SAFETY IMPLICATIONS OF ROADWAY DESIGN AND MANAGEMENT: NEW EVIDENCE AND INSIGHTS IN THE TRADITIONAL AND EMERGING (AUTONOMOUS VEHICLE) OPERATING ENVIRONMENTS

by

Sikai Chen

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Lyles School of Civil Engineering

West Lafayette, Indiana

August 2019

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Samuel Labi, Chair

Lyles School of Civil Engineering

Dr. Kumares C. Sinha

Lyles School of Civil Engineering

Dr. Paul V. Preckel

Department of Agricultural Economics

Dr. Lefteri H. Tsoukalas

School of Nuclear Engineering

Dr. Richard de Neufville

Institute for Data, Systems, and Society, Massachusetts Institute of Technology

Approved by:

Dr. Dulcy M. Abraham

Head of the Graduate Program

This is dedicated to my family, my amazing professors and mentors, and my friends, without whom this work would not have been accomplished.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude and appreciation to my major advisor, Dr. Samuel Labi, for his continuous support throughout my entire Ph.D. studies. Dr. Labi is extremely kind and patient and always provides tremendous help to his students when they have hard times. He is not only an inspiring mentor but also a caring friend, who kept me on the right path and helped me grow both professionally and personally.

I would also like to thank my wonderful Ph.D. committee members at Purdue University, Dr. Kumares C. Sinha, Dr. Paul V. Preckel, and Dr. Lefteri H. Tsoukalas, for the time and effort they set aside to instruct me and for their insights and comprehensive feedback on my doctoral research. My gratitude also goes to another member of my Ph.D. committee, Dr. Richard de Neufville of MIT who provided valuable guidance. Their suggestions always were insightful and helped me improve my research significantly. I am blessed to have been advised by a committee that instilled in me the virtues of good hard work and an interdisciplinary research background in transportation, infrastructure, systems, economics, operations research, computer science and artificial intelligence. Their advice and guidance will benefit me throughout my life.

In addition, I would like to acknowledge the Office of Interdisciplinary Graduate Programs at Purdue University for the Bilsland Fellowship, which supported my doctoral studies.

Finally, my deepest gratitude goes to my family, especially my wife Yue Leng, who also will complete her Ph.D. soon, and to all my friends. I would have never been able to make it this far without them. No words can express what their endless love, encouragement, and support meant to me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Author: Chen, Sikai. PhD
Institution: Purdue University
Degree Received: August 2019
Title: Safety Implications of Roadway Design and Management: New Evidence and Insights in the Traditional and Emerging (Autonomous Vehicle) Operating Environments
Committee Chair: Samuel Labi

In the context of highway safety factors, road geometrics and pavement condition are of particular interest to highway managers as they fall within their direct control and therefore can be addressed through highway projects. In spite of the preponderance of econometric modeling in highway safety research, there still remain areas worthy of further investigation. These include 1) the lack of sufficient feedback to roadway preservation engineers regarding the impacts of road-surface condition on safety; 2) the inadequate feedback to roadway designers on optimal lane and shoulder width allocation; 3) the need for higher predictive capability and reliability of models that analyze roadway operations; and 4) the lack of realistic simulations to facilitate reliable safety impact studies regarding autonomous vehicles (AV). In an attempt to contribute to the existing knowledge in this domain and to throw more light on these issues, this dissertation proposes a novel framework for enhanced prediction of highway safety that incorporates machine learning and econometrics with optimization to evaluate and quantify the impacts of safety factors. In the traditional highway operating environment, the proposed framework is expected to help agencies improve their safety analysis. Using an Indiana crash dataset, this dissertation implements the framework, thereby 1) estimating the safety impacts of the road-surface condition with advanced econometric specifications, 2) optimizing space resource allocations across highway cross-sectional elements, and 3) predicting the fatality status of highway segments using machine learning algorithms. In addition, this dissertation discusses the opportunities and the expected safety impacts and benefits of AV in the emerging operating environment. The dissertation also presents a proposed deep learning-based autonomous driving simulation framework that addresses the limitations of AV testing and evaluation on in-service roads and test tracks.

CHAPTER 1. INTRODUCTION

1.1 Background

1.1.1 Prelude

Highway safety is influenced by multiple factors (e.g., highway geometry, driver characteristics, vehicle characteristics, and environment), and crashes are typically caused by an interaction of more than one of these factors (Sinha and Labi, 2007). Regarding the geometric characteristics, highways in most countries face problems with deficient design, due not to deliberately poor original design but rather to a continuing legacy of old designs that served users well under conditions long past. In several countries, a significant number of highways have safety deficiencies arising from inadequate road geometry, driver information deficiencies, lack of passing opportunities, inadequate shoulders, poor condition of pavement, and traffic conflicts due to driveways (Paterson, 1987; Silva and Liautaud, 2011; AASHTO, 2011; FHWA, 2006; 2014; 2016). Most of the deficient roads were designed and built to standards that became outdated vis-à-vis current design policies or operating conditions.

Financial constraints make it impossible to reconstruct (or improve) all highways to make them consistent with current established standards or national requirements. At the same time, agencies do not have the option of allowing highways to operate under their current geometric deficiencies as that would cause increased crashes, vehicle operating costs, delays, and inconveniences. Thus, driver frustration and lost time are prevalent, exacerbating safety problems that already exist due to inadequate geometry. For this reason, highway agencies seek to focus on areas identified as needing urgent safety intervention. The development of models that help predict crashes therefore are helpful in this regard.

The need to continually investigate the relationships between road conditions and road characteristics is underscored by the recent and continuing trends in the highway administration environment, which include ever-increasing population (hence, highway demand), advancements in computer technology, and legislation calling for operational accountability of highway assets in terms of their mobility, safety, and state of good repair. Of these needs, safety continues to be a major focus of highway agencies (Sinha and Labi, 2007). The application of crash prediction models that can adequately account for changing operational trends including emerging technologies, is expected to lead to more productive use of highways, better allocation of safety resources, and more effective management of crash risks.

1.1.2 Crash Causes

Roadway crashes are caused by several factors that are related to the vehicle, the driver, the natural or built-up environment, the road pavement and geometrics, the highway policy and legislation, or a combination thereof (Sinha and Labi, 2007). With regard to the road engineering factors, the literature is replete with studies that have investigated the safety impacts of road geometrics, such as lane width, shoulder width, vertical and horizontal alignment (Zegeer and Council, 1994; Farah et al., 2009; Labi et al., 2017), traffic operating conditions such as traffic volume, traffic stream composition (Ma et al., 2008; Ayati and Abbasi, 2011), and weather conditions (Huang et al., 2004). Of these factors, those associated with road geometrics and condition are of particular interest to highway asset managers as they fall within their direct control and can be addressed through highway projects. The American Society of Civil Engineers (ASCE) in 2014 advocated a significant, sustained effort to reduce traffic crashes and related deaths and injuries through improvements in all aspects of highway system performance and advocated for the reduction of crashes by "auditing existing roadway systems to identify roadway hazards and safety improvement opportunities and implementing highway and other engineering-related improvements proven effective in reducing the potential for, and severity of, traffic crashes."

1.1.3 The True Cost of Crashes

The adverse impacts of traffic crashes on society should not be underestimated. In 2015, it was reported that approximately 1.4 million lives are lost worldwide each year due to highway vehicle crashes, making traffic crashes the ninth leading cause of death (World Health Organization, 2015). In the U.S., traffic crashes were the fourth leading cause of death for all ages and the leading cause of death among young people (ages 1 to 44) (U.S. Centers for Disease Control and Prevention, 2016). Every year, more than 30,000 people in the U.S. lose their lives due to highway vehicle crashes (USDOT, 2004); specifically, this number was 37,461 in 2016, which was an increase of 5.6 percent from 2015 (NHTSA, 2017). Crashes and the resulting deaths and related injuries are an enormous strain on society; and it was estimated that the total economic and societal cost of highway crashes in the U.S. is $836 billion (Blincoe et al., 2015).

Given the significant economic and social costs of highway accidents and the potential for cost-effective implementation of safety countermeasures, highway agencies continue to seek more knowledge about the relationships between highway accidents and road geometry, pavement condition, and traffic. It is expected that such knowledge can facilitate effective identification of hazardous areas and, overall, enhance decision-making based on strategic safety-related performance indicators aligned with a clear vision of the societal benefits of safety improvements (Aguero-Valverde, 2013; Qiao et al., 2018).

1.1.4 Efforts to Reduce Crashes

In a bid to reduce the frequency of traffic crashes and their impact on society, several initiatives have been undertaken, ranging from policies (regarding the driver, vehicle, and speed limit) to physical interventions (road improvements) to regional, national, and international programs. An example of one such initiative was launched by the Swedish government in 1997, "Vision Zero," which aimed to eliminate vehicle-related deaths in Sweden. "Vision Zero" was subsequently adopted in the U.S. (Tingvall, 1997), and recently the United Nations set "Vision Zero" as one of the agency's goals, seeking to reduce global road traffic fatalities by 50% by 2020 (UNHCR, 2016). For "Vision Zero" to be reasonably successful and other ambitious safety improvement goals to be achieved, it is imperative to continue investigating the main factors involved in traffic crashes so that prediction of fatal crashes can be refined. The highway safety literature is replete with the work of many researchers who have not only attempted to predict crash frequency and severity but also have studied the relationships between highway crashes and crash factors (Zegeer et al. 1994; Milton et al. 1998; Lee et al. 2002; Pulugurtha et al. 2013; Ma et al. 2016; Chen et al. 2017a; 2017b; Labi et al. 2017; Tang et al. 2018).

In order to reduce crashes efficiently and effectively, it is necessary to have a comprehensive understanding of the relationships between crash occurrence and crash factors (e.g., the expected impact of widening the width of the shoulder on a certain roadway segment with respect to the number of injury crashes). Therefore, having reliable models is essential as they will provide researchers and agencies necessary information about these relationships.

1.1.5 Understanding the Crash Experience through Modeling

The abundant highway safety research literature pertaining to methodological developments and increasing model sophistication over the last several decades have spawned improvements in the predictive capability and reliability of models (Mannering and Bhat, 2014). This is especially true from an econometric modeling perspective. Specifically, researchers have put forth significant and sustained efforts to account for various types of unobserved heterogeneities. For example, when modeling highway safety data, researchers have considered a wide range of factor categories that affect highway crashes, such as the attributes of the driver, vehicle, enforcement levels, natural environment, and road engineering features (Sinha et al., 1981; Sinha and Labi, 2007) as well as the complex interactions among them (Park, 2016; Mannering et al., 2016). Unfortunately, the information collected in crash databases does not span this entire spectrum of accident factors. As such, efforts to model accident frequency and severity are often proceeded with tacit or explicit admission that some accident factor information is missing in the data that, if available, could have helped better explained the accident experience. This problem of unobserved heterogeneity is further exacerbated by the spatial and temporal correlation of accident data and the correlation across the different levels of accident severity. Failure to account for these correlations may lead to invalid inferences and unreliable parameter estimation (Lee et al., 2015).

Safety researchers have risen to this challenge, however, as highway safety research has continued to advance the state of the art regarding analytical methods that can characterize and address the effect of such unobserved factors. For example, over the past few decades, there has been an upsurge in the development and application of statistical multivariate techniques and approaches for joint modeling of accident frequencies (Hauer, 2010; El-Basyouny and Sayed, 2009; Aguero-Valverde, 2013; Bijleveld, 2005; Ma and Kockelman, 2006; Ma et al., 2008; Song et al., 2006; Wang et al., 2011; Chen et al., 2017a; Chen et al., 2017b). Also, a number of research studies have used a variety of statistical tools to address other sources of possible errors in modeling accident data (El-Basyouny and Sayed, 2009; Barua et al., 2016; Chen and Tarko, 2014; Park et al., 2015; Dinu and Veeraragavan, 2011; El-Basyouny et al., 2014a, Garnowski and Manner, 2011; Dong et al., 2015; El-Basyouny et al., 2014b; Lee et al., 2015; Venkataraman et al., 2011; Barua et al., 2016; Chen et al., 2019).These efforts have gradually spawned substantial improvements in the predictive capability and reliability of models.

1.2 Problem Statement

Despite the preponderance of econometric modeling in the highway safety data analysis research area, several important issues remain to be answered and further investigated:

1) Insufficient feedback to roadway preservation engineers regarding the impacts of road-surface condition on safety. Nearly 30% of annual highway fatalities are caused by poor condition of pavements (Zimmerman and Larson, 2005). Past studies have attempted to investigate both the direct and indirect effects of road-surface condition on safety; however, the direction of pavement impact (positive vs. negative) has been rather ambiguous and additional research was recommended.
2) Insufficient feedback for roadway design on optimal lane and shoulder width allocation. It is important to not only identify new and innovative approaches to improving safety on two-lane highways, but also to optimize the implementation of appropriate safety countermeasures. A case in point is determination of the optimal width of lanes and shoulders for roadways. The term “optimal width” suggests that both the cases of narrow lanes and shoulders and very wide lanes and shoulders are undesirable; and somewhere between these extremes, an optimal width exists. At one extreme, it has long been recognized that small lanes and widths are undesirable (widening the travel lanes or/and their shoulders has a positive impact in terms of reductions in crashes of various types, particularly roadway departure crashes and horizontal curve crashes). Given the apparent challenge but potential benefits associated with the optimal allocation of lane and shoulder widths within a given total roadway width (TRW), it is crucial that public highway agencies can apply consistent and objective methodologies to a variety of allocation scenarios to ascertain the optimal allocation of resources while providing the greatest safety benefit to the road users.
3) Need for continued search for higher predictive capability and reliability for analyzing roadway operations. Econometric modeling techniques are becoming more sophisticated; for example, when modeling crash counts, negative binomial and Poisson models have become the common choice in the literature. Researchers continue to incorporate advanced elements into their models, such as random effect, random parameter, multivariate analysis, seemingly unrelated equations, etc. However, there is a lack of approaches that can really do the following: a) assess whether the current econometric modeling techniques are adequate, b) analyze the

necessity to keep improving these modeling techniques, and c) assess the extent to which they are improved to assure they are devoid of issues such as over-fitting.

In addition, state-of-the-art algorithms in data mining and machine learning, despite their demonstrated efficacy in other disciplines, are barely found in highway safety research literature. These algorithms typically have superior prediction power (but weaker interpretation power) compared to econometric models. Therefore, it can be argued that they should be considered in crash prediction analysis.

It is noted that in a sense, econometric and regression models may be considered as machine learning algorithms because regression coefficients are “learned” from the data. In this dissertation, machine learning algorithms exclude econometric methods but include only Support Vector Machine, Random Forests, Bayesian Network, and Deep Learning algorithms.

4) Lack of literature on testing and evaluating the impacts of autonomous vehicles (AVs) on safety. In addition, AVs are expected to eliminate human errors in driving. Some factors previously found significant may not be significant in the AV era (e.g., driver’s gender, driver’s age, etc.). Due to these impacts, highway agencies need to consider the implementation of new safety-related designs, policies, and regulations that will impact AV operations. These implementations need to be conducted and evaluated in a safe setting to ensure that the AV does not pose undue safety hazards and produces the prospective benefits.

Also, AVs contain technology that will help collect data and provide data that may not have been accessible in the past (e.g., real-time data on weather conditions, traffic, speed of vehicles involved in crashes, etc.).

1.3 Study Objectives

In light of the above problem statement, the objectives of this dissertation are:

1) Create a framework for enhanced prediction of highway safety considering the design factors.
2) In the traditional operating environment (without AVs), apply the proposed framework to do the following:
    a) Estimate the impacts of the road-surface condition on safety
    b) Optimize the space resource allocation across highway cross sectional elements
    c) Predict the fatality status of highway segments

3) In the emerging operating environment (with AVs):
    a) Discuss the safety factors and impacts for adopting AVs
    b) Discern the current limitations of testing and evaluating AVs on in-service roads and test track
    c) Propose and demonstrate a deep learning-based framework for testing AVs using simulation that can help facilitate more realistic AV safety studies in a simulated environment.

1.4 Organization of the Dissertation

The remainder of this dissertation proceeds as follows. Chapter 2 introduces an enhanced crash prediction framework and describes the model types (for both econometric models and machine learning models) and their attributes. Implementation of the framework as well as the case studies are presented in Chapter 3, along with the evidence found regarding crash factors and roadway design in the traditional environment. Chapter 4 discusses the safety factors and the impacts in the era of AV. Chapter 5 proposes and demonstrates a simulation framework that enable researchers to test and evaluate the emerging AV technology in a safe environment. Finally, Chapter 6 summarizes and concludes this dissertation and presents the study's limitations and avenues for future work. Figure 1.1 presents the general structure of this dissertation.

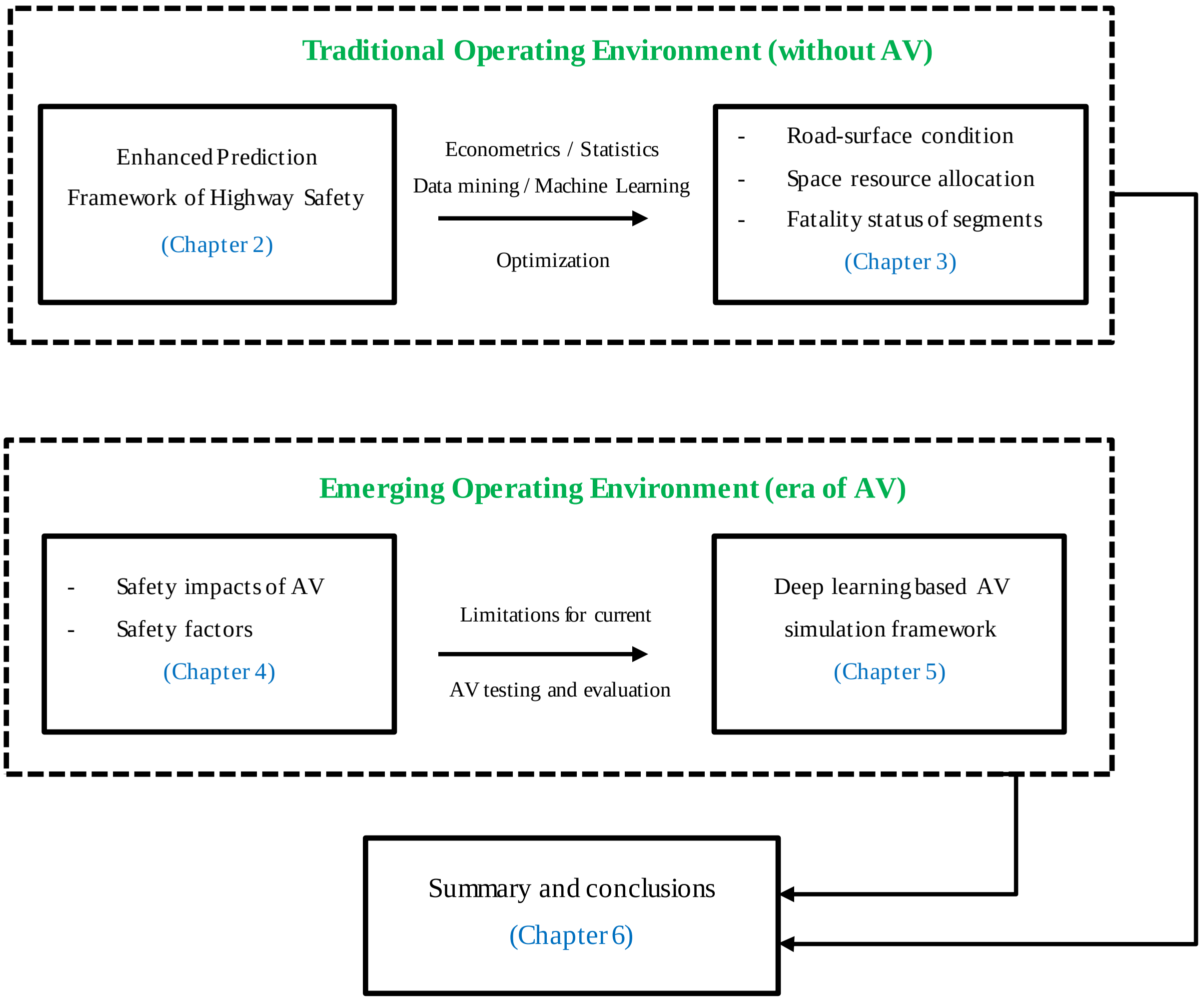


Figure 1.1 General structure of the dissertation

CHAPTER 2. FRAMEWORK FOR ENHANCED PREDICTION OF HIGHWAY SAFETY CONSIDERING THE DESIGN FACTORS

2.1 Introduction

To deepen the understanding of the impacts of highway safety factors, this dissertation proposes a framework to enhance the prediction of highway crashes. The proposed framework includes machine learning algorithms and econometric models to quantify the impacts of safety factors, which can be further incorporated into an optimization framework with engineering economic analysis to help highway agencies improve their safety-related decision-making process.

Figure 2.1 illustrates the general flow of the proposed framework. Part I of the framework consists of collecting and synthesizing information pertaining to highway safety, which includes crash data, road pavement and geometrics, traffic information, and other factors such as weather, driver age and gender, etc. This part of the framework is expected to change a lot in the era of AV for the following reasons: 1) AVs are expected to drastically affect safety factors; for example, some of the factors, such as driver age and sobriety, which significantly affect roadway safety at the present time may not be significant in the future once AV attains high market penetration. In addition, some of the new factors, including the level of automation and the types of AV control algorithm, may appear as significant safety factors. 2) AVs are expected to collect and report data automatically. As a result, information such as vehicle speed at the time of a crash, which is difficult to collect, is expected to be available for researchers. Chapter 4 of this dissertation will discuss the safety benefits and impacts of AV in detail. Part 2 of this framework entails data analysis. The discussion includes the use of advanced econometric and statistical modeling techniques for diagnostics and the use of machine learning and data mining algorithms for making reliable predictions. Part 3 of the framework consists of the results and feedback provided by the improved roadway safety modeling for highway agencies to improve their decision-making process regarding design, operation, and maintenance. Both the diagnostic and predictive models will be assessed using appropriate evaluation metrics. This assessment will also evaluate the quality of the current data and provide feedback for future collection. The elements of this framework are discussed in detail below.

The developed framework is expected to help agencies improve their safety analysis process so they can predict and evaluate crashes more reliably and identify the correct influence of the crash factors. With better prediction capabilities, the framework can help predict long-term safety performance and support funding requests for safety projects and for safety resource allocation. Highway safety projects include construction, reconstruction, expansion, rehabilitation, or maintenance of guardrails, crash barriers, median cables, road signs, pavement markings, rumble strips, etc. With better characterization of the relationship between roadway factors and crashes, such information can be provided to designers for guidance in design audits and reviews, possible revisions in lane width and shoulder width, and threshold pavement conditions in the agency's design and preservation manuals.

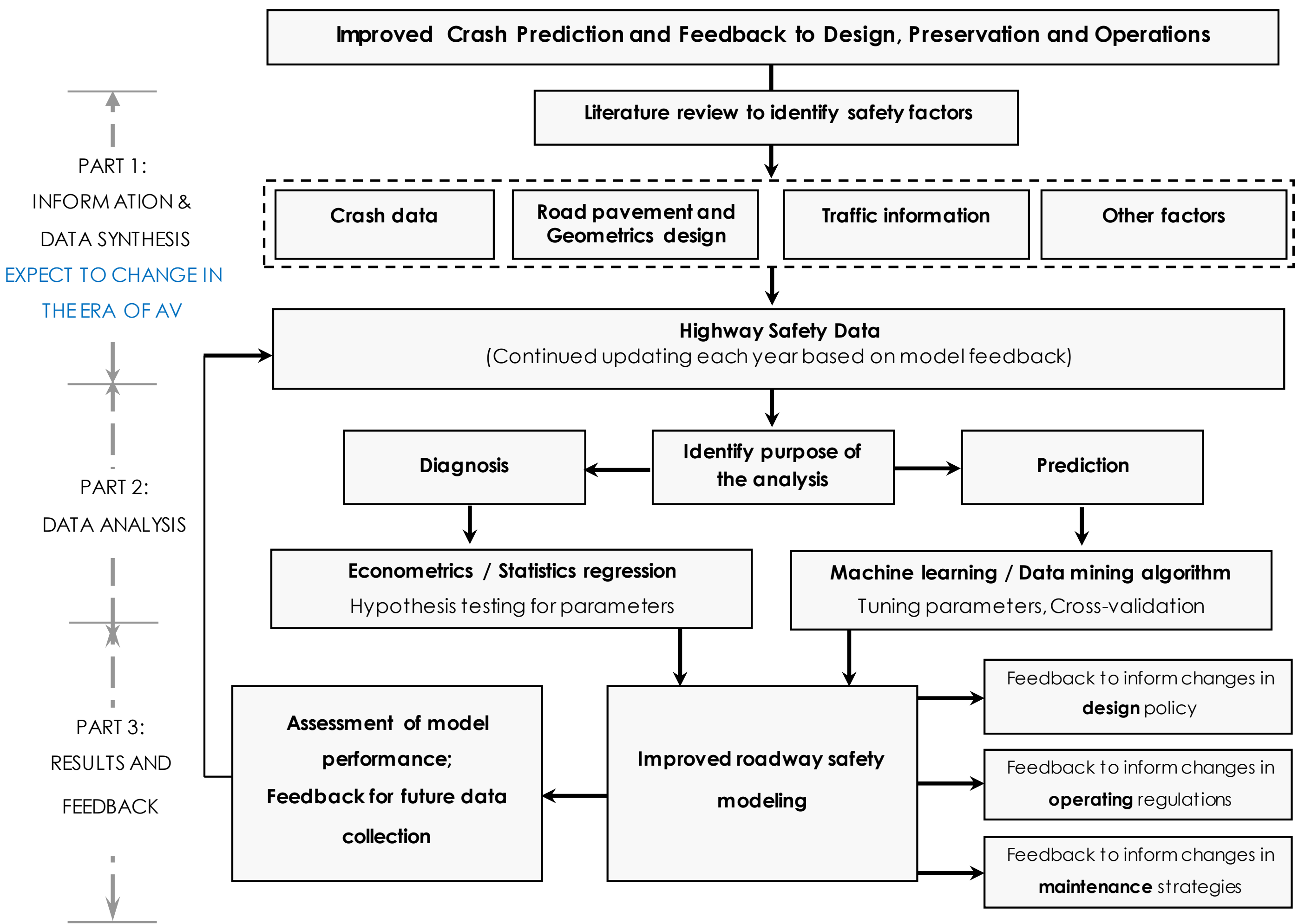


Figure 2.1 General flow of the framework

2.2 Highway Safety Data

The literature suggests that crash factors are directly related to the driver, the vehicle, the natural or built-up environment, the road pavement, the geometrics (and traffic volume), or a combination of these factors (Sinha and Labi, 2007). In this dissertation, the crash factors will be categorized as follows:

1) Road pavement, geometrics design. These factors are typically within direct control of highway agency and can be mitigated through highway project implementation.
2) Traffic information. Passenger car, truck (autonomous passenger car, autonomous truck). Traffic is considered partially under the control of the highway agency through regulations, tolling, etc.).
3) Other factors such as driver age, weather condition, vehicle condition, driving behavior, etc.

Driving behaviors, weather conditions, regulations, culture, and other factors vary across regions, and the modeling results of a certain area therefore may not be applicable to other areas. To replicate the analysis, collecting data pertaining to the future study area is recommended.

2.3 Econometric and Statistical Regression of Crash Data

2.3.1 Negative Binomial (NB) Regression for Count Data

In the literature, most of the studies regarding highway safety analysis used econometric and statistical regression models as their major approach to evaluate the impacts of safety factors (independent variables) on crashes (dependent variables). For example, one of the most popular econometric models is Negative Binomial (NB) regression, which is a commonly applied model form for analyzing crash count data (Gkritza and Mannering, 2008; Milton et al., 2008; El-Basyouny and Sayed, 2009; Labi, 2011; Dinu and Veeraragavan, 2011; Garnowski and Manner, 2011; Venkataraman et al., 2011; 2013; Wu et al., 2013; Xiong and Mannering, 2013; Behnood et al., 2014; Chen and Tarko, 2014; Russo et al., 2014; El-Basyouny et al., 2014a; Barua et al., 2016; Chen et al., 2016; Heydari et al., 2016; Serhiyenko et al., 2016; Mothafer et al., 2017; Gomes et al., 2017). The NB model has a general function form (Washington et al., 2011):

$\lambda_{ik} = \text{EXP}(\boldsymbol{\beta}_k \boldsymbol{X}'_{ik} + \varepsilon_{ik})$

where,

$\boldsymbol{X}_{ik}$= (1, X_{1k}, X_{2k}, … , X_{Nk}), is the vector of independent variables

$\boldsymbol{\beta}_k$= (β_{0k}, β_{1k}, … , β_{Nk}), the vector of coefficients

λ_{ik} is the expected number of crashes in i^{th} road segment and k^{th} severity level (e.g. fatal, injury, no-injury)

ε_{ik} is the unobserved error term for i^{th} road segment and k^{th} severity level

For this type of data, each observation i is a road segment, which contains the number of crashes (sometimes by crash severity or crash pattern) and information regarding road geometric design, traffic, pavement condition, and other factors.

2.3.2 Unobserved Heterogeneity

From the modeling perspective, the error term ε_{ik} appears to be the most intriguing element. It consists of unobserved information (heterogeneity). Since it is almost impossible to have the perfect dataset due to lack of information, researchers typically expend great effort in making the error term ε_{ik} as small as possible to calibrate coefficients $\boldsymbol{\beta}_k$ as precise as possible (it is assumed that the errors in data collection are negligible or under control compared to the heterogeneity). These efforts include choosing different distributions for $\boldsymbol{\beta}_k$ and/or ε_{ik} through the modeling procedure as well as other modeling advancements.

Some advancements for minimizing the error term include estimation of multivariate/bivariate crash count models (Heydari et al., 2016; Serhiyenko et al., 2016), incorporation of random parameters in crash models (Venkataraman et al., 2014; Coruh et al., 2015; Barua et al., 2016; Bhat et al., 2017; Waseem et al., 2019); accounting for spatial/temporal correlation in crash models (Chiou et al., 2014; Chiou and Fu, 2015; Hong et al., 2016; Saeed et al., 2019a) and estimation of latent class, finite mixture, two-state Markov switching, mixed logit and neural network models (Malyshkina et al., 2009; Park and Lord, 2007; Behnood et al., 2014; Behnood and Mannering, 2016; Zeng et al., 2016). These improvements arose from concerted efforts to identify and address a number of complex issues associated with highway crash data and analysis, which included unobserved heterogeneity, spatial and temporal correlation, correlated collision types (Mannering and Bhat, 2014; Chiou and Fu, 2015; Hong et al., 2016; Mannering et al., 2016; Amoh-Gyimah et al., 2017; Behnood and Mannering, 2017a; Huang et al., 2017; Qiao et al., 2018). For example, failure to address any unobserved heterogeneity in the data could result in biased, inefficient, or inconsistent parameter estimates when a traditional fixed-parameter model is specified, consequently leading to incorrect inferences (Washington et al., 2011). Taking a similar position, Mannering et al. (2016) argued that neglecting any unobserved heterogeneity (thus, assuming implicitly that the effects of observable factors are the same across all observations) may cause model specification problems, biased and inefficient parameters, and ultimately, erroneous predictions.

In this vein, a number of research studies over the last decade have used a variety of statistical tools to account for the unobserved heterogeneity across observations in crash data (Gkritza and Mannering, 2008; Milton et al., 2008; El-Basyouny and Sayed, 2009; Dinu and Veeraragavan, 2011; Garnowski and Manner, 2011; Venkataraman et al., 2011; 2013; 2014; Wu et al., 2013; Xiong and Mannering, 2013; Behnood et al., 2014; Chen and Tarko, 2014; Russo et al., 2014; El-Basyouny et al., 2014a; Barua et al., 2016; Mothafer et al., 2017; Gomes et al., 2017).

The use of random-parameter models to account for unobserved heterogeneity is prevalent in recent studies (Gkritza and Mannering, 2008; Milton et al., 2008; Chen and Tarko, 2014; Russo et al., 2014; Venkataraman et al., 2014; Coruh et al., 2015; Behnood and Mannering, 2015; 2016; 2017b; Sarwar et al., 2017; Seraneeprakarn et al., 2017; Anderson and Hernandez, 2017; Bogue et al., 2017; Saeed et al., 2017; Ahmed et al., 2017). This is because random parameter models can explicitly account for heterogeneity across observations that are mainly due to unobserved explanatory traffic, environmental, roadway design, vehicle characteristics, driver behavior, and other related factors. This technique allows every estimated parameter in the model to vary across each individual observation according to an analyst-defined continuous distribution and thus requires a parametric distribution. Since it is quite possible that the individual coefficients in the model will follow different distributions, using a parametric approach could potentially cause statistical problems (Mannering et al., 2016). Analysts find that it is often impractical to consider all these factors simultaneously because highway crashes are often the outcome of a combination of factors, such as those related to the driver, vehicle, roadway, and natural environment and the road (engineering) as well as the complex interactions among them (Sinha and Labi, 2011; Mannering et al., 2016).

2.3.3 Multivariate Analysis

Crash data are inherently multivariate. Therefore, it can be problematic to model crashes within each crash severity level separately without accounting for possible correlations between the crashes across the severity levels. Using a univariate modeling approach for the correlated crash counts can lead to less precise estimates of the risk factors associated with the different severity levels. In this regard, using multivariate modeling approaches to model crash frequency jointly across the different crash severity levels has gained attention in recent years (Bijleveld, 2005; Ma and Kockelman, 2006; Song et al., 2006; Park and Lord, 2007; Ma et al., 2008; Wang et al., 2011; El-Basyouny et al., 2014a, 2014b; Lee et al., 2015; Abay and Mannering, 2016; Barua et al., 2016; Serhiyenko et al., 2016; Zeng et al., 2016; Heydari et al., 2016; 2017; Sarwar et al., 2017). Multivariate models have been found to adequately account for correlations among the different levels of crash severity, which seem to emerge from a variety of sources including data on traffic collisions that involve multiple occupant injuries from the same crash. See Mannering et al. (2016) for more detailed information. In this example, the occupants would likely suffer different levels of injury (crash severity), but the unobserved factors that influence the injury severity levels would be correlated (Eluru et al., 2010; Abay et al., 2013; Yasmin et al., 2014; Russo et al., 2014; Ma et al., 2017). The unobserved factors may impact the multiple crash counts of different severity levels simultaneously for each roadway segment under consideration (Mannering et al., 2016); therefore, estimating separate univariate models could cause statistical problems. To avoid these problems, the concept of random parameters in a multivariate modeling framework to accommodate the unobserved heterogeneity in the correlated crash data is presented in Chapter 3.

2.4 Machine Learning and Data Mining Methods

Machine learning algorithms are known for their superior predictive power and their ability to deal with the problems related to high dimensionality (Hastie et al., 2009). In recent years, the field of transport studies has seen an explosion of interest in machine learning and data mining methods. However, in the highway safety analysis area, these state-of-the-art methods have not been well explored. Some general description of machine learning and data mining methods are discussed below (Hastie et al., 2009; James et al., 2013).

2.4.1 k-Fold Cross-validation

The idea of cross-validation is to select the appropriate hyper-parameters for the classifiers to perform well not only in training but also in validation and testing to avoid overfitting and underfitting (Hastie et al., 2009).

This approach involves randomly dividing the data into k folds (groups) of equal size. Then, for the first round, the first fold is used as a validation dataset, and the classifier is fitted onto the remaining folds. After training, the mean squared error MSE_1 is computed on the first fold. This procedure is repeated k times (e.g., for the second round, the second fold is used as a validation dataset, and the same classifier is fitted onto the remaining folds. After training, the mean squared error MSE_2 is computed on the second fold, etc.) This process results in k estimated validation errors: MSE_1, MSE_2, . . . , MSE_k. Computing the average of these values, the k-fold cross-validation (CV_k) estimate is obtained:

$$CV_k = \frac{1}{k}\sum_{i=1}^{k} MSE_i$$

For each set of hyper-parameters, CV_k is computed using the above procedure. The final classifier uses the set of hyper-parameters that yields the smallest CV_k.

k = 5 or k = 10 is typically preferred in practice. When k is set to equal the number of observations n, it becomes a special case Leave-One-Out Cross-Validation (LOOCV), which requires fitting the machine learning method n times (n refers to number of observations). This process potentially can be computationally expensive. Cross-validation is a general approach that can be applied to almost any machine learning method. Some learning algorithms have computationally intensive fitting procedures. Therefore, performing LOOCV may pose computational problems, particularly if n is extremely large. In such cases, performing 5-fold or 10-fold cross-validation could be much more feasible.

2.4.2 Support Vector Machine

Support vector machine (SVM) is a supervised machine learning algorithm that analyzes data based on statistical learning theory. In general, if the data (with different labels) can be perfectly separated using a hyperplane, SVM can help establish the maximal margin hyperplane (the hyperplane that is farthest from the training observations) (Scholkopf and Smola, 2001; Hastie et al., 2009; James et al., 2013). Consider the construction of a maximal margin hyperplane based on n training observations $\mathrm{x}_1, \mathrm{x}_2 \ldots \mathrm{x}_n \epsilon R^p$ and with corresponding labels $\mathrm{y}_1, \mathrm{y}_2 \ldots \mathrm{y}_n \epsilon \{-1,1\}$. The maximal margin hyperplane solution can be established using optimization techniques as follows:

$$
\begin{aligned}
&\text{Maximize } M(\beta_0, \beta_1 \ldots \beta_p, \varepsilon_1, \ldots \varepsilon_n) \\
&\qquad \text{s.t. } \textstyle\sum_{j=1}^{p} \beta_j = 1 \\
&\qquad \mathrm{y}_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}) \geq M(1 - \varepsilon_i) \quad \forall i = 1, \ldots, n \\
&\qquad \varepsilon_i \geq 0, \ \ \forall i = 1, \ldots, n \qquad \sum_{i=1}^{n} \varepsilon_i \leq C
\end{aligned}
$$

where C is a non-negative constant, M is the width of the margin, $\beta_0, \beta_1 \ldots \beta_p$ are the coefficients that define the separating hyperplane in p dimensional space, and $\varepsilon_1, \ldots \varepsilon_n$ are the slack variables that allow individual observations to stay on the "wrong" side of the margin.

For data that are not linearly separable, a kernel function $K(x, \mathrm{x}_i)$ can be used to replace the hyperplane function, which transforms the input space into a high-dimensional space using a nonlinear transformation defined by an inner product function. In this dissertation, two popular kernel functions (radial basis function [RBF] and polynomial) are applied and compared:

$$\text{RBF kernel:} \quad K(x, x') = exp(-\gamma||x - x'||^2)$$
$$\text{Polynomial kernel:} \quad K(x, x') = (x^T \cdot x' + 1)^d$$

where γ is a hyper-parameter that controls the shape of the RBF kernel and d is the degree of the polynomial kernel. For fixed values of C and γ or d, there is always a globally optimal solution to the optimization problem. In practice, these hyper-parameters are generally selected using cross-validation.

2.4.3 Random Forests

Random forests (Breiman, 2001) is a substantial modification of bagging (a general procedure for reducing the variance of a machine learning model (James et al., 2013) that builds a large collection of de-correlated trees and then averages them. In many cases, random forests perform in a way similar to boosting (a general procedure for improving model prediction by combining many “weak” models (Hastie et al., 2009)), however, random forests are generally simpler to train and tune. The essential idea behind bagging is to reduce the variance by generating many sub-datasets that are used to train approximately unbiased models. Decision trees are ideal candidates for bagging due to their extraordinary flexibility and ability to capture complex interaction structures in the data. If trees grow sufficiently deep, they have relatively low bias. The random forests algorithm is provided in Hastie et al. (2009) and James et al. (2013):

- Train the random forests:

  For b = 1 to B:

  (a) Draw a bootstrap sample Z* of size N from the training data.

  (b) Grow a random forest tree T_b to the bootstrapped data by recursively repeating the following steps for each terminal node of the tree until the minimum node size n_{min} is reached:

  i. Select m variables at random from the p variables.

  (m can be any integer from 1 to p. Common choices are $m = \sqrt{p}$ or $m = log_2(p)$.)

  ii. Pick the best variable/split-point among the m variables.

  iii. Split the node into two daughter nodes.

  Output the ensemble of trees $\{T_b\}_1^B$.

- Make a prediction at a new point x:

  Let $C^b(x)$ be the class prediction of the b^{th} random forest tree.

  Then, *predicted class* = *majority vote* $\{C^b(x)\}_1^B$.

2.5 Metric of Model Performance

Using appropriate metrics is crucial in model performance evaluation. A common error in highway safety literature for evaluating models is using only accuracy (for a machine learning model), ρ^2 or likelihood (for econometric regression) to evaluate and select models.

Imagine a case where 90% of the highway segments do not have fatal crashes and the remaining 10% do. One can easily propose a completely useless model with a high accuracy of

90%: no matter what information (e.g., geometric design, traffic volume) is provided, the model always predicts "the segment does not have fatal a crash." Therefore, for highway safety data analysis, the accuracy of a model by itself is generally not a good metric for evaluating model performance while using machine learning algorithms. Three metrics for model performance evaluation (Precision, Recall, and F1 score) are introduced here and recommended for use (Hastie et al., 2009; James et al., 2013).

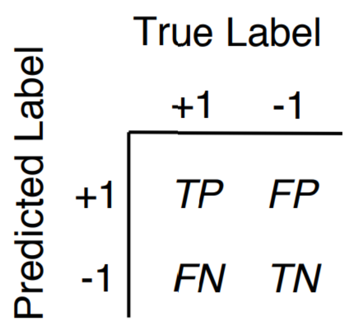


Figure 2.2 Definitions of true positive (TP), false positive (FP), false negative (FN), and true negative (TN)

Based on the definitions of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) (Figure 2.2), the Precision, Recall, and F1 score are defined as:

$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}}$ (% positive predictions that are correct)

$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}$ (% positive instances that are predicted positive)

$\text{F1} = 2\frac{\text{Precision}\cdot\text{Recall}}{\text{Precision+Recall}}$ (Harmonic mean of precision and recall)

Chapter 3 discusses the advantages of using Precision, Recall, and F1 score in tuning machine learning models to analyze fatal highway crashes.

Regarding performance metrics for econometric models, ρ^2 and likelihood are good representations of a model's overall fit to the data. However, from a statistics perspective, adding increasingly more independent variables or higher order terms (e.g. $x^2, x^3, \ldots, x^n$) always increases a model's overall fit to the data. In this sense, since there is no penalty for the increased number of independent variables, using only ρ^2 or likelihood in model selection may lead to issues such as overfitting. Therefore, in the model selection process, Akaike information criterion (AIC) (Akaike, 1987), Bayesian information criterion (BIC) (Burnham and Anderson, 2004), or other model performance metrics are suggested because they penalize large numbers of independent variables. Chapter 3 presents examples of the use of AIC in evaluating advanced econometric models.

When selecting machine learning or econometric regression models, the trade-off between prediction accuracy and model interpretability should be considered. Of the many commonly used econometric and machine learning methods, some are less flexible than others. For example, linear regression is a relatively inflexible approach because it can generate only linear functions such as lines in a two-dimensional space or a linear hyperplane in a high-dimensional space. Other methods, such as decision tree, are considerably more flexible because they can generate a much more complicated and multi-dimensional discrete function.

If inference is the main interest, then restrictive models are much more interpretable. For example, when inference is the goal, a linear model may be a good choice because it is quite easy to understand the relationship between the dependent variable Y and the independent variables X_1, X_2,…X_p. On the other hand, flexible approaches, such as a decision tree, will generate a complicated function where it will be relatively difficult to interpret the effects of any independent variable on the dependent variable.

Generally, regression type models are often considered less flexible compared to machine learning algorithms for analyzing highway safety data. Due to their functional forms, regression models are typically continuous, monotone, and differentiable. With regard to econometric models, the marginal effects and elasticity for regression functions can be easily computed. As a result, econometric and regression models have high interpretation power. In contrast, machine learning algorithms usually have relatively lower interpretation power but higher prediction power due to the flexibility in their function forms. Based on the purpose of their specific application, highway agencies should consider this trade-off when selecting appropriate models for predicting or analyzing crashes.

2.6 Network-Level and Project-Level Decision Making Using Optimization

After analyzing safety data with machine learning and econometric models, the sub-framework of optimization programming introduced in this section is recommended for helping highway agencies improve the decision-making process at both the network level (Labi et al., 2006; Lamptey et al., 2008; Murillo Hoyos et al., 2015) and the project level (Irfan et al., 2011; Labi et al., 2017). For example, in network level decision-making, machine learning models are generally preferred as econometric regression models compute the average or overall response values by their definition, which will lead to relatively low performance compared to machine learning models insofar as predicting individual cases. Machine learning algorithms generate the inputs of predicted "dangerous/high risk" road segments for the optimization sub-framework. Using optimization, highway agencies can determine the most cost-effective intervention strategies to improve their funding allocations. For example, if they seek to be conservative and cover as many potential "dangerous/high risk" road segments as possible, machine learning models should be tuned based on recall; if funding is extremely limited, and agencies seek to avoid wasting any funds on false positive "dangerous/high risk" road segments, machine learning models should be tuned based on precision; otherwise, F1 score should be used as the model performance metric in order to achieve a balanced decision. See Section 3.4 for more details.

Regarding project-level decision-making, econometric models are generally preferred due to their high interpretation power. Using calibrated regression coefficients, highway agencies can evaluate the impacts of changes in road safety factors (e.g., widening lane width and improving pavement condition). For example, Section 3.3 describes a scenario where the outputs from econometric models are used as inputs for the optimization sub-framework. In that section, life-cycle analysis is incorporated into the optimization model to determine the optimal distribution of lane and shoulder width on a two-lane highway with a restricted right of way.

2.7 Chapter Summary

This chapter presents a proposed framework to enhance prediction of highway crashes. With a comprehensive literature review, the chapter described the elements of the proposed framework, which included advance econometrics and machine learning techniques. Then, the concepts and reasons for using appropriate performance metrics for model assessment were explained. Finally, the chapter discussed the trade-offs between econometrics and machine learning models, and how to incorporate them within an optimization framework to improve the network-level or project-level decision making process.

CHAPTER 3. SAFETY IMPLICATIONS OF ROADWAY DESIGN AND MANAGEMENT — OLD FINDINGS AND NEW EVIDENCE IN THE TRADITIONAL ENVIRONMENT

3.1 Introduction

In response to the growing public demand for safer roadways, transportation agencies have developed a variety of safety programs, roadway safety education, accelerated highway renewal, and innovative technology to improve roadway safety design and management (Bonneson et al., 2007). Some of the traditional methods to evaluate the safety effect for a certain design are based on the laws of physics, before-after studies, and comparisons (e.g., two types of different safety components). However, highway safety is generally influenced by multiple factors that include highway geometry, driver characteristics, roadway surface condition, vehicle characteristics, and the environment. Typically, crashes are caused by an interaction of more than one of these factors. Due to these complicated interactions, traditional methods may not able to conclude accurate results. As a result, highway engineers tend to follow design standards, such as AASHTO (2011), and rely on policies to guide them in the design phases. The underlying assumption is that these design standard and policies are expected to "warrant safe roadways." Generally, this assumption regarding safety is valid; however, with the continual changes occurring in the driving culture, traffic environment, vehicle characteristics, emerging technology, etc., traditional approaches for design and management may not be able to "warrant safe roadways," and the corresponding safety assumption may be violated.

Fortunately, over the past decade, the quality of crash data has been improving, which has created opportunities in developing advanced statistical models and data mining techniques. Using these advancements, this dissertation proposes a framework to enhance the prediction of highway crashes to help agencies estimate the safety impacts of design factors and management more precisely. In addition, from the roadway safety perspective, engineering factors are considered particularly relevant because 1) engineering factors are within the direct control of highway agencies and 2) enhancements in engineering factors can help reduce the crash contributions of the other crash factors (Sinha and Labi, 2007). Taking these into account, and with the opportunities discussed in Chapter 1, this dissertation demonstrates the implementation of the proposed prediction framework using an Indiana crash dataset for three different scenarios:

a) Estimating the impacts of the road-surface condition with advanced econometric models
b) Identifying the optimal lane and shoulder width on a two-lane rural highway using multiobjective optimization
c) Predicting fatal highway segments with machine learning algorithms

The original data for the case studies are from a 2006 study by Labi (2006) that analyzed several geometric factors on rural two-lane roads in a bid to ascertain their relative effects on crash reduction. Labi's study used a variety of data sources to formulate models to estimate the number of crashes. Those datasets included:

- Roadway Inventory Dataset (RID), a comprehensive database of the geometric, operational, and pavement conditions for road segments on the Indiana state highway network.
- Route Alignment Dataset (RAD), which contains information on the horizontal and vertical alignments (including information on segment curvatures) for the all Indiana state highways.
- Indiana Crash Dataset (ICD), which is a listing of the individual crashes and the corresponding circumstances on Indiana highways based on data received from the Indiana State Police. The records from 1997 to 2000 were utilized for the crash prediction models.
- Road Sections Dataset (RSD), which is a geospatial database consisting of the RID and various state and local jurisdictional boundaries, with the goal of assigning each crash and corresponding road segment to an urban or rural categorization.

Data for the agency cost estimates were collected from two sources: 1) information on shoulder construction and maintenance costs were from a 2001 study at Iowa State University and 2) lane construction and maintenance costs were from the Wisconsin and Washington DOT. Data from Midwestern states were particularly valuable because they generally share similar conditions with Indiana that are related to the economy, highway operations, weather, and terrain.

3.2 Impacts of Road-Surface Condition on Safety

Recent studies have begun to shed more light on the crashes experienced on rural roads by examining the influence of a road's pavement surface condition. In a bid to contribute to this growing body of knowledge and to facilitate comprehensive evaluation of pavement maintenance projects, this case study explored the safety effects of the pavement condition of rural roads. It tested the hypotheses that pavement roughness generally has a non-trivial residual impact on safety outcomes and that the magnitude and direction of these impacts differ across road segments. To explore these hypotheses, the case study contained crash frequency models for three levels of crash severity and five levels of road surface condition. The developed models used the multivariate random parameters negative binomial (MRPNB) specification to account for the unobserved heterogeneity and correlation among the different levels of crash severity.

3.2.1 Literature Review

Several recent studies found significant variation in the effects of roughness or overall pavement condition on crash frequency and severity (Lamptey, 2004; Labi, 2006; Labi, 2011; Mayora and Piña, 2009; Li et al., 2013; Buddhavarapu et al., 2013; Anastasopoulos, 2016). A school of thought contends that crash propensity increases as pavement condition improves because a very good pavement condition is generally associated with higher speeds and subsequently higher crash counts. Other researchers maintained that such effects have yet to be proven beyond a reasonable doubt (Harwood et al., 2003), a position that was seemingly corroborated by Agent et al. (2004) whose analysis using Kentucky highway data found no difference in crash frequency before and after pavement resurfacing. Al-Masaeid (1997), on the other hand, indicated that pavement condition, expressed in terms of the international roughness index (IRI) had significant but rather opposite effects on single- and multiple-vehicle crashes. NCHRP Project 17-9(2) reported that resurfacing had a negative effect on safety in certain states and a positive effect in others. Earlier work by Cleveland (1987) seemed rather prescient, suggesting that the direction of impact was rather ambiguous and that additional research was needed. Drivers tend to behave less carefully on roads with adequate levels of service where they perceive a low level of risk and more carefully on roads with inadequate levels of service where there is a high level of perceived risk. Therefore, it seems plausible that a high level of pavement condition could exhibit either direction of impact depending on driver behavior: a positive net impact where adequate friction and surface condition outweighed the positive risk compensation of drivers, and a negative net impact where the negative risk compensation of drivers outweighed adequate friction and surface condition. Further, the net direction of impact may be positive or negative depending on factors such as driver (heterogeneous) behavior in situations of good or poor pavement condition, the level of the poorest pavement condition in the analysis dataset.

It is generally hypothesized that crash propensity decreases as pavement condition improves from poor (high IRI) to good (low IRI); however, it is important to explore the variable effects of the pavement condition on safety that could arise due to the heterogeneous response in driver behavior (Mannering et al., 2016). This variation in response or adjustment/non-adjustment of driving behavior to the perceived level of risk, left unobserved, can vary from one road segment to the next, particularly in the case of poor pavement condition. Most past studies suggested a unidirectional relationship between pavement condition and safety. However, the possibility of risk-compensation or offset-hypothesis effects, where driver behavior changes in response to the level of perceived risk (Winston et al., 2006), exists. In the case of poor pavement conditions, drivers may or may not adjust their behavior. Where drivers make adjustments, they drive more carefully in response to a poor pavement condition (higher level of perceived risk) but less carefully in the case of a pavement in good condition which makes them feel safer. Investigating this issue is important to achieving this study's main objective of exploring the impact of pavement condition on rural roadway safety across the various levels of pavement condition. These levels, in terms of pavement roughness (IRI in/mile), may be defined as: Excellent (60 – 100), Good (101 – 125), Good ~ Fair (126 – 150), Fair (151 – 200) and Poor (≥ 200) based on the thresholds established in the pavement management literature (INDOT, 2000; Sinha et al., 2005). In addressing the issue, this dissertation intends to produce models that reduce the bias and inconsistency in the estimation process and to enhance the reliability of highway safety prediction.

3.2.2 Case Study

3.2.2.1 Introduction

In this case study, the multivariate random parameter negative binomial (MRPNB) model was chosen for analyzing the pavement condition impacts on highway safety. The overdispersion parameters confirmed that the negative binomial models were preferred over their Poisson counterparts. Winkelmann (2000), Winkelmann (2008), and Shi and Valdez (2014) described the general framework of the multivariate negative binomial model with the expected number of crashes, λ, in i^{th} road segment and k^{th} severity level. In this dissertation, these levels are fatal, injury, and no-injury):

$$\lambda_{ik} = \mathrm{EXP}(\boldsymbol{\beta}_k \boldsymbol{X}'_{ik} + \varepsilon_{ik}) \qquad k = 1,2,3$$

where, $\boldsymbol{X}_{ik}$= (1, X_{1k}, X_{2k}, … , X_{Nk}), is the vector of independent variables

$\boldsymbol{\beta}_k$= $(\beta_{0k}, \beta_{1k}, \ldots, \beta_{Nk})$, the vector of coefficients

EXP(ε_{ik}) is a multivariate gamma-distributed error term with mean 1 and variance α^{-1}.

The negative binomial is appropriate where α is significantly different from 0, otherwise, the Poisson model is preferred (Washington et al., 2011).

In addition, the error term ε_{ik} is multivariate and is normally and independently distributed with zero mean, variance $\boldsymbol{\sigma^2}$, and correlation ρ based on the unstructured correlation covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{21} & \sigma_1\sigma_3\rho_{31} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 & \sigma_2\sigma_3\rho_{32} \\ \sigma_3\sigma_1\rho_{31} & \sigma_3\sigma_2\rho_{32} & \sigma_3^2 \end{bmatrix}$$

It is important to account for the influence of unobserved heterogeneity; for example, the effects of individual independent variables may be different across road segments. In this case, the estimated parameter should not be specified as a fixed value but instead should be made to follow a distribution such as normal, uniform, etc. To do so, random parameters were introduced into this multivariate modeling framework. Greene (2008) described estimation procedures for incorporating random parameters in Poisson and negative binomial count-data models. The estimable parameters are then written as:

$$\boldsymbol{\beta_k} = \boldsymbol{\beta} + \boldsymbol{\omega}_k$$

where $\boldsymbol{\omega_k}$ is a randomly-distributed term (e.g., a normally-distributed term with mean 0 and variance σ^2). A random parameter should be kept in the model only if its estimated σ^2 is significantly greater than zero and if the likelihood ratio test also recommends using a model with random parameters.

The simulation-based maximum likelihood method to maximize the simulated log-likelihood function, with $\sigma \neq 0$ allowing for heterogeneity, was used for estimating the random parameters multivariate negative binomial models. Two hundred (200) Halton draws typically provide an efficient distribution of draws for numerical integration (Milton et al., 2008; Russo et al., 2014), and therefore was used in this dissertation to estimate the random parameters.

The Akaike information in this case study was found to be smaller than that of the comparison models and thus provided a superior fit and possessed parameter estimates that closely converged to the true model parameters.

To evaluate the statistical superiority of the alternative model specifications (the fixed vs. random parameters multivariate negative binomial, hereinafter referred to as the fixed model f and the random parameters model r, respectively) likelihood ratio tests (Washington et al., 2011) was carried out for the model representing pavements in poor condition:

$$\chi^2 = -2\left[LL(\beta_f) - LL(\beta_r)\right]$$

where, $\mathrm{LL}(\beta_f)$ is the log-likelihood at convergence of the fixed parameters multivariate negative binomial model and $\mathrm{LL}(\beta_r)$ is the log-likelihood at convergence of the random parameters multivariate negative binomial model.

Moreover, the models were estimated separately for the different levels of pavement condition: excellent, good, good~fair, fair, and poor. The likelihood ratio test for parameter transferability, which was carried out to investigate the transferability of the results over all five levels of pavement condition, is given as (Washington et al., 2011):

$$\chi^2 = -2\left[LL(\beta_t) - LL(\beta_i) - LL(\beta_j)\right]$$

where, $LL(\beta_t)$ is the log-likelihood at convergence of the model estimated with the data from all the pavement condition levels, $LL(\beta_i)$ is the log-likelihood at convergence of the model using pavement condition level i data, and $LL(\beta_j)$ is the log-likelihood at convergence of the model using pavement condition level j data.

The marginal effect reflects the impact of a one-unit change in an independent variable on the dependent variable (in this case, the expected frequency of non-injury, injury, or fatal crashes). For crash severity level k, the marginal effect of the m^{th} independent variable is calculated as:

$$\mathrm{ME}_{X_{mk}}^{\lambda_{mk}} = \frac{\partial \lambda_k}{\partial X_{mk}} = \beta_{mk} EXP(\boldsymbol{\beta_k X_k'})$$

where λ_k is the estimated model for crash severity level k; X_{mk} is the m^{th} independent variable; β_{mk} is the corresponding estimated coefficient; $\boldsymbol{X_k'}$ are $\boldsymbol{\beta_k}$ are the vectors of the independent variables and the corresponding estimated coefficients, respectively.

3.2.2.2 Data Statistics

The studied database of 1,524 highway segments was developed by merging two independent datasets: a road pavement condition dataset and a road safety dataset. The combined dataset contained the historical vehicle crash data of rural highways in Indiana over a three-year period. Based on the pavement condition, the data were further divided into subsets depending on the level of pavement condition in terms of the International Roughness Index (IRI): Excellent ($60<\mathrm{IRI}\leq100$), 681 observations; Good ($100<\mathrm{IRI}\leq125$), 324 observations; Good ~ Fair ($125<\mathrm{IRI}\leq150$), 111 observations; Fair ($150<\mathrm{IRI}\leq200$), 99 observations; Poor ($\mathrm{IRI}>200$), 67 observations. Each of these subsets contained the pavement condition and road geometric information such as lane width, inside and outside shoulder widths, and curvature. In addition, the crashes on each road segment were categorized into three levels of crash severity: fatal, injury, and no-injury.

The pavement surface condition indicator, IRI, is measured in units of inches per mile (in/mile) or meters per kilometer (m/km), which is defined technically as the deviations of a pavement surface from a true planar surface with characteristic dimensions that affect vehicle dynamics, ride quality, dynamic loads, and drainage (Sayers and Steven, 1998). Roughness is an important pavement characteristic because it affects the dynamics of moving vehicles, increasing the wear on vehicle parts and the handling of a vehicle, and therefore has significant impacts on vehicle operating costs and safety, as well as the comfort and speed of travel. It also increases the dynamic loadings imposed by vehicles on the surface, accelerating the deterioration of the pavement structure (Sayers et al., 1986). IRI is typically measured annually using pavement profiling laser sensors. IRI is the accumulated vertical deviations of the pavement surface along a longitudinal path divided by the distance traveled by the vehicle during the measurement (such as kilometers or miles). For this reason, the IRI units are meters per kilometer (m/km) or inches per mile (in/mi). The IRI of a new pavement can be as low as 60 in/mi, and for a pavement in poor condition, as high as 300 in/mi (Sayers et al., 1986; Chang et al., 2009). IRI, which often serves as the basis for pavement preservation decisions, is collected on a regular basis by most highway agencies worldwide and is believed to adequately represent user perception of road quality. In the U.S., the FHWA adopted IRI for the state highway system because it uses a standardized procedure, indicates consistency across jurisdictions, represents an objective measurement, and has gained acceptance worldwide as a universal measure of pavement surface condition. Table 3.1 presents the summary statistics of the key variables used in the analysis.

Table 3.1 Descriptive statistics of the key variables

Variable Description	Mean	Std Dev	Minimum	Maximum
Pavements in Excellent Condition				
$(60 < \text{IRI} \leq 100)$				
Total number of fatal crashes	0.217	0.472	0	3
Total number of injury crashes	3.800	4.125	0	41
Total number of no-injury crashes	13.638	13.165	0	110
IRI (in/mile)	80	11.092	60.500	100
Segment length in miles	5.901	2.599	0.560	16.610
Average annual daily traffic of passenger cars (in 10,000s)	1.419	1.457	0.027	12.865
Natural logarithm of average daily trucks	7.160	1.073	3.701	9.867
Number of lanes	3.289	1.490	2	7
Lane width in ft.	12.407	1.889	8.290	26.520
Median width in ft.	23.834	28.205	0	99
Outside shoulder width in ft.	6.642	3.494	0	13
Inside shoulder width in ft.	1.887	2.532	0	18
Average vertical curve grade (%)	1.143	1.713	0	6.446
Road segment with 2 lanes	56.82%			
US road	33.92%			
SR road	40.67%			
Interstate	25.40%			
Pavements in Good Condition				
$(100 < \text{IRI} \leq 125)$				
Total number of fatal crashes	0.138	0.363	0	2
Total number of injury crashes	3.296	3.414	0	21
Total number of no-injury crashes	11.802	11.399	0	88
IRI (in/mile)	109	6.880	100.500	125
Segment length in miles	5.837	2.558	0.560	16.610
Average annual daily traffic of passenger cars (in 10,000s)	0.945	1.168	0.029	12.446
Natural logarithm of average daily trucks	6.808	0.971	3.779	9.834
Number of lanes	2.858	1.376	2	7
Lane width in ft.	12.514	2.283	8.290	26.520
Median width in ft.	16.829	28.373	0	99
Outside shoulder width in ft.	5.953	3.515	0	12
Inside shoulder width in ft.	1.106	1.981	0	18
Average vertical curve grade (%)	1.229	1.719	0	6.753

Table 3.1 continued

Road segment with 2 lanes	71.60%			
US road	44.38%			
SR road	40.18%			
Interstate	15.44%			
Pavements in Good ~ Fair Condition				
(125 < IRI ≤ 150)				
Total number of fatal crashes	0.144	0.481	0	4
Total number of injury crashes	3.144	3.476	0	22
Total number of no-injury crashes	11.945	10.340	0	51
IRI (in/mile)	137	7.251	125.500	150
Segment length in miles	5.730	2.538	1	15.140
Average annual daily traffic of passenger cars (in 10,000s)	1.041	1.319	0.031	10.198
Natural logarithm of average daily trucks	6.845	1.028	3.786	9.635
Number of lanes	3.144	1.520	2	7
Lane width in ft.	12.053	1.685	9	25.160
Median width in ft.	19.862	27.981	0	99
Outside shoulder width in ft.	6.385	3.509	0	12
Inside shoulder width in ft.	1.540	2.088	0	9
Average vertical curve grade (%)	1.109	1.692	0	5.838
Road segment with 2 lanes	63.06%			
US road	52.25%			
SR road	40.54%			
Interstate	7.20%			
Pavements in Fair Condition				
(150 < IRI ≤ 200)				
Total number of fatal crashes	0.212	0.498	0	2
Total number of injury crashes	3.667	3.016	0	14
Total number of no-injury crashes	12.101	9.695	0	55
IRI (in/mile)	174	13.572	151	200
Segment length in miles	5.434	2.224	0.620	9.910
Average annual daily traffic of passenger cars (in 10,000s)	0.954	0.718	0.066	3.068
Natural logarithm of average daily trucks	6.940	0.873	4.607	8.434
Number of lanes	3.454	1.501	2	5
Lane width in ft.	12.163	2.003	8.750	24
Median width in ft.	24.302	25.577	0	75.870

Table 3.1 continued

Outside shoulder width in ft.	6.813	3.584	0.300	13
Inside shoulder width in ft.	2.055	2.460	0	13
Average vertical curve grade (%)	1.495	1.798	0	6.753
Road segment with 2 lanes	51.51%			
US road	65.66%			
SR road	27.27%			
Interstate	7.07%			
Pavements in Poor Condition				
(IRI > 200)				
Total number of fatal crashes	0.089	0.286	0	1
Total number of injury crashes	2.716	3.450	0	21
Total number of no-injury crashes	8.895	8.884	0	54
IRI (in/mile)	235	34.847	200.500	365
Segment length in miles	5.127	1.960	1.120	9.510
Average annual daily traffic of passenger cars (in 10,000s)	0.779	0.681	0.124	3.569
Natural logarithm of average daily trucks	6.769	0.745	5.226	8.585
Number of lanes	2.761	1.308	2	5
Lane width in ft.	12.859	2.281	10.390	26.440
Median width in ft.	12.603	22.228	0	60
Outside shoulder width in ft.	5.476	3.396	1	14
Inside shoulder width in ft.	1.007	1.895	0	9.240
Average vertical curve grade (%)	1.569	1.706	0	5.948
Road segment with 2 lanes	74.62%			
US road	44.77%			
SR road	50.74%			
Interstate	4.47%			

3.2.2.3 Model Estimation Results and Discussion

Tables 3.2 through 3.6 present the estimation results for the MRPNB models developed separately for pavement segments in excellent, good, good~fair, fair, and poor condition; and Table 3.7 reports the AIC and other goodness-of-fit values for all the models. Based on these values, the estimated models were found to have a superior fit and their parameter estimates closely converged to the true model parameters. Moreover, Table 3.8 presents the results for the likelihood ratio tests for examining parameter transferability. These tests were performed to assess the appropriateness of estimating the models separately for different levels of pavement condition.

The variables found to be statistically significant at a 10% level of significance are listed in Tables 3.2 through 3.6. Likelihood-ratio tests were carried out to ensure that the inclusion of a variable in the model led to an improvement in the overall model fit at a given level of confidence (see Washington et al. (2011) for a detailed description of the likelihood-ratio tests). The signs of the estimated coefficients in the model were found to be intuitive. In addition, Table 3.6 presents the results for the likelihood ratio test to establish the level of confidence that one of the two alternative models (fixed parameters vs random parameters) was statistically superior to the other. Due to the MRPNB model functional form, the model coefficients appear as exponents. For example, holding all other variables in a model at constant values, a U.S. road segment with an excellent pavement condition is expected to have the number of injury crashes to increase by 20.13% ($e^{0.1833} - 1 = 0.2013$) compared to a non-U.S. Road segment in similar condition.

The pavement condition is expressed in terms of its surface roughness (a lower roughness reflects better pavement condition). In the models for each category or range of pavement condition (except the model representing the poor condition), the pavement condition variable was found to have a fixed parameter estimate (i.e., fixed impacts on the expected crash frequency) across the road segments in that condition range. However, in the model representing the segments in poor pavement condition, the pavement condition variable was found to have a random parameter estimate. This same variable also was found to have a normally distributed random parameter in the model for the injury and non-injury crash severity levels.

Regarding the fatal and injury crash models, the interaction term representing the interaction of the segment length and the pavement condition (IRI) was found to significantly affect crash severity on road segments with pavements in excellent condition but also was found to impact only fatal crashes in the case of road segments with pavements in good condition.

Moreover, for pavements in good~fair pavement condition, the interaction term also turned out to be a significant factor for fatal and no-injury crashes. The results in Tables 3.2 through 3.4 show that an increase in either of these two or both in the interaction term had an increasing effect on the expected crash frequency in the models indicated. This interaction term is an interesting and unusual variable but highly intuitive and suggests that an "extended" road, that is, one that has long distance between its terminal points (intersections) with poor pavement condition, is more likely to have an even higher crash frequency compared to a similar short-length road. In other words, the sum of the effects of the pavement condition and road length combined is greater than the effect of the sum of the pavement condition only and road length only.

To explore the effect on safety across the different functional classes of rural roads and the different number of lanes on a road segment, several indicator variables were created and explored in the models. These included the U.S. road indicator, state road indicator, interstate road indicator, two-lane road indicator, three-lane road indicator, four-lane road indicator, five-or-more lane road indicator, and multi-lane (>2) road indicator. Some of these results provided greater insight into the model interpretations. Compared with state roads and interstates, U.S. roads were found generally to have a higher expected injury-crash frequency where the pavement condition is excellent, good, good~fair, or poor; a higher no-injury crash frequency where the pavement condition is excellent; and a higher expected fatal-crash frequency where the pavement condition is good or fair. Moreover, compared to non-state roads, state roads with poor pavement condition tend to have a higher frequency of injury crashes. In addition, rural two-lane road pavements in good condition tend to have less injury crashes compared to rural multi-lane roads in good condition; on the other hand, rural two-lane road pavements in poor condition tend to have a significantly lower expected frequency of no-injury crashes compared to rural multi-lane roads in poor condition. These results seem to be indicative of risk compensation, which is discussed in the next section. Regarding rural road pavements in excellent condition, a higher number of lanes was associated with increased frequency of injury and no-injury crashes compared with those with fewer lanes, which is indicative of driver behavior such as overtaking.

When the pavement condition is poor, an increase in the IRI (further decrease in pavement condition) would generally increase the expected crash frequency, but the coefficient was relatively smaller compared to the effect observed in the models developed for the higher levels of pavement condition. This result was observed in this dissertation for the injury and no-injury models developed for pavements in this condition category.

Regarding the model for injury crashes, the estimated random coefficient of IRI followed a normal distribution with a mean value of 0.00936 and standard deviation of 0.0142. The results indicate that for 25.48% of the observations, the pavement condition parameter had a negative sign, which means that for 25.48% of the observations, an increase in IRI (or decrease in pavement condition) was associated with a decrease in crashes and for 74.52% of the observations, an increase in IRI was associated with an increase in crashes. Regarding the model for no-injury crashes, the corresponding mean value of the normally distributed parameter coefficient for pavement condition was 0.00443 and the standard deviation was 0.0201. In addition, in 41.27% of the observations, the pavement condition parameter had a negative sign, which means that for 41.27% of the observations, an increase in IRI (or decrease in pavement condition) was associated with a decrease in crashes and for 58.73% of the observations, an increase in IRI was associated with an increase in crashes.

The random parameter estimate for poor pavement condition suggests that the impact of the surface condition on crash frequency was indeed bi-directional and the specific direction of the increasing pavement condition depends on the characteristics of the road segment in question. The negative portion of the density function of the condition parameter suggests that for those observations (road segments) that yielded this result, a higher IRI (further lowering of the pavement condition) would decrease the expected crash frequency. This finding can be attributed to the risk compensation effect where drivers generally become more careful where they perceive greater risks in their operating environment. Such heightened caution in driving behavior likely results in lower crash occurrence. The random parameter of IRI in the poor pavement condition model also served as evidence of the heterogeneous nature of driver behavior (i.e., in response to greater risk in the driving environment posed by poor pavement condition, some drivers drive more cautiously. This behavior varied across the different road segments. Generally, this finding is consistent with the concepts known as the Peltzman effect, risk homeostasis, or offset hypothesis, whereby humans tend to engage in less risky behavior when they perceive a riskier environment (Winston et al., 2006; Labi, 2006; Labi, 2016).

Another interesting finding for the road segments with pavements in poor condition is that relatively fewer road geometry factors were significant in the crash frequency model compared to the model for segments with pavements in fair-excellent condition. This may happen because the factors related to road geometry (e.g., lane width and shoulder width) become less influential in affecting the crash frequency when the pavement condition is poor.

Table 3.2 Multivariate negative binomial model for pavements in excellent condition

Variables	**coefficient estimates**	**t-stat**
Constant [FAT]	–4.1887	–3.78
Constant [INJ]	–4.2057	–10.44
Constant [NINJ]	–2.9520	–7.29
Segment length in miles [NINJ]	0.1019	9.18
Natural logarithm of average daily trucks [FAT]	0.3832	2.09
Natural logarithm of average daily trucks [INJ]	0.6436	10.80
Natural logarithm of average daily trucks [NINJ]	0.5816	11.33
Average annual daily traffic of passenger cars (in 10,000) [FAT]	0.1903	2.06
Lane width in ft. [FAT]	–0.0813	–1.95
Outside shoulder width in ft. [FAT]	–0.0514	–2.36
Outside shoulder width in ft. [INJ]	–0.0525	–2.68
Outside shoulder width in ft. [NINJ]	–0.0224	–2.60
IRI (in/mile) [NINJ]	0.0052	2.47
Interaction between segment length and IRI (Length $\times$ IRI/100) [FAT][1]	0.1181	3.49
Interaction between segment length and IRI (Length $\times$ IRI/100) [INJ][1]	0.1264	9.03
Average vertical curve grade [INJ]	0.0220	2.30
Average vertical curve grade [NINJ]	0.0259	2.99
Inside shoulder width in ft. [FAT]	–0.1538	–2.34
Inside shoulder width in ft. [INJ]	–0.1660	–5.27
Inside shoulder width in ft. [NINJ]	–0.1732	–6.15
Median width in ft. [INJ]	–0.0114	–4.11
Median width in ft. [NINJ]	–0.0065	–2.51
Number of lanes on the road segment [INJ]	0.2516	4.12
Number of lanes on the road segment [NINJ]	0.1981	3.37
Functional class US (1 if US road, 0 otherwise) [INJ]	0.1833	2.40
Functional class US (1 if US road, 0 otherwise) [NINJ]	0.1468	2.16
Covariance matrix		
σ^2 [FAT]	5.1666	
σ^2 [INJ]	2.5042	
σ^2 [NINJ]	0.9943	
ρ[FAT/INJ]	0.7063	
ρ[FAT/NINJ]	0.1404	
ρ[INJ/NINJ]	0.1543	

FAT, fatal crashes; INJ, injury crashes; NINJ, no-injury crashes.
Variables are statistically significant at 90% degree of confidence.
[1] (divided by 100 in order to scale the interaction term. The scaling helps to address the convergence issue during the maximum likelihood optimization procedure)

Table 3.3 Multivariate negative binomial model for pavements in good condition

Variables	**coefficient estimates**	**t-stat**
Constant [FAT]	–6.3110	–3.81
Constant [INJ]	–1.2977	–1.48
Constant [NINJ]	–2.1190	–2.63
Segment length in miles [INJ]	0.1875	9.20
Segment length in miles [NINJ]	0.1657	10.01
Natural logarithm of average daily trucks [FAT]	0.5641	2.05
Natural logarithm of average daily trucks [NINJ]	0.3912	6.05
Average annual daily traffic of passenger cars (in 10,000) [INJ]	0.4310	6.85
Average annual daily traffic of passenger cars (in 10,000) [NINJ]	0.0563	1.12
Lane width in ft. [FAT]	–0.2024	–2.23
IRI (in/mile) [INJ]	0.0126	1.62
IRI (in/mile) [NINJ]	0.0066	2.18
Interaction between segment length and IRI (Length $\times$ IRI/100) [FAT][1]	0.4233	6.48
Inside shoulder width in ft. [FAT]	–0.2536	–1.83
Inside shoulder width in ft. [INJ]	–0.1980	–3.33
Median width in ft. [NINJ]	–0.0084	–4.08
Two-lane road segment (1 if two-lane, 0 otherwise) [INJ]	–0.4746	–2.12
Functional class US (1 if US road, 0 otherwise) [FAT]	0.5996	3.08
Functional class US (1 if US road, 0 otherwise) [INJ]	0.2246	2.34
Covariance matrix		
σ^2 [FAT]	0.5658	
σ^2 [INJ]	0.8647	
σ^2 [NINJ]	1.0594	
ρ[FAT/INJ]	0.5683	
ρ[FAT/NINJ]	0.1214	
ρ[INJ/NINJ]	0.1530	

FAT, fatal crashes; INJ, injury crashes; NINJ, no-injury crashes.
Variables are statistically significant at 90% degree of confidence.
[1] (divided by 100 in order to scale the interaction term. The scaling helps to address the convergence issue during the maximum likelihood optimization procedure)

Table 3.4 Multivariate negative binomial model for pavements in good~fair condition

Variables	**coefficient estimates**	**t-stat**
Constant [FAT]	18.6371	3.11
Constant [INJ]	–5.3488	–2.71
Constant [NINJ]	–4.3847	–2.97
Segment length in miles [INJ]	0.1465	3.81
Natural logarithm of average daily trucks [FAT]	1.9246	1.98
Natural logarithm of average daily trucks [INJ]	0.9527	4.28
Natural logarithm of average daily trucks [NINJ]	0.7243	7.01
Average annual daily traffic of passenger cars (in 10,000) [FAT]	0.9085	2.18
Average annual daily traffic of passenger cars (in 10,000) [INJ]	0.1158	2.20
Lane width in ft. [FAT]	–0.2116	–2.96
Lane width in ft. [INJ]	–0.1485	–2.37
Lane width in ft. [NINJ]	–0.1225	–2.32
Outside shoulder width in ft. [INJ]	–0.1264	–3.19
Outside shoulder width in ft. [NINJ]	–0.0456	–1.70
IRI (in/mile) [INJ]	0.0102	2.37
IRI (in/mile) [NINJ]	0.0197	2.02
Interaction between segment length and IRI (Length $\times$ IRI/100) [FAT][1]	0.1543	1.60
Interaction between segment length and IRI (Length $\times$ IRI/100) [NINJ][1]	0.1048	4.97
Average vertical curve grade [INJ]	0.0792	1.62
Average vertical curve grade [NINJ]	0.0591	1.55
Median width in ft. [FAT]	–1.7975	–2.05
Functional class US (1 if US road, 0 otherwise) [INJ]	0.5052	2.66
Covariance matrix		
σ^2 [FAT]	1.0811	
σ^2 [INJ]	0.7644	
σ^2 [NINJ]	0.5638	
ρ[FAT/INJ]	0.7933	
ρ[FAT/NINJ]	0.4389	
ρ[INJ/NINJ]	0.3276	

FAT, fatal crashes; INJ, injury crashes; NINJ, no-injury crashes.

Variables are statistically significant at 90% degree of confidence.

[1] (divided by 100 in order to scale the interaction term. The scaling helps to address the convergence issue during the maximum likelihood optimization procedure)

Table 3.5 Multivariate negative binomial model for pavements in fair condition

Variables	coefficient estimates	t-stat
Constant [FAT]	5.5748	1.56
Constant [INJ]	–9.9081	–4.69
Constant [NINJ]	–7.3271	–5.13
Segment length in miles [INJ]	0.1337	3.33
Segment length in miles [NINJ]	0.1292	4.09
Natural logarithm of average daily trucks [INJ]	1.3755	4.18
Natural logarithm of average daily trucks [NINJ]	1.0171	5.15
Average annual daily traffic of passenger cars (in 10,000) [FAT]	0.4703	1.59
Average annual daily traffic of passenger cars (in 10,000) [INJ]	0.6878	2.05
Average annual daily traffic of passenger cars (in 10,000) [NINJ]	0.4530	2.09
Lane width in ft. [FAT]	–0.7279	–2.14
Lane width in ft. [INJ]	–0.0975	–2.40
Outside shoulder width in ft. [INJ]	–0.0797	–1.58
IRI (in/mile) [INJ]	0.0121	2.08
IRI (in/mile) [NINJ]	0.0140	3.11
Median width in ft. [INJ]	–0.0129	–1.78
Median width in ft. [NINJ]	–0.0202	–5.94
Functional class US (1 if US road, 0 otherwise) [FAT]	1.2756	1.67
Covariance matrix		
σ^2 [FAT]	1.4732	
σ^2 [INJ]	1.2243	
σ^2 [NINJ]	0.8036	
ρ[FAT/INJ]	0.7312	
ρ[FAT/NINJ]	0.4312	
ρ[INJ/NINJ]	0.3016	

FAT, fatal crashes; INJ, injury crashes; NINJ, no-injury crashes.
Variables are statistically significant at 90% degree of confidence.

Table 3.6 Multivariate random parameters (RPs) negative binomial model for pavements in poor condition

Variables	coefficient estimate	t-stat	Std (RPs)	t-stat (RPs)
Constant [FAT]	–6.2923	–2.48		
Constant [INJ]	–6.5937	–4.35		
Constant [NINJ]	0.9012	2.84		
Segment length in miles [NINJ]	0.1718	4.32		
Natural logarithm of average daily trucks [INJ]	0.6727	4.43		
Average annual daily traffic of passenger cars (in 10,000) [NINJ]	0.3389	2.16		
IRI (in/mile) [FAT]	0.0156	1.61		
IRI (in/mile) [INJ] (random parameter with normal distribution)	0.00936 (25.48% negatively signed)	2.42	0.0142	4.75
IRI (in/mile) [NINJ] (random parameter with normal distribution)	0.00443 (41.27% negatively signed)	1.55	0.0201	3.88
Average vertical curve grade [NINJ]	0.0911	2.02		
Two-lane road segment (1 if two-lane, 0 otherwise) [NINJ]	–0.5453	–2.47		
Functional class US (1 if US road, 0 otherwise) [INJ]	1.0636	2.26		
Functional class SR (1 if State road, 0 otherwise) [INJ]	1.2761	2.72		
Covariance matrix				
σ^2 [FAT]	0.5949			
σ^2 [INJ]	0.6363			
σ^2 [NINJ]	1.0035			
ρ[FAT/INJ]	0.6471			
ρ[FAT/NINJ]	-0.0562			
ρ[INJ/NINJ]	0.0271			
$LL(\beta_f)$	-358.64			
$LL(\beta_r)$	-355.40			
$\chi^2 = -2[LL(\beta_f) - LL(\beta_r)]$	6.48			
Degree of freedom	2			
Critical value for 95% level of confidence: $\chi^2_{0.05,2}$	5.99			

FAT, fatal crashes; INJ, injury crashes; NINJ, no-injury crashes.
Variables are statistically significant at 90% degree of confidence.

The effects of specific variables were also investigated, and Table 3.9 presents the marginal effects of each estimated parameter. A marginal effect is defined as the change in a dependent variable due to a one-unit change in an independent variable. In this dissertation, the marginal effect reflects the impact of a unit change in each variable (Tables 3.2 through 3.6) on the expected frequencies of fatal, injury, and no-injury crashes on a road segment. For example, the results of the estimated models indicate that if the annual average daily traffic of passenger cars increases by 10,000, the following could be expected: an average increase of 0.0199 fatal crash frequency on road segments with pavements in excellent condition; increased injury and no-injury crashes by 1.3197 and 0.4805, respectively, on road segments with pavements in good condition; and increased fatal and injury crashes by 0.0125 and 0.4188, respectively, on road segments with pavements in good~fair condition. In addition, a similar increase in passenger car volume can be expected to cause increases of 0.0641, 2.034, and 7.5445 in fatal, injury, and no-injury crashes, respectively, on road segments with pavements in fair condition; and 5.7135 in no-injury crashes on road segments with pavements in poor condition.

This case study also investigated the marginal effects of truck traffic. If the average daily truck traffic (in the form of its natural logarithm) increased by one unit, the resulting impacts on the crash frequency were estimated from the model results. These marginal effects of truck traffic were examined across the different pavement conditions and across the different crash severity levels. The marginal impacts of truck traffic on crash frequency generally ranged from 1.7165 to 4.0677 for injury crashes, 3.3386 to 16.9394 for no-injury crashes, and 0.0265 to 0.0551 for fatal crashes, depending on the pavement condition. For road segments with pavements in excellent condition, the marginal impacts of truck traffic were 0.04, 1.7165, and 5.9864 for fatal, injury, and no-injury crashes, respectively. For segments with pavements in good condition, the effects were 0.0551 and 3.3386 for fatal and no-injury crashes, respectively; and for segments with pavements in good~fair condition, the effects were 0.0265, 3.4454, and 7.9374 for fatal, injury, and no-injury crashes, respectively. In addition, for segments with pavements in fair condition, the marginal effects were 4.0677 and 16.9394 for injury and no-injury crashes, respectively; and for segments with pavements in poor condition, the marginal effect of truck traffic was 2.4289 for injury crashes. These results are consistent with engineering expectation, and also past research determined that truck traffic, compared with passenger car traffic, has a greater impact on the expected crash frequency (Konduri et al., 2003).

This case study confirmed the findings of past research regarding the impact of highway geometry on rural highway safety and the variation of these impacts across the different levels of crash severity (Zegeer et al., 1994; Council and Stewart, 1999; Fitzpatrick, 2000; Lamptey, 2004; Prato et al., 2014; Labi et al., 2017). In addition, it was determined that segments with wide lanes generally have fewer crashes compared to those with narrower lanes. In terms of rural highway safety, a 1 ft. increase in lane width was found to reduce the expected fatal crashes by 0.0085, 0.0198, 0.0029, and 0.0991 on road segments with pavements in excellent, good, good~fair and fair condition, respectively. In addition, for segments with pavements in good~fair condition, the expected injury and no-injury crashes decreased by 0.5370 and 1.3424, respectively; and for segments with pavements in fair condition, the expected injury crash frequency tended to decrease by 0.2883 with a 1 ft. increase in lane width. Regarding the outside shoulder, the results suggest that a 1 ft. increase in width would reduce fatal crashes by 0.0054, injury crashes by 0.14, and no-injury crashes by 0.2306 on road segments with pavements in excellent condition; reduce injury crashes by 0.4571 and no-injury crashes by 0.4997 on road segments with pavements in good~fair condition; reduce injury crashes by 0.2357 on road segments with pavements in fair condition; and would have no significant safety impact on road segments with pavements in good and poor conditions. The results also suggest that a wider inside shoulder had little impact on road segments with pavements in conditions below the "good" level. However, this feature had significant impacts on segments with pavements in excellent or good condition: for roads with pavements in excellent condition, it was found that a 1 ft. increase in the inside shoulder width resulted in a 0.0161, 0.4427, and 1.7827 reduction in expected fatal, injury, and no-injury crashes, respectively. For roads with pavements in good condition, this increment resulted in a 0.0248 and 0.6062 reduction in expected fatal and injury crashes, respectively. This case study also determined that regarding pavements in poor condition, the median width was not a significant variable in the crash models. However, a 1 ft. increase in median width was associated with the following: reduction of fatal crashes by 0.0247 at segments with pavements in good~fair condition, reduction of injury crashes by 0.0304 and 0.0381 at segments with pavements in excellent or fair condition; and reduction of no-injury crashes by 0.0669 and 0.0717 at segments with pavements in excellent or good condition. In addition, a lower average vertical curve grade was found to be associated with reduction of the expected crash frequency for injury and no-injury crashes but not for fatal crashes at segments with pavements in excellent and good~fair condition; and lower expected no-injury crashes at road segments with pavements in poor condition.

These results are consistent with past research, which concluded that median and shoulder widths have variable impacts on vehicle crashes with regard to the different levels of crash severity (Knuiman et al., 1993; Shankar et al., 1997; Labi, 2011; Anarkooli et al., 2017).

Table 3.7 Goodness-of-fit of the multivariate random-parameter negative binomial models

	Excellent Pavement Condition (60<IRI≤100)	**Good Pavement Condition** (100<IRI≤125)	**Good ~ Fair Pavement Condition** (125<IRI≤150)	**Fair Pavement Condition** (150<IRI≤200)	**Poor Pavement Condition** (IRI>200)
$LL(C)$	–4545.65	–2019.47	–740.21	–688.49	–411.34
$LL(\beta)$	–3113.70	–1616.23	–627.69	–576.56	–355.40
$McFadden\ pseudo\ \rho^2$	0.315	0.201	0.152	0.164	0.136
Akaike information criterion (AIC)	6277.40	3270.46	1299.38	1189.12	736.81
N	681	324	111	99	67

$LL(C)$ = Restricted Log-likelihood; and $LL(\beta)$ = Log-likelihood at convergence

Table 3.8 Likelihood ratio tests for parameter transferability

	Excellent Pavement Condition	**Good Pavement Condition**	**Good ~ Fair Pavement Condition**	**Fair Pavement Condition**	**Poor Pavement Condition**
Excellent Pavement Condition	---	$\chi^2 = 463.27$ DOF = 25	$\chi^2 = 387.69$ DOF = 23	$\chi^2 = 775.73$ DOF = 23	$\chi^2 = 833.16$ DOF = 16
Good Pavement Condition	---	---	$\chi^2 = 274.44$ DOF = 20	$\chi^2 = 413.21$ DOF = 18	$\chi^2 = 442.97$ DOF = 15
Good ~ Fair Pavement Condition	---	---	---	$\chi^2 = 141.54$ DOF = 21	$\chi^2 = 243.15$ DOF = 18
Fair Pavement Condition	---	---	---	---	$\chi^2 = 178.59$ DOF = 15
Poor Pavement Condition	---	---	---	---	---

$\chi^2_{c,0.05,25}$ =37.65248413, $\chi^2_{c,0.05,23}$ = 35.17246163, $\chi^2_{c,0.05,21}$ = 32.67057334, $\chi^2_{c,0.05,20}$ = 31.41043284, $\chi^2_{c,0.05,18}$ = 28.86929943, $\chi^2_{c,0.05,16}$ = 26.2962276, $\chi^2_{c,0.05,15}$= 24.99579014. For all the cases (between different pairs of pavement condition), $\chi^2 > \chi^2_{c,0.05,DOF}$, the null hypothesis that parameters are the same for various pavement condition is rejected. The transferability tests indicate that the models should be estimated separately on the basis of various levels of pavement condition.

Table 3.9 Marginal effects of the estimated MRPNB regression parameters

Variable Description	**Excellent Pavement Condition (60<IRI≤100)**	**Good Pavement Condition (100≤IRI<125)**	**Good ~ Fair Pavement Condition (125<IRI≤150)**	**Fair Pavement Condition (150<IRI≤200)**	**Poor Pavement Condition (IRI>200)**
Segment length in miles [FAT]	0.0099	0.0451	0.0029	–	–
Segment length in miles [INJ]	0.2697	0.5741	0.5298	0.3954	–
Segment length in miles [NINJ]	1.0489	1.4141	1.5734	2.1518	2.8964
Natural logarithm of average daily trucks [FAT]	0.04	0.0551	0.0265	–	–
Natural logarithm of average daily trucks [INJ]	1.7165	–	3.4454	4.0677	2.4289
Natural logarithm of average daily trucks [NINJ]	5.9864	3.3386	7.9374	16.9394	–
Average annual daily traffic of passenger cars (in 10,000) [FAT]	0.0199	–	0.0125	0.0641	–
Average annual daily traffic of passenger cars (in 10,000) [INJ]	–	1.3197	0.4188	2.034	–
Average annual daily traffic of passenger cars (in 10,000) [NINJ]	–	0.4805	–	7.5445	5.7135
Lane width in ft. [FAT]	-0.0085	-0.0198	-0.0029	-0.0991	–
Lane width in ft. [INJ]	–	–	-0.537	-0.2883	–
Lane width in ft. [NINJ]	–	–	-1.3424	–	–
Outside shoulder width in ft. [FAT]	-0.0054	–	–	–	–
Outside shoulder width in ft. [INJ]	-0.14	–	-0.4571	-0.2357	–
Outside shoulder width in ft. [NINJ]	-0.2306	–	-0.4997	–	–
IRI (in/mile) [FAT]	0.0007	0.0024	0.0001	–	0.0011
IRI (in/mile) [INJ]	0.0199	0.0386	0.0369	0.0358	0.0338
IRI (in/mile) [NINJ]	0.0535	0.0563	0.2817	0.2332	0.0747
Average vertical curve grade [INJ]	0.0587	–	0.2864	–	–
Average vertical curve grade [NINJ]	0.2666	–	0.6477	–	1.5359
Inside shoulder width in ft. [FAT]	-0.0161	-0.0248	–	–	–
Inside shoulder width in ft. [INJ]	-0.4427	-0.6062	–	–	–
Inside shoulder width in ft. [NINJ]	-1.7827	–	–	–	–
Median width in ft. [FAT]	–	–	-0.0247	–	–
Median width in ft. [INJ]	-0.0304	–	–	-0.0381	–
Median width in ft. [NINJ]	-0.0669	-0.0717	–	-0.3364	–
Number of lanes on the road segment [INJ]	0.671	–	–	–	–
Number of lanes on the road segment [NINJ]	2.039	–	–	–	–
Two-lane road segment (1 if two-lane, 0 otherwise) [INJ]	–	-1.4531	–	–	–
Two-lane road segment (1 if two-lane, 0 otherwise) [NINJ]	–	–	–	–	-9.1932
Functional class US (1 if US road, 0 otherwise) [FAT]	–	0.0586	–	0.1737	–
Functional class US (1 if US road, 0 otherwise) [INJ]	0.4889	0.6877	1.827	–	3.8403
Functional class US (1 if US road, 0 otherwise) [NINJ]	1.511	–	–	–	–
Functional class SR (1 if State road, 0 otherwise) [INJ]	–	–	–	–	4.6075

3.2.2.4 Summary of Results of the MRPNB Model

The model results suggest that for pavements in fair or good condition, the surface condition parameter had fixed effects on the crash frequency, irrespective of the crash severity level. However, for pavements in poor condition, the surface condition variable in the crash model had a significant random parameter that was normally distributed. The positive portions of the parameter density function suggest that higher roughness (poorer condition) generally increased the expected crash frequency, likely because drivers may lose control of their vehicles. The negative portions suggest that within that condition range, higher surface roughness is generally associated with a lower expected crash frequency because drivers are generally likely to drive more carefully on pavements in very poor condition (a manifestation of risk-compensation behavior). These developed models can help highway engineers quantify not only the safety benefits of road resurfacing projects but also the safety consequences of worsening road surface conditions arising from deferring the pavement maintenance.

The phenomenon of risk-compensation behavior found in this dissertation may vary in the future, particularly in the emerging operating environment in the era of AV because risk-compensation behavior is based in human psychology. Therefore, it may not exist in the fully AV era when the vehicle is completely controlled by AI and algorithms.

3.3 Optimizing the Space Resource Allocation across Highway Cross Sectional Elements

3.3.1 Introduction

Lane and shoulder widths are important highway safety factors because wider lanes and shoulders generally are known to help reduce crashes. In practice, however, due to the physical limitations of the right-of-way or statutory restrictions, it may be the case that the existing overall width of the roadway (lanes plus shoulders) cannot exceed a certain value. For a given overall roadway width constraint, wider lanes mean there will be smaller space for shoulders, and vice versa. In such situations, it is not certain if a cross-sectional configuration with wide lanes and narrow shoulders is safer compared to one with narrow lanes and wide shoulders. Considering the dichotomous effect of shoulder or lane widening as established in the safety literature, one of the greatest identified challenges in optimizing lane and shoulder width for two-lane rural roadways is the quest to balance the area for recovery as well as for accommodating driver recklessness. Addressing this issue is important for agencies that seek to identify optimal widths to evaluate their existing lane and shoulder design practices. Clearly, for a given total roadway width (TRW), the tradeoff between the shoulder and lane widths should be evaluated to maximize the user benefits (safety) without unduly incurring excessive agency life-cycle costs of construction and preservation. The TRW allocations to the lane and shoulder, expressed as the ratio of the lane width to the shoulder width, generally depend on the TRW, lane and shoulder pavement material types and costs, and traffic volume, among other factors. To optimize the lane and shoulder width allocations for a given TRW, this dissertation proposes a framework that formulates the problem as an optimization problem with the objective of minimizing the total (agency and user) life-cycle cost of the roadway. First, the relationships are established between the lane/shoulder widths and the resulting user costs (crashes) and between the lane/shoulder widths and the associated agency costs of lane/shoulder construction and preservation. The sensitivity of the optimal solution to different evaluation inputs (e.g., the relative weights between the agency and user costs) is also analyzed. Using the developed framework, the dissertation presents several decision support charts that can be used by highway agencies to determine the optimal lane and shoulder widths under a given set of conditions including the highway functional class and the total available roadway width. The flexibility of the proposed optimization framework to accommodate the risk compensation effects also is demonstrated.

3.3.2 Literature Review (General)

Users of transport facilities and other physical infrastructure are invariably subjected the risks of personal injury, fatality, and property damage arising from their use of such facilities. While some risks are generally unavoidable, effective safety management can lead to their reduction with subsequent benefits to all parties directly or indirectly involved with the usage or administration of these infrastructures. The key elements of safety management include (Wegman and Hagenzieker, 2010; Lamptey et al., 2011): the establishment of safety performance indicators (Hale, 2009; Gitelman 2010; Øien et al., 2011; Reiman and Pietikäinen, 2012; Hinze et al., 2013; Podgorski, 2015), measurement or prediction of safety performance under a given set of physical conditions and operational characteristics (Cozzani and Salzano, 2004; Kjellen, 2009; Hauer, 2010; Sgouru et al., 2010), identification and effectiveness assessment of safety prescriptions for a given safety hazard (Robson et al., 2007), prioritization of safety investments (Law et al., 2006), and the measurement and consideration of the costs borne by each infrastructure stakeholder in the provision or use of the safety prescriptions (Cigna, 2008; Pellicer et al., 2014). Another key element is the development of safety policies regarding the infrastructure design (dimensions) in a manner that considers all the above issues and thus enhances infrastructure operations from a safety standpoint (Lamptey et al., 2010).

Determination of the optimal width of lanes and shoulders is a key issue for highway agencies. The term "optimal width" suggests that the cases of both narrow lanes and shoulders and very wide lanes and shoulders are undesirable; and therefore somewhere between these extremes, an optimal width exists. At one extreme, it has long been recognized that small lanes and widths are undesirable: widening the travel lanes or/and their shoulders has a positive impact in terms of reductions in crashes of various types, particularly roadway departure crashes and horizontal curve crashes.

Given the apparent challenge but potential benefits associated with the optimal allocation of lane and shoulder widths within a given total roadway width (TRW), it is crucial that highway agencies apply consistent and objective methodologies to a variety of allocation scenarios to ascertain the optimal allocation of resources while providing the greatest safety benefit to the road users. This is the overarching goal of this dissertation and is further discussed in later sections.

In presenting and demonstrating a methodology to address the allocation of shoulder and lane width (or evaluation of existing distributions) for rural two-lane roads, this section first reviews the literature on the relationships between crash frequencies and lane or shoulder width, crash costing, and lane and shoulder width allocation policy. This section then describes the data and the study methodology, presents its results in the context of existing design policy, and ends with a discussion of the implications for practice and possible future research in this area.

3.3.3 Literature Review (Crash Factors)

On both extremes of the relationship between safety feature intensity (such as lane width) and crash experience, it has long been recognized that it is generally undesirable to have lanes and shoulders that are very narrow or very wide. The extant literature review below shows that past research established that very narrow lanes and shoulders deny drivers the needed opportunity for maneuverability and provide smaller distances between oncoming vehicles. On the other hand, relatively recent research contended that very wide lanes or shoulders not only imbue drivers with a false sense of security such that they may tend to operate their vehicles more recklessly but also are costly to build and maintain. The literature review also examines past research that investigated the appropriate widths of highway lanes and shoulders.

Increased Safety resulting from Lane or Shoulder Widening

Zegeer (1980) analyzed two-lane roads concerning geometries, accidents, and traffic volumes. The analysis was based on the accident rates for various combinations of lane and shoulder widths and established that for roads without shoulders, the crash rate is over 80% lower than when the lane widths increase from 2.1 m to 3.7 m. For roads with shoulders, accident rates were found to be lower for wider lanes. Also, Zegeer et al. (1994) established that wide lanes and shoulders help reduce roadway crashes (Figure 3.1).

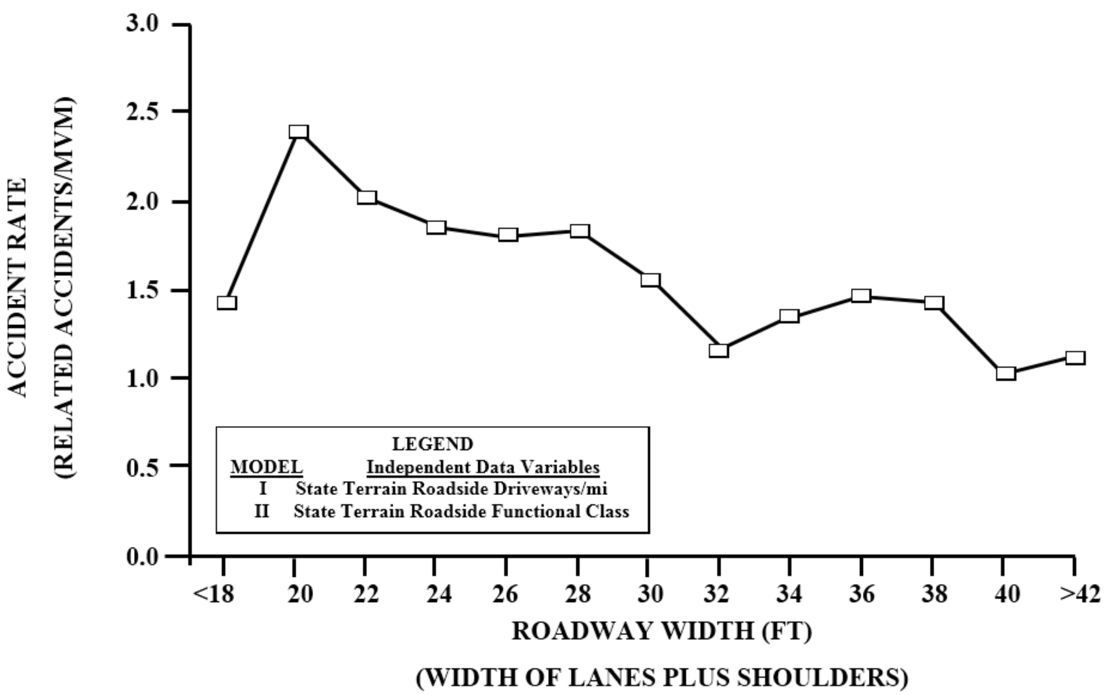


Figure 3.1 Crash rates by total roadway width (Zegeer, 1980)

Hadi et al. (1995) found that greater shoulder width was associated with lower crash rates on rural two-lane highways on the Florida state highway system. Abboud and Bowman (2001) argued that the additional construction cost of 4-ft. shoulders on two-lane rural highways (compared with the existing 2-ft. shoulders) was not justified by any statistically significant reduction in crashes. Fitzpatrick et al. (2005) and Geni and Peng (2008) documented the benefits of widening the travel lanes or/and their shoulders in terms of reductions in crashes of various patterns, particularly roadway departure crashes and horizontal curve crashes. In Pennsylvania, Gross and Jovanis (2007) investigated the safety impacts of shoulder-widening projects on rural two-lane highways and found that for shoulder widening of 2 ft. and 8 ft., the crash reduction ranged from 13–43% for unpaved shoulders and 16–59% for paved shoulders. In Wisconsin, Örnek and Drakopoulos (2007) examined state highway sections with shoulder widths ranging from 1 ft. to 10 ft. and found that the crash rates were lower for sections with wider shoulders; however, shoulder widths beyond 10 ft. were found to yield no further reductions in the crash rates. Using the empirical Bayes method for a before–after comparison, Wu et al. (2015) studied the safety impacts of narrow pavement widening projects on rural two-lane roads in Texas. They investigated how wider lanes and shoulders affected collisions by type and severity and showed that widening resulted in a 31.5% reduction in total crashes, a 35.7% reduction in run-off-the-road crashes, a 55.4% reduction in head-on crashes, and a 29.5% reduction in fatal and injury crashes. Park (2016) investigated the safety effect of shoulder widening on 17 rural multilane roadway sections and determined that this treatment resulted in reductions of 12% of crashes of all severity levels, 18% of injury crashes, and 21% of severe crashes. Regarding run-off-the-road crashes, they found that the treatment yielded reductions of 25% for all crash severities, 28% for injury crashes, and 31% for severe crashes.

Many studies in the literature on the tradeoffs between lane width and shoulder width adopted case-control methods that involved the use of "odds ratios" (i.e., the probability of a crash on a road segment with nonstandard lane and shoulder widths divided by the probability of a crash on a similar road segment with standard widths). Using Pennsylvania data spanning 1997 to 2001 and a matched case-control design, Gross and Jovanis (2007) estimated the safety effectiveness of lane and shoulder widening on rural two-lane roads and established odds ratio estimates using a base case of a 12-ft. lane and a 6-ft. shoulder. They found that the crash likelihood generally decreased with increasing lane and shoulder widths, except for width values on the extreme upper end of the scale where crash likelihood increased slightly. A similar study using data from Pennsylvania and Washington showed that increases in both lane and shoulder width generally had positive impacts (reduced crashes), but the tradeoffs between the two elements for a fixed pavement width were unclear (Gross et al., 2009). Gross et al. (2009) further evaluated the safety effectiveness of various lane-shoulder width configurations for fixed total paved widths as a countermeasure for roadway departure crashes and found that wide lanes (9.75–10.97 m or 32–36 ft) were generally associated with fewer crashes compared to narrow lanes (7.92–9.14 m (26–30 ft). Gross and Donnell (2011) undertook a similar approach, comparing crash modification factors (CMFs) developed using case-control studies to CMFs developed using more traditional cross-sectional models. They used data from Pennsylvania to estimate the CMFs associated with improvements in shoulder width, by comparing the results of the case-control studies with those from negative binomial econometric models developed for the cross-sectional data and found that varying the lane width had inconsistent impacts on highway crash reduction; nevertheless, increased shoulder width (up to 10 ft. of additional width) consistently resulted in fewer crashes and, consequently, smaller CMFs. In a similar study sponsored by Kansas DOT, Schrock et al. (2011) used the CMF concepts to compare the safety effectiveness of widening shoulders and adding passing lanes on rural two-lane roads and found that adding passing lanes was more economically efficient compared to shoulder widening.

Dong et al. (2015) examined the issue of highway geometrics and safety from the specific perspective of trucks. The authors categorized truck-involved crashes by collision vehicle types since passenger cars and large trucks differ in dimensions, size, weight, and operating characteristics and determined that the significant factors of such crashes included the lane width and the right-side shoulder width, among other factors.

No Change or Decreased Safety Resulting from Lane or Shoulder Widening

Several studies (Milton and Mannering, 1998; Sawalha and Sayed 2001) found that reducing the lane width led to lower speeds and accident frequency, but the ultimate effect depended on the lane width and road class (Fildes and Lee, 1993). In Alberta, Canada. El-Assaly and Hempsey (2000) concluded that no correlation or dependency existed between the width of a paved two-lane highway and the casualty collision rates. Garder (2006) used crash data on rural two-lane highways in Maine to study the safety impacts of shoulder attributes (width and surface material type). Their results indicated that a wider shoulder led to a higher percentage of crashes producing serious injuries (23% for highways with shoulders at least 7 ft. wide compared to 18% for highways with narrower shoulders). Rosey et al. (2009) used data from field tests and laboratory simulations to determine that reducing the lane width had no impact on speeds but did induce the participants to drive closer to the center of the road.

Evaluation of Existing Widths or Identification of Optimal Widths

Researchers have long sought to optimize aspects of highway geometrics besides lane and shoulder width as well as using other performance metrics besides safety. Jha and Schonfeld (2004) developed a highway alignment optimization model using a geographic information system (GIS), based on criteria including agency cost, community (environmental cost), and the user costs of travel time, vehicle operations, and accidents. Also, the use of optimization as a tool for decision support and evaluation of existing design or operational policy is widespread in transportation safety policy and practice, which is demonstrated by the numerous contexts, including geometric design (Cedar and Eldar, 2002), project-level selection of safety countermeasure (Mishra et al., 2015), network-level resource allocation (Ahmed, 1983; Melachrinoudis and Kozanidis, 2002; Sinha et al., 1981; Brown, 1980, Lamptey et al., 2010). Other researchers such as Farah et al. (2009) applied optimization to address other aspects of rural two-lane road safety, such as long-term work programming and scheduling.

Political and budgetary priorities also have been a significant driving force behind research into the effects of lane and shoulder widths on crashes in recent years. Minnesota DOT (MNDOT) completed a study a few years ago to assess the benefits and costs if the lanes of low-volume rural roads were 11 ft. wide instead of the conventional 12 ft. They found that for roads with daily traffic volumes less than 1,200, the reduced construction costs of narrower lanes would greatly offset the additional user cost (higher number of crashes) due to the narrow lanes. Their finding that there were fewer crashes on low-volume rural two-lane roads with 11 ft. wide lanes compared to 12 ft. wide lanes, was subsequently corroborated by van Schalkwyk and Washington (2008). The MNDOT study established the potential cost savings in rehabilitation and design standards and also highlighted the underlying intriguing question regarding the relative importance of the competing interests associated with lane and shoulder widening, namely, should each dollar of an agency's cost of construction and maintenance of lanes and shoulders be valued the same as each dollar of the safety cost incurred by the road users? Gross et al. (2009) argued that for a fixed pavement width, there was a general safety benefit associated with wider lanes and narrower shoulders. For example, for total paved widths of 7.92 to 9.75 m (26 to 32 ft.), a 3.66-m (12 ft.) lane provides the optimal safety benefit; and for a 10.36-m (34 ft.) total paved width, 3.35-m (11 ft.) lanes provide the optimal safety benefit.

In the emerging AV operating environment, there could be an opportunity for more narrow lanes and shoulders. Currently, there are lanes (and sometimes even shoulders) that are much wider than vehicles to provide space to minimize the possible interactions with vehicles in adjacent lanes. These interactions are typically caused by human errors in driving. Having narrower lanes and shoulders is expected to make it possible to either have more lanes using the same amount of space or require less land for right-of-way, which consequently could be expected to increase mobility on the highways or reduce agency costs. It may be noted that it is important to evaluate the safety impacts of having narrow lanes and shoulders in the context of AV operations.

3.3.4 Case Study

This case study identified the optimal distribution of a fixed width of the total (lane + shoulder) using nonlinear programming techniques. The development of an optimization framework based on the life-cycle agency cost and life-cycle user (safety) cost for the roadway is described here. The objective is to minimize the total life-cycle cost. The constraints are that the sum of the lane and shoulder widths equals the total available width, and the lane width cannot be smaller than 10 ft. The decision variables are the lane and shoulder widths in feet. Also, the sensitivity of the optimal solutions with respect to the relative weight between the agency and user cost dollars was studied for the different levels of the fixed width of a roadway. The case study concludes with a discussion of the use of optimization in evaluating the relationship between lane and shoulder width and suggests several promising areas of future research.

3.3.4.1 Methodology

To optimize the allocation between the lane width and the shoulder width for a fixed total width of roadway, there are a number of factors that must be considered. The life-cycle costs refer to any costs incurred by the agency or the user throughout the service life of an asset (in this case, the rural two-lane highway). It may be appropriate to have a framework where the analyst has the flexibility to assign relative weights, as the analyst deems appropriate, to the agency cost and user cost. The objective is to minimize the weighted sum of these two cost classifications. The two decision (input) variables are the lane and shoulder widths. The following sections describe in detail the methods used to arrive at the total agency and the safety life-cycle cost estimates and how these estimates are combined to yield an overall life-cycle cost for the highway.

Crash prediction modeling

To compute the life-cycle safety cost for different lane and shoulder width combinations, it was first necessary to derive a set of models which can be used to estimate the expected number and type of crashes. This was completed for each of the two-lane rural road classes. For this dissertation, the following broad grouping of the crash severity levels was used: property damage only (PDO) crashes, and fatal + injury crashes (which includes all crashes with injuries, fatalities, or some combination of the two). Fatal and injury crashes were combined as one category due to the very limited number of fatal crashes (as evidenced by Figure 3.2), which precluded the development of a separate statistically significant functional relationship for the fatal level of crash severity. If enough data were available to allow estimation of a statistically significant crash prediction model for fatal crashes, this model could easily accommodate such a situation.

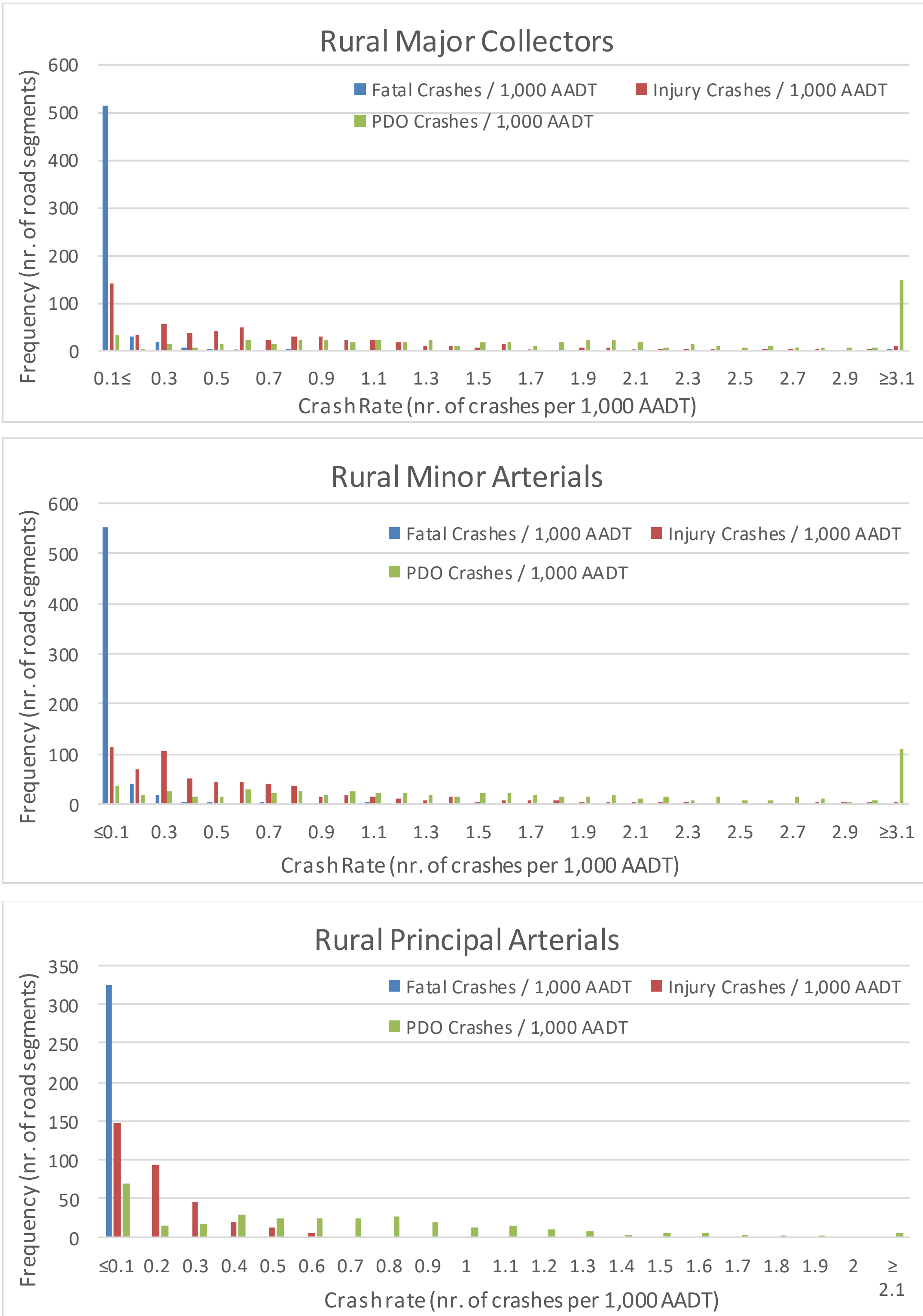


Figure 3.2 Distributions of crash rates for different crash severity levels across the three functional classes

Due to the differences in the geometric and operational characteristics across the different functional classes that constitute rural two-lane highways, this case study used separate crash models for the three different classifications of rural two-lane highways: rural major collectors, rural minor arterials, and rural principal arterials. Principal arterials primarily service interstate trips and routes with heavier volumes so while it is rare to see principal arterial routes with only two lanes, they were still considered in this dissertation. Rural minor arterials service long-distance intrastate trips, connect urbanized areas, and are more likely to be two-lane highways compared with principal arterials. Rural major collectors often service shorter intrastate and inter-county trips, and combined with lower volumes, are more likely to be designated along two-lane rural highway routes (FHWA, 2013). Because this dissertation is concerned with assessing the tradeoff between the cost effectiveness of lane and shoulder widths on Indiana highways in particular, the equations previously-developed for rural two-lane highways by Labi (2011), presented below, were used for the crash estimation component of the final objective function. It should be noted that these equations are shown and used here for demonstration only. In future replications of this analysis, these equations may be replaced by any safety performance equation relevant to the region in question or the site under consideration.

<u>Rural major collectors:</u>

PDO Crash Model

$$\mu_{\mathrm{PDO}} = \exp[-4.06689 + 0.8706 * LN(LENGTH) + 0.6259 * LN(AADT) - 0.0617 * LW - 0.0119 * SW - 0.0190 * FR + 0.0163 * ARAD + 0.1100 * AGRAD]$$

Fatal+Injury Crash Model

$$\mu_{\mathrm{fatal+injury}} = \exp[-6.6231 + 0.9237 * LN(LENGTH) + 0.8526 * LN(AADT) - 0.0928 * LW - 0.0321 * SW - 0.0156 * FR + 0.0262 * ARAD + 0.0541 * AGRAD]$$

<u>Rural minor arterials:</u>

PDO Crash Model

$$\mu_{\mathrm{PDO}} = \exp[-3.81056 + 0.9981 * LN(LENGTH) + 0.6363 * LN(AADT) - 0.0960 * LW - 0.0317 * SW - 0.0127 * FR + 0.0722 * ARAD + 0.0944 * AGRAD]$$

Fatal+Injury Crash Model

$$\mu_{fatal+injury} = \exp[-6.4612 + 0.9964 * LN(LENGTH) + 0.8255 * LN(AADT) - 0.0896 * LW - 0.0511 * SW - 0.0330 * FR + 0.0580 * ARAD + 0.0866 * AGRAD]$$

Rural principal arterials:

PDO Crash Model

$$\mu_{PDO} = \exp[0.8921 + 0.7097 * LN(LENGTH) + 0.2409 * LN(AADT) - 0.1128 * LW - 0.0676 * SW - 0.0624 * SI + 0.0553 * ARAD + 0.0646 * AGRAD]$$

Fatal+Injury Crash Model

$$\mu_{fatal+injury} = \exp[-2.2394 + 0.9231 * LN(LENGTH) + 0.4499 * LN(AADT) - 0.0837 * LW - 0.0943 * SW - 0.1962 * SI + 0.0364 * ARAD + 0.0489 * AGRAD]$$

For these crash models, the detailed descriptions of the individual variables, along with their corresponding units of measure, are as follows:

μ_k = number of crashes of particular type k (fatal + injury, or PDO); Length = Segment length in mile; AADT = Annual average daily traffic; LW = Lane width in ft.; SW = Shoulder width in ft. ARAD = Average horizontal curve radius in miles; AGRAD = Average vertical curve grade (%) FR = Pavement friction number; SI = Present serviceability index of the pavement.

The length of the sample road segment is 10 miles, the AADT is 8,000, and the pavement friction number is 3.5. Additionally, the pavement serviceability index is 3, the average horizontal curve radius is 2.6 miles, and the average vertical grade is 3%. The example values to be used in this analysis are average values from the rural two-lane road database and were provided for demonstrating the proposed methodology. For any given in-service road segment of interest, these values could be smaller or larger than the values shown. Based on these values, the proposed framework was demonstrated to yield the optimal lane and shoulder width combination for different total road widths as well as for different road classes. This was done by minimizing the total agency and safety life cycle costs.

Agency life cycle cost

The agency life cycle cost is comprised of several different components: first, the shoulder and lane costs, and for each of these two road features, the construction costs (incurred in the initial year) and the preservation costs (incurred over the rest of life). Regarding the surface material types, this dissertation considered the case where the same material type was used for the lane and shoulders (hot-mix asphalt (HMA)). The preservation cost component of agency life cycle cost is that which is required annually or periodically to preserve the lanes and the shoulders, which can range from low-level routine work, such as crack sealing on shoulders and lanes, to high-level treatments that provide a new wearing surfaces, such as microsurfacing or slurry seal. Using data from Iowa DOT, the average contract bid prices for calendar year 2000, and converted to year 2012 with the U.S. Army Corps of Engineers Civil Works Construction Cost Index (USACE, 2016), the following lane and shoulder construction and maintenance costs were determined:

Table 3.10 Agency construction and maintenance costs for HMA paved lanes and shoulders ($ per ft-width per mile)

	Construction Cost	Maintenance Cost
Lane	$197,101	$493
Shoulder	$27,710	$39

** Lane reference cost values given for 12-ft lanes; this was divided by 12 to obtain the unit cost.*
** Shoulder reference cost values given for 3-ft shoulders; this was divided by 3 to obtain the unit cost.*

User cost of safety over life cycle

The user cost of safety over the road segment life cycle can be represented by the annual economic cost due to different types of crashes on the road segment. As described in the crash modeling section above, these crashes are often grouped by crash severity, and for the purposes of this dissertation, two broad groups were considered: fatal + injury, which comprises all crashes in which the worst injury is a fatality or incapacitating injury; and property damage only (PDO), where there is no fatality or injury.

Pursuant to these crash severity type groupings, the crash prediction models were used to estimate the annual number of crashes on a 10-mile road segment for the different highway functional classes. After the number of crashes was estimated, a dollar value representing the economic cost of these crashes was applied to determine the total annual safety cost. These dollar values were derived from the Indiana Five Percent Report (Chen et al., 2012), converted to year 2012 using the Consumer Price Index (CPI), and are shown in Table 3.11

Table 3.11 Estimated unit crash cost values ($ per crash)

Crash Severity	**Crash Cost**
Fatal + Injury	$509,086
PDO	$5,244

(Chen et al., 2012)

Optimization framework

The objective of this part of the dissertation was to minimize the agency and safety life cycle costs by changing the lane and shoulder width within given constraints (e.g., total road width and minimum lane width). In mathematical notation, this is written as follows:

Decision variables:

LW, SW

Objective function:

Minimize $T(LW, SW) =$

$$2\{w_{agency}[(CC_l * LW + CC_s * SW)L + \frac{(1+i)^N - 1}{i(1+i)^N}(MR_l * LW + MR_s * SW)L]$$
$$+ w_{user}\frac{(1+i)^N - 1}{i(1+i)^N}\sum_1^k c_k \mu_k(LW, SW)\}$$

Subject to:

$2LW + 2SW = TRW$

$LW \geq MLW$

where:

T = total life cycle cost in $

CC_l = unit initial construction cost for lane in $ per ft-width per mile

CC_s = unit initial construction cost for shoulder in $ per ft-width per mile

MR_l = unit annual maintenance and rehabilitation cost for lane in $ per ft-width per mile

MR_s = unit annual maintenance and rehabilitation cost for shoulder in $ per ft-width per mile

c_k = crash cost for crash severity k (fatal + injury, or PDO)

μ_k = number of crashes of crash severity k (fatal + injury, or PDO)

w_{agency}, w_{user} = weight of agency cost dollar and user cost dollar, respectively.

LW, *SW* = lane and shoulder width, respectively, in ft.

L = section length in miles

TRW = total roadway width

MLW = minimum lane width

N = analysis period in years

i = interest rate

The above problem is nonlinear. Also, it can be shown that for the crash models presented above, the objective is a strictly convex function of *LW* and *SW*. Given that the constraints are linear, it is known that a local optimum for this problem will be globally optimal and unique (see, e.g., Hadley, 1964, p. 91). In this case study, the General Algebraic Modeling System (Brooke et al., 2005) was used to compute the optimal solution.

This section's agency life-cycle cost provides details on the computation of the initial construction cost and the maintenance and rehabilitation costs. The section on the user cost of safety over the life-cycle provides details about how the annual crash costs were calculated. Only lane width and shoulder width were considered. Hence, the sum of the lane width and the shoulder width for both directions should equal the TRW. Note that it is not reasonable to have a highway section without an adequate driving lane so there was a lower bound on lane width indicating that the lane width should not be less than 10 ft. A uniform series present worth factor was used to convert future costs, which occur uniformly over an extended period, to a present value. For illustration, a 20-year analysis period (N=20) was assumed, which is the nominal service life of a hot-mix asphalt pavement in Indiana (INDOT, 2013), and a 5% discount rate was used in this case study.

3.3.4.2 Results and Discussion

Results for each functional class

The proposed optimization framework was applied to different highway functional classes with TRWs varying from 26 to 46 ft., which captured most existing width situations on two-lane rural highways in the state. The optimization results are shown in Appendix A and are illustrated in Figure 3.3. For each given TRW, the results of the sensitivity analysis show that the effect varied the ratio of agency cost weight and user cost weight (e.g., a ratio of 1 indicates that the agency and user cost were equally important).

Figure 3.3 presents the optimization results in terms of the shoulder and lane width allocations for the different functional classes and TRWs. For minor arterials and major collectors, the shapes of the optimal widths as functions of the relative weight on the user cost relative to the agency costs were qualitatively similar. When the weight on the user cost was extremely high, the optimum had a shoulder width of zero (the lanes constituted the entire TRW). This was considered intuitive because the contribution of the increased width to crash reduction was always higher for the lane compared to the shoulder (from the crash prediction models). When the weight of the user cost was very low, the optimum had a lane width of 10 ft., and the shoulder took up the remaining TRW. This was intuitive because construction and maintenance costs for shoulders are much lower compared to those of lanes. Between these extremes, there is a range for the ratio of the weight on the agency cost divided by the user cost, which is specific to the TRW and the road functional class; over this range, the optimal width of the lane transitioned from its maximum value of TRW/2 to its minimum feasible value of 10 ft. Regardless of the TRW value, these transitions occurred at lower values for the ratio of agency-to-user cost ratio for minor arterials compared to major collectors.

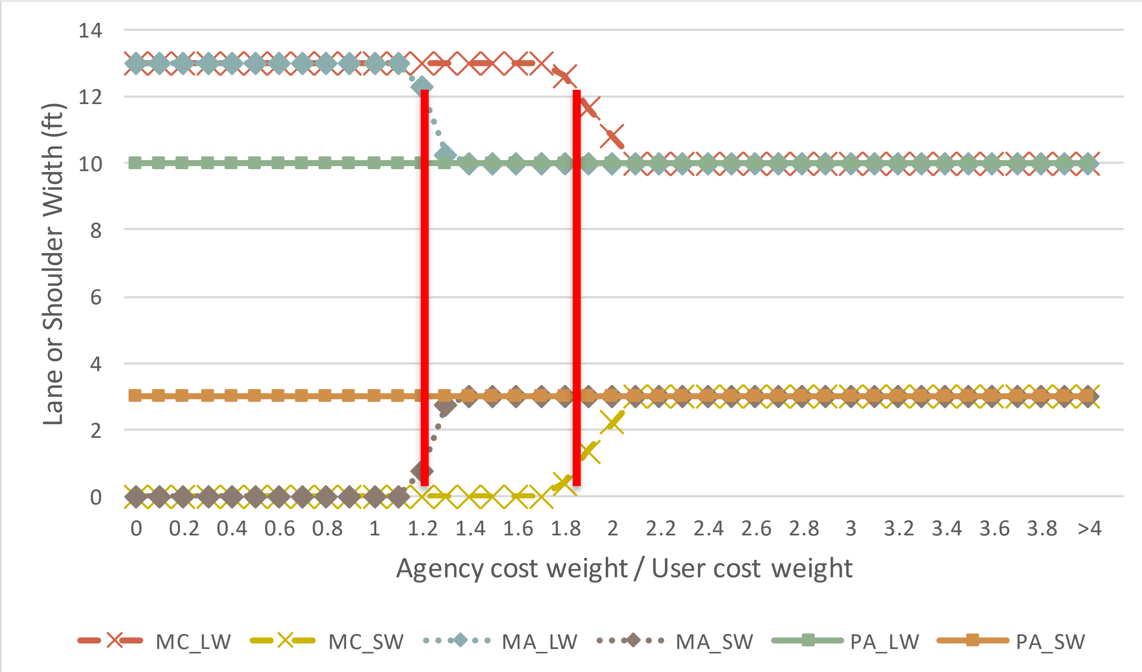
The situation was different for principal arterials, with the optimum having a lane width of 10 ft. and the rest of the TRW taken up by the shoulder, which was due to two factors. First, the crash prediction model for fatal+injury crashes (which had a much higher unit cost compared to PDO crashes), places greater value on wide shoulders than on wide lanes for the purpose of crash reduction, unlike the other two functional classes, specifically regarding the fatal+injury crash prediction model for rural principal arterials. The coefficient of the lane width variable was −0.0837, which had a smaller absolute value compared to the corresponding coefficient of the shoulder width variable, −0.0943. As a result, the absolute value of the marginal effects for lane width was smaller than for the shoulder width. Therefore, the optimization framework was shown to prefer shoulder widening to lane widening. This was intuitive because wide shoulders are more likely to reduce crashes than wide lanes, potentially because even though a wide lane increases the separation between vehicles, it tends to encourage higher vehicle speeds. In addition, the construction and maintenance costs of shoulders are much lower compared to lanes. There was no other penalty for very wide lanes or shoulders because the models used in this case study do not account for the risk compensation that would have indicated increasing crashes for very wide lanes or shoulders. Therefore, from both the user and agency perspectives, having narrow lanes and shoulders that are as wide as possible is preferable. For principal arterials, the optimization framework tends to provide results with minimum lane widths.

The crash prediction models and case studies were used and presented here only for purposes of illustration and can be replaced by other models to fit specific conditions. In addition, using different pavement materials may change the optimal lane and shoulder width ratio. For example, if an unbound shoulder is used instead of a paved shoulder, the resulting lower cost of shoulder construction will benefit the shoulder (i.e., the optimal allocation will call for a wider shoulder and a narrower lane). However, if a longer lifecycle is used in the analysis, the opposite will likely result because an unbound shoulder generally has higher maintenance and rehabilitation costs.

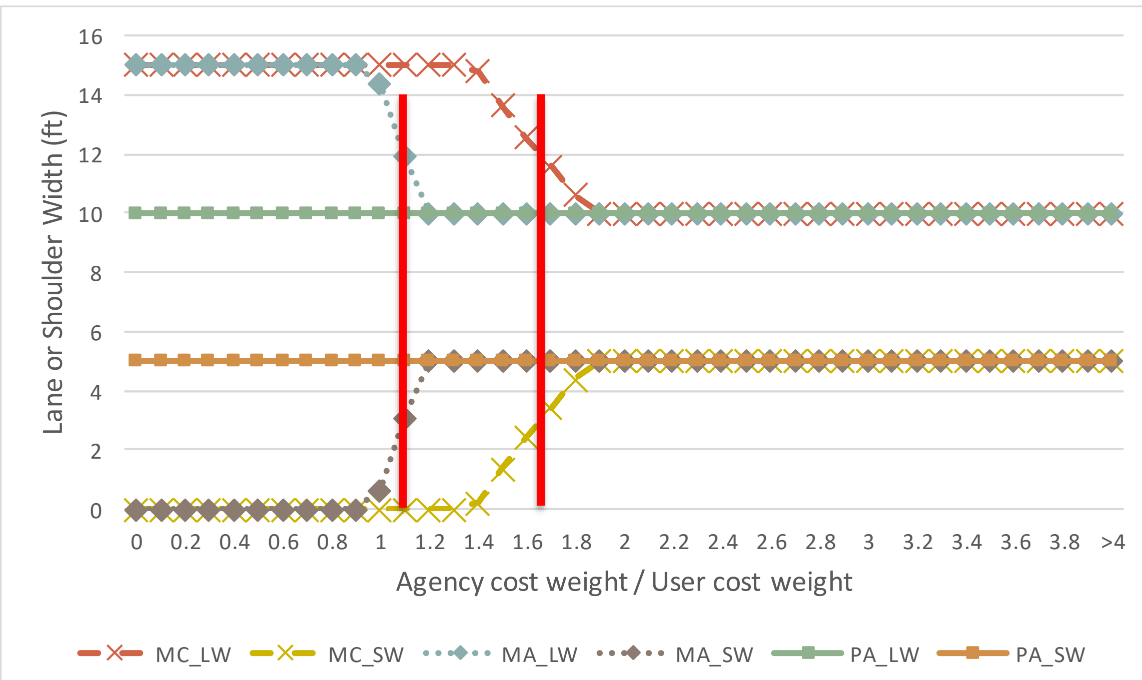
Comparison with prior recommendations

The sensitivity analysis of the agency cost and user cost ratios given by the optimization framework would be useful for evaluating existing practices. For example, the results of the optimal lane width analysis can be compared with the lane and shoulder width specifications in highway design manuals, such as *Policy on Geometric Design of Highways and Streets* (AASHTO, 2011). The AASHTO manual recommends that rural highways should be designed with lanes 12 ft. wide, but in some cases, an 11 ft. lane is acceptable if the alignment and safety record are satisfactory. In the case study, all the variables and parameters in the optimization framework stayed within the range where the crash prediction models were estimated; therefore, the result shows that the case study's highway agency inherently had an agency-to-user cost ratio of 1.3–1.9 for rural major collectors and 0.9–1.2 for rural minor arterials. In other words, for rural major collectors and rural minor arterials, the agency considered the agency cost to be more important than the user cost by a factor ranging from 1.3 –1.9 and 0.9–1.2, respectively. The generally higher weight for the agency cost was chosen because that cost is borne directly by the agency in the form of cash payments for construction and maintenance while the user costs are not felt by the agency and often are indirect and in some cases, even intangible. The vertical red lines in the figures represent the current practice 12 ft. lane width. Figure 3.5 presents examples of the optimal ratio for lane and shoulder widths for different TRWs and different w_{agency}/w_{user} ratios and across different road functional classes.

Assuming a w_{agency}/w_{user} ratio of 1:1 (a reasonable ratio which stays within the range in the discussion above), the total life cycle costs of the optimal solutions were compared with the current practice 12 ft. lane width suggested by AASHTO. Table 3.12 and Figure 3.4 show the total life cycle benefits across the road functional classes for different TRWs if the optimal lane/shoulder ratio is applied instead of the current practice 12 ft. lane width. If a different ratio is applied in the sensitivity analysis, the expected benefits vary depending on the difference between the selected ratio and the current practice ratio corresponding to the 12 ft. lane width (represented by the vertical red bars). The larger difference indicates the higher life-cycle benefits. The discount rate was also found to have small impacts on the optimal lane and shoulder width ratio. Generally, an increase in the discount rate favored wider lanes at the expense of shoulders. The change in ratio was smaller than 10% for any change in the discount rate over the range from 2.5% to 10%.



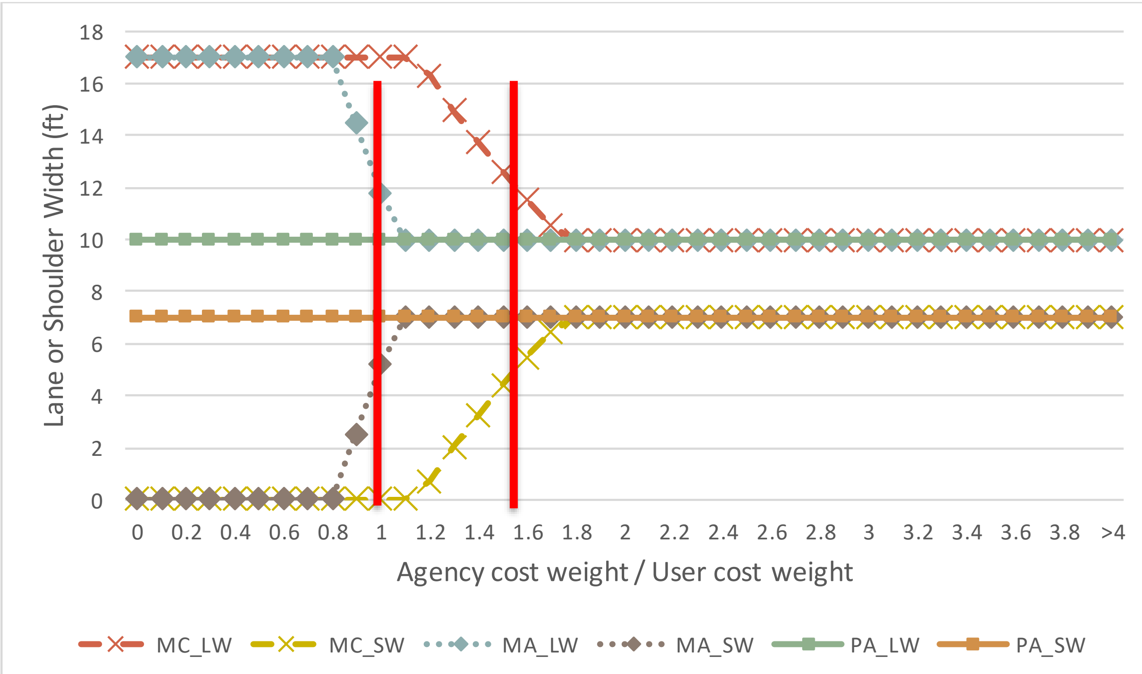
(a) TRW = 26 ft.



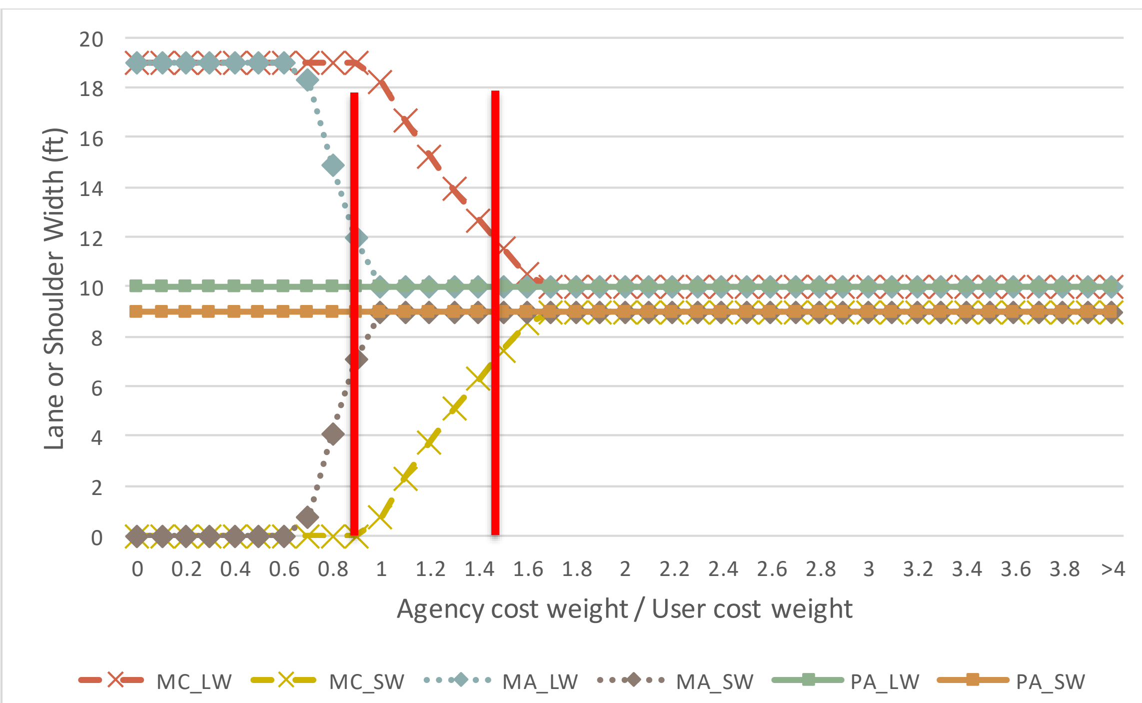
(b) TRW = 30 ft.

Figure 3.3 Optimal lane and shoulder widths across road functional classes for different TRWs

Figure 3.3 continued



(c) TRW = 34 ft.



(d) TRW = 38 ft.

Table 3.12 Total life cycle benefits (LCB) across road functional classes for different TRWs

TRW \ Total LCB (M$)	Major Collector			Minor Arterial			Principal Arterial		
	Current practice	Optimal	Benefit	Current practice	Optimal	Benefit	Current practice	Optimal	Benefit
26 ft.	1.516	1.491	0.025	1.578	1.571	0.007	1.792	1.699	0.093
30 ft.	1.465	1.411	0.054	1.485	1.481	0.004	1.582	1.493	0.089
34 ft.	1.418	1.357	0.061	1.401	1.401	0	1.409	1.324	0.085
38 ft.	1.374	1.326	0.048	1.327	1.323	0.004	1.268	1.186	0.082
42 ft.	1.333	1.301	0.032	1.262	1.251	0.011	1.153	1.073	0.080
46 ft.	1.296	1.276	0.020	1.204	1.187	0.017	1.060	0.981	0.079

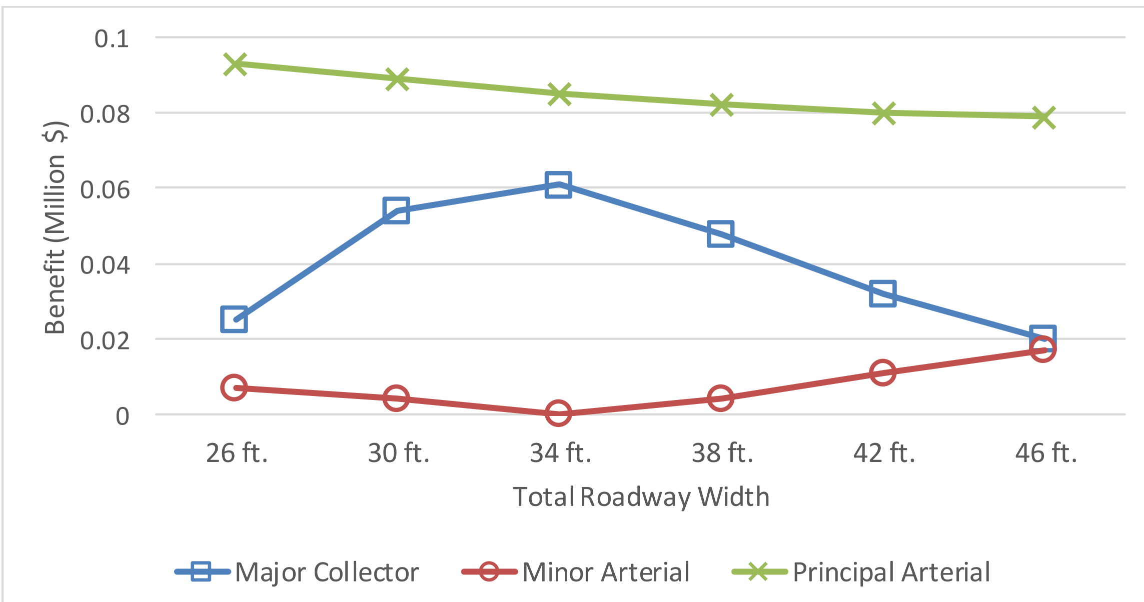


Figure 3.4 Total life-cycle benefits across road functional classes for different TRWs

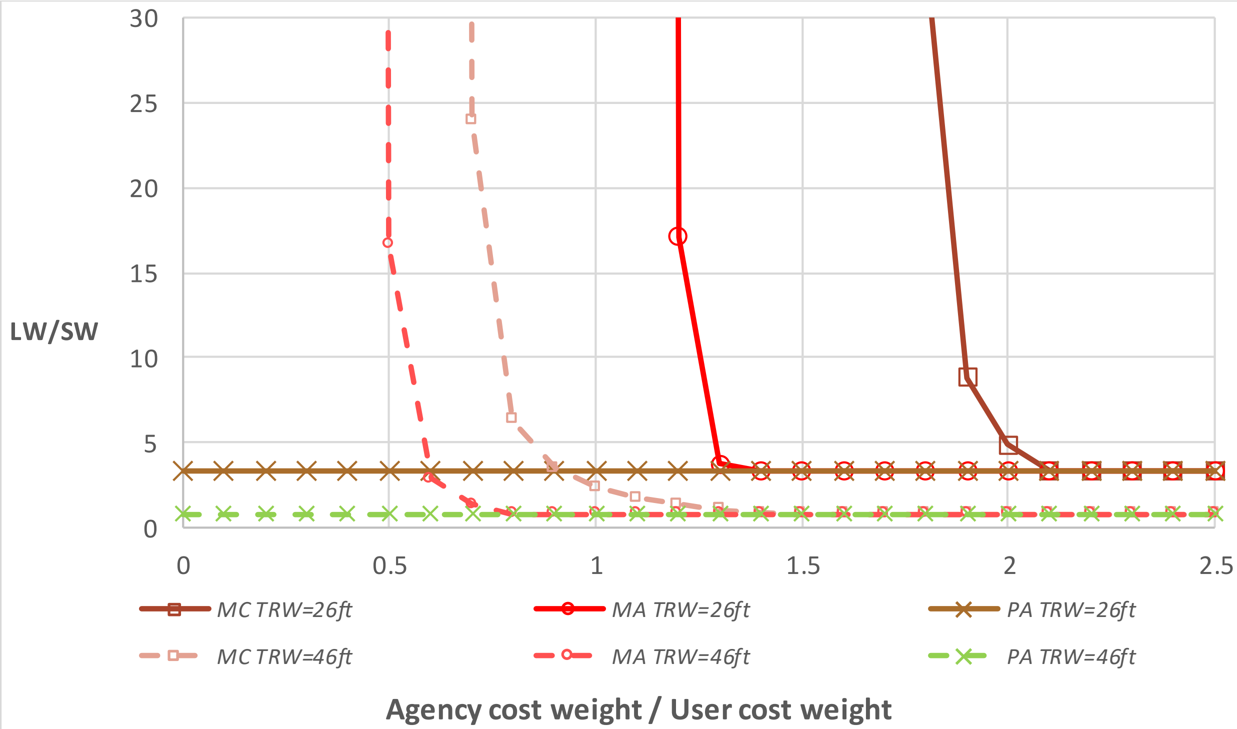


Figure 3.5 Optimal lane and shoulder width ratio by w_{agency}/w_{user} ratio across road functional classes for different TRWs (26 ft. and 46 ft.)

Optimal lane-shoulder width diagrams that consider risk compensation effects

Thus far, the crash prediction models used to demonstrate the proposed framework in this dissertation are not sensitive enough to capture the anomalous safety effects of excessively narrow or excessively wide lanes and shoulders. As the results for the "normal-range" widths show (Figure 3.3), these optimal widths plateaued after a certain point due only to the higher agency costs of very wide lanes or shoulders and not due to the diminishing safety benefits of excessive dimensions. Within these "normal" ranges of lane or shoulder widths, the crash prediction models exhibit the traditional trend (decreasing crashes with increasing width). Risk compensation (also referred in certain literature as risk homeostasis, the Peltzman effect, or offset hypothesis) is the seemingly unintuitive phenomenon where narrow lanes or shoulders have fewer crashes compared to their normal-range widths, and very wide lanes and shoulders have more crashes compared to their normal-range widths (Mannering, 2009; Labi, 2016). This has been attributed to careful driver behavior on segments with very narrow lanes or shoulders and/or reckless behavior on segments with very wide lanes or shoulders. The question is whether the incorporation of such anomalous behavior would yield optimal allocations of the lane and shoulder widths that differ from those obtained using the traditional crash prediction models.

Specifically, the hypothesis is that, using crash prediction models that incorporate risk compensation will result in further penalization of very wide shoulders or lanes by considering other consequences besides their higher agency costs. The overall penalty is likely to be less severe for the shoulder compared to the lane, which means that the optimal solutions will generally have slightly wider shoulders and slightly narrower lanes compared to solutions that do not consider risk compensation. This dissertation investigated this hypothesis for rural major collectors 46 ft. wide.

Figure 3.6 illustrates the difference between the crash prediction functions with and without the risk compensation effects.

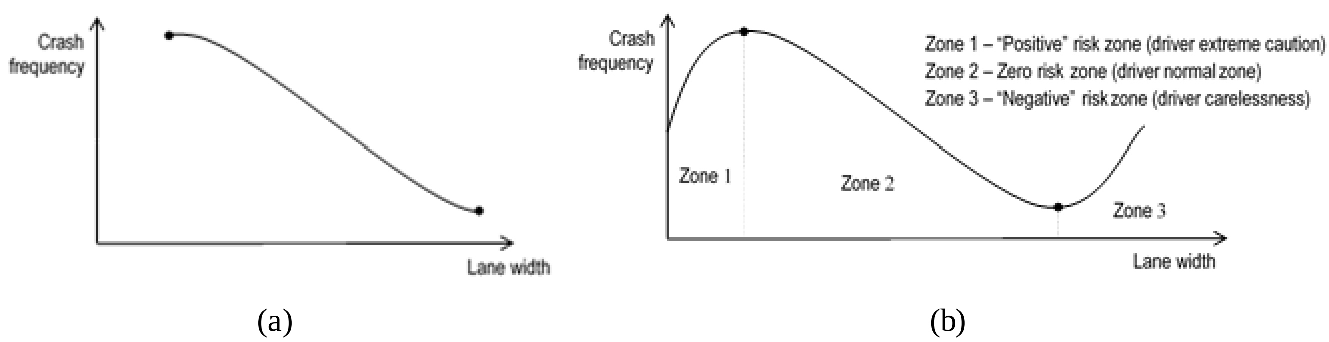


Figure 3.6 Comparison of the crash prediction model without and with risk compensation, rural major collectors, 46 ft. total road width

Figure 3.6(a) depicts a strictly decreasing function, such as the crash prediction equations presented in Section 3.3.4.1, which ignores the anomalous cases of very narrow or very wide lanes where the anomalous effects of risk compensation become manifest. The function in Figure 3.6(b) depicts a third-order polynomial that accounts for the risk compensation effects at the left and right tails.

To demonstrate the effect of risk compensation on the optimal solution (optimal lane-shoulder width allocations), the traditional crash prediction model is replaced by a third-order polynomial, based on Newton's method (Chapra and Canale, 1998). The polynomial function was used to develop a crash prediction model that shows the anomalous effects of very narrow or very wide lanes and shoulders, hence accommodating the risk compensation effects at both extremes of the curve. The analysis was carried out for PDO crashes on rural major collectors with a TRW of 46 ft. and an 8 to 22 ft. range of lane widths. For the lower bound of the lane width, 8 ft. was used instead of 10 ft. to highlight the impact of lower extreme widths.

When the ratio of agency-to-user cost was small, consideration of risk compensation yielded smaller lane-width allocations compared to when risk compensation was not considered. In addition, as the agency-to-user cost weight ratio increased (i.e., as agency cost became more important than user cost), the lane-width allocations in both with- and without-risk compensation cases converged to the lower bound of 8 ft. due to the higher agency cost of lane widening compared to shoulder widening.

When risk compensation was considered, an interesting phenomenon occurred when the agency-to-user cost ratio increased to a certain point (in this case 0.6), namely, the optimal lane width converged sharply to the lower bound. This convergence took place because when the agency cost became increasingly important to the point where the agency cost began to dominate the model, the framework chose the optimal solution as the minimum lane width choice. This was exacerbated by the fact that there was little incentive to protect the user benefits (safety) and very little additional penalty for further reductions in the lane width at that region of the curve because the crash prediction function no longer increased strictly with the decreasing lane width. In general, using crash prediction models that account for the risk compensation effects will further penalize very wide lanes or shoulders beyond just their higher costs. The overall penalty will likely be less severe for the shoulder compared to the lane, which means that the optimal solutions will generally have slightly wider shoulders and slightly narrower lanes compared to solutions that do not consider risk compensation.

Earlier in Figure 3.1, it was shown that Zegeer et al. (1994) demonstrated that when the TRW was greater than 20 ft. (i.e., 10 ft. for each direction), the accident rate generally decreased as the TRW increased. Hence, with the constraint that at least 10 ft. of the roadway width is allocated to the lane on each direction, as in this dissertation, the models and results obtained without considering risk compensation can be considered appropriate for the decision-making process for national or state highways because these highways have road geometries consistent with the domain of geometries where drivers do not exhibit the same risk compensation behavior with respect to lane and shoulder width.

3.4 Efficacy of Machine Learning in Road Safety Analysis: Predicting the Fatality Status of Highway Segments

3.4.1 Introduction

Prediction models for traffic crashes are a key element of safety management systems because these models help forecast crashes and thus identify hazardous locations. Advancements in modeling techniques and new model performance metrics in the existing literature provide opportunities to further enhance the predictive capabilities of crash prediction models, particularly for fatal crashes. This part of the dissertation uses two machine learning modeling techniques (MLMTs) – support vector machine (SVM) and random forests (RF) – to predict whether a highway segment can be considered fatal based on road engineering factors, which are the only factor category within direct control of highway agencies. Arguing that traditional model evaluation metrics can be misleading, this dissertation uses the techniques of precision, recall, and F1 score for model evaluation and validation (MEV) for the following reasons. Using precision yields conservative predictions of fatal segments, which reduces false positives and ultimately minimizes wasting scarce safety resources on highway segments where such resources may be unwarranted. Using recall yields aggressive prediction of fatal segments, which reduces false negatives and ultimately minimizes the denial of needed safety investments at deserving segments. SVM with radial basis function (RBF) kernel was found to yield the best model. This section discusses how the availability of funding within an agency can guide the choice of the MEV technique. Overall, this dissertation demonstrates that MLMTs can help highway agencies identify which highway links have a propensity for fatal crashes and thereby provide information that can help develop fatal crash countermeasures at those locations.

3.4.2 Literature Review

To identify the factors that influence the likelihood of fatality or serious injury for vehicle occupants, past studies categorized crash factors as follows: driver attributes, vehicle condition, enforcement, natural environment, road engineering features, and the complex interactions among these factors (Zeeger et al. 1994; Prato et al., 2014; Mannering et al., 2016). Of these factors, road engineering features have been the focus of most of the past research for two reasons: (a) a wellengineered road "forgives" the adversities imposed by non-engineering factors and (b) unlike non-engineering factors, engineering factors are within the direct control of the highway agency.

The safety literature suggests that most crash prediction models belong to the family of regression models, particularly negative binomial (NB) and Poisson regressions (Milton and Mannering, 1998; Caliendo, 2007). The attractiveness of econometric regression modeling for safety analysis can be explained by the distributional-property nature of crash data (Milton and Mannering, 1998). Probably the most popular fatal crash regression models are zero-inflated NB (ZINB) and zero-inflated Poisson (ZIP) models, owing to their capability in predicting sparse outcomes, because fatal crash data typically contain a preponderance of zeros due to the relatively infrequent occurrence of crashes of this severity. In addition, because an excessive number of zeros in a dataset can cause the variance of the outcome to exceed the mean, the ability of the ZINB models to account for overdispersion in the data arising from the prevalence of zeros in fatal crash data is important (Lee and Mannering, 2002).

Even though the ZINB and ZIP models are suitable for predicting phenomena with sparse outcomes, like all other regression models, they are built with the assumption that the mathematical relationship that exists between the independent and dependent variables is known *a priori*. In cases when this assumption is violated, the outcomes of these models are likely to yield incorrect inferences (Chang, 2005). Additionally, while regression models minimize the error term by accounting for heterogeneity across observations (Washington et al., 2011), excessive error minimization without post-elaboration validation could raise issues related to overfitting. These limitations of regression models have continued to motivate the search for other modeling techniques for crash prediction purposes.

Several artificial neural network (ANN) models have been developed for civil and infrastructure engineering research (Adeli, 2001; Ghosh-Dastidar and Adeli, 2003; Jiang and Adeli, 2007; Panakkat and Adeli, 2009; Yao et al., 2017; Cao and Dai, 2017; Zhang et al., 2018; Nabian and Meidani, 2018). Specifically, researchers have adopted ANN models for predicting motor vehicle crashes (Chang, 2005; Chong et al., 2004; Abdelwahab and Abdel-Aty, 2002; Pande and Abdel-Aty, 2008). However, ANNs have been criticized for their tendency to overfit and have been classified as black-box networks compared to other common machine learning techniques (Yu and Abdel-Aty, 2013). Due to ANN's limitations, researchers have advocated for the use of support vector machine (SVM) models for predicting traffic crashes (Yu and Abdel-Aty, 2013; Li et al., 2012; Chen et al., 2016). SVM has been proven to address problems with overfitting while providing better or similar performance compared to ANN models, particularly for medium-size datasets (Li et al., 2008). SVM models have been shown to be powerful in predicting and classifying problems such as the probability of roadway crash severity levels (Li et al., 2012). Similarly, RF models have been shown to be useful for classification and regression. RF is a classic ensemble method where predictions are made by not only a single model but a group of models (Breiman, 2001; Pham et al. 2010).

3.4.3 Case study

3.4.3.1 Introduction

This case study proposes prediction models to determine which highway segments in a network can be expected to have at least one fatality based on their geometric design, traffic volume, or pavement condition. The case study used the concepts of precision, recall, and F1 score to measure the performance of prediction models developed using MLMTs versus econometric techniques. This comparison was made in order to investigate the notion that econometric models are less flexible than machine learning algorithms (James et al., 2013). These model evaluation metrics were used because the traditionally used metric of accuracy can be misleading (Provost et al., 1998; Davis and Goadrich, 2006). Using precision, recall, and F1 score curves, machine learning models and traditional econometric models were compared for their efficiency and reliability in predicting fatal highway segments. It is anticipated that these proposed models can help highway agencies identify fatal highway segments and thereby better allocate their resources (safety funding, labor, equipment, etc.) to the most fatal segments.

3.4.3.2 Literature Review

Several past studies predicted vehicle crashes and related fatality and injury severity levels using statistical models because such models help promote understanding of the data. Milton and Mannering (1998) used negative binomial regression as a statistical model for predicting crash frequency considering highway geometrics. Donnell and Mason (2006) used negative binomial regression models to predict the frequency of median barrier crashes. Caliendo et al. (2007) used

Poisson, negative binomial, and negative multinomial regression models to predict crashes for multi-lane roads in Italy considering geometric and traffic factors.

Zero-inflated regression models have been widely applied in the literature for predicting data with an excessive number of zeros. In addressing zero inflation, Lee and Mannering (2002) used ZIP and ZINB models to analyze the impact of roadside features on the frequency and severity of run-off-road crashes. Lord et al. (2004) provided guidelines for selecting the appropriate type of regression model for crash analysis and concluded that Poisson and NB models provide only approximations of crashes. Poisson models perform well under homogenous conditions, and NB models perform well under overdispersed conditions (Washington et al., 2011).

Researchers also have adopted various modeling techniques to account for heterogeneity across observations, which is not addressed by traditional NB and Poisson models. Milton and Mannering (2008) used mixed logit to model highway crash severities while capturing segment-specific heterogeneity. Kim et al. (2010) also used mixed logit to model pedestrian injury severity. Malyshkina et al. (2009) encouraged the use of Markov switching instead of zero-inflated models to predict vehicle crash frequencies. Some researchers also have used multivariate modeling to address the heterogeneity across observations. For example, to predict crash counts, Ma et al. (2008) proposed a multivariate Poisson-lognormal regression model using the Bayesian method as an improvement over the traditional counterpart univariate model. Park et al. (2007) used multivariate Poisson-lognormal models to jointly model crash frequency by severity. Ma et al. (2016) observed that the distribution of crash rates was skewed and developed a lognormal hurdle model to capture the heterogeneity of the data.

Machine learning paradigms have been adopted as a promising modeling technique to avoid violating the assumptions in regression models for predicting vehicle crashes. The main approaches for machine learning have been neural networks (Mussone et al., 1999; Abdelwahab and Abdel-Aty, 2002), Bayesian neural networks (Riviere et al., 2006; Xie et al., 2007), and SVM (Li et al., 2008). SVM is one of the most widely adopted machine learning methods due to its capability of forecasting phenomena with relatively little data (Alenezi et al., 2008; Rajasekaran et al., 2008). Chong et al. (2005) compared neural networks, SVM, decision trees, and a concurrent hybrid model involving decision trees and neural networks to model injury severity during head-on front-impact point collisions and concluded that the hybrid model had the highest accuracy. Li et al. (2008) compared SVM and NB models for predicting motor vehicle crashes and suggested that SVM models are faster to implement and perform similarly to, if not better than, back-propagation neural networks (BPNN).

Additionally, Li et al. (2012) compared the performance of SVM and ordered probit models in predicting injury severity and found that the SVM model had a higher percentage of correct predictions compared to the ordered probit model. Yu and Abdel-Aty (2013) used SVM models to evaluate the real-time crash risk of vehicles and suggested that SVM models can achieve good predictive performance even with a small sample size. Dong et al. (2015) used SVM to evaluate crash risk while accounting for cross-zonal spatial correlations. Also, work has been done using RF models for crash prediction. Pham et al. (2010) used RF to separate traffic situations into non-crash and pre-crash conditions using data collected from loop detectors.

This case study contributes to the existing literature in a number of ways. First, it focused on the crash severity level that is the most difficult to predict, fatal crashes, and does so using machine learning techniques. There was a gap in the literature regarding the identification of fatal segments exclusively using machine learning techniques. Additionally, to assess the efficacy of the developed model, this dissertation used the concepts of precision, recall, and F1 score as performance measures for highway crash prediction models. Furthermore, this dissertation focused on road engineering factors, which are the only crash factors that are within direct control of highway agencies. This focus allows agencies to react more quickly and directly to the problem of fatal segments. Finally, this case study provides learning curves that be used to evaluate the quality of the data used to develop crash prediction models. Using these learning curves, researchers can determine whether the dataset they use to develop their models is sufficient or whether a higher quality/quantity (with more variables or observations) of data can be used to further improve the models.

3.4.3.3 Methodology

This case study used a machine learning approach (SVM and RF) to develop a fatal crash prediction function. The methodology also includes a comparison of the predictive efficacy of machine learning models and zero-inflated models and logistic regression. Generally, econometric models are considered less flexible than machine learning algorithms (James et al., 2013). Due to their functional forms, econometric models are typically continuous, monotone, and differentiable. Marginal effects and elasticities can be easily computed for econometric models. As a result, econometric models usually have high interpretative power. In contrast, machine learning algorithms usually have superior predictive power due to the flexibility of their function forms.

3.4.3.4 Econometric Models for Comparison

For comparison purposes, traditional econometric models commonly used in analyzing traffic crashes were implemented in this dissertation using the same dataset used for the machine learning algorithms. These models include zero-inflated Poisson regression and logistic regression. Note that some researchers argue that econometric models can be considered “statistical machine learning” techniques. For example, one may argue that logistic regression could behave similarly to SVM in some cases if kernel functions are adopted and regularization is applied to control model parameters. In such cases, the main difference between logistic regression and SVM is the techniques’ different loss functions. However, in this case study, the econometric models refer to regression models without kernel functions and regularization. The main concepts of zero-inflated regression and logistic regression are briefly described in Washington et al. (2011) and Washington et al. (2014).

a) Zero-Inflated Poisson/Negative Binomial Regression

Due to the count nature of crash frequency, two forms of count models (Poisson and negative binomial) were considered in this dissertation. As determined from past research, these model forms differ in their assumptions of the conditional mean and variance of the dependent variable. The Poisson model requires the mean and variance to be equal (i.e., E(crash frequency) = Var(crash frequency)). If these two values are not equal, then the data are either underdispersed or overdispersed and the estimated parameters become biased. In such cases, the negative binomial model is recommended.

There is also a class of count models that considers the possibility that count data are generated from two distinct processes: a normal count state (when crashes occur according to a count process, such as Poisson or negative binomial) and a zero-count state (when crashes do not occur). Models that consider this dual-state process are known as zero-inflated models because the observed “zero” is inflated by the possibility that a zero-count state exists. Zero-inflated models have been applied elsewhere with considerable success, particularly in the literature on motor vehicle crash frequency (Washington et al., 2011; Chin and Quddus, 2003).

In the case of highway crashes, the zero-inflated Poisson model assumes that the crash frequency y_i at section i is as follows:

$y_i = 0$ with probability $p_i + (1 - p_i)EXP(-\lambda_i)$

$y_i = y$ with probability $\frac{(1-p_i)EXP(-\lambda_i)\lambda_i^y}{y!}$, y = 1, 2, 3…

where p_i is the probability of being in the zero state, y is the number of crashes, and λ_i is the expected crash frequency.

The zero-inflated negative binomial regression model follows a similar formulation:

$y_i = 0$ with probability $p_i + (1 - p_i)\left[\frac{1/\alpha}{(1/\alpha)+\lambda_i}\right]^{1/\alpha}$

$y_i = y$ with probability $(1 - p_i)\frac{\Gamma((1/\alpha)+y)u_i^{1/\alpha}(1-u_i)^y}{\Gamma(1/\alpha)y!}$, y = 1, 2, 3…

where $u_i = \frac{(1/\alpha)}{(1/\alpha)+\lambda_i}$, Γ denotes the γ distribution and all other terms are as previously defined. The function that describes the probability of being in the zero state p_i is typically based on the assumption of normality (probit model); therefore, $p_i \approx$ Normal [$\boldsymbol{\beta}\mathbf{Z}_i$], where $\boldsymbol{\beta}$ is a vector of estimable parameters and $\mathbf{Z}_i$ is a vector of independent variables that determine the zero-state probability. It should be noted that the influential factors of one process are not necessarily the influential factors of the other (Washington et al, 2011). In other words, an independent variable that affects highway crash frequency does not necessarily influence the probability that a highway section is in a zero crash state.

To statistically determine the appropriate count model to be used for modeling (Poisson, negative binomial, ZIP, or ZINB), the Vuong statistic was used, which is defined as follows (Vuong, 1989):

$$m_i = LN\left(\frac{f_1(y_i|X_i)}{f_2(y_i|X_i)}\right)$$

where $f_1(y_i|X_i)$ is the probability density function of model 1 and $f_2(y_i|X_i)$ is the probability density function of model 2. Using this statistic, with *n* being the total number of observations (highway segments), the statistic for testing model 1 versus model 2 is as follows (Washington et al., 2010):

$$V = \frac{\sqrt{n}[\frac{1}{n}\sum_{i=1}^{n} m_i]}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(m_i - \frac{1}{n}\sum_{i=1}^{n} m_i)^2}}$$

The Vuong statistic is asymptotically standard normal, and if the absolute value of *V* is less than $V_{critical}$ (1.65 for a 95% confidence level, one-tailed *t* distribution), the test does not support the selection of one model over another and the test is inconclusive. Large positive values of *V* greater than $V_{critical}$ favor model 1 over model 2, whereas large negative values support model 2 over model 1. Shankar et al. (1997) provided model selection guidelines based on possible values of Vuong test and overdispersion parameter (α) *t*-statistic (using the 95% confidence level):

Use ZIP or Poisson when $V < -1.65$ and α's *t*-statistic < 1.65
Use NB when $V < -1.65$ and α's *t*-statistic > 1.65
Use ZIP when $V > 1.65$ and α's *t*-statistic < 1.65
Use ZINB when $V > 1.65$ and α's *t*-statistic > 1.65

b) Logistic Regression

A logistic regression model is used to predict the probabilities of the possible outcomes of a dependent variable given a set of independent variables. In this part of the dissertation, the dependent variable is the probability (P_i) that segment *i* is a fatal segment:

$$P_i = \frac{EXP[\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots + \beta_p X_{p,i}]}{1 + EXP[\beta_0 + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots + \beta_p X_{p,i}]}$$

where β_0 is the model constant, p is the number of independent variables, and $\beta_1, \ldots, \beta_p$ are unknown parameters corresponding to the independent variables $X_1, \ldots, X_p$. These unknown parameters are typically estimated using maximum likelihood methods.

The probability that a highway segment is a fatal segment is calculated, and a threshold probability T is used to make the final assessment. For example, highway segment i is identified as a fatal segment if $P_i \geq T$. This threshold is typically tuned using a validation dataset.

3.4.3.5 Model Performance Metrics, Model Selection, and Learning Curves

A common error in evaluating the performance of highway safety models is to use only their predictive accuracy. For example, less severe crashes (e.g., property damage only and injury crashes) are typically much more frequent than fatal crashes. Out of the 1,601 highway segments in the study database, 1,528 (95.44%) of the segments had no fatal crashes and 51 segments had only 1 fatal crash (see Figure 3.7). Only 10, 5, 5, 0, 1, and 1 segments had 2, 3, 4, 5, 6, and 7 fatal crashes, respectively. Since most of the highway segments had no fatal crashes, one can easily propose a non-informative and superfluous model with a high accuracy of 95.44% which will predict that a given segment will not have a fatal crash irrespective of any provided information (e.g., geometric design, traffic volume, etc.). Therefore, the accuracy of a model by itself is generally not an appropriate metric for evaluating model performance in terms of predicting fatal crashes. In this part of the dissertation, the metrics for model performance evaluation are precision, recall, and F1 score. The definitions of precision, recall, and F1 score are discussed in Chapter 2.

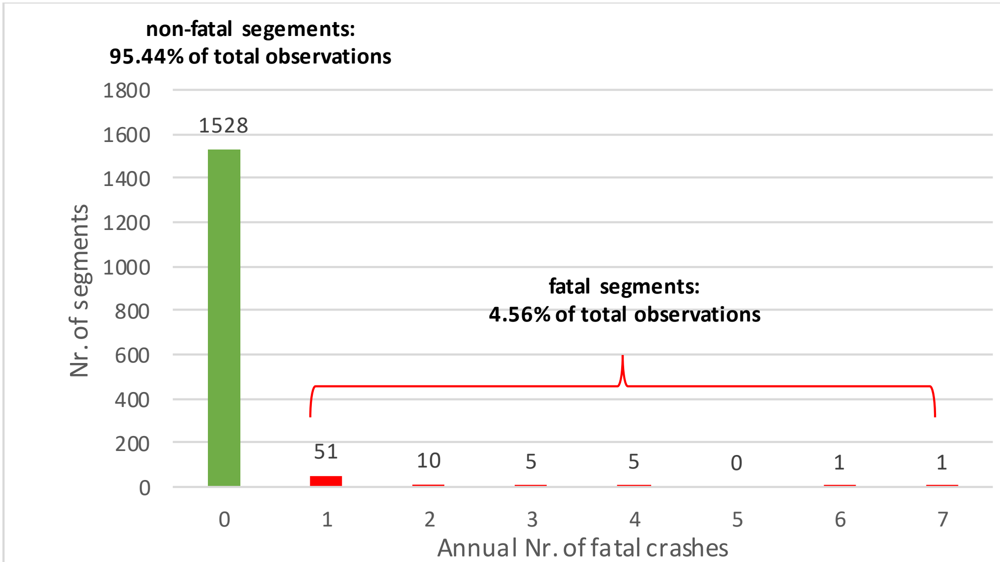


Figure 3.7 Distribution of fatal crashes

For each of the classifiers, several plots are presented in this dissertation: the precision, recall, and F1 score versus the hyper-parameters for both the training and validation data. In the cross-validation process, the average values of precision, recall, and F1 score for the training and validation folds were computed and plotted. For each type of modeling technique, the classifiers with the best performance in the cross-validation process were elected into the final pool of potential best models. Then, multiple hypothesis tests were formulated with Bonferroni correction to determine the final selected model for this dissertation. The precision, recall, and F1 score of the final selected model on the testing data are shown.

The learning curves showing the relationship between the model performance and the number of training samples are presented below. These curves can help highway agencies understand the effects of training sample size. For example, they can provide highway agencies with useful information to evaluate (a) how much data are needed for the classifier and (b) whether additional data can help train the classifier and enhance predictive accuracy. This information can help agencies make projections to evaluate whether additional data can help train the classifiers and help the agency make better decisions.

3.4.3.6 Model Implementation and Results

The goal of the model was to build classifiers that can be used to predict the fatality status of a given highway segment. The fatality status is defined as follows:

- Fatal segment: there is at least one fatal crash on the highway segment in a given year
- Non-fatal segment: there are no fatal crashes on the highway segment in a given year

It is expected that this analysis can help highway agencies in two ways. First, if a classifier's performance is acceptable, it can help highway agencies determine which highway segments can be expected to be fatal. Then, highway agencies can further investigate these fatal segments through site visits to identify the appropriate safety countermeasures (e.g., guardrails, pavement rehabilitation, etc.) and to allocate available resources (e.g., safety funding). Second, if a classifier's performance is not acceptable, it may indicate that the current features are not adequate to determine whether a highway segment is fatal or non-fatal. Such a result suggests that highway agencies should collect other information (e.g., weather conditions, driver ages, etc.) to identify fatal segments. This result also indicates that information on highway design, traffic volume, and pavement condition alone may not be adequate for reliable predictions.

3.4.3.7 Classification Goal in Terms of Performance Metrics

In training machine learning classifiers, the hyper-parameters can be fine-tuned, and final classifiers can be selected based on the precision, recall, or F1 score. The choice of performance metric depends on the objective of the agency. Selection based on precision encourages the classifiers to predict fatal segments conservatively and thereby reduces the number of false positive predictions. As a result, highway agencies can potentially prevent wasting funds on highway segments that are predicted as fatal but that in reality are not fatal. Selection based on recall encourages the classifiers to predict fatal segments aggressively, thereby reducing false negative predictions. In such cases, the number of false negative predictions is reduced at the cost of a greater number of false positive predictions most of the time. As a result, a greater number of truly fatal highway segments can be detected by the classifier at the expense of a potentially greater number of false positive predictions of fatal segments. If funding is limited, highway agencies should choose classifiers that have high precision. If funding is not a significant concern, agencies should choose classifiers that have high recall. The F1 score, by definition, is the harmonic mean of precision and recall so this metric is recommended for highway agencies when balanced decisions are preferred. In addition, the receiver operating characteristics (ROC) curve (James et al., 2013) can be used as a supplemental performance metric for evaluating machine learning classifiers. With respect to ROC, the overall performance of a classifier is given by the area under the (ROC) curve (AUC), where the larger the AUC, the better the classifier. When the differences in classifier performance in terms of precision, recall, and F1 score are not statistically significant, the AUCs of classifiers can be compared.

Prior to training the classifiers, the features are standardized by removing the mean and scaling to the unit variance. During training and testing, the data are centered and scaled independently on each feature by computing the relevant statistics. This process can improve both the data fitting and prediction performance because it normalizes the dissimilar units and magnitudes of the different features (Li et al., 2008). The standardization is accomplished using the following equation:

$$x_{ip}' = \frac{x_{ip} - mean(x_{ip})}{std(x_{ip})} \qquad \forall i, p$$

where x_{ip}' is the standardized value of observation i and feature p and x_{ip} is the original value of observation i and feature p.

3.4.3.8 Implementation of Machine Learning Models

General Procedure

Initially, the dataset is divided into two parts: training and validation data (80% of the observations) and testing data (20% of the observations). The training and validation data are used to tune the hyper-parameters through k-fold cross-validation using the criteria discussed in Section 2.4.1. Three groups of models are trained:

(1) Linear SVM, SVM with polynomial kernel of degree 2, 3, and 4

(2) SVM with RBF kernel

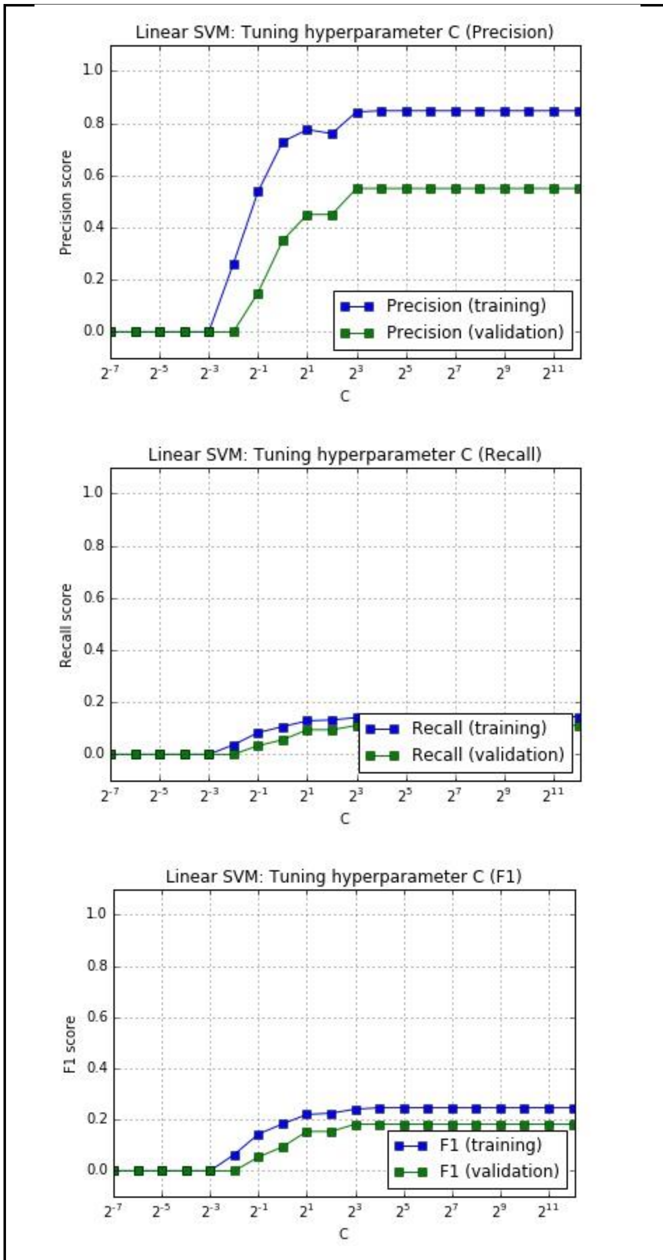
(3) RFs

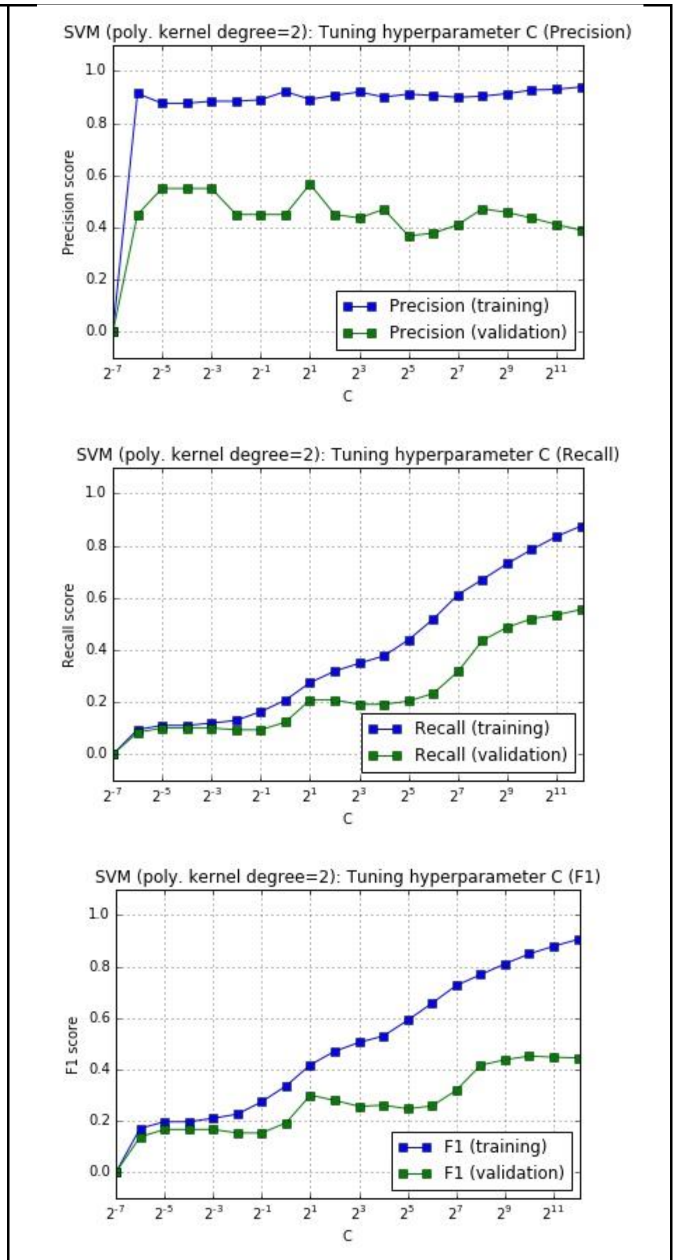
For SVM, the RBF kernel is much more flexible than the linear or polynomial kernels because it maps the features onto an infinite dimensional space (Friedman et al., 2001). If the hyper-parameters are tuned properly, SVM with RBF kernel is expected to outperform SVM with linear or polynomial kernels, considering that the correlation and causation between the dependent variable (fatal highway crashes) and independent variables (geometric design, traffic, and pavement condition) are extremely complicated. In this dissertation, however, SVM with linear or polynomial kernels were trained and validated because the aim was to determine whether the correlation and causation of fatal crashes with other factors can be quantified to an understandable level (e.g., to the fourth order of feature interactions). When any SVM model with a first to fourth order of polynomial kernel performs similarly to SVM with RBF kernel, it is an indication that fatal crashes are highly correlated with the geometric design, traffic, and pavement condition up to their fourth order of interactions. This information is expected to be useful for improving the econometric models in the literature because even though machine learning algorithms have superior predictive power compared to econometric models, econometric models are easier to interpret.

Tuning Hyper-parameters for Linear SVM, SVM with Polynomial Kernel of Degree 2, 3, and 4 (Group 1)

The hyper-parameter tuned in group (1) is C. Parameter C is a regularization term that provides a way to control overfitting: a large value of C discourages any positive ε_i (i.e., slack variables, see Section 2.3.2) and may lead to an overfit wiggly boundary in the original feature space, while a small value of C encourages the boundary to be smoother (Hastie et al., 2009). The best model in this group is SVM with polynomial kernel degree 3 with a C value of 32. It has an average precision of 0.6869, an average recall of 0.6446, and an average F1 score of 0.6265. (Note that these values are the average values for the k-fold cross-validation. That is, the equation $\mathrm{F1} = 2\frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}}$ does hold for individual values but may not hold for the average value.)

Figure 3.8 presents the precision, recall, and F1 score versus the values of the hyper-parameter. The ROC curve for the best model in this group is shown in Figure 3.9, where the area under the curve is 0.7724.

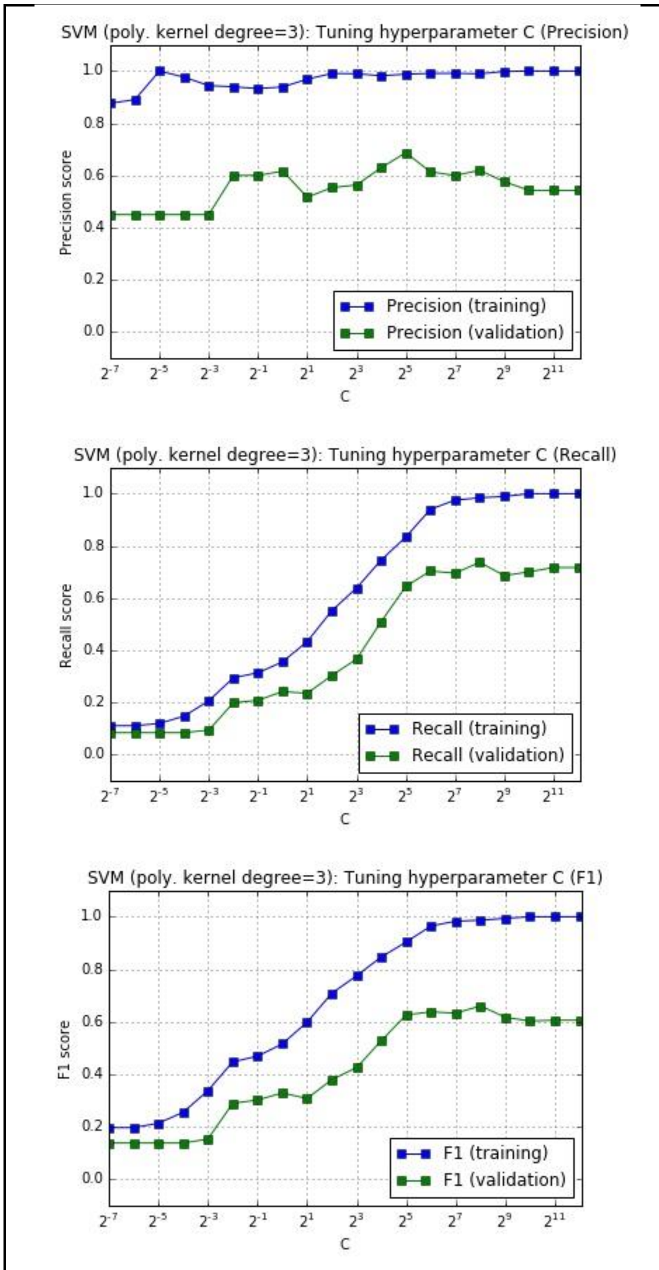




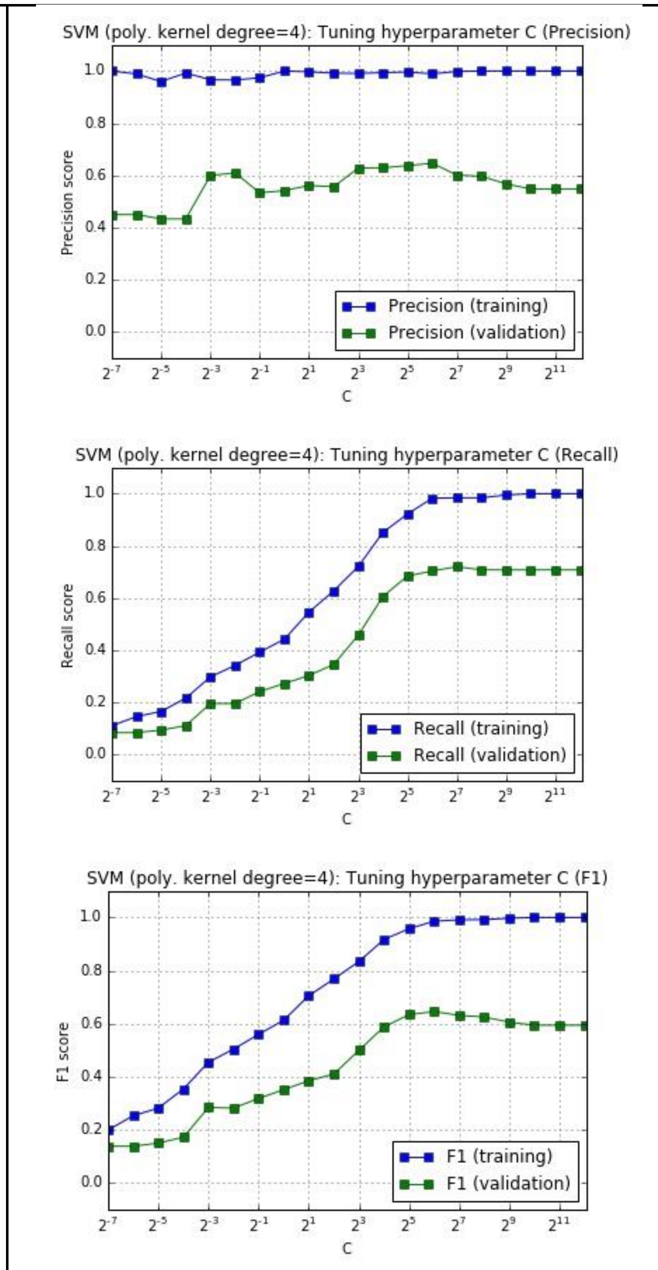
(a) Linear SVM	(b) SVM with polynomial kernel of degree 2

Figure 3.8 Tuning hyper-parameters for linear SVM, SVM with polynomial kernel of degrees 2, 3, and 4 (examples shown here)

Figure 3.8 continued



(c) SVM with polynomial kernel of degree 3



(d) SVM with polynomial kernel of degree 4

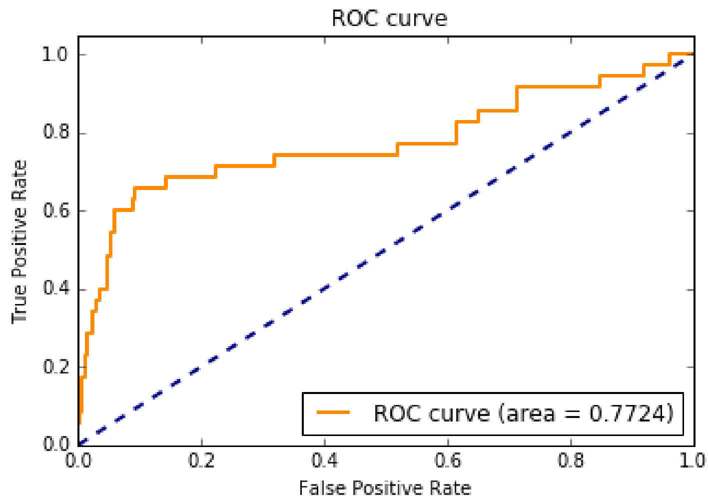
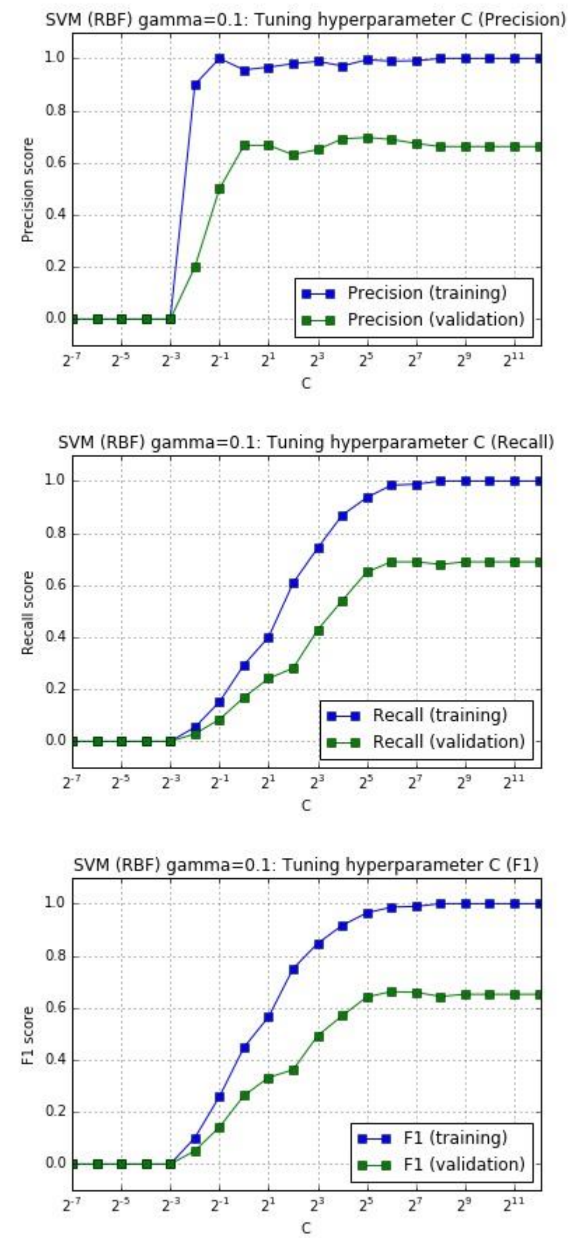


Figure 3.9 ROC curve (with AUC) for SVM with polynomial kernel degree 3 (Group 1 best model)

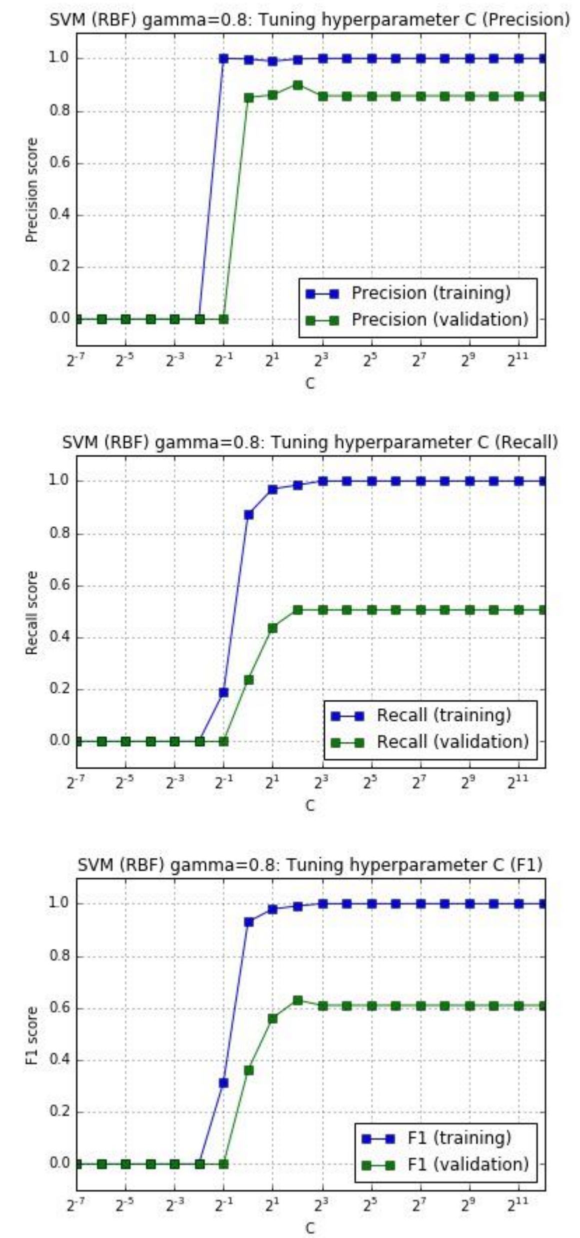
Tuning Hyper-parameters for SVM with RBF Kernel (Group 2)

The hyper-parameters tuned in group (2) are C and γ. Parameter C is a regularization term that provides a way to control overfitting: a large value of C discourages positive ε_i (i.e., slack variables, see Section 2.3.2) and may lead to an overfit wiggly boundary in the original feature space, while a small value of C encourages the boundary to be smoother (Hastie et al., 2009). γ is a positive constant that controls the Euclidean distance between observations: as γ increases, generally the fit becomes more nonlinear, which may lead to an overfit wiggly boundary (James et al., 2013). The best model in this group is SVM with RBF kernel with a C value of 4 and a γ value of 0.8. It has an average precision of 0.9016, an average recall of 0.5054, and an average F1 score of 0.6298.

Figure 3.10 presents the precision, recall, and F1 score versus the values of the hyper-parameters. The ROC curve for the best model in this group is shown in Figure 3.11, where the area under the curve is 0.9228.



(a) SVM with RBF kernel: γ = 0.1



(b) SVM with RBF kernel: γ = 0.8

Figure 3.10 Tuning hyper-parameters for SVM with RBF kernel (examples shown here)

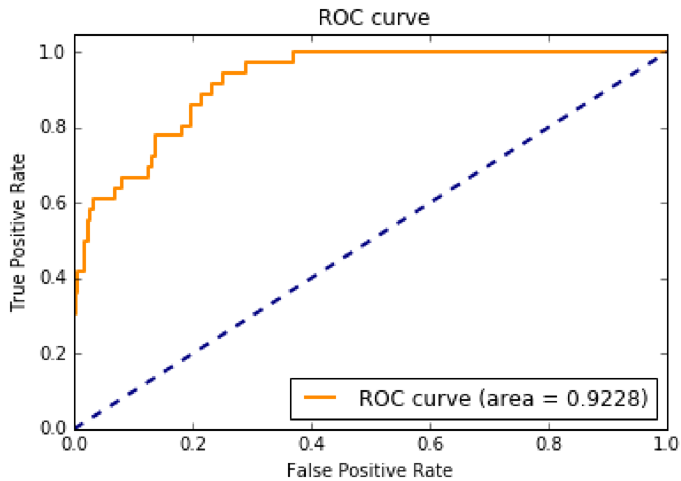
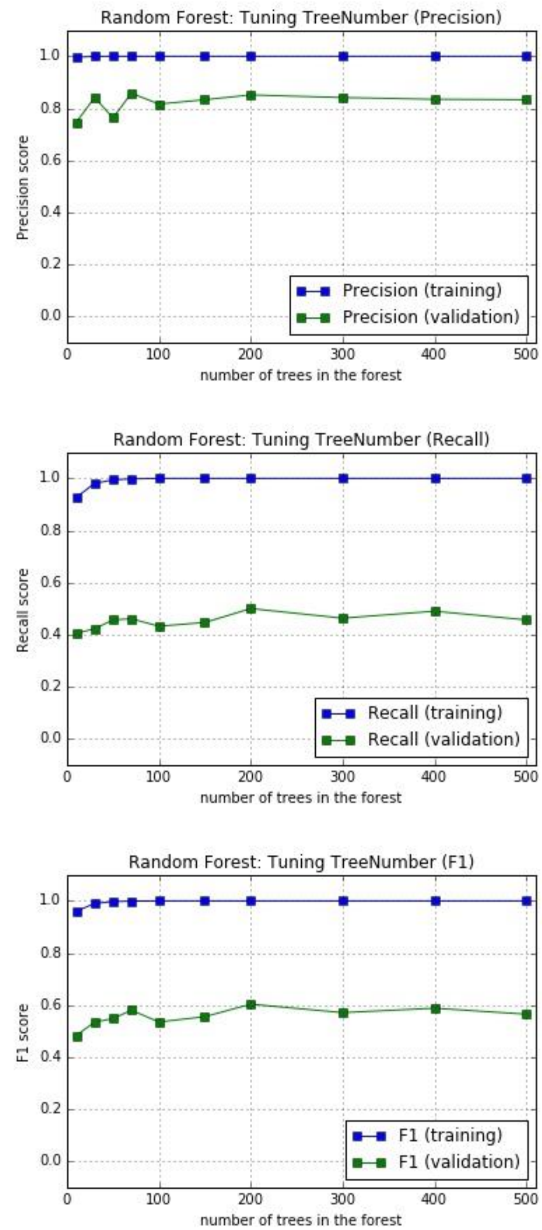


Figure 3.11 ROC curve (with AUC) for SVM with RBF kernel

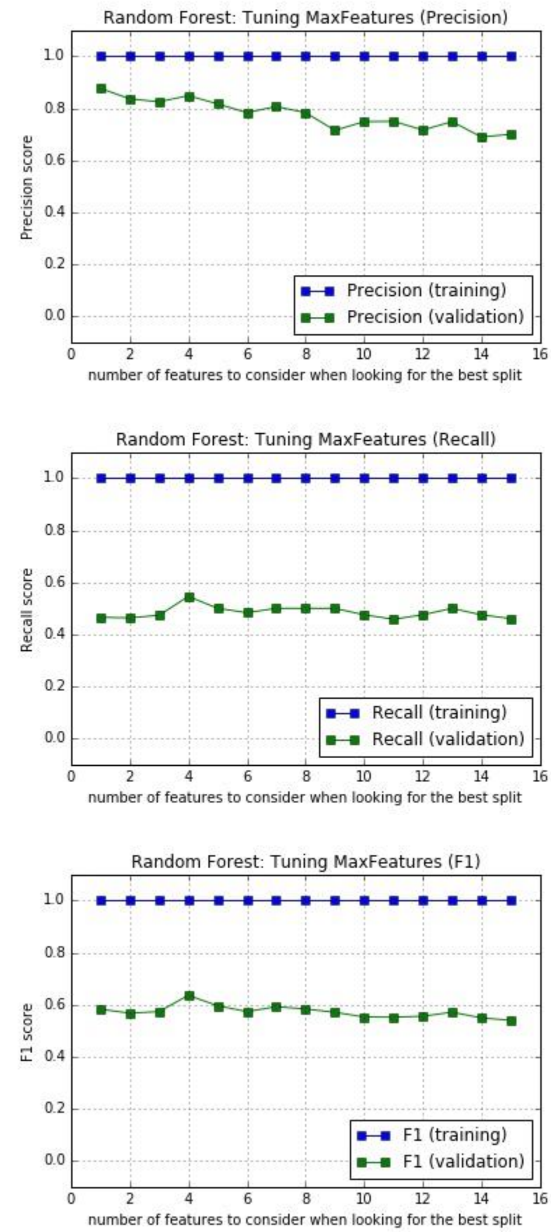
Tuning Hyper-parameters for Random Forests (Group 3)

Figure 3.12 shows that when the number of trees exceeds 200, the validation performance tends to become stable. Increasing the number of trees generally improves the performance by reducing the variance at the expense of the computing time (James et al., 2013). In practice, the best value for the maximum number of features m to be considered at each split for building the tree structure depends on the problem (Hastie et al., 2009), although popular choices include $\sqrt{p}, \log(p), p/3$, etc. The value of m was chosen to be 1 in this part of the dissertation. The best model in this group is RF with 300 trees, 1 maximum feature, and a maximum depth of 30. It has an average precision of 0.8766, an average recall of 0.4661, and an average F1 score of 0.5831.

The ROC curve for the best model in this group is shown in Figure 3.13, where the area under the curve is 0.8984.



(a) Tuning number of trees in the forest



(d) Tuning number of features to consider at splits

Figure 3.12 Tuning hyper-parameters for RF (examples shown here)

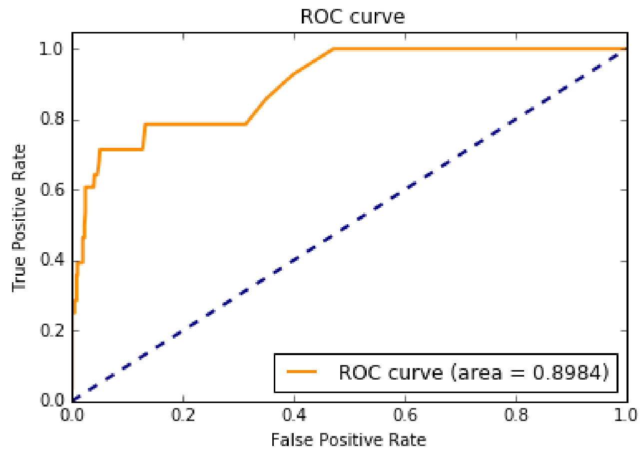


Figure 3.13 ROC curve (with AUC) for RF

Final Selection of Machine Learning Classifiers using Hypothesis Tests

Based on the cross-validation process, SVM with RBF kernel seemed to perform better than the other two types of models in terms of average performance with the validation data. For example, the validation precision values for SVM with RBF kernel, SVM with polynomial kernel, and RFs are 0.9016, 0.6869, and 0.8766, respectively. The AUC values for SVM with RBF kennel, SVM with polynomial kernel, and RFs are 0.9228, 0.7724, and 0.8984, respectively. Multiple hypothesis tests with Bonferroni correction were carried out to ascertain whether the performance differences were significant. It was found that SVM with RBF kernel and RF performed significantly better than SVM with polynomial kernel. However, the performance difference between SVM with RBF kernel and RF was not statistically significant. Figure 3.14 shows the performance of the final selected machine learning models with the testing data.

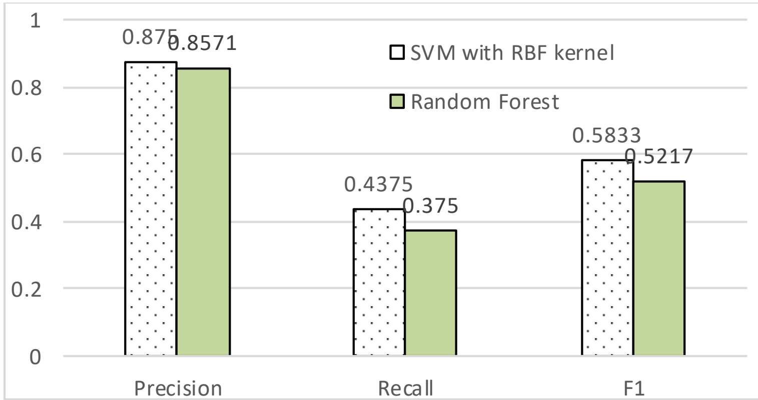
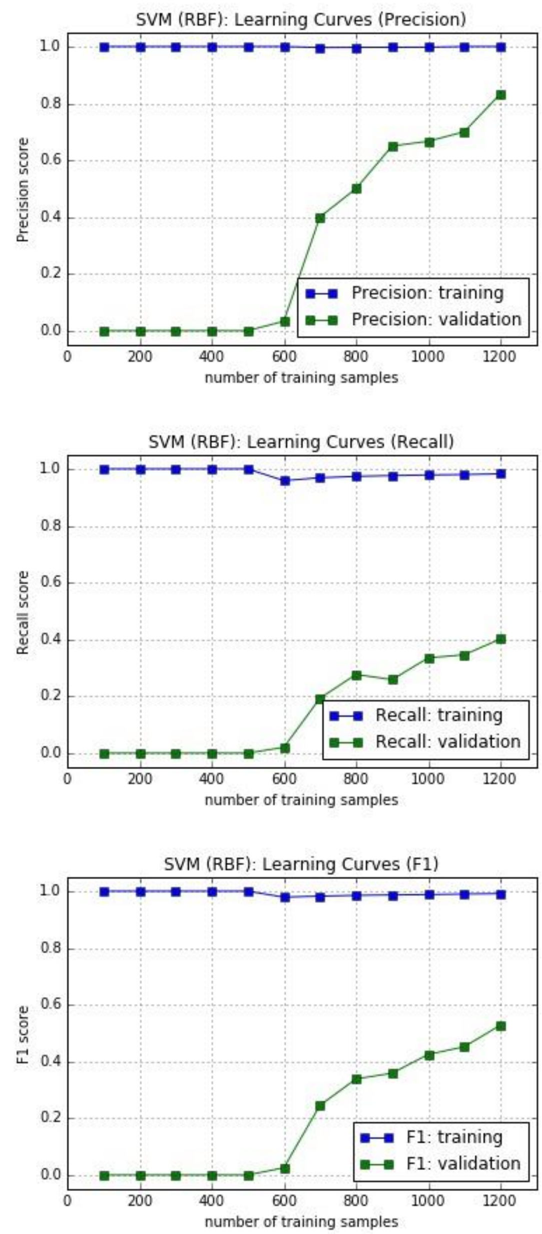


Figure 3.14 Performance of final selected machine learning models with testing data

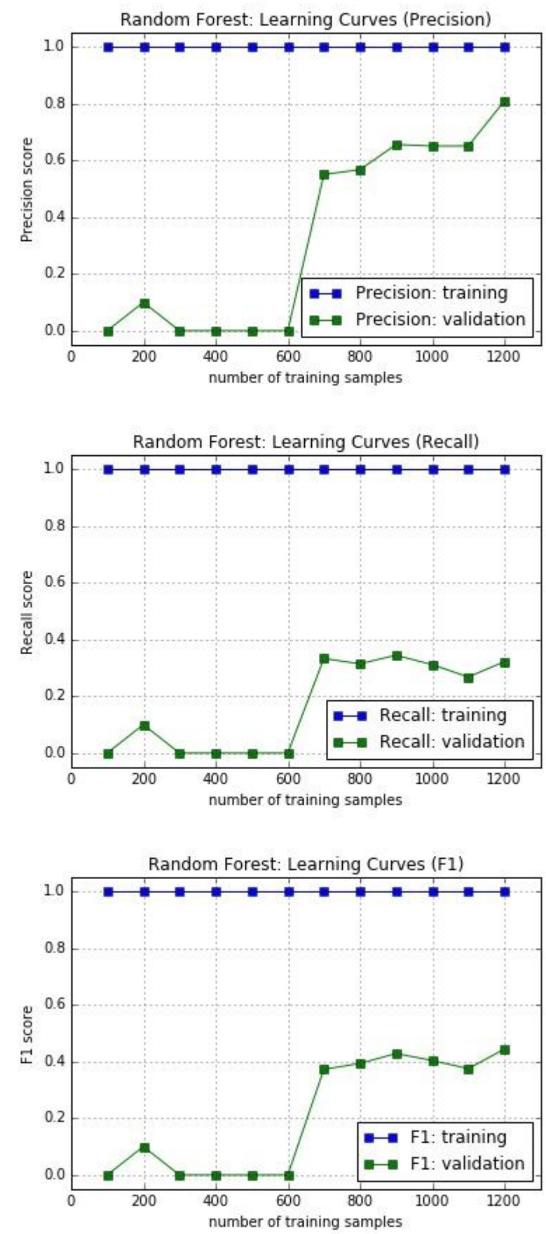
In addition, the significant difference between SVM with RBF kennel and SVM with polynomial kernel indicated that the relationship between fatal crashes and geometric design, traffic, and pavement condition is extremely complicated. The results show that the fourth order of interactions between these features is not capable of providing a reliable prediction of fatal crashes.

Learning Curves for the Final Selected Classifiers

The learning curves for the final selected models are shown in Figure 3.15. For both SVM with RBF kernel and RF, the learning curves show that there is still potential to improve the models' performance by acquiring additional data. The number of observations is considered sufficient when the training and validation curves converge to each other, that is, when the gap between the training and validation curves tends to remain constant as more observations are fed into the model. In addition, in terms of precision, highway agencies require at least 900 and 1,200 observations to train these models to reach precision values of approximately 0.6 and 0.8, respectively.



(a) Learning curve for SVM with RBF kernel



(b) Learning curve for RF

Figure 3.15 Learning curves for final selected machine learning models

3.4.3.9 Implementation of Econometric Regression Models

The preliminary diagnostic analysis of the data (Figure 3.7) suggested that very few road segments had fatal crash counts greater than zero. The maximum count of fatal crashes for any segment was seven. These characteristics, from the perspective of model specification, suggested two actions. First, the Poisson model form, instead of the negative binomial form, should be used for modeling because both the expected crash frequency and its variance were close to 0 and, more importantly, were equal, which was confirmed by the insignificant overdispersion parameter for these data. Second, due to the preponderance of road segments with zero fatal crashes, a ZIP model was preferred to model fatal crashes, which was confirmed by the Vuong statistic (Vuong, 1989).) In addition, a logistic regression function also was estimated for this analysis to evaluate the safety level of highway segments for comparison purposes. Table 3.13 and Figure 3.16 present the estimation results of the econometric models and the corresponding performance of the models with the testing data, respectively.

Table 3.13 Estimation results of econometric models

Model type	**Variables**	**Coef.**	**t-stat**	**P-value**
ZIP	Constant	–0.981	–7.72	<0.001
	*Segment length (miles)	–0.490	–5.06	<0.001
	LN(Average annual daily traffic)	0.978	7.74	<0.001
	Median width (ft.)	–0.009	–2.29	0.022
	Left side shoulder width (ft.)	–0.080	–1.98	0.047
	$\rho^2 = 0.153$			
	Vuong stat = 4.08			
Logistic Regression	Constant	-7.588	-9.89	<0.001
	Functional class SR (1 if State road, 0 otherwise)	-1.594	-2.91	0.004
	IRI (in/mile)	0.006	1.97	0.049
	Segment length (miles)	0.254	6.33	<0.001
	Lane width (ft.)	0.466	2.90	0.004
	Average annual daily traffic	0.000028	3.15	0.002
	Left side shoulder width (ft.)	-0.1529	-1.86	0.063
	Average vertical curve grade	0.2449	3.44	0.001
	$\rho^2 = 0.311$			

Variables are statistically significant at 90% degree of confidence.
**Variable estimated for determining zero-state probability*

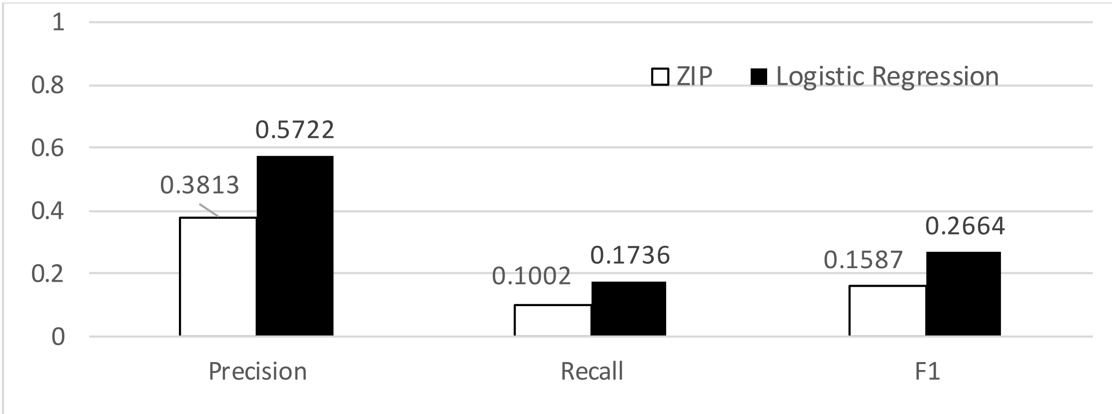


Figure 3.16 Performance of econometric models with testing data

3.4.4 Summary of Results for the Machine Learning and Regression Models

In selecting between machine learning and econometric regression models, the trade-off between prediction and interpretability should be considered. Of the many typically applied econometric and machine learning methods, some are less flexible than others. For example, linear regression is a relatively inflexible approach because it can only generate linear functions such as lines in two-dimensional space or linear hyperplanes in high-dimensional space. Other methods, such as RFs, are considerably more flexible because they can generate a much more complicated discrete function. If the main interest is statistical inference, then inflexible models are much more interpretable. For example, when inference is the goal, linear models may be preferred because it is quite easy to understand the relationship between the dependent variable and the independent variables through estimated coefficients, marginal effects, or elasticities. In contrast, flexible approaches, such as RFs, generate a complicated function that is relatively difficult to interpret in terms of the effects of any independent variable on the dependent variable.

Figure 3.17 juxtaposes the contents of Figure 3.14 and Figure 3.16 to compare the performance of the machine learning algorithms and econometric models. Econometric regression models, by definition, compute the average or overall response values (Jiang and Sinha, 1989), which leads to relatively low performance in predicting individual cases compared to machine learning models. As expected, machine learning algorithms outperform econometric models by a large margin in predicting fatal highway segments.

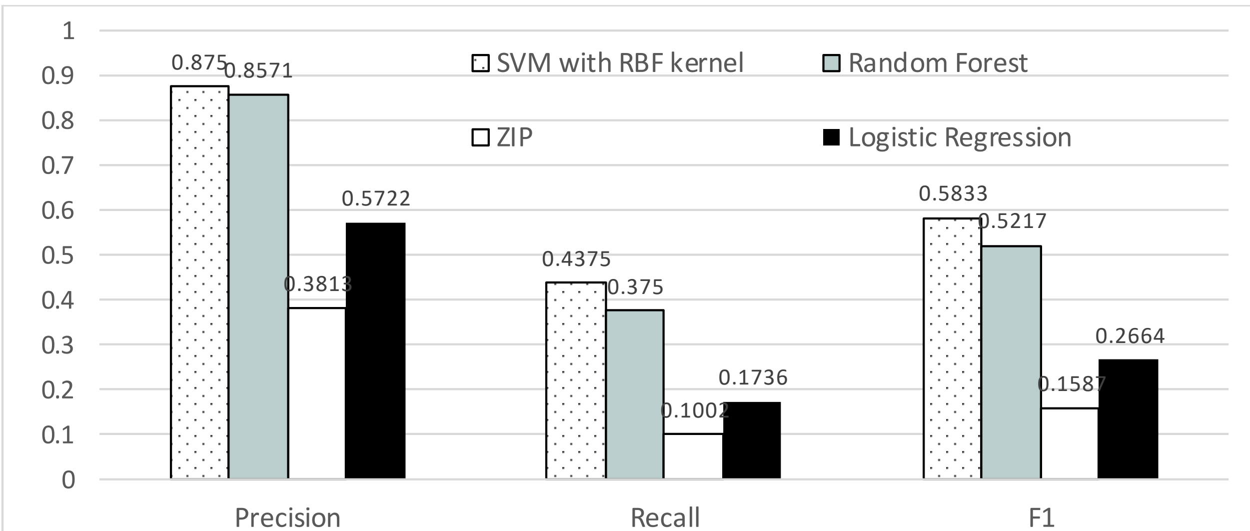


Figure 3.17 Comparison of performance of machine learning and econometric models with testing data

CHAPTER 4. INSIGHTS INTO EMERGING OPERATING ENVIRONMENT - HIGHWAY SAFETY IN AN ERA OF AUTONOMOUS VEHICLE

4.1 Introduction

The previous chapters of this dissertation addressed various aspects related to enhanced prediction of highway safety, including a proposed framework for enhanced prediction of highway safety. By focusing on crash factors that are directly under highway agency's control (i.e., roadway engineering factors), this dissertation demonstrated the application of the framework in providing feedback from the preservation phase (safety impacts of road-surface condition), design phase (optimal space allocation across highway cross-sectional elements), and operations phase (reliable prediction of fatal crashes).

Notwithstanding the noble intentions of "forgiving" designs, humans, for ages, have been the cause of most highway crashes. As far back as the 1970s, it was found that humans were a probable cause of 92% of accidents (Treat et al., 1979). In the 1980s, it was reported that human actions were a sole or contributory factor in 90% to 95% of all accidents (Rumar, 1985). In the 1990s, studies reported that 45% to 75% of all crashes were caused by human (driver) error (Aberg and Rimmo, 1998; Hankey et al, 1999). At the turn of the millennium, researchers continued to find that driver error dominated the crash causes (Wierwille et al., 2002). Salmon et al. (2005) found that driver error contributed to as much as 75% of all roadway crashes. As shown in Figure 4.1 (FHWA, 1995), the distribution of road accident causes in the U.S. was heavily skewed toward the driver. If this chart is representative of the current distribution of crash causes, then only 7% of crashes are currently not attributed to the driver. The underlying causes of driver-attributed crashes include driver age extremes, state of sobriety, fatigue, inexperience, or distractedness, such as texting while driving. Another possible cause of the predominance of driver errors is that driving is vision-based. In other words, although other mechanisms of information delivery are relevant to drivers, the most critical information for the driving task comes from human vision (Gordon and Lindberg, 2015). Macadam (2003) stated that human driving is 90% visual. Because human driving is heavily vision-based, humans are unable to fully recognize the potential hazards in their driving environment.

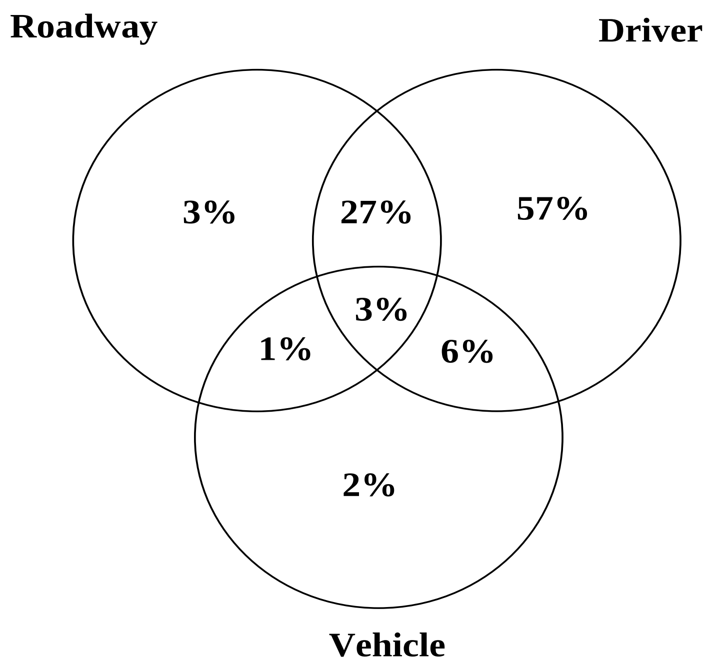


Figure 4.1 Distribution of road accident causes in the U.S. (FHWA, 1995).

At the current time, several nations are investing heavily in safety initiatives, including Vision Zero, a multi-national road traffic safety that seeks to eliminate all road traffic fatalities or serious injuries. However, safety-inspired initiatives and improvements in highway policy and design (such as wider lanes and shoulders) and vehicle technology advances (such as antilock brakes, airbags, and collision-avoidance systems) seem to have had limited effects due to the effect of human risk-compensation behaviors (Winston et al., 2006; Labi, 2016). Therefore, as long as human drivers operate vehicles, it is difficult to see how such safety initiatives can reach their targets.

4.2 The Design Safety Conundrum in the Current Highway Environment

In the traditional operating environment, highways are designed, maintained, and operated to serve human drivers, who are subject to errors due to distraction, impairment, fatigue, inexperience and other factors. The goal is to provide “forgiving highways” that mitigate the effects of bad driving.

4.3 The Emerging Era of AVs and Safety Implications

A driverless vehicle is one that is capable of navigating without a human driver. A driverless vehicle may be controlled remotely by an off-site operator with complete data on the driving environment and therefore makes the driving decisions for the vehicle. In this case, the vehicle is driverless but not autonomous. Examples of this include prospective systems for driverless taxis (Greenblatt and Saxena, 2015) and shared mobility systems where a central planner/dispatcher makes the vehicle's driving control decisions remotely (Bischoff and Maciejewski, 2016). Another class of driverless vehicles includes those that are autonomous because they are capable of self-control without external intervention. Such driverless vehicles are capable of (a) receiving stored data about the driving environment from a control office or server, interpreting the data to establish in real-time a 3D characterization of the driving environment, and making its own driving decisions; (b) monitoring the driving environment in its immediate vicinity via video, lidar, GPS, and/or odometry, interpreting the data in real-time, and making its own driving decisions without external input regarding the interpretation and driving decisions; or (c) a combination of (a) and (b). Therefore, based on the above definitions, it is clear that while AVs inherently possess driverless capability, not all driverless vehicles are autonomous.

In the U.S., the National Highway Traffic Safety Administration (NHTSA) and SAE established an official classification system for defining the different levels of autonomy (NHTSA, 2013a). At Level 0, the driver completely controls the vehicle at all times; at Level 1, the individual vehicle controls (e.g., electronic stability control and automatic braking) are automated. At Level 2, at least two controls can be automated in unison, such as adaptive cruise control in combination with lane keeping. At Level 3, in certain conditions, the driver can fully cede control of all safety-critical functions to the vehicle, and the vehicle can sense when conditions require the driver to retake control and provides a sufficiently comfortable transition time for such a transfer of control. At Level 4, the vehicle performs all the safety-critical functions for the entire trip, including parking, and the driver is not expected to control the vehicle at any time. At Level 5, all functions are undertaken by the AV with no possibility of driver input. It is clear that fully autonomous mobility, that every vehicle is at Level 5 autonomy, will not happen all at once. There will be a transition phase when roadways are expected to host a mix of vehicles at different levels of autonomy (Saeed et al., 2019b; 2019c).

Technological developments towards AVs and their levels of autonomy have proceeded at a rapid pace, much faster than the establishment and promulgation of operational regulations associated with AVs (Miller, 2018). At the current time, most newly manufactured vehicles incorporate some advanced features that are part of this trend: automatic anti-lock braking systems, electronic stability control systems, rear-movement collision avoidance and warning systems, and lane-guidance systems (Mikusova, 2017). It is speculated that over the next decade, driving functions will increasingly become more automated and that AV operations will be permitted for certain vehicle classes and at certain road classes (Saeed, 2019). According to a U.S. DOT report, fully automated vehicles could be commercially available by 2034 (USDOT, 2014).

The potential impacts of autonomous vehicles include changes in the need for (or dimensions of) certain types of existing infrastructure and the introduction of new physical or cyber infrastructure (Labi et al., 2015). Other impacts include the benefits of enhanced mobility, particularly for children, the elderly, and the disabled, and relieving travelers from driving and navigation chores and freeing up in-vehicle travel time for other activities. The Volpe Center and the Rand Corporation discussed the effect of AVs on demand, mobility, accessibility, congestion, land use, energy, and emissions, among others (Smith et al., 2015; Anderson et al., 2016). In addition, it has been forecast that AVs will significantly reduce the need for parking and will likely promote business models for shared transportation. Of the various prospective benefits of AVs, safety improvements have been touted as the most significant in terms of certainty and magnitude. The next section addresses the prospective safety impacts of AVs.

The safety benefits of automated vehicles are paramount. The potential for automated vehicles to save lives and reduce injuries is rooted in one critical and tragic fact: 94 percent of serious crashes are due to human error. Automated vehicles have the potential to remove the human error factor from the crash model, which will help protect drivers and passengers, as well as bicyclists and pedestrians.

AVs are expected to greatly reduce crashes, and their safety benefits are already being seen at the low levels of vehicle autonomy (levels 2 and 3). An NHTSA-sponsored research study determined that the electronic stability control feature already implemented in vehicles prevented over 2,200 fatalities between 2008 and 2010 (Srivinski, 2011). It is envisioned that the sensors and algorithms in driverless vehicles will continue to advance to a level where they can reliably handle anomalous roadway conditions and situations including inclement weather, work zones, and unusual or unexpected obstacles in the roadway. With well-tested safety technology, AVs can react much more quickly to events on the road because the sensors and algorithms are able to respond faster than human perception and reaction. In addition, sensors and algorithms, if programmed and tested properly, have a higher reliability of detection compared to human drivers as human drivers may not always pay attention to everything in the driving environment.

However, based on surveys conducted by OnePoll, people have safety concerns about AVs related to pedestrian safety (56%), passenger safety (51%), rise in accidents (49%), connectivity and backhaul network failure (35%) and cyber-attacks on personal information (29%) (Scroxton, 2019).

4.4 Safety Factors in the Emerging Operating Environment

In the AV operating environment, the traditional analysis of crash factors may become less and less important or even eventually have no impact, as the crash factors that significantly affect highway safety in the traditional operating environment may no longer have significance. For example, some of the human factors, such as driver's gender and age, are expected to have less impacts with higher penetration of fully AVs. With the adoption of advanced autonomous driving technologies, human error in driving may be eliminated completely in the future. As a result, the focus of highway design could be shifted from "forgiving highways" to other factors, including enhanced mobility, efficient land use, etc.. In addition, roadway design changes may be required to support AV operations. Therefore, the safety impacts of the engineering factors are expected to change as well, especially the new crash factors related to AV operations, algorithms, etc. This section specifies and discusses the crash factors and data that need to be collected for analyzing roadway crashes in the emerging AV operating environment. When the data are available in the future, analysis need to be carried out using the framework proposed in this dissertation for enhanced prediction of highway safety and therefore for quantifying the safety impacts of these crash factors.

4.4.1 Roadway Engineering Factors

In the pre-AV era, roadways are being designed as "forgiving highways" to promote safety by compensating for human driving errors. With the adoption of AVs, it is therefore expected that the safety impacts of several current roadway engineering factors will need to be evaluated to facilitate AV traffic operations, such as, lane and shoulder width, horizontal and vertical alignment, superelevation, and stopping sight distance.

- Lane and shoulder width: Since AVs are expected to have precise positioning capabilities, several studies indicate that the current roadway lane and shoulder widths could be reduced (Somers and Weeratunga 2015; Hayeri et al., 2015). As a result, additional lanes could be created for human driving vehicles (HDVs) as well as dedicated bus/truck lanes, which could lead to enhanced mobility. However, the safety impacts of reducing lane and shoulder widths need to be evaluated thoroughly when relevant data is available.
- Horizontal and vertical alignment, and superelevation: For AV control, AI algorithms, such as reinforcement learning, can help AVs perform at their optimal speed and wheel control for turning. This need creates opportunities for new designs for roadway alignment including horizontal and vertical alignment and superelevation, which may result in enhanced safety and mobility.
- Sight distance: Sight distance allows the driver to estimate the conditions in the environment and make appropriate responses. Sight distance requires the driver's awareness of their surroundings and driving properly. Usually, the design of sight distance is very conservatively calculated. For example, stopping distance, which is the sum of two distances (the distance traveled during perception plus reaction time and the distance to stop the vehicle) is designed with lower deceleration and slower perception reaction time than normally expected from the driver. As most AVs are expected to be connected and autonomous vehicles (CAVs), they are expected to drive safer than human in situations where the sight distance is limited with the help of connectivity (V2I, V2V, V2X).

4.4.2 Driver Characteristics

In the traditional environment, driver characteristics are considered impactful in roadway safety because crashes are influenced by driver age and gender, experience, alcohol, fatigue etc. (Islam and Mannering, 2006; Sinha and Labi, 2007). For example, Kweon and Kockleman (2003) found males (ages 18-55 years) were more likely to get involved in an accident than their female counterparts. For older age groups, however, males were less likely to have an accident than female drivers. In addition, drivers in the younger and older age groups tended to have higher crash rates in comparison. Moreover, intoxicated drivers were far more likely to have higher crash rates (crashes per VMT) than sober drivers. The authors found that approximately 1/3 of the roadway fatalities involved intoxicated drivers. As a result, the adoption of AVs is expected to have higher safety impacts for the groups that are more likely to have crashes (such as young and middle-aged males, older age females, and intoxicated drivers). Moreover, the connection between highway safety, particularly speeding, and the characteristics of the drivers has attracted a lot of research attention as personality traits are more definitive indicators of speeding. Several studies found that drivers with a "sensation seeking" tendency were more likely to exhibit risky behaviors such as speeding (e.g. Jonah et al., 2001, Greaves and Ellison, 2010). Other researchers who focused on the "classic" personality types, showed that people with Type-A personalities tended to speed (Tay et al., 2003). In turn, drivers who were altruistic and averse to risk were unlikely to speed (Machin and Sankey, 2008, Greaves and Ellison, 2010)

From the human factor perspective, operating AVs may involve a variety of human-machine interactions (HMIs), such as driver response to warning signs, take-over requests, and other driving assistance systems. Education of drivers will be required for operating AVs safely. Inadequate education may lead to unsafe driving as drivers may not be able to perform HMI tasks properly. From the technology perspective, AVs need to be "educated" properly through their AI systems as well. This "AV education" is necessary for AVs to understand the appropriate HMIs, driving environment, policy and regulation, vehicle control, etc. For example, it would be more efficient if AVs are educated to interact with drivers in human-like ways. In human communications, explanations and expectations are crucial in trust establishment. People find it difficult to trust an agent that acts unexpectedly and unpredictably. Researchers have found that AVs might be able to reduce this negative impact by giving more information or explanation regarding the specific actions they took (Yang et al., 2018). On the contrary, poor "AV education," may cause dangerous AV maneuvers on roadways. For instance, when people are talking in the vehicle, human conversation can be misunderstood by AVs as a signal when it contains certain "messaging" words. Moreover, voice recognition in AI might need to be trained for every user, which can be a good thing, but also may lead to false detection.

4.4.3 Policy

Policy, as a crash factor, refers to the laws and regulations associated with driving and includes sobriety laws that regulate the alcohol limit for truck drivers, seat belt and helmet laws, etc. In highway safety, one of the most controversial policy factors is the speed limit. Crash severities and rates can be affected by modifying speed limits or establishing speed differentials by vehicle class. These changes typically depend on multiple other factors such as highway functional class, severity type of crash, current speed limits, etc. In the era of AV, policies regarding speed limits may have different safety impacts compared to the traditional operating environment because AV speed will be controlled by AI algorithms instead of a human driver, which is expected to eliminate speeding issues. Due to improved control of the vehicle and its speed, speed limits could be increased for new roadway designs, especially for AV-dedicated lanes to increase safety and mobility. In addition, in the emerging operating environment with AVs, potential changes regarding the policy factors will provide opportunities for platooning of passenger cars and heavy-duty vehicles to increase roadway safety, save fuel, and add roadway capacity. With a higher market penetration, policy factors are expected to have a higher impact on roadway safety.

4.4.4 Traffic Heterogeneity

In the traditional environment, particularly at developing countries, the traffic is highly heterogeneous, consisting of vehicles of wide-ranging static and dynamic characteristics. These vehicles include motorcycles, cars (jeeps and small vans), large passenger vans, and small four-wheeled goods vehicles, buses, trucks, and bicycles. In addition, sometimes, a small percentage of animal-drawn vehicles also can be seen on the roads. All these different types of vehicles share the same road space. Past researchers have sought to optimize the utilization of available infrastructure in such a way that the carrying capacity (number of persons transported) of the roadway is enhanced. Cox (1975) recommended the use of exclusive bus lanes and found that the assignment of special lanes to buses did not adversely affect the level of service of the vehicular traffic with decreasing travel time and stops and increasing speed.

With the development of AVs, the future vehicular traffic, including human-driving and self-driving vehicles, would greatly enhance highway capacity. Autonomous vehicles offer a wide variety of potential benefits. One commonly-discussed benefit is improving traffic operations (decreased congestion, decreased delay, and improved efficiency). AVs are expected to perform differently from human-driven vehicles in many aspects, such as operating with smaller headways, shorter reaction times, and higher speeds than human-driven vehicles (Bierstedt et al., 2014). Researchers believe that with enough cooperation and penetration, vehicle automation could enhance traffic flow in the future (Vander Werf et al., 2002; Shladover, 2009).

Nevertheless, there is still a slow transition towards automated systems. It is likely that many human-driven and autonomous vehicles will appear at the same time on the roadways. As most AV operating technologies and regulations are still in development, there is limited publicly available empirical data on how they will behave in heterogeneous traffic flow. The impacts of partial vehicle automation and driver assistance systems have not been completely elucidated in practice (Milanés et al., 2014). These uncertainties consist of the efficiency and safety trade-off (Hoogendoorn et al., 2014), the interactions between automated and non-automated vehicles (Calvert et al., 2016), traffic flow stability (Kesting and Treiber, 2008), and many human factors (Jamson et al., 2013).

Based on the literature, self-driving vehicles will see a significant increase in the vehicle fleet share by 2020 but may not be greater than 25% of all vehicles by 2030 (Calvert et al., 2017). Some studies proposed that the relatively low-level AVs would have limited influence compared to high share levels. Other researchers, who were focused on the effect of automatic driving vehicles on traffic flow, found that in the heterogeneous free flow, AVs indeed could reduce the probability of traffic breakdown even with low penetration (10%-20%). At the same time, AVs could increase the threshold flow rate for spontaneous traffic breakdowns as well as the maximum and minimum capacities of free flow at the bottleneck (Kerner, 2016).

4.4.5 Vehicle-Related Factors

Vehicle Mechanical Perspective:

Vehicle design features (such as size, weight, and shape) affect crash severity and crash frequency (Torrão, et al., 2012; Everett et al., 2014). In addition, fatal crashes occurred at a higher rate in small passenger cars than in heavier or larger vehicles. Recently-developed cars with new design features and safety equipment could provide better protection in crashes compared to older models, thus tending to reduce crash severities and frequencies (Sinha and Labi, 2007). Unfortunately, drivers in these newer cars seem to drive more aggressively, thus canceling out the designed benefits of safety features (Winston et al., 2006). For occupants, transit vehicles, such as buses, are likely to have low crash frequency per mile as well as low injury rates. Large vans and sport utility vehicles were reported to have a high rate of rollover crashes. Pedestrians, bicyclists, and motorcyclists are more likely to be injured in a crash (Sinha and Labi, 2007).

For AVs, AI algorithms can interact with the vehicle's mechanics. For example, besides designing a controller that drives an AV, the design process also applies algorithms to advanced driver-assistance systems (Giarratana, 2016). As a result, the controller is made of individual controlling modules aimed to serve specific tasks (e.g., a feedforward steering controller to track the vehicle's path accurately and stably). Moreover, these modules can be used independently as well. For example, using a steering feedback module alone could provide assistance when the driver is losing control (NVIDIA, 2018). The features of individual modules and how they advance the robustness of the entire AV system is being studied.

Vehicle Cyber perspective:

Several-recently developed motor vehicles have advanced features that could prevent drifting into adjacent lanes or making unsafe lane changes, brake automatically when the front car stops or slows suddenly, and warn other vehicles behind them when backing up (Urmson et al., 2008). These safety systems combine hardware (sensors, cameras, and radar) and software technologies to help recognize safety risks and warn the driver to prevent accidents (LeBlanc et al.,1996). The goal of future automotive technology is to deliver automated driving systems which can handle all driving tasks.

Traditional vehicle manufacturers like Audi, Ford, Mercedes, and Nissan, and new tech companies including Google, Tesla and Uber, are currently in a race to produce the first fully autonomous self-driving vehicle (Greenblatt, 2016). In general, the methods they use are similar. Essentially, an AV should be able to perform three actions to take over from a human driver: to perceive, to think, and to act. These actions are realized by the assistance of high-tech devices including computers, cameras, and controllers (Birdsall, 2014).

1) Sensor technologies

The perception (or sight) of an AV includes several parts that have their relative advantages and disadvantages (Menke, 2017). Cameras are placed on the roof and/or other spots to surround the vehicle for 360º vision. Although cameras have enough resolution and can recognize colors, they cannot detect the range or distance to the objects. Moreover, their performance is influenced by lighting and unpleasant weather (Lee et al., 2013). Light Detection and Ranging (LiDAR) devices with a constantly rotating laser beam, are much better at measuring the position and shape of items in a round view, but they are costly, sensitive to snow or rain, and could be affected by other self-driving vehicles' laser sources (Cho et al., 2014; Hecht, 2018). Besides cameras and LiDAR equipment, driverless vehicles often are equipped with traditional radars, which are useful in fog, snow, and rain and can determine the relative speeds of other cars or moving objects accurately. But they are not helpful at finding the shape and size of the object and can only see very fuzzily, and thus are not very useful in identifying the object (Jenn, 2019). Radars also are not efficient in lane detection. Because of these two major limitations, radars are not used as a primary sensor for self-driving cars but rather serve as a backup in adverse weather (Patole et al., 2017).

The inconsistency between human driving and what the machine is programmed to do will lead to conflicts between human drivers and AVs during the transition period (Montgomery et al., 2018.). While human driving overwhelmingly is based on eyesight, vision is not useful for self-driving vehicles, which are informed by cameras, lidar, radar, and other sensors (Anderson et al., 2014). These sensors determine other patterns and frequencies of emission and so form a perceptual "world" that is different from human perception. These sensors are integrated together such that the AVs observe (i.e., gather information in pattern-recognition assemblages) the environment in a very distinct way than humans do (Hancock et al., 2019). The difference associated with human and machine abilities and the related variability of attribution, indicates AVs and human drivers are far away from integrating completely, which may bring safety issues (Hancock et al., 2019). For example, it is difficult for a human to differentiate whether vehicles are human-driven or self-driving. Then, when passing on multilane highways, this uncertainty can result in dangerous attribution errors. For example, a human driver could wrongly assume that an overtaken car has advanced 360º machine vision whereas it is driven by another driver who in fact has blind spots (Vanderbilt, 2017).

2) AI Algorithms

A significant area in self-driving development and testing is artificial intelligence (AI). Deep learning, which mimics neuron activity, is believed to be the most important technology behind autonomous driving AI (Pan et al., 2017). Deep learning is helpful in voice search, speech and voice recognition, image recognition and processing, movement detection, and data analysis. As a result, AVs can detect and recognize pedestrian traffic, other cars, and traffic signals and can follow mapped-out routes (Chen et al., 2017). For safe operation, AVs require very advanced computers to deal with all the sensor-collected data in real time (Wu, 2018). For example, some intelligent and efficient algorithms that are trained to recognize items in the environment have been developed that can differentiate motorcycles versus bicycles as well as other objects, such as traffic lights, vehicles, pedestrians, and obstacles. After identifying which objects surround the car and how far away they are, the computers also determine how to react and then executes the decision (Lu et al., 2010). For example, when encountering a closed lane due to road construction, the computer would figure out the options: it can either immediately switch lanes or slow down first to let the other vehicles in the next lane pass. Here, different AI systems may have different decisions to make, which may trigger some safety issues.

Before any AV can safely navigate on the roadway, engineers must first test and validate the AI algorithms and other software that enable the vehicle to drive itself (Yan et al., 2016). AV systems must be subjected to huge numbers of scenarios to gain confidence in their safety. Many challenges of how AI can teach road safety to AVs will arise around edge case scenarios, where rarely-occurring, interdependent and ultra-specific sets of circumstances that are hard to predict come into play (Litman, 2017). AI-powered AVs must be able to respond properly to the incredibly diverse situations they could experience, such as emergency vehicles, pedestrians, animals, and a virtually infinite number of other obstacles (Ramos et al., 2017).

3) Security of AV system

Quick and stable connectivity between self-driving vehicles and the outside environment, such as cloud infrastructure, will allow more efficient signal transmission. The cutting-edge 5G wireless technology, which ensures high speed connections and data exchange, is expected to improve connectivity among vehicles and provide a wide range of services (Cheng et al., 2017; Chen et al., 2018). There are a few protocols that will regulate how AVs connect to their surroundings. The comprehensive term is Vehicle-to-Everything (V2X) (Abou-zeid), which involves a) Vehicle-to-Infrastructure (V2I) communication, which is a wireless exchange of information with the surrounding road infrastructure to operate within the bounds of the speed limits, traffic lights, and signage and also helps manage fuel consumption and prevent crashes; and b) Vehicle-to-Vehicle (V2V) communication, which allows operations safety within traffic situations and helps prevent collisions (Amoozadeh et al., 2015).

4.4.6 Natural Environment Factors

Environmental conditions, such as low visibility, rain, or snow, strong winds, icy roadway, crossing animals, etc. are important factors of roadway collisions (Sinha and Labi, 2007; Andrey, 2010).

In the emerging operating environment, AVs must perform regardless of the weather, road, or lighting conditions (Sivak and Schoettle, 2015). There is no feasible way to physically road test vehicles in all these situations nor is road testing sufficiently controllable, repeatable, exhaustive, or fast enough. Testing self-driving vehicles safely in a realistic simulation environment is important. Combining simulated miles with actual road miles in the data center is the key to testing and validating AVs (Grazioli et al., 2017).

4.4.7 Enforcement Factors

Patrol frequency, driver education, and license restrictions improve safety in general (Williams, 2007). Generally, severe penalties for traffic infringement are likely to encourage operator responsibility and thus improve traffic safety as well (Ivers et al., 2009). In the emerging operating environment, AVs need to be able to properly interact with patrols to ensure operation safety. Malfunctioning AI or control systems that are not able to recognize patrol and response accordingly may decrease roadway safety and cause severe highway crashes (Reimer, 2014; Yang and Coughlin, 2014).

4.4.8 AV Decision-Making Ethics as a Crash Factor

For the past two decades, AV has demonstrated its success statistically by traveling thousands of miles with low crashes rates on real roads (Kalra and Paddock, 2016). Paradoxically, the weakness of automated systems is the same as their strength: they involve no tactical human decision-making. This may lead to controversial ethical issues, such as should AV prioritize protecting its own passengers or a pedestrian in an unavoidable crash? Unfortunately, humans probably have no clue on how AV would react to such novel situations. Humans have empathy for understanding how other people behave when facing boundary scenarios by imagining themselves in certain situations and how they would respond. AVs at present, however, have no such foresight. At the current time, it is difficult to forecast how AVs would react to these situations, thus giving rise to a social science challenge (Goodall, 2014).

4.5 AV Testing and Related Safety Problems

As with all new technologies, it is important to carry out field-testing of AVs to ensure that they do not pose undue safety hazards to the public and that they produce the prospective benefits. Such a testing period will facilitate the training of AVs to operate in non-deterministic driving environments, including during the transition phase where both AVs and traditional vehicles share the road. AVs will need to learn to accommodate the capricious and often unpredictable habits of human drivers of traditional vehicles. AVs have been tested mostly at dedicated test facilities largely devoid of human-driven vehicles.

It can be argued that the true capabilities of AVs can be best determined on in-service roads and not on test tracks, as the latter often fail to adequately replicate the driving environment. For this reason, government policy and legislation in many countries has supported (or at least permitted) the testing and use of driverless vehicles on public roads. The continued and expanded permitting of AV testing on public roads is a welcome development because it will facilitate reliable training of AVs in varied and realistic driving environments. Unfortunately, there have been a few AV crashes at some of these public testing locations. While AV crash rates are much lower compared to traditional vehicles, the vast media attention they garner may be negative and seems to have damaged some public trust in AVs (Knight, 2018).

It has been established that it is dangerous to test self-driving cars on actual roads, particularly at night, and these vehicles need enhanced warning systems that quickly detect, interpret, and react to potentially hazardous situations. All the reported instances of AV crashes had humans at the wheel but that was not enough to prevent the crashes. Clearly, having a human driver at the wheel is no guarantee of safety. These high-profile crashes have led to a need for AV testing that does not pose a safety hazard to the public. Such testing can be carried out in a safe environment using AV test tracks or driving simulation. Unfortunately, as discussed above, test tracks are limited in what they can replicate as far as real driving environments due to cost constraints. The most promising recourse, therefore, is computer simulation of AV driving. To address the limitations of testing AVs on in-service road and test tracks, Chapter 5 proposes and demonstrates a deep learning-based simulation framework for autonomous driving.

CHAPTER 5. A DEEP LEARNING BASED SIMULATION FRAMEWORK FOR AUTONOMOUS DRIVING

5.1 Introduction

AV stakeholders continue to seek assurance of the safety performance of this new technology through AV testing on in-service roads, AV-dedicated road networks, and AV test tracks. However, recent AV-related fatalities on in-service roads may have exacerbated public skepticism and eroded some public trust in the safety of AV operations. Further, test tracks are unable to characterize adequately the real-world driving environment. For this reason, driving simulators continue to serve as an attractive means of AV testing. However, in most AV driving simulators, the AV operation is based on commands external to the vehicle and embedded in the code for the driving environment. To address the simulation shortfalls associated with this approach, this part of the dissertation develops a deep Convolutional Neural Network – Long Short-Term Memory (CNN-LSTM) algorithm for self-driving simulation. This algorithm observes and characterizes the AV's driving environment and controls the AV movement in the driving simulation. The CNN part extracts the features that use transfer learning to introduce human prior knowledge, and the LSTM part uses temporal information to process the extracted features and incorporates temporal dynamics to predict driving decisions. AV also may use an external server with a database containing road environment data as an additional source of information. It is acknowledged that different driving simulators differ in their functions and their capabilities to access driving-environment data. Therefore, to make it sufficiently flexible to facilitate replication by other researchers who use driving simulators, the algorithm has been designed and demonstrated using only image data of the driving environment as input because roadway image data are easily and readily accessible from the screen of any driving simulator. The proposed algorithm was tested using The Open Racing Car Simulator (TORCS) test track platform and was found to be able to mimic human driving decisions with a high degree of accuracy.

The objective of this part of the dissertation was to develop, for self-driving simulation purposes, an algorithm that directly observes and characterizes the AV's driving environment and controls the AV movement in the driving simulation. First, the various ways for classifying vehicle simulation are presented. Next, the methodology is explained, including the self-driving computational model that involves transfer of prior knowledge and the incorporation of temporal information. Then, the implementation of the computational model is presented, including the collection and balancing of the training data, the results, and how the developed algorithm was validated. Finally, its contribution in the context of AV simulation and avenues for possible future work in this domain are discussed.

5.2 Classes of Computer Simulation

Simulation refers to the abstraction of a system so that it is represented by a replica that is more amenable to study (Labi, 2014). Simulation is particularly useful when the system is too large to be studied; it is too expensive or impractical to create scenarios; the real-life system is inaccessible; it is dangerous, unethical, or unacceptable to use the real-life system; or the real-life system does not yet exist (Sokolowski and Banks, 2009). In such situations, simulation is useful for acquiring insights and describing or predicting behavior over time when the real-life system is subjected to different internal or external conditions or different courses of action.

Simulation is used in a wide range of disciplines including the military, where it is too dangerous or costly for trainees to interact with the real world (Aldrich, 2004). It is used in engineering, medicine, and the social sciences for the tasks of testing, training and education, monitoring and optimizing the performance of systems, and predicting future system failure or diagnosing the reasons for failure (Hartman, 1996; Adeli and Jiang, 2008; Adeli and Kim, 2009). For example, in driving simulation, researchers study the naturalistic behaviors of drivers by providing them with a lifelike experience of the driving environment with features related to the road width, alignment, friction, traffic conditions, and weather conditions (Kandhai et al., 2011). Driving simulators mimic the internal and external conditions of a real vehicle in a virtual driving environment as the user (driver) undergoes a realistic experience of sitting in a seemingly real vehicle and driving along a seemingly real roadway.

In the next three subsections, a number of simulation classifications are described as a prelude to a discussion of the context in which this part of the dissertation makes a contribution to existing knowledge or practice.

5.2.1 Simulation Classification by Simulation Player and Environment

The Modeling and Simulation Coordination Office of the U.S. Department of Defense (DOD) categorizes training simulation programs by whether the players or the environment are being simulated (Table 5.1). In Live simulation (RVPE I), both the players and the environment are real as is the case for AVs operating on test tracks or in-service roads. In Virtual Type 1 simulation (RVPE II), the players are real while the environment is simulated; in Virtual Type 2 simulation (RVPE III), the players are simulated but the environment is real; and in Constructive simulation (RVPE IV), both the players and the environment are simulated. Most driving simulators belong to the constructive RVPE category of simulation, where both the player (vehicle) and the (driving) environment are simulated.

Table 5.1 Real-virtual player-environment (RVPE) or DOD classification of simulation (Labi, 2014)

Category of Training Simulation	Simulation Entity	
	Players	**Environment**
Live (RVPE I)	Real	Real
Virtual Type 1 (RVPE II)	Real	Simulated
Virtual Type 2 (RVPE III)	Simulated	Real
Constructive (RVPE IV)	Simulated	Simulated

5.2.2 Simulation Classification by Source of Information and Source of Vehicle Control (SIVC)

AV simulation also can be categorized based on the source of the information they use to characterize their driving environment and the source of the vehicle control (Table 5.2). The driving environment refers to the road geometry and other roadway characteristics and the locations, dimensions, and other features of the roadside and traffic. These features include lane widths, curves, shoulder locations, traffic signs, traffic signals, and lane markings, as well as visibility, other vehicles, bicycles, and pedestrians.

Table 5.2 Simulation classification by source of information and source of the AV control

		Source of information	
		External server with database on road environment	AV’s sensors (video, lidar, GPS)
Source of the AV control	Entity external to the AV*1	SIVC I	SIVC II
	The AV*2	SIVC III	SIVC IV

**1 Driverless but not autonomous. *2 Autonomous*

The sources of information about the driving environment may include:

- A central office or server that provides the AV with stored and/or real-time information on the driving environment.
- Sensing technology (camera, lidar, sensors, laser-based technology, etc.) that is used by the AV to characterize its immediate driving environment in real time.

The sources of vehicle control may include:

- A person or computer in a central office that controls the AV’s driving. This is the situation that is expected for shared autonomous taxis (a concept that is still hypothetical) where the environment is completely connected in terms of vehicle-to-everything (V2X) communication.
- The AV controls the driving by itself, through interpretation of data it receives from the central office or the AV’s onboard sensors using AI tools that include machine learning, deep learning, fuzzy control, etc.

This part of the dissertation posits that the source of vehicle control, as explained above, is the basis upon which a simulation of autonomous driving can be considered a more representative simulation or otherwise. Therefore, a more representative simulation of autonomous driving is defined herein as a simulation where the source of vehicle control is the AV itself (i.e., SIVC III and IV). In contrast, in simulation where the source of vehicle control is some entity external to the vehicle (i.e., SIVC I and II) as can be expected in the case of driverless taxis (Bischoff and Maciejewski, 2016; Greenblatt and Saxena, 2015), then the simulation cannot really be considered an adequate simulation of AV driving as defined in this dissertation.

In other words, simulation of autonomous driving is only a good representative of actual AV driving when it mimics real-life autonomous driving where the driving task is carried out by real-time on-site collection and interpretation of data on the immediate driving environment. For example, from the control perspective, a real-life autonomous vehicle, such as Tesla's partial autopilot system (Dikmen and Burns, 2018), can be categorized as SIVC IV, or possibly, a combination of SIVC III and IV.

Regarding the source of information, most AVs use their sensors to obtain information regarding the driving environment (e.g., the distance to roadside infrastructure, traffic signals, lane markings, other vehicles, pedestrians, etc.). In the case of connected vehicles, the information regarding the driving environment is provided fully or partially (SIVC III in Table 5.2) through vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), or vehicle-to-everything (V2X) communication. Regarding the source of vehicle control, a more representative autonomous driving (not route planning) is realized by the AV itself through its AI system.

In contrast, the autonomous driving systems that are represented in the driving simulators at most universities and research institutions may not adequately represent simulation of autonomous driving as defined in this dissertation. Such simulations are similar to video games, where both the players and their environment are programmed together in the same computer code; in these simulations, the AV does not need to obtain driving-environment information through its sensors because it obtains all such information directly through the same computer code. In such cases, the simulated autonomous driving is realized not by the AV itself (through its AI) but by entities external to the vehicle, such as preprogrammed rules (e.g., pre-specified or pre-randomized paths) or pre-defined functions (e.g., equations that take into account the stored data in the computer program regarding the driving environment, such as lane widths, number of lanes, distance to nearest pedestrian, traffic signal, etc.).

The concept of a more representative simulation of autonomous driving is indeed essential. As discussed previously, one of the benefits of simulating a system is that the simulation does not pose a safety hazard to the public. Through driving simulations, researchers can investigate the AV user's reaction time to hazards and warnings, user behavior, etc., in a safe setting that does not cause any risk to users. However, the current autonomous driving simulators may not be adequately mimicking the autonomy of real-life AVs and may not be yielding outcomes that adequately predict the expected impacts of prospective real-life AV operations.

In summary, the DOD or RVPE simulation classification suggests that, currently, most driving simulators belong to the constructive category of simulation, where both the player (vehicle) and the driving environment are simulated. From another perspective (and therefore, within each of the DOD classification categories), AV simulations can be further categorized based on the source of information they use to characterize their environment (in-vehicle vs. external or AV sensors) and the source of vehicle control (an external entity or the AV itself). Further, most driving simulators belong to the category where the source of vehicle control is some entity external to the vehicle rather than the AV itself. Such simulations cannot be considered as adequately representing simulation of autonomy.

5.2.3 Simulation Classification by Level of Driving Decision (LODD)

It is also possible to categorize the simulation of driving operations by the level of the driving decisions that the AV algorithms are permitted or designed to make:

- Strategic: high-level driving decisions/tasks, such as route choice (LODD I)
- Tactical: mid-level driving decisions/tasks, such as whether to overtake (LODD II)
- Operational: low-level real-time driving decisions/tasks (LODD III)

Classifying simulation by the level of driving decision is an important consideration in categorizing next-generation vehicles. For example, regarding shared transportation, a central office makes strategic decisions or assigns tasks (route choices), while operational tasks (e.g., accelerating, stopping, braking to avoid a pedestrian) are made by the vehicle. In the coming era of vehicles that are autonomous, connected, and shared, simulations of the operations of these vehicles will need to fully recognize and duly account for the level of decision-making assigned to the vehicle and to other entities.

The next section presents an AV driving simulation computational model that is based on the constructive category of simulation and the category where the source of control for the AV is the vehicle itself, uses sensors to characterize its driving environment, and undertakes operational tasks and decisions (similar to level 2-3 of autonomy).

5.3 Methodology

5.3.1 A Proposed Self-driving Computational Model

The aim of the proposed self-driving computational model is to realize the SIVC IV and LODD III class of simulation for AV operational decisions, where the AV uses sensors (video, lidar, GPS) as the source of information and controls itself through its AI algorithm. This class of simulation is the most similar to a real-life scenario, and, in this dissertation, is realized by applying the proposed computational model on simulated real-life driving data from The Open Racing Car Simulator (TORCS) (Wymann, 2015). A successful computational model for AV simulation is one that can drive the host car in a way that mimics closely manually-operated human driving behavior within a simulated environment, such as TORCS. TORCS was chosen in this dissertation to demonstrate the framework because it is an open source software and therefore is readily available to researchers seeking to replicate this dissertation's methodology. Researchers may choose software and games such as Aimsun, PreScan, etc. depending on their needs. For example, Martinez et al., 2017 developed and tested 3D-CNN for autonomous driving in Grand Theft Auto V, one of the most popular commercially available games.

The motivation for focusing on building the self-driving computational model for SIVC IV simulation instead of a combination of SIVC III and IV was that it was sought to make the proposed computational model flexible enough to be replicated by universities and research institutes that own driving simulators. This goal is important because while different driving simulators have different capabilities and functions and differ in their ability to access information regarding the driving environment, they all have the same ability to generate images of the driving environment on their driving simulator screens.

Figure 5.1 illustrates the proposed computational model. The input is a real-time frame simulated by TORCS. This input is fed into and processed by the deep CNN-LSTM to generate the corresponding output for the frame. The output refers to the driving decisions regarding the use of the accelerator, brake, and wheel controls for this input frame. The decisions are used to drive the host car in the TORCS environment. The entire procedure then moves to the next frame and is repeated. This is a real-time process. In this dissertation, 10 frames were simulated and processed by the computational model per second. Table 5.3 presents the nine classes of driving decisions used in this dissertation.

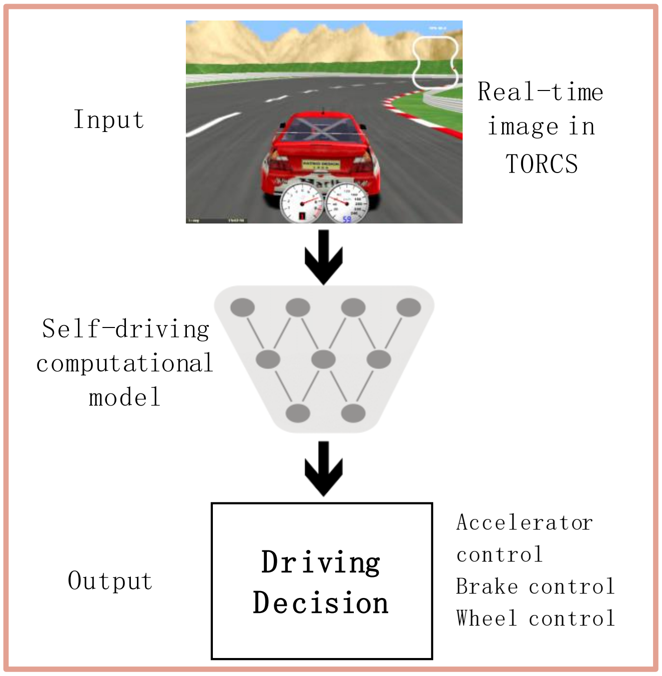


Figure 5.1 General self-driving computational model for the AV

Table 5.3 Classes of driving decisions

Driving decision class	Output labels
Accelerate	1
Brake	2
Turn Left	3
Turn Right	4
Turn Left + Accelerate	5
Turn Left + Brake	6
Turn Right + Accelerate	7
Turn Right + Brake	8
No wheel, accelerator and brake control	9

5.3.2 Deep CNN-LSTM Algorithm

This section explains the proposed deep CNN-LSTM algorithm in detail. The main components in the proposed algorithm include data processing, a convolutional neural network (CNN) for feature extraction, a (LSTM network (Hochreiter et al., 1997) that processes the extracted features and incorporates temporal dynamics to predict driving decisions, and a fully connected classification layer. Figure 5.2 shows the architecture of the proposed deep CNN-LSTM algorithm used in the computational model.

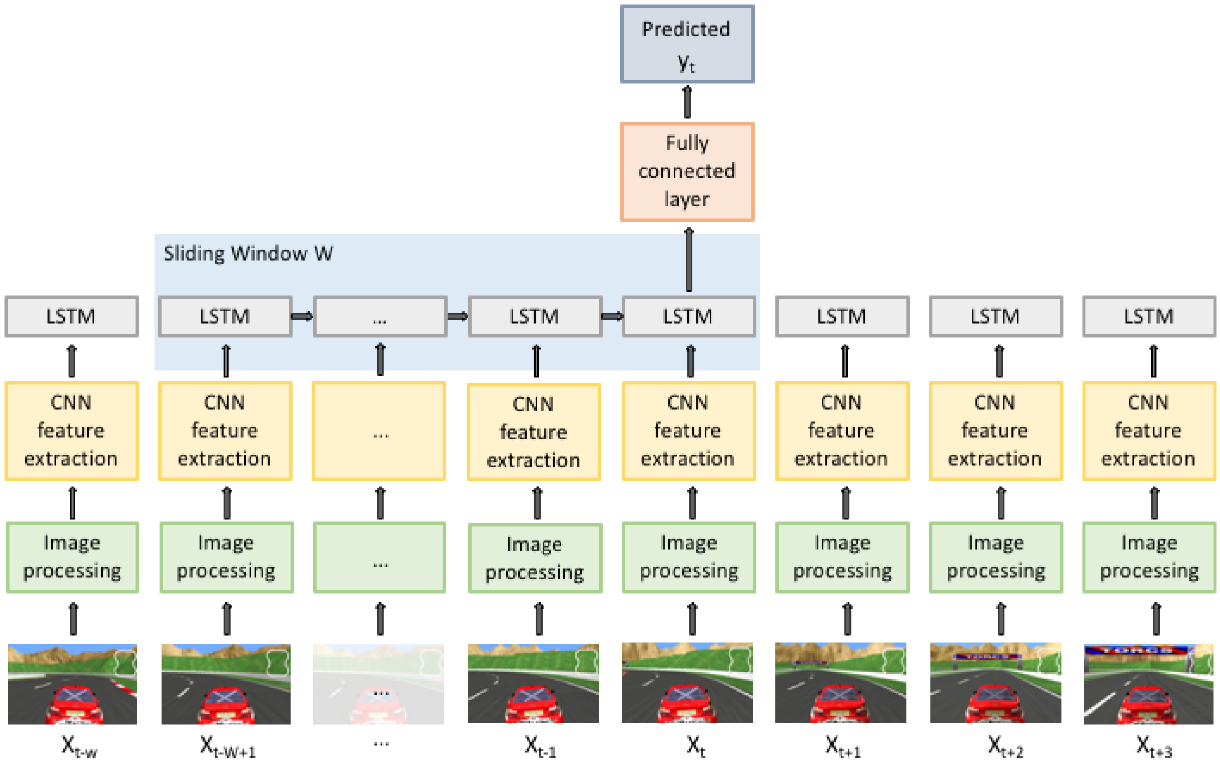


Figure 5.2 General architecture of the deep CNN-LSTM algorithm used in the proposed computational model

5.3.2.1 Data Processing

A common problem in real-life applications of deep learning-based classifiers is that some of the classes (i.e., driving decisions) have a significantly higher number of examples in the training set compared to other classes. This difference is referred to as class imbalance. There are numerous examples of this phenomenon in other fields, including medical diagnosis and fraud detection, where this issue is highly significant because the frequency of one class (e.g., cancer patients) can be 1,000 times lower than that of other classes (e.g., healthy patients). It has been established that class imbalance can have significant detrimental effects on the training of a neural network (Buda, 2018). A high disparity between the frequencies of majority and minority classes as well as minority classes with very low frequencies have a negative effect on the performance of the resulting classifiers. Furthermore, the effect of imbalanced data is significantly stronger for tasks with higher complexity.

To solve this problem, the most straightforward and common approach is to use sampling methods. These methods operate on the data itself (rather than the neural network) to increase its balance. The concept of random majority under-sampling, which describes the removal of a random portion of examples from the majority classes, is widely adopted and has been proven to work well (Japkowicz, 2002). Under-sampling results in having the same number of examples in each class. Examples are removed randomly from majority classes until all the classes have the same number of observations (referred to as "examples" in the literature).

Normalizing the input data as well as the propagated data between the network layers will smoothen and stabilize the training process of deep neural networks (Ioffe and Szegedy, 2015; Lin et al., 2017). The data are normalized using the following equation:

$$x_{ip}' = \frac{x_{ip} - mean(x_{ip})}{std(x_{ip})} \qquad \forall i, p$$

where: x_{ip}' is the normalized input value of observation i and feature p, and x_{ip} is the original value of observation i and feature p.

In addition, horizontal flipping augmentation (He et al., 2018) is performed on the training data. All the frames are flipped horizontally to obtain their mirror frames. This helps to smoothen the algorithm learn turning decisions for driving on both directions of the tracks and to make the AV self-driving system more robust.

5.3.2.2 Deep Neural Network

Deep neural networks (or deep learning) is a burgeoning area of AI that allows computational models composed of multiple processing layers to learn representations of data with multiple levels of abstraction (Adeli, 2001; Hinton et al., 2006; LeCun, 2015). Since 2006, deep learning methods have been applied widely in various fields such as autonomous driving, natural language processing, computer vision, and drug discovery to enhance the state of the art and the state of the practice in these fields. Specifically, deep convolutional nets have brought about breakthroughs in processing images, video, speech, and audio (LeCun, 2015).

In the civil and infrastructure engineering domain, researchers have adopted this technique to detect cracks in infrastructure (Jiang and Adeli, 2007; Zhang et al., 2018), manage and automatize construction projects (Adeli, 2001; Ghosh-Dastidar and Adeli, 2003; Luo et al., 2018; Fang et al., 2018), investigate highway safety (Chang, 2005; Pande and Abdel-Aty, 2008), analyze traffic and transportation networks (Yao et al., 2017; Nabian and Meidani, 2018), predict earthquakes (Panakkat and Adeli, 2009; Less and Adeli, 2010), and evaluate structural reliability (Dai and Cao, 2017).

5.3.2.3 Transferring Prior Knowledge with Classic CNN

Besides the applications of CNN in the domain of civil engineering, one of the most popular tasks is object recognition using an Imagenet dataset. Imagenet contains approximately 1.2 million images of 1,000 different classes (Krizhevsky et al., 2012). To date, several well-known network structures in the literature have been proven to be efficient and effective in recognizing a large variety of objects to a high degree of accuracy.

Deep CNN, which has made considerable progress in recent years, allows researchers to tackle complex problems efficiently and effectively. However, the amount of training data and training time required for deep learning algorithms typically far exceed those for traditional machine learning algorithms. Transfer learning has attracted a lot of attention in the past decade due to its efficiency and effectiveness in solving complicated problems and saving computational resources. The concept of transfer learning is derived from the human learning process, namely, humans typically do not learn everything from scratch. They normally leverage and transfer their knowledge (particularly when the knowledge has been proven to be useful) from previously-learnt domains to newer domains. In this dissertation, the idea of transfer learning is adopted to test if prior knowledge learned from image recognition tasks are transferable to learning driving policy.

Figure 5.3 shows an example of feature extraction with Inception network. Initially, the data on the image (the road environment that faces the AV), is processed and then fed into the network. In this process, information (knowledge) such as edge, shape, and other high-level features, are extracted and mapped to feature vectors as the output of the network. To do this, the last three layers (the dropout, fully connected, and softmax layers) need to be removed from the original Inception network.

The intermediate features learned were found to have universal expressiveness across multiple domains, which motivated this dissertation to leverage the power of pre-trained networks and apply transfer learning from object classification to learning driving policy. Two of the most popular models, the Inception network (Szegedy et al., 2015) and the ResNet-50 network (He et al., 2016), along with a baseline CNN, were compared in this dissertation.

The Inception network is a widely adopted image classification model built by the Google Brain team. This network has been shown to attain greater than 78.1% accuracy on the ImageNet dataset and achieved 21.2% top 1 and 5.6% top 5 error rates (i.e., the percentage of test examples for which the correct class was not in the top 1 and top 5 predicted classes, respectively) for single-frame evaluation with a computational cost of 5 billion multiply-adds per inference and using less than 25 million parameters. In one of the most recent applications of the Inception v3 network structure, a research team from New York University trained the network to classify different types of lung cancer (Coudray et al., 2018).

ResNet is a widely used deep convolutional neural network framework for image recognition invented by He et al. (Microsoft Research) in 2015. This model won first place in the ILSVRC 2015 classification competition with a top-5 error rate of 3.57% and won first place in ILSVRC and COCO 2015 competition in ImageNet Detection, ImageNet localization, COCO detection, and COCO segmentation. In this dissertation, ResNet-50, rather than other versions of ResNet, was adopted in transfer learning to reduce the run time of the model and allowed the framework to be replicated using a personal computer.

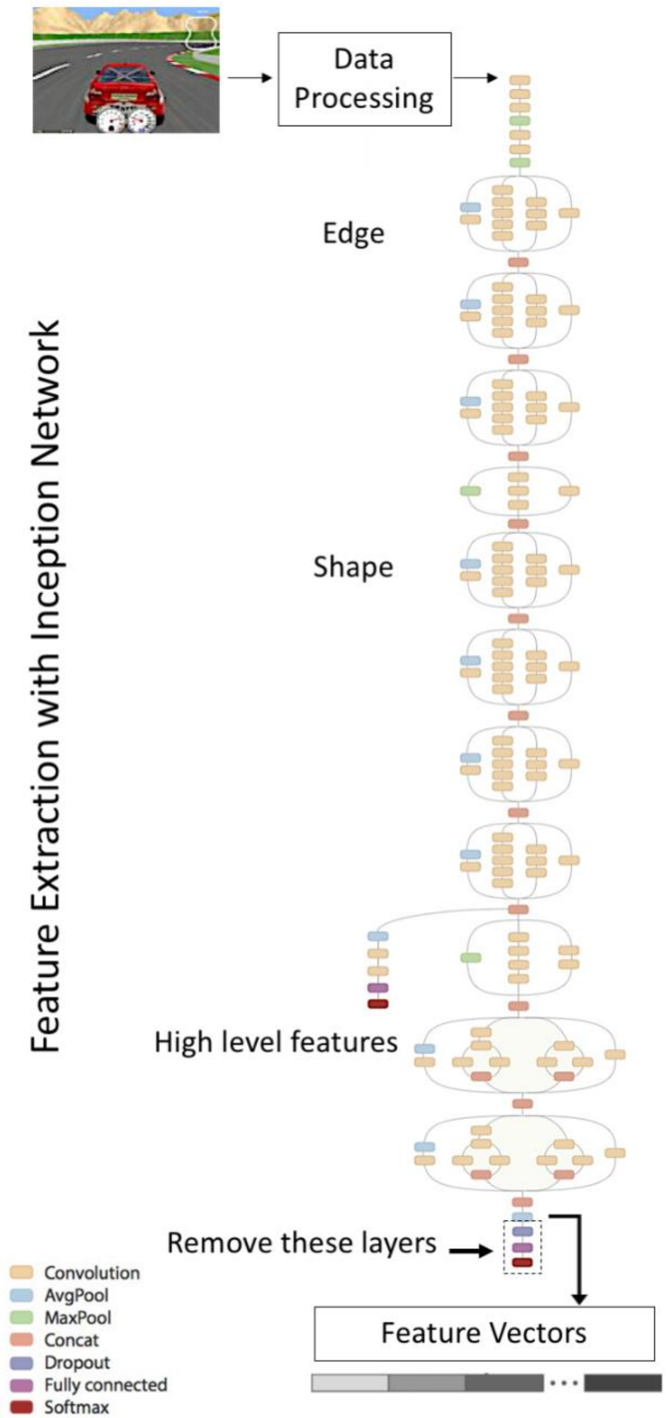


Figure 5.3 Example of feature extraction process using inception network

5.3.2.4 Incorporating Temporal Information with LSTM

The autonomous driving decision is not fully independent at each time step. For example, in making driving decisions, human drivers typically consider information in the past and previous actions. Similarly, an autonomous driving system ideally does not rely purely on the information at the current time step but also incorporate information from the past. After extracting features from the CNN, these features are processed by a LSTM network which takes information of the previous driving environment into account. The LSTM network processes w feature vectors in a sliding window, which indicates that the driving decision at a time step is predicted with w of past observations $(\boldsymbol{X_{t-w+1}}, \boldsymbol{X_{t-w+2}}, \dots, \boldsymbol{X_t})$. The size of the sliding window w controls the length of memory used to make the current driving decision. In general, a smaller w leads to shorter memory, which tends to react faster in terms of driving decision-making but learns less from the past information. On the other hand, a larger w generally corresponds with smoother driving as it processes more information from the past. However, it requires more time to train and make driving decisions.

The output of the LSTM is then fed into a fully connected network with the softmax layer to predict driving decisions. The softmax layer (Friedman and Rob, 2010) serves as a generalization of the logistic regression for multi-class classification. When the class label y has more than two values, i.e., $y \in \{1,2,3, \dots, \mathrm{c}\}$, for input data $\boldsymbol{x}$, the probabilities that $\boldsymbol{x}$ belongs to each class is:

$$\mathrm{h}_{\theta}(\mathbf{x}) = \begin{bmatrix} p(y^{(i)} = 1|\boldsymbol{x};\boldsymbol{\theta}) \\ p(y^{(i)} = 2|\boldsymbol{x};\boldsymbol{\theta}) \\ \vdots \\ \vdots \\ p(y^{(i)} = c|\boldsymbol{x};\boldsymbol{\theta} \end{bmatrix} = \frac{1}{\sum_{j=1}^{c} e^{\boldsymbol{\theta}_j^T \boldsymbol{x}}} \begin{bmatrix} e^{\boldsymbol{\theta}_1^T \boldsymbol{x}} \\ e^{\boldsymbol{\theta}_2^T \boldsymbol{x}} \\ \vdots \\ \vdots \\ e^{\boldsymbol{\theta}_c^T \boldsymbol{x}} \end{bmatrix}$$

Where $\theta_1, \theta_2, \dots \theta_c$ are the parameters for the softmax layer. After the training is complete, the probabilities that input data $\boldsymbol{x}$ belong to label j can be computed using the equation above, and the class label assigned to this input is:

$$\mathrm{y} = j = \operatorname*{argmax}_{j} \frac{e^{\boldsymbol{\theta}_j^T \boldsymbol{x}}}{\sum_{j=1}^{c} e^{\boldsymbol{\theta}_j^T \boldsymbol{x}}}$$

5.3.2.5 Recovering from Possible Data Mismatch

This work involves making a sequence of decisions given a sequence of observations over time. The learning approach is to train a classifier to predict a human driver's behavior based on the training data of the observations (input) and the decisions (output) performed by the person. However, since the classifier's current prediction affects future input observations during execution of the learned policy, which seriously violates the independent and identically distributed assumption made by most statistical learning approaches and leads to poor performance as a result. This is intuitive because as soon as the classifier makes a mistake, it may encounter completely different observations than those in the training set, leading to a compounding of errors.

In order to make the test trajectory distribution the same as the distribution learned from the driving expert's demonstration, the initial policy trained by the learning approach described earlier is fine-tuned, as follows:

Initialize D_0 and P_0 to be the initial training set and the initial policy. Then, perform iteratively the following steps:

For each training track:

For i = 1 to N:

1. Train a neural network to output a policy P using data D (input image data D_{image} + driving maneuver data D_{label}) generated by a human driver.
2. Run the policy to generate more observation $D_{\text{image+}}$ and execute the decision Y_{p} using current policy P.
3. Manually label the observation $D_{\text{image+}}$ and generate $D_{\text{label+}}$.

If $D_{\text{label+}} \neq D_{\text{p}}$ (incorrect decision):

$D' = \{D_{\text{image+}}, D_{\text{label+}}\}$.

4. Aggregate $D_{\text{i}} = D' \cup D_{\text{i}-1}$.

At the end of each episode, the model is saved back to the previous classifier to ensure lifelong and continual learning. To improve sample efficiency, only the incorrect decisions are saved. In this case, the method iteratively bridges the gap between the testing trajectory distribution and that learned during the training by having a human supervise the observation generated by the current policy. This iterative approach thus solves the distribution mismatch problem.

5.4 Implementing the Computational Model

This section of the dissertation demonstrates the implementation of the proposed self-driving simulation computational model. First, the data collection and balancing process are described. Second, the process of training the deep transfer learning algorithm is introduced and a number of parameters that affected the training are discussed. Next, the results of this demonstration are presented using the curves generated from the training process. Finally, the effectiveness of the method for the implementation demonstration is verified.

5.4.1 Data

5.4.1.1 Collecting Training Data

The experimental datasets were collected by manually driving the host car on four tracks (shown in Figure 5.4) in TORCS to serve as a benchmark for the training. A Python program was used to record the screen on a frame-by-frame basis via screenshots and each frame was matched to the corresponding driving decision (Table 5.3) by recording the keyboard action. The screenshots taken while driving the host car were resized to 160×90 resolution. The frame rate was set to 10 frames per second. The host car was driven by a human player on each track multiple times, and extreme driving cases (e.g., running off the road or travelling at dangerous speeds while turning) were intentionally created to collect data that would make the self-driving system more robust. A total of 137,500 frames were collected.

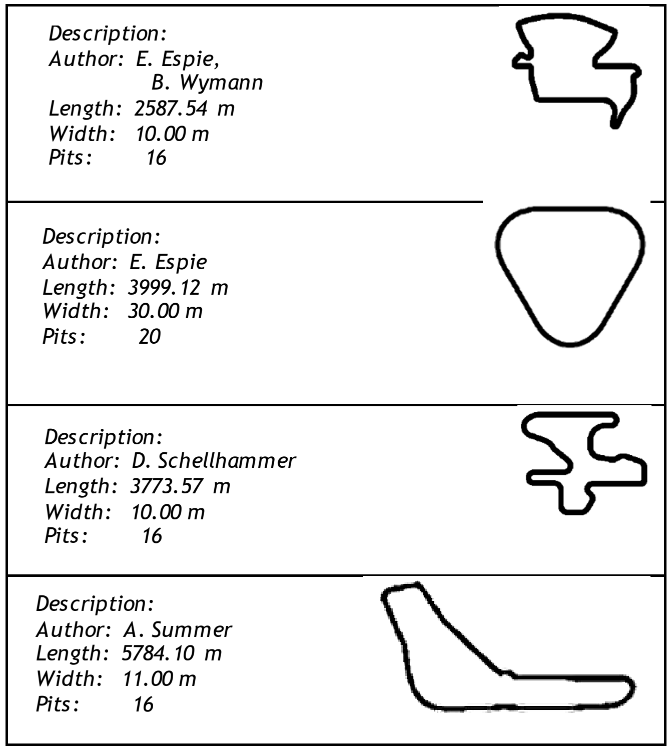


Figure 5.4 The TORCS tracks on which training data were collected

5.4.1.2 Balancing the Data

For the dataset initially collected (without balancing), the frames in the “accelerate” class were the majority (nearly 70% of the data) for the following two reasons. First, as shown in Figure 5.4, the host car was driving on a relatively straight line for these tracks most of the time; and second, due to the fact that there was no granular control for the accelerator (i.e., the keyboard binary control could either input "full accelerate" or "no accelerate"), the host car needed to "accelerate" to keep moving forward. However, feeding the model directly with this unbalanced dataset may jeopardize the training process because the deep neural network strives to maximize predictive accuracy during the training process. Therefore, it will quickly learn to predict only “accelerate” (or “go straight”) to obtain the 70% accuracy without taking much effort to really learn how to drive properly. This certainly is not the outcome one would seek from a self-driving system, and the random majority under-sampling method was applied to solve this issue. A significant disadvantage of this method is that it discards a portion of the data available. However, in this dissertation, the driving data were easily obtained in the simulated TORCS environment. Therefore, discarding very similar examples of majority class data (e.g., driving on a straight track without turning) is not considered wasteful. In addition, additional data were collected specifically for minority classes, such as the host car approaching or negotiating a curve. After balancing the data, 42,633 frames were used to develop the network.

5.4.2 Settings of the CNN-LSTM network structure

Five alternative networks were developed and their corresponding performance levels were compared:

1) Baseline CNN+FCN: The baseline CNN had six convolutional layers. The kernel size for the first 3 layers was 4×4, and for the last three layers was 2×2. The dimension of the six layers were 32, 32, 48, 48, 64, and 64. The activation function used was ReLu (Nair and Hinton, 2010). Batch normalization and dropout layers were added after each convolutional layer. The CNN was followed by a fully connected network (FCN) with a softmax layer to predict the driving decisions.
2) Inception+FCN: This network used the Inception network to extract features. The extracted feature vectors were then fed into a FCN with a softmax layer to predict the driving decisions.
3) ResNet-50+FCN: This network used the ResNet-50 network to extract features. The extracted feature vectors were then fed into a FCN with a softmax layer to predict the driving decisions.
4) Inception+LSTM+FCN: This network used the Inception network to extract features. The extracted feature vectors were then fed into a LSTM network. The LSTM, which had 64 hidden units and sliding window size $w = 4$ produced the best performance on the dataset. The LSTM was followed by a FCN with a softmax layer to predict the driving decisions.
5) ResNet-50+LSTM+FCN: This network used the ResNet-50 network to extract features. The extracted feature vectors were then fed into a LSTM network. Similar to 4), the LSTM, which had 64 hidden units and sliding window size $w = 4$, produced the best performance on the dataset. The LSTM was followed by a FCN with a softmax layer to predict the driving decisions.

5.4.3 Training of the Model Based on the Collected Data

The training was carried out using TensorFlow on a hardware platform with the following specifications: Overclocking 5.0 GHz Intel i9 processor, 32 GB DDR4/2666 MHz memory, and

Nvidia GeForce GTX 1080. TensorFlow is an open-source software platform that allows computation to be carried out on multiple GPUs or CPUs (Tensorflow, 2016). The platform allows users to create large and complex deep neural networks and perform training effectively. The parameters used for training the networks are briefly discussed below.

5.4.3.1 Batch Size

The batch size was set as 128. The batch size controls the number of samples to be propagated through the network at a time. Typically, gradients computed during the training process are more accurate for larger batch sizes compared to smaller ones. A smaller batch size requires less memory. In addition, the use of small batch sizes has been shown to avoid local minima and improve generalization performance and optimization convergence in the training of deep neural networks (LeCun et al., 2012; Keskar et al., 2016).

5.4.3.2 Learning Rate

The Adam optimization algorithm (Kingma and Ba, 2014) was applied to train the networks. The initial learning rate was set as 0.001 with a momentum decay of 0.9. In general, a high learning rate may make it difficult for a network to converge, particularly if the network is trained with a small batch size. However, a very low learning rate requires many training epochs to achieve convergence.

5.5 Results

To measure the accuracy of the driving decisions made by the self-driving computational model, a validation dataset was constructed, which consisted of the last 10% samples of the data collected on each training track (i.e., data collected from a certain part of the track will not be included in training). The reason for choosing the last 10% instead of random selected 10% of the data is to avoid data leak. For example, if 10% of random selected data is used as validation set, some of the validation information will be very similar to the training data, which can lead to low generalization ability for the model. The model was trained using an accuracy metric that indicated how close the predicted driving decision was to the real driving decision for each frame. The higher the accuracy, the closer the predicted decision was to the human driver's decision. Table 5.4 shows the training and validation accuracies for the five networks described in Section 5.4.2.

Table 5.4 Comparison of training and validation accuracy across different network structures.

Network Structure	Training Accuracy	Validation Accuracy
Baseline CNN+FCN	80.33%	73.21%
Inception+FCN	84.19%	79.76%
ResNet-50+FCN	83.67%	81.04%
Inception+LSTM+FCN	86.94%	85.67%
ResNet-50+LSTM+FCN	87.12%	84.01%

During the data collection process, it was observed that human drivers may make some driving decisions randomly. For example, for a person who drives the host car twice on the same track, it is impossible for the wheel, accelerator, and brake decisions to be identical in both driving sessions. Therefore, it is impossible to reach 100% accuracy, even when training data are used.

Comparing Inception+FCN and ResNet-50+FCN with Baseline CNN+FCN, it was observed that by using features extracted from a pretrained network with Imagenet data, the network performance increased significantly compared to training a CNN from scratch. This result indicates that prior information learned from image recognition task is transferable to learning driving policy.

With LSTM, the accuracies for Inception and ResNet networks were further enhanced to 85.67% and 84.01%, respectively, which indicates that incorporating temporal information with recurrent neural networks helps improve driving decisions as human drivers also use information from the past while driving.

The final selected networks were Inception+LSTM+FCN and ResNet-50+LSTM+FCN. They were evaluated in the TORCS driving environment. When the proposed self-driving models were used to operate a vehicle on the TORCS track, the only information it accessed was the front-facing image fed to them. Within the simulation, the model outputs can be visualized and used as input to control the keyboard through a Python program that operates the vehicle. The self-driving model was first tested on the four tracks where the training data were collected (Figure 5.4). It was observed that the models drove the host car on these four training tracks satisfactorily without any collisions or dangerous driving situations. Then, the models were tested on four "new" tracks, that is, tracks that were not used in the model training (Figure 5.5). Again, it was observed that the models drove the host car satisfactorily without collisions or near collisions. In some scenarios with sharp turns, it was observed that the host car sometimes slightly overshot its target path at first while turning but then quickly recovered to the appropriate path. In general, the computational models were shown to be capable of producing a smooth driving experience.

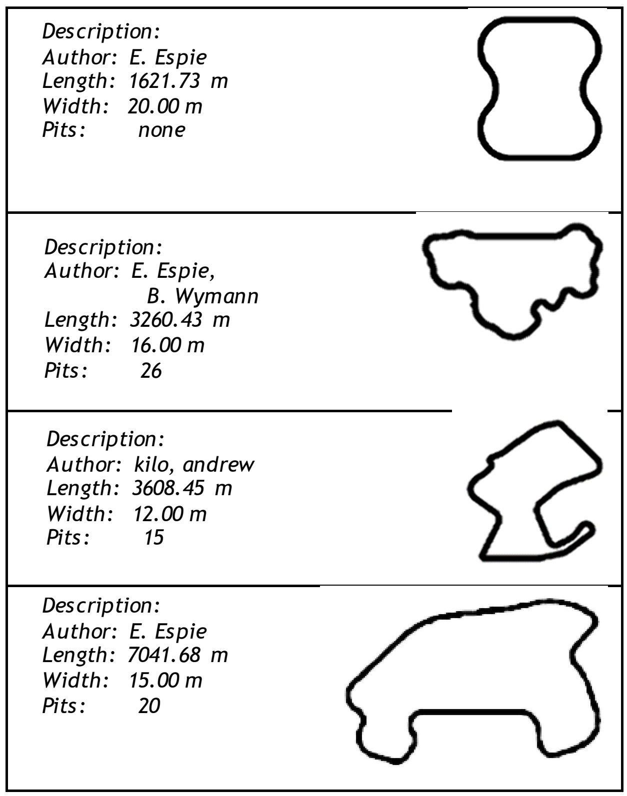


Figure 5.5 TORCS tracks used for testing the self-driving model

CHAPTER 6. SUMMARY AND CONCLUSIONS

6.1 Introduction

This chapter presents the conclusions and key findings of the dissertation with respect to highway safety in both traditional and emerging operating environments are discussed. In addition, this chapter identifies limitations and some possible avenues for the future work.

6.2 The Traditional Operating Environment

6.2.1 Impacts of Road-Surface Condition on Safety

Chapter 3.2 investigated the impacts of pavement condition and other factors including geometric features and traffic characteristics on rural highway safety across different road surface conditions. This investigation was conducted for three levels of crash severity – fatal, injury, and no-injury. The data represented hundreds of in-service pavement sections, and the level of highway safety at each section was expressed in terms of the number of each crash type along a given section each year within the specified analysis period. Crash data are inherently multivariate in nature; therefore, modeling each severity level separately, while not accounting for the likely correlation between the crash counts across the crash severity levels, can be expected to be problematic. Using a univariate modeling approach for the correlated crash counts can lead to less precise estimates for the risk factors associated with different severity levels. In this regard, for joint modeling of crash frequency, a multivariate modeling approach was used for the analysis. In order to account for the unobserved heterogeneity and correlation among the three levels of crash severity, MRPNB models were estimated for each of the three subclasses of rural highways (Interstates, State roads and U.S. roads) as these road classes generally differ from each other in terms of geometric standards, pavement quality, and traffic volume.

The results were consistent with engineering intuition and also with a-priori indications from past research. A number of traffic and road geometric covariates were found to significantly influence the number of crashes, at various severity levels, and across the various pavement surface conditions of rural highways. For roads with pavements in poor condition, fewer road geometry factors were found to be significant compared to those with pavements in fair-excellent condition. The results generally supported the hypothesis that crash propensity generally decreases as pavement condition improves from poor (high IRI) to good condition (low IRI). However, this part of the dissertation also established that for pavements in poor condition, the effect of further reductions or increases in condition can have a dichotomous effect on the number of crashes.

In addition, the results show that for most crash types and most pavement condition categories, the impact of pavement condition on rural road safety is generally similar in direction but not in magnitude. In certain cases, the direction of influence was found to be opposite. Specifically, the results indicate that state roads with pavements in poor condition have a greater frequency of injury crashes compared to non-state roads in similar condition. In addition, two-lane highways with pavements in poor condition have a lower frequency of no-injury crashes compared to highways with four or six lanes. For roads with pavements in excellent condition, more lanes are associated with a higher frequency of injury and no-injury crashes.

The models representing each pavement condition suggest that the safety effects of pavement condition differ across various levels of pavement condition; with the exception of the poor pavement condition model, the models suggest that a fixed-effects parameter is appropriate for the pavement condition variable (IRI). In the poor pavement condition model, the results indicate that the pavement condition variable has a normally distributed random parameter and therefore leads to an interpretation that reflects both opposing effects (i.e., increased pavement condition within that condition range causes a higher frequency of crashes in certain situations and a lower frequency in other situations). The observation of the latter effect is not only reflective of the phenomenon of risk compensation but also reinforces the notion that driver behavior in response to operating conditions can be significantly heterogeneous.

For each crash severity level, this part of the dissertation confirmed the results of recent studies that the effect of pavement condition on safety differs across the different pavement conditions. For no-injury and injury crashes, it was determined that the higher the road functional class (from state roads to interstates), the greater the sensitivity of safety to changes in the pavement condition; however, this dissertation observed the opposite trend for fatal crashes.

Due to the demonstrated impact of pavement condition on highway safety, an agency's pavement preservation program can provide an excellent opportunity to substantially reduce the number of roadway crashes. According to Zimmerman and Larson (2005) and Mahoney et al.

(2006), by incorporating safety measures into its pavement management systems as well as its long-range plans, an agency can utilize a systematic approach for identifying and prioritizing pavement preservation needs. The conceptual frameworks in these two studies argued for the need to incorporate safety performance in long-range pavement planning and management (Lamptey et al., 2005; Lamptey et al., 2010; Murillo-Hoyos et al., 2015; Anastasopoulos et al., 2016). The results of this part of the dissertation could provide input toward such applications. Furthermore, the developed models can help highway engineers quantify not only the safety benefits of rural road resurfacing projects but also the safety consequences of worsening road surface conditions arising from maintenance that is denied or delayed.

There is a wealth of knowledge to be gained by continuing the research in this area. First, future work could investigate a greater range of possibilities of interactions between the variables. Second, future work could acquire and incorporate more data on other crash factors besides the engineering factors. Third, the work could be extended to urban highways where a greater number of factors exist that can influence highway crashes. Finally, future work could extend the analysis presented in this part of the dissertation to account for both spatial and temporal heterogeneity within the multivariate random parameters framework.

6.2.2 Optimizing Space Allocation across Highway Cross Sectional Elements

Chapter 3.3 developed a framework for determining the optimal allocation of shoulder and lane widths on two-lane rural highways. The rural road functional classes considered are: major collectors, minor arterials, and principal arterials. For each functional class, first, a set of crash prediction models was used to estimate the expected number of fatal+injury and PDO crashes, based upon predefined geometric and operating conditions. These crash results were combined with the cost of constructing and maintaining the lane and shoulder for different combinations of lane and shoulder width under a given total roadway width constraint. Then the agency and user costs were assigned relative weights to achieve the optimal distribution of total roadway width between the lane and the shoulder such that total agency costs and user safety costs over life cycle were minimized using nonlinear programming.

For minor arterials and major collectors, the resulting shapes of the optimal solutions (funnel diagrams) as functions of the relative weight of user cost to agency cost, are similar. For high user cost weights, it is observed that the optimum has a shoulder width of zero (the lanes constitute the entire TRW). This is because in the crash prediction models used, wider lanes contribute more to crash reduction compared to wider shoulders. For low weights of the user cost, it is seen that the optimum has a lane width of 10 ft., and the shoulder takes up the remaining TRW. This is because the shoulder construction and maintenance costs are much lower compared to the lanes. Between these two extremes of user cost weight, there is a range for the ratio of the weight on agency cost divided by user cost, which is specific to the TRW and the road functional class; over this range, the optimal width of the lane transitions from its maximum value of TRW/2 to its minimum feasible value of 10 ft. Irrespective of the TRW value, these transitions occur at lower values of the agency-user cost weight ratio for minor arterials compared to major collectors.

With regard to principal arterials, the results are somewhat different, with the optimum being a lane width of 10 ft. and the rest of the TRW taken up by shoulder. This is because unlike the other two functional classes, the principal arterial roads' crash prediction model for fatal+injury crashes places greater premium on wide shoulders than it does on wide lanes, in terms of crash reduction. The absolute value of the marginal effects for lane width is small compared to the shoulder width; thus the optimization result favors shoulder widening to lane widening. This is coupled with the fact that fatal and injury crashes have a much higher unit cost compared to PDO crashes. Furthermore, agency costs (of shoulder construction and maintenance) are much lower compared to the lanes, and there is no other penalty for very wide lanes or shoulders, for example, the models used in this study do not account for the risk compensation that would have indicated increasing crashes for very wide lanes or shoulders. Therefore, from both the user and agency perspectives, having narrow lanes and shoulders that are as wide as possible is preferable. The optimization framework tends to provide results with minimum lane width for principal arterials.

The sensitivity analysis on the agency cost and user cost ratios given by the optimization framework were found to be useful in evaluating existing practice. AASHTO standards generally recommend 12 ft. wide lanes for rural highways. The results present the optimal ratio for lane and shoulder widths, for different TRWs, different agency-user cost weight ratios, and across different road functional classes. The discount rate was also found to have small impacts on optimal lane and shoulder width ratio. Generally, an increase of discount rate will favor wider lanes at the expense of shoulders. The crash prediction models and case studies are used and presented here only for purposes of illustration; they can be replaced by other models to fit for specific conditions.

In this dissertation, the consequences of a widened shoulder or lane were measured in terms of the agency cost of construction and maintenance, and the user safety benefits (reduced crashes). Future studies could examine other cost and benefit types associated with road widening. For example, the benefits of lane or shoulder widening could be made to include improvement to structural integrity (Wu et al., 2015); this is due not to any increase in pavement thickness but to increased protection of the road edges, thus reducing road edge breakoffs that cause lane or shoulder drop-offs (defined as the uneven edge or vertical height differential between the paved lane and the unpaved shoulder, or the paved shoulder and the surrounding ground (FHWA, 2006). Another potential benefit of shoulder widening is its potential use for hard-shoulder running (Ma et al., 2016) in times of recurrent or incident-induced congestion.

There are several potentially restrictive assumptions which were made in this analysis that could be addressed in future research. For instance, in the agency life-cycle cost analysis, this study assumed that the installation and maintenance costs for paved shoulders are lower than those for paved lanes. This would reflect the practice of an agency first paving the travel lanes to a certain depth, and then paving the shoulders to a shallower depth in order to reduce overall pavement costs. In reality, if the agency were to pave the lanes and shoulders to the same depth using a single pass of an HMA paver, then this assumption, and the corresponding construction costs, may need to be revised. Also, this study used fixed unit construction and maintenance costs that are independent of the added width. In reality, for each mile, the unit cost ($/ft.) is likely to decrease as added width increases due to scale economies. In addition, this study used deterministic inputs including agency and user costs (safety performance, costs of construction, maintenance, and rehabilitation). Future studies could duly consider stochastic values of these inputs. Also, the analysis was demonstrated using crash prediction models from a specific state that may not reflect operating and economic conditions that exist at other locations. Nonetheless, the framework presented in this dissertation is sufficiently flexible for use at locations that have different values of these inputs and models, and is also accommodates the effects of risk compensation.

6.2.3 Efficacy of Machine Learning in Road Safety Analysis: Predicting the Fatality Status of Highway Segments

Chapter 3.4 explored the application of machine learning algorithms to predict highway crash fatalities. The goal of the part of the dissertation was to develop machine learning models to predict fatal highway segments and to compare these models with zero-inflated and logistic econometric models using precision, recall, and F1 score as performance measures. Both the machine learning and econometric models were constructed using engineering factors as inputs. Engineering factors were chosen because they are the crash factors that can be manipulated by highway agencies. Simply using historical data from fatal highway segments may be considered sufficient in certain cases when making predictions for a static network. For systems that are changing (e.g., where traffic flow is increasing or decreasing due to regulations and new routes, pavement deterioration is occurring, new highways and lanes are being constructed, and so on), machine learning models developed in this study are able to make predictions with a high level of confidence. For this part of the dissertation, several machine learning models were trained and their performance evaluated. In addition, this part of the dissertation proposed the use of precision, recall, and F1 score instead of accuracy to select models for predicting fatal highway segments based on the specific needs of highway agencies. In general, when agencies are less restrictive with funding and the goal is to improve as many potential fatal segments as possible, then a model with high recall is recommended. Otherwise, a model with high precision is recommended because high-precision models can help agencies be more conservative with their funding through reductions in the number of false positive predictions. The F1 score is the harmonic mean of precision and recall and is suggested for agencies that prefer greater balance.

Both the final selected models, SVM with RBF kernel and random forests, exhibit good performance in terms of precision (above 0.85 with the testing data), while their recall values are relatively low. The results can be considered reasonable from a highway engineering standpoint. This is because the features used in the analysis (e.g., highway geometric design, pavement condition, and road functional class) are able to determine whether a highway segment is safe or dangerous in terms of design (leading to high precision). However, these features are unable to identify safe or dangerous highway segments in terms of other factors, such as the presence of drunk drivers or bad weather (leading to low recall). In addition, the learning curves for the final selected models suggest that there is still potential to further improve the models' performance if additional data are included. For future work, more data (observations and crash conditions) are planned to be collected to further improve the efficiency and reliability of fatal crash predictions.

The proposed machine learning algorithms have been proven to be more efficient in identifying fatal highway segments. This performance can help provide meaningful feedback for highway operations: for segments that are predicted by high-precision machine learning models to be fatal, a highway agency can further investigate these segments through roadway audits as a first step toward prescribing a countermeasure.

It is important to note that driving behaviors, weather conditions, regulations, culture, and other factors vary across regions, and therefore the modeling results of a certain area may not be applicable to other areas. As such, in order to replicate the analysis, it is recommended that the requisite design and operational data pertaining to any future study area be collected and analyzed using the methodology presented in this part of the dissertation. In addition, future study may also investigate other data sources including the second Strategic Highway Research Program (SHRP 2) Safety Data, which is administered by the Transportation Research Board of the National Academies and includes naturalistic driving data and provides information regarding vehicle speed, acceleration and braking, lane position, near-crashes, and so on.

6.3 The Emerging Operating Environment

6.3.1 Contributions and Concluding Remarks

In the operating environment, an era of AV, the traditional analysis of roadway safety factors may become less and less important. Safety factors that significantly affect highway safety in the traditional operating environment may not be significant in the future. Due to the potential changes in safety factors and the resulting uncertainty, it is important to carry out field testing on AVs to ensure their safety performance, particularly in the current era where recent AV crashes have somewhat eroded public trust and confidence in AVs. An impediment to field testing, however, is the inability of test tracks to fully and adequately replicate the real-world driving environment. This part of the dissertation therefore proposes a simulation methodology to facilitate reliable testing of AVs to ascertain their safety performance and other prospective benefits. As explained earlier in this part of the dissertation, existing driving simulators belong to the constructive category of simulation (where both the vehicle and the driving environment are simulated) and the category where the AV does not independently and externally characterize the driving environment. That is, the AV operation is based on commands external to the vehicle already embedded in the code for the driving environment. Such simulation, where the source of vehicle control is some entity external to the vehicle (i.e., SIVC I and II in Table 5.2), the simulation may not be considered as an adequate representation of AV driving as defined in this part of the dissertation. In an earlier section, this part of the dissertation argued that a more representative simulation of autonomous driving refers to a simulation where the source of vehicle control is the AV itself (i.e., SIVC III and IV in Table 5.2).

SIVC IV, where the AV uses sensors (video, lidar, GPS) as the source of information and controls itself (through its AI system), is the category that most closely resembles a real-life situation. Category SIVC III, where the AV uses an external server with a road environment database as the source of information and controls itself, is also similar to a real-life scenario in terms of connected and autonomous vehicles (CAVs); however, this scenario is not the focus of this part of the dissertation. In the future, as more data become available (e.g., data that mimic V2V, V2I, V2X communication), the data can easily be integrated into the proposed model to enhance it further.

To address the simulation shortfalls associated with traditional simulation (SIVC I and II), this part of the dissertation proposes a self-driving computational model that enables SIVC IV simulation. The dissertation proposes a deep Convolutional Neural Network – Long Short-Term Memory (CNN-LSTM) algorithm. This algorithm observes and characterizes the AV's driving environment, and controls the AV movement in the driving simulation. The CNN part extracts features that use transfer learning to introduce human prior knowledge, and the LSTM part uses temporal information to process the extracted features, and incorporates temporal dynamics to make driving decisions. The dissertation demonstrated the developed self-driving computational model using only image data of the driving environment as input. The proposed model was tested using TORCS and exhibited an ability to predict human-like driving decisions reliably. The AV may also use an external server with a database containing road environment data such as a GPS as an additional source of information on the roadway.

The developed approach can be considered flexible enough to be adopted or replicated by universities and research institutes that have their own simulators. This is motivated by the fact that while different driving simulators have different functions and different capabilities to access driving-environment data, every simulator has access to image data on its screen. Other types of data (e.g., data that mimic V2V, V2I and V2X communication, etc.) can easily be integrated into the proposed computational model to realize SIVC III simulation.

The proposed computational model for simulators can help to overcome the challenges caused by the scarcity of data, particularly in the domain of civil and infrastructure engineering.

With the proposed model, researchers will be able to collect data and conduct CAV-related research on human factors, education, smart infrastructure (e.g., V2I communication), roadway design, and so on. From the human factor perspective, the model provides researchers with enhanced ability to study CAV passenger behavior, reactions and reaction times to different types of warning systems (while passengers perform different tasks), passenger comfort with different driving styles (e.g., different distances between vehicles, speed controls, etc.), and so on. From an education perspective, researchers will be able to determine the necessary training and investigate different ways to train CAV users based on their age, gender, education level, etc., and children will be provided an opportunity to experience CAVs for educational purposes in a safe environment.

In addition, the computational model, when implemented, can play an important role in new roadway design for CAVs. With the adoption of CAVs, it is expected that various levels of current roadway design features will need to be evaluated to facilitate AV traffic operations. These design factors include lane and shoulder width, horizontal and vertical alignment, cross section, superelevation, stopping sight distance, entrance and exit design, and so on. In real-life, physically changing the levels of these design features incrementally and assessing the resulting impacts on AV operations, will be not only impractical and hazardous but also prohibitively expensive. In a simulation platform, it is quick and easy. It is also necessary to evaluate alternative proposed new designs from a human factor perspective because CAV users may not be comfortable with certain design features.

Furthermore, from a technology perspective, the computational model will need to be retrained with the new roadway designs and will require a large amount of data that will not be available before the new designs are implemented. In sum, it is suggested that potential new roadway designs to accommodate AVs need to be studied in driving simulators using the “more representative” simulation computational model introduced in this part of the dissertation prior to their adoption in real life highway systems.

This dissertation has strived to make a few significant contributions to existing knowledge regarding the simulation of AV operations in a simulated environment. Traditional research into autonomous driving system development has focused on specific elements of the driving tasks such as lane detection (Wang et al., 2004; Assidiq et al., 2008), object detection (Huang et al., 2004), path planning (Chu et al., 2012), lane change (Naranjo et al., 2008). Subsequent advancements in deep learning algorithms and enhanced modern computing power have allowed researchers to realize end-to-end learning frameworks for autonomous driving decisions. For example, Chen et al., (2015), Bojarski et al., (2016), Martinez et al., (2017) and Dosovitskiy et al., (2017) proposed deep convolutional neural networks that process image data to realize autonomous driving.

However, the past models could be improved further in a number of areas: a) incorporation of the driver's prior general knowledge of driving environments. Human drivers make certain assumptions and follow certain rules while driving, and are able to use prior knowledge such as the expected positions of lane markings relative to their vehicle. However, the previously proposed models are expected to learn from scratch, which is not efficient and effective for researchers that lack computing power and data; b) consideration of additional temporal information (not only the current information) to make driving decisions. There is a need to accommodate the reality that human drivers typically use information from the past (speed and visual information collected a few seconds ago, for example) along with the current information, to make driving decisions.

In this dissertation, deep convolutional neural network-long short-term memory (CNN-LSTM) model is proposed to improve the simulation of autonomous driving. The methodology, unlike its predecessors, allows for incorporation of prior knowledge using transfer learning of classic and powerful CNN (e.g., Inception, ResNet, etc.). This is done to improve driving decision-making and to save computing and training resources. The Long Short-Term Memory (LSTM) network structure duly considers temporal information (past and present). LSTM processes transferred knowledge from classic CNNs; that way, the proposed framework can utilize past information to make driving decisions. In addition, an improved data collection process is proposed such that it addresses the data mismatch problem (where the test trajectory distribution deviate from the training trajectory distribution) which may lead to navigation paths of the AV that stray outside the lanes in the testing environment.

6.3.2 Limitation and Possible Directions for Future Research

In this dissertation, driving decisions (Table 5.3) were implemented using a keyboard. Therefore, the host car was able to perform only the following actions at any given time: full acceleration, braking, and turning, due to the binary keyboard inputs. On the keyboard, a key representing left turn, for example, would make the car turn full left turn only (maximum left turn of the front wheels) when pressed. In reality, however, this is not the case because drivers with a rotary steering wheel can turn the wheel to any angle between zero and the full-turn angle, not just the full turn angle only. The same is true for the driver's control over the accelerator and brake: Using limited driving data to train the model will teach the self-driving system to drive using binary control of the host car's turns (no turn vs. full turn), accelerations (no acceleration vs. full acceleration), and braking (no brake vs. full brake). This limitation can be addressed in future work by introducing more fluid control. Instead of using a keyboard to control the movement of the host car, a rotary wheel with accelerator/brake pedals could be used. With such increased fluidity of control, the computational model can be expected to smoothen the self-driving of the host car.

The driving environment in this part of the dissertation was a relatively simple one, in that there were no other cars on the simulated track. In reality, the AV driving environment is expected to be populated by a mélange of vehicle types, some AV, others traditional. In such scenarios, it is necessary to address situations that require a 360-degree panoramic characterization of the driving environment. For example, when the host vehicle moves to another lane to overtake another vehicle, it is difficult to ascertain whether it is safe to complete the overtaking maneuver by returning to the initial lane because the host vehicle is unable to "see" the overtaken vehicle. To simulate a real-world environment, in which AVs typically have 360-degree cameras, it is suggested that the simulator use 360-degree screens to make the driving environment more realistic. In addition, Future research may consider incorporating microsimulation (which is able to simulate complicated traffic network) with driving simulation, or consider to use multiple driving simulators (autonomous or manual modes) in the same simulated environment to test the interactions between human driving vehicles and autonomous vehicles.

Future research can also apply reinforcement learning in certain situations to let the self-driving system learn to drive on its own. The reason for this proposed future modification is that during the demonstration, it was observed that when there were very sharp turns and the host car was driving at high speeds, the human driver was not able to prepare and react quickly and effectively, which led to crashes in the simulation. It is expected that using reinforcement learning will enable the AI algorithm to learn through failure and ultimately identify and adopt the optimal turning angles and speeds. Other related refinements that could be made to the proposed self-driving computational model include additions of sub-models for object detection, hazard warning, and speed controls.

To enhance further the framework, the proposed deep learning algorithm can be fused with other algorithms. Furthermore, after more granular input controls are implemented in the simulation, real-world data (real driving data on highways, local roads, etc.) collected using smartphone cameras can be used to test the TORCS-based model developed in this part of the dissertation. For example, the wheel and speed control data from real-world highway driving can be compared with the predictions made by the deep neural network.

6.4 Overall Summary

This dissertation began with a comprehensive discussion of the literature regarding highway safety, crash factors, and modeling techniques in the traditional operating environment, followed by a discussion of the opportunities that exist to enhance crash prediction models to provide better feedback with respect to preservation phase, design phase and operations phase. Then, a framework for enhanced prediction of highway safety was introduced for evaluating and quantifying the safety impacts of road-surface condition, optimal lane and shoulder width allocation, and reliable crash prediction model to predict highway fatalities. Finally, with the insights gained from the traditional operating environment, this dissertation discussed the opportunities and the expected safety impacts and benefits of adopting AV into the operating environment. To address the limitations of testing and evaluating the safety impacts of AV on in-service roads and test tracks, this dissertation proposed and demonstrated an autonomous driving simulation framework that can enable researchers to test and evaluate these impacts in a safe environment.

APPENDIX

Optimal lane & shoulder widths across road functional classes for different TRWs

		Major Collector		Minor Arterial		Principal Arterial	
TRW	W_{AGENCY}/W_{USER}	Lane	Shoulder	Lane	Shoulder	Lane	Shoulder
26 ft.	0 to 1.1	13	0	13	0	10	3
	1.2	13	0	12.282	0.718	10	3
	1.3	13	0	10.262	2.738	10	3
	1.4 to 1.7	13	0	10	3	10	3
	1.8	12.62	0.38	10	3	10	3
	1.9	11.674	1.326	10	3	10	3
	2	10.78	2.22	10	3	10	3
	>2.1	10	3	10	3	10	3
30 ft.	0 to 0.9	15	0	15	0	10	5
	1	15	0	14.348	0.652	10	5
	1.1	15	0	11.941	3.059	10	5
	1.2	15	0	10	5	10	5
	1.3	15	0	10	5	10	5
	1.4	14.785	0.215	10	5	10	5
	1.5	13.643	1.357	10	5	10	5
	1.6	12.576	2.424	10	5	10	5
	1.7	11.573	3.427	10	5	10	5
	1.8	10.627	4.373	10	5	10	5
	>1.9	10	5	10	5	10	5
34 ft.	0 to 0.8	17	0	17	0	10	7
	0.9	17	0	14.469	2.531	10	7
	1	17	0	11.812	5.188	10	7
	1.1	17	0	10	7	10	7
	1.2	16.289	0.711	10	7	10	7
	1.3	14.964	2.036	10	7	10	7
	1.4	13.738	3.262	10	7	10	7
	1.5	12.597	4.403	10	7	10	7
	1.6	11.529	5.471	10	7	10	7
	1.7	10.526	6.474	10	7	10	7
	>1.8	10	7	10	7	10	7
38 ft.	0 to 0.6	19	0	19	0	10	9
	0.7	19	0	18.277	0.723	10	9
	0.8	19	0	14.905	4.095	10	9
	0.9	19	0	11.938	7.062	10	9
	1	18.261	0.739	10	9	10	9
	1.1	16.683	2.317	10	9	10	9
	1.2	15.243	3.757	10	9	10	9
	1.3	13.919	5.081	10	9	10	9
	1.4	12.692	6.308	10	9	10	9
	1.5	11.551	7.449	10	9	10	9
	1.6	10.483	8.517	10	9	10	9
	>1.7	10	9	10	9	10	9

REFERENCES

AASHTO. (2011). A policy on geometric design of highways and streets. 6th Ed., America Association of State Highway and Transportation Officials, Washington, DC.

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., & Kudlur, M. (2016). Tensorflow: a system for large-scale machine learning. OSDI 16, 265-283.

Abay, K. A., & Mannering, F. (2016). An empirical analysis of risk-taking in car driving and other aspects of life. Accident Analysis and Prevention 97, 57-68.

Abay, K., Paleti, R., & Bhat, C. (2013). The joint analysis of injury severity of drivers in two-vehicle crashes accommodating seat belt use endogeneity. Transportation Research Part B 50, 74-89.

Abboud, N.K., & B. L. Bowman. (2001). Evaluation of two- and four-ft. shoulders on two-lane state routes. ITE Journal, 34–39.

Abdelwahab, H., & Abdel-Aty, M. (2002). Artificial neural networks and logit models for traffic safety analysis of toll plazas. Transportation Research Record: Journal of the Transportation Research Board(1784), 115-125.

Aberg, L., & Rimmo, P. (1998). Dimensions of aberrant driver behavior. Ergonomics 41(1), 39–56.

Abou-zeid, H., Pervez, F., Adinoyi, A., Aljlayl, M., & Yanikomeroglu, H. Cellular. V2X Transmission for Connected and Autonomous Vehicles: Standardization, Applications, and Enabling Technologies.

Adeli, H. (2001). Neural networks in civil engineering: 1989–2000. Computer-Aided Civil and Infrastructure Engineering, 16(2), 126-142.

Adeli, H., & Jiang, X. (2008). Intelligent Infrastructure: Neural Networks, Wavelets, and Chaos Theory for Intelligent Transportation Systems and Smart Structures, CRC Press, Boca Raton, FL.

Adeli, H., & Kim, K. (2009). Wavelet-based vibration control of smart buildings and bridges. CRC Press, Boca Raton, FL.

Agent, K., Pigman, J., & Green, E.R. (2004). Effect of pavement resurfacing on safety, Research Report KTC-04-02/SPR 257-03-IF., University of Kentucky, Lexington, KY.

Aguero-Valverde, J. (2013). Multivariate spatial models of excess crash frequency at area level: Case of Costa Rica. Accident Analysis and Prevention 59, 365-373.

Ahmed, N.U. (1983). An analytical decision model for resource allocation in highway maintenance management. Transportation Research Part A 17A(1), 133–138.

Ahmed, A., Saeed, TU., Murillo-Hoyos, J., & Labi, S. (2017). Pavement repair marginal costs – accounting for heterogeneity using random-parameters regression. Journal of Infrastructure Systems. 23(4)

Akaike, H. (1987). Factor analysis and AIC. In Selected Papers of Hirotugu Akaike (pp. 371-386). Springer, New York, NY.

Aldrich, C. (2004). Simulations and the future of learning: an innovative (and perhaps revolutionary) approach to e-learning. Pfeifer–Wiley.

Alenezi, A., Moses, S. A., & Trafalis, T. B. (2008). Real-time prediction of order flowtimes using support vector regression. Computers & Operations Research, 35(11), 3489-3503.

Alex Scroxton. (2019). The nightmare driving test: How to make sure self-driving cars are safe. https://www.computerweekly.com/feature/The-nightmare-driving-test-How-to-make-sure-self-driving-cars-are-safe. Retrieved May,10, 2019.

Al-Masaeid, H. (1997). Impact of pavement condition on rural road accidents, Canadian Journal of Civil Engineering 24(4), 523-531.

Al-Masaeid, H., Sinha, K., & Kuczek, T. (1993). Evaluation of safety impact of highway projects. Transportation Research Record 1401, 9-16.

Amoh-Gyimah, R., Saberi, M., & Sarvi, M. (2017). The effect of variations in spatial units on unobserved heterogeneity in macroscopic crash models. Analytic Methods in Accident Research 13, 28-51.

Amoozadeh, M., Raghuramu, A., Chuah, C. N., Ghosal, D., Zhang, H. M., Rowe, J., & Levitt, K. (2015). Security vulnerabilities of connected vehicle streams and their impact on cooperative driving. IEEE Communications Magazine, 53(6), 126-132.

Anarkooli, A. J., Hosseinpour, M., & Kardar, A. (2017). Investigation of factors affecting the injury severity of single-vehicle rollover crashes: a random-effects generalized ordered probit model. Accident Analysis and Prevention, 106, 399-410.

Anastasopoulos, P.Ch., Sarwar, M.T., & Shankar, V.N. (2016). Safety-oriented pavement performance thresholds: accounting for unobserved heterogeneