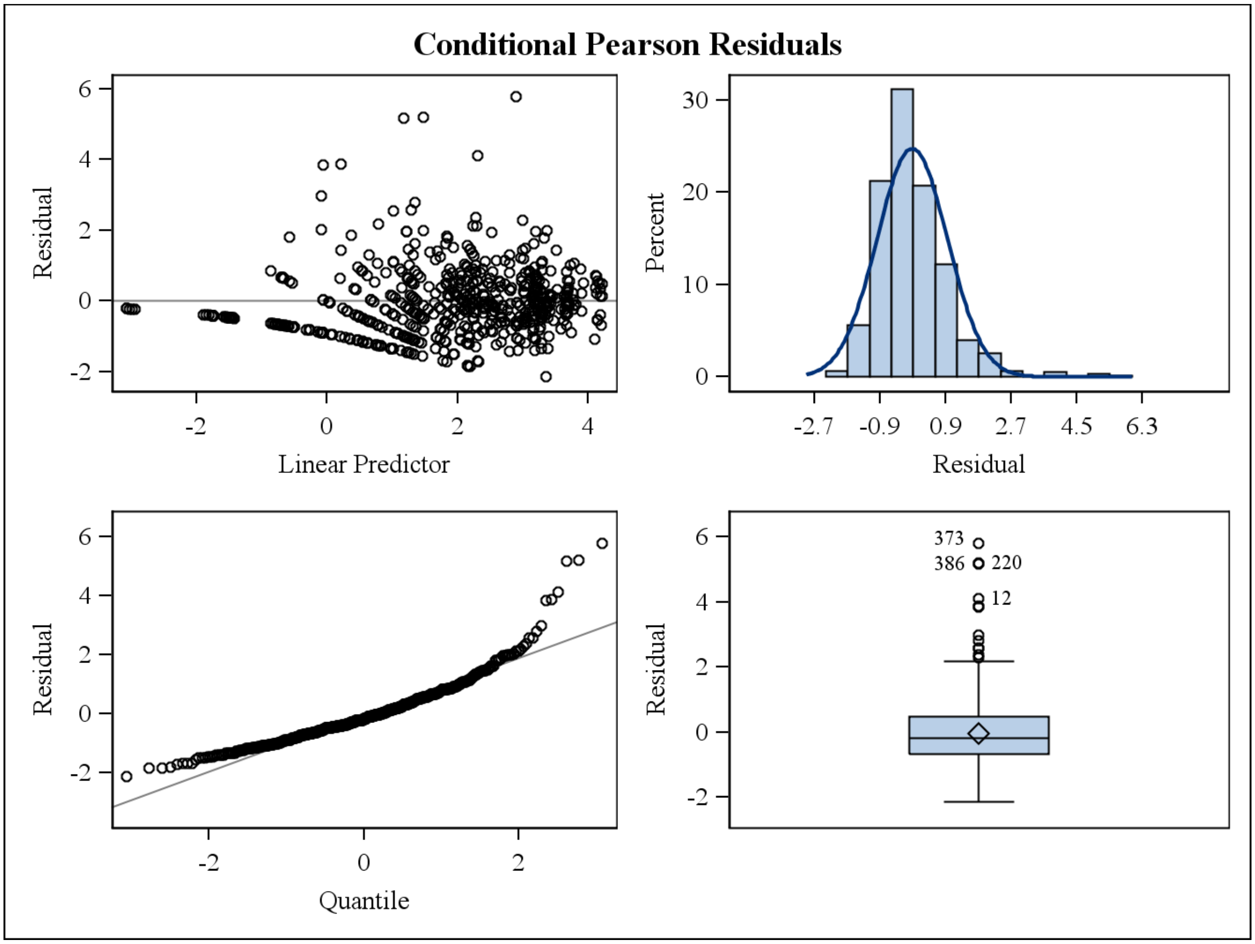
Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3516.51
AIC (smaller is better)	3526.51
AICC (smaller is better)	3526.61
BIC (smaller is better)	3539.54
CAIC (smaller is better)	3544.54
HQIC (smaller is better)	3531.78

Fit Statistics for Conditional Distribution	
-2 log L(Choques_totales | r. effects)	3179.85
Pearson Chi-Square	544.83
Pearson Chi-Square / DF	0.93

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.7666	0.1321
Scale	0.1827	0.02227

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-10.5764	1.1450	98	-9.24	<.0001
logAADT	1.1657	0.1095	482	10.65	<.0001
Int_signal	0.1541	0.07003	482	2.20	0.0283


Conditional Pearson Residuals
Residual
Linear Predictor
Percent
Residual
-2.7
-0.9
0.9
2.7
4.5
6.3
Residual
Quantile
Residual
373
386
220
12

3.10.2 INJURY-M4

Model 3- Injury crashes (Table 6.4 and 6.9)

Model 4 Information	
Data Set	LB.ARTERIAS2GLMSELECT
Response Variable	No_fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	599
Number of Observations Used	583

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Likelihood	3489.57
AIC (smaller is better)	3499.57
AICC (smaller is better)	3499.68
BIC (smaller is better)	3512.60
CAIC (smaller is better)	3517.60
HQIC (smaller is better)	3504.84

Fit Statistics for Conditional Distribution	
-2 log L(No_fatal | r. effects)	3154.71
Pearson Chi-Square	548.49
Pearson Chi-Square / DF	0.94

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.7636	0.1315
Scale	0.1829	0.02236

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-10.8457	1.1561	98	-9.38	<.0001
logAADT	1.1895	0.1105	482	10.76	<.0001
Int_signal	0.1441	0.07026	482	2.05	0.0408

3.10.3 FATAL-M4

Model 3- Fatal crashes (Table 6.4 and 6.10)

Model 3 Information	
Data Set	LB.ARTERIAS2GLMSELECT
Response Variable	Fatal
Response Distribution	Negative Binomial
Link Function	Log
Variance Function	Default
Offset Variable	loglength
Variance Matrix	Not blocked
Estimation Technique	Maximum Likelihood
Likelihood Approximation	Laplace
Degrees of Freedom Method	Containment

Number of Observations Read	599
Number of Observations Used	583

Convergence criterion (GCONV=1E-8) satisfied.

The GLIMMIX Procedure

Fit Statistics	
-2 Log Likelihood	706.43
AIC (smaller is better)	716.43
AICC (smaller is better)	716.53
BIC (smaller is better)	729.46
CAIC (smaller is better)	734.46
HQIC (smaller is better)	721.70

Fit Statistics for Conditional Distribution	
-2 log L(Fatal | r. effects)	612.49
Pearson Chi-Square	456.24
Pearson Chi-Square / DF	0.78

Covariance Parameter Estimates		
Cov Parm	Estimate	Standard Error
Section_ID	0.5899	0.2139
Scale	0.04435	0.1826

Solutions for Fixed Effects					
Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-10.2481	2.0431	98	-5.02	<.0001
logAADT	0.7677	0.1933	482	3.97	<.0001
Int_signal	0.2213	0.09458	482	2.34	0.0197

4. SCRIPT SAS FREEWAY AND ARTERIAL MODELS

```
ods rtf;

PROC IMPORT OUT= WORK.autopistas

        DATAFILE= "C:\Users\402063747\Desktop\sasautopistas.xlsx"

        DBMS=EXCEL REPLACE;

   RANGE="Sheet2$";

   GETNAMES=YES;

   MIXED=NO;

   SCANTEXT=YES;

   USEDATE=YES;

   SCANTIME=YES;

RUN;


libname lb "C:\Users\INCI\Documents\ERIKA\Erika\Tesis - Marzo 3 2014\Modelos SAS\Dataset";

run;

data lb.autopistas2;

set autopistasstep1 ;

logAADT=log(AADT);

loglength=log(length);

loglengthAADT=log(AADT)+log(length);

No_fatal= (Choques_Totales-Fatal);

APKPY=Choques_Totales/length;

apkpyaadt=apkpy/(AADT*365)*1000000;

Curves_km=Curves_A/lenght;

if Section_ID=10000000 then delete;

if Section_ID= 21000000 then delete;

if lanes=2 then delete;

if lanes=>7 then lanes=7;
```

```
if curves_A<>0 then group_curves=1;
if curves_B<>0 then group_curves=2;
if curves_C<>0 then group_curves=3;
if curves_D<>0 then group_curves=4;
if curves_E<>0 then group_curves=5;
if curves_F<>0 then group_curves=6;
if grades_A<>0 then group_grades=1;
if grades_B<>0 then group_grades=2;
if grades_C<>0 then group_grades=3;
if grades_D<>0 then group_grades=4;
if grades_E<>0 then group_grades=5;
if grades_F<>0 then group_grades=6;
if curves_A_km<>0 then group_curves_km=1;
if curves_B_km<>0 then group_curves_km=2;
if curves_C_km<>0 then group_curves_km=3;
if curves_D_km<>0 then group_curves_km=4;
if curves_E_km<>0 then group_curves_km=5;
if curves_F_km<>0 then group_curves_km=6;
if grades_A_km<>0 then group_grades_km=1;
if grades_B_km<>0 then group_grades_km=2;
if grades_C_km_<>0 then group_grades_km=3;
if grades_D_km<>0 then group_grades_km=4;
if grades_E_km<>0 then group_grades_km=5;
if grades_F_km<>0 then group_grades_km=6;
if section_ID=. then delete;
if at_grade_signal<>0 then delete;
if at_grade_other<>0 then delete;
```

```
if green<>0 then delete;
label type_terrain='type of terrain';
label group_curves= 'degree of curvature';
label group_grades= 'gradient percent';
label lanes= 'number of lanes';
label rural_urban= 'rural_urban code';
label vert_align= 'vertical alignment';
label horz_align='horizontal alignment';
label design_speed='design speed';
label speed_limit='speed limit';
label urban_code= 'urban code';
run;

proc format;
        value type_terrain 0= "Not aplicable"
                                        1= "Level"
                            2= "Rolling"
                3= "Mountainous";

   value group_curves 1= "0.0-3.4"
                    2= "3.5-5.4"
                                        3= "5.5-8.4"
                                        4= "8.5-13.9"
                                        5= "14.0-27.9"
                                        6= "28+";

        value group_grades 1= "0.0-0.4"
```

```
2= "0.5-2.4"
3= "2.5-4.4"
4= "4.5-6.4"
5= "6.5-8.4"
6= "8.5+ ";

value vert_align 0="Not required"
1= "All grades an vertical curves meet minimun design standards appropiate for the terrain"
2= "Some grades an vertical curves are belown design standards appropiate for the terrain"
3= "Infrequent grades an vertical curves that impair distance or affect the speed of trucks"
4="Frequent grades and vertical curves that impair sight distance or severely affect the speed of trucks";

value horz_align 0="Not required"
1="Apropiate design"
2= "some curves below apropiate design"
3= "Infrequent curves with design speeds less speed limit"
4="several curves with design speeds less speed limit";

value county    1="San Juan"
2= "Arecibo"
3= "Aguadilla"
4= "Mayaguez"
5= "Ponce"
6= "Guayama"
7= "Humacao";
```

```
value shoulder_type 1="None"
                2= "Surfaced"
                3= "Stabilized"
                4= "Combination"
                5= "Earth"
                6= "Barrier curb";

value median_type 1= "Curbed"
                2= "Positive barrier"
                3= "Unprotected"
                4= "None";

value lanes   4="4"
              6="6"
              7="Multilane";
run;

title 'model 1';
proc glimmix method=laplace data=lb.autopistas2;
where apkpy<60;
where section_ID<>.;
class section_ID county lanes  shoulder_type group_grades;
model Choques_totales= logAADT lanes shoulder_type / dist=nb link= log offset=loglength solution
htype=3;
random section_ID/solution;
run;
proc glimmix method=laplace data=lb.autopistas2;
where apkpy<60;
```

```
where section_ID<>.;
class section_ID county lanes  shoulder_type group_grades;
model No_fatal= logAADT lanes shoulder_type / dist=nb link= log offset=loglength solution htype=3;
random section_ID/solution;
run;
proc glimmix method=laplace data=lb.autopistas2;
where apkpy<60;
where section_ID<>.;
class section_ID county lanes  shoulder_type group_grades;
model fatal= logAADT lanes shoulder_type / dist=nb link= log offset=loglength solution htype=3;
random section_ID/solution;
run;
ods rtf close;

ods rtf;
title 'model 2';
proc glimmix method=laplace data=lb.autopistas2;
where apkpy<60;
where section_ID<>.;
class section_ID county lanes  shoulder_type group_grades;
model Choques_totales= logAADT shoulder_type / dist=nb link= log offset=loglength solution htype=3;
random section_ID/solution;
run;
proc glimmix method=laplace data=lb.autopistas2;
where apkpy<60;
where section_ID<>.;
class section_ID county lanes  shoulder_type group_grades;
```

```
model No_fatal= logAADT shoulder_type / dist=nb link= log offset=loglength solution htype=3;

random section_ID/solution;

run;

proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

where section_ID<>.;

class section_ID county lanes  shoulder_type group_grades;

model fatal= logAADT shoulder_type / dist=nb link= log offset=loglength solution htype=3;

random section_ID/solution;

run;

ods rtf close;


ods rtf;

title 'model 3';

proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

where section_ID<>.;

class section_ID county lanes  shoulder_type group_grades;

model Choques_totales= logAADT shoulder_type lanes county/ dist=nb link= log offset=loglength solution
htype=3;

random section_ID/solution;

run;


proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

class section_ID county lanes  shoulder_type group_grades;

model No_fatal= logAADT shoulder_type lanes county/ dist=nb link= log offset=loglength solution
htype=3;
```

```
random section_ID/solution;

run;


proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

where section_ID<>.;

class section_ID county lanes  shoulder_type group_grades;

model fatal= logAADT shoulder_type lanes county/ dist=nb link= log offset=loglength solution htype=3;

random section_ID/solution;

run;

ods rtf close;


ods rtf;

title 'model 4';

proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

where section_ID<>.;

class section_ID county lanes  shoulder_type group_grades;

model Choques_totales= logAADT shoulder_type county/ dist=nb link= log offset=loglength solution
htype=3;

random section_ID/solution;

run;


proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

where section_ID<>.;

class section_ID county lanes  shoulder_type group_grades;

model No_fatal= logAADT shoulder_type county/ dist=nb link= log offset=loglength solution htype=3;
```

```
random section_ID/solution;

run;


proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

where section_ID<>.;

class section_ID county lanes  shoulder_type group_grades;

model fatal= logAADT shoulder_type county/ dist=nb link= log offset=loglength solution htype=3;

random section_ID/solution;

run;

ods rtf close;


ods rtf;

title 'model 5';

proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

where section_ID<>.;

class section_ID county lanes  shoulder_type group_grades;

model Choques_totales= logAADT county/ dist=nb link= log offset=loglength solution htype=3;

random section_ID;

run;

proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

where section_ID<>.;

class section_ID county lanes  shoulder_type group_grades;

model No_fatal= logAADT county/ dist=nb link= log offset=loglength solution htype=3;

random section_ID;
```

```
run;
proc glimmix method=laplace data=lb.autopistas2;
where apkpy<60;
where section_ID<>.;
class section_ID county lanes  shoulder_type group_grades;
model fatal= logAADT county/ dist=nb link= log offset=loglength solution htype=3;
random section_ID/solution;
run;
ods rtf close;


ods rtf;
title 'model 6';
proc glimmix method=laplace data=lb.autopistas2;
where apkpy<60;
where section_ID<>.;
class section_ID county lanes  shoulder_type group_grades;
model Choques_totales= logAADT/ dist=nb link= log offset=loglength solution htype=3;
random section_ID;
run;
proc glimmix method=laplace data=lb.autopistas2;
where apkpy<60;
where section_ID<>.;
class section_ID county lanes  shoulder_type group_grades;
model Choques_totales= logAADT/ dist=nb link= log offset=loglength solution htype=3;
random section_ID county;
run;
proc glimmix method=laplace data=lb.autopistas2;
```

```
where apkpy<60;

where section_ID<>.;

class section_ID county lanes  shoulder_type group_grades;

model No_fatal= logAADT/ dist=nb link= log offset=loglength solution htype=3;

random section_ID;

run;

ods rtf;

proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

where section_ID<>.;

class section_ID county lanes  shoulder_type group_grades;

model fatal= logAADT/ dist=nb link= log offset=loglength solution htype=3;

random section_ID/solution;

run;

ods rtf close;


ods rtf;

title 'model 7';

proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

class section_ID county lanes  shoulder_type group_grades speed_limit design_speed;

model Choques_totales= logAADT  speed_limit design_speed county/ dist=nb link= log offset=loglength
solution htype=3;

random section_ID;

run;

proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

class section_ID county lanes  shoulder_type group_grades speed_limit design_speed;
```

```
model No_fatal= logAADT  speed_limit design_speed county/ dist=nb link= log offset=loglength solution
htype=3;

random section_ID;

run;


proc glimmix method=laplace data=lb.autopistas2;

where apkpy<60;

class section_ID county lanes  shoulder_type group_grades speed_limit design_speed;

model fatal= logAADT speed_limit / dist=nb link= log offset=loglength solution htype=3;

random section_ID/solution;

run;
```

Arterias

```
ods rtf;

PROC IMPORT OUT= WORK.arteriasstep1

         DATAFILE= "C:\Users\INCI\Documents\ERIKA\Erika\Tesis - Marzo 3 2014\Archivos excel\Segmentos\Excel para SAS\sasarteriasstep1.xlsx"

         DBMS=EXCEL REPLACE;

    RANGE="'2004-2009 Arterias (618)$'";

    GETNAMES=YES;

    MIXED=NO;

    SCANTEXT=YES;

    USEDATE=YES;

    SCANTIME=YES;

RUN;

libname lb"C:\Users\INCI\Documents\ERIKA\Erika\Tesis - Marzo 3 2014\Modelos SAS\Dataset";

run;

data lb.arterias2glmselect;

set arteriasstep1;

logAADT=log(AADT);

loglength=log(length);

loglengthAADT=log(length)+log(AADT);

No_fatal= (Choques_Totales-Fatal);

APKPY=Choques_Totales/length;

APKPYAADT=APKPY/(AADT*365)*1000000;

if 0.54>int_signal=>0 then group_int=1;

if 1.37>int_signal=>0.54 then group_int=2;

if int_signal=>1.37 then group_int=3;

if 0.67>int_others=>0 then group_intothers=1;

if 1.47>int_others=>0.67 then group_intothers=2;
```

```
if int_others=>1.47 then group_intothers=3;

if Section_ID=10000000 then delete;

if Section_ID= 21000000 then delete;

if Section_ID=. then delete;

if lanes=2 then delete;

if lanes=>7 then lanes=7;

if curves_A<>0 then group_curves=1;

if curves_B<>0 then group_curves=2;

if curves_C<>0 then group_curves=3;

if curves_D<>0 then group_curves=4;

if curves_E<>0 then group_curves=5;

if curves_F<>0 then group_curves=6;

if grades_A<>0 then group_grades=1;

if grades_B<>0 then group_grades=2;

if grades_C<>0 then group_grades=3;

if grades_D<>0 then group_grades=4;

if grades_E<>0 then group_grades=5;

if grades_F<>0 then group_grades=6;

if curves_A_km<>0 then group_curves_km=1;

if curves_B_km<>0 then group_curves_km=2;

if curves_C_km<>0 then group_curves_km=3;

if curves_D_km<>0 then group_curves_km=4;

if curves_E_km<>0 then group_curves_km=5;

if curves_F_km<>0 then group_curves_km=6;

if grades_A_km<>0 then group_grades_km=1;

if grades_B_km<>0 then group_grades_km=2;

if grades_C_km_<>0 then group_grades_km=3;
```

```
if grades_D_km<>0 then group_grades_km=4;

if grades_E_km<>0 then group_grades_km=5;

if grades_F_km<>0 then group_grades_km=6;

label type_terrain='type of terrain';

label group_curves= 'degree of curvature';

label group_grades= 'gradient percent';

label lanes= 'number of lanes';

label rural_urban= 'rural_urban code';

label vert_align= 'vertical alignment';

label horz_align='horizontal alignment';

label design_speed='design speed';

label speed_limit='speed limit';

label urban_code= 'urban code';

label group_int='signal intersections per kilometer';

label group_intothers='others intersections per kilometer';

label int_signal='number of signal intersections';

label int_sign='number of sign intersection';

label int_others='number of other intersection';

run;


proc format;

        value type_terrain 0= "Not aplicable-Urban"

                                                   1= "Level"

                                      2= "Rolling"

                 3= "Mountainous";


   value group_curves 1= "0.0-3.4"
```

```
              2= "3.5-5.4"
                          3= "5.5-8.4"
                          4= "8.5-13.9"
                          5= "14.0-27.9"
                          6= "28+";

value group_grades 1= "0.0-0.4"
              2= "0.5-2.4"
                          3= "2.5-4.4"
                          4= "4.5-6.4"
                          5= "6.5-8.4"
                          6= "8.5+ ";

value vert_align 0="Not required"
        1= "All grades an vertical curves meet minimun design standards appropiate for the terrain"
                          2= "Some grades an vertical curves are belown design
standards appropiate for the terrain"
                          3= "Infrequent grades an vertical curves that impair distance or
affect the speed of trucks"
                          4="Frequent grades and vertical curves that impair sight
distance or severely affect the speed of trucks";

value horz_align 0="Not required"
            1="Apropiate design"
                          2= "some curves below apropiate design"
                          3= "Infrequent curves with design speeds less speed limit"
                          4="several curves with design speeds less speed limit";

value county     1="San Juan"
```

```
                    2= "Arecibo"
                                   3= "Aguadilla"
                                   4= "Mayaguez"
                                   5= "Ponce"
                                   6= "Guayama"
                                   7= "Humacao";

    value shoulder_type 1="None"
                     2= "Surfaced"
                                       3= "Stabilized"
                                       4= "Combination"
                  5= "Earth"
                  6= "Barrier curb";

 value median_type 1= "Curbed"
                 2= "Positive barrier"
                                   3= "Unprotected"
                                   4= "None";

 value group_int 1= "0-0.53"
                2= "0.54-1.36"
                               3= "1.37+";

 value group_intothers 1= "0-0.66"
                  2= "0.67-1.46"
                                 3= "1.47+";
run;
```

```
ods rtf;

title 'model 1';

          proc glimmix data=lb.arterias2glmselect method=laplace;

          where apkpy<100;

          class section_ID county lanes group_int;

          model Choques_Totales= county logAADT  group_int/dist=nb link=log offset=loglength solution
htype=3;

          random section_ID;

          run;

          proc glimmix data=lb.arterias2glmselect method=laplace;

          where apkpy<100;

          class section_ID county lanes group_int;

          model No_fatal= county logAADT  group_int/dist=nb link=log offset=loglength solution htype=3;

          random section_ID;

          run;

          proc glimmix data=lb.arterias2glmselect method=laplace;

          where apkpy<100;

          class section_ID county lanes group_int;

          model Fatal= county logAADT  group_int/dist=nb link=log offset=loglength solution htype=3;

          random section_ID;

          run;


title 'model 2';

          proc glimmix data=lb.arterias2glmselect method=laplace;

          where apkpy<100;

          class section_ID county lanes group_int;

          model Choques_Totales= county logAADT/dist=nb link=log offset=loglength solution htype=3;

          random section_ID;
```

```
	run;
	proc glimmix data=lb.arterias2glmselect method=laplace;
	where apkpy<100;
	class section_ID county lanes group_int;
	model No_fatal= county logAADT/dist=nb link=log offset=loglength solution htype=3;
	random section_ID;
	run;
	proc glimmix data=lb.arterias2glmselect method=laplace;
	where apkpy<100;
	class section_ID county lanes group_int;
	model Fatal= county logAADT/dist=nb link=log offset=loglength solution htype=3;
	random section_ID;
	run;

title 'model 3;
	proc glimmix data=lb.arterias2glmselect method=laplace;
	where apkpy<100;
	class section_ID county lanes group_int;
	model Choques_Totales=logAADT county Int_signal/dist=nb link=log offset=loglength solution
htype=3;
	random section_ID;
	run;
	proc glimmix data=lb.arterias2glmselect method=laplace;
	where apkpy<100;
	class section_ID county lanes group_int;
	model No_fatal= logAADT county Int_signal/dist=nb link=log offset=loglength solution htype=3;
	random section_ID;
	run;
```

```
        proc glimmix data=lb.arterias2glmselect method=laplace;
        where apkpy<100;
        class section_ID county lanes group_int;
        model Fatal= logAADT county Int_signal/dist=nb link=log offset=loglength solution htype=3;
        random section_ID;
        run;

title 'model 4';
        proc glimmix data=lb.arterias2glmselect method=laplace;
        where apkpy<100;
        class section_ID county lanes group_int;
        model Choques_Totales=logAADT Int_signal/dist=nb link=log offset=loglength solution htype=3;
        random section_ID;
        run;
        proc glimmix data=lb.arterias2glmselect method=laplace;
        where apkpy<100;
        class section_ID county lanes group_int;
        model No_fatal= logAADT Int_signal/dist=nb link=log offset=loglength solution htype=3;
        random section_ID;
        run;
        proc glimmix data=lb.arterias2glmselect method=laplace;
        where apkpy<100;
        class section_ID county lanes group_int;
        model Fatal= logAADT Int_signal/dist=nb link=log offset=loglength solution htype=3;
        random section_ID;
        run;
```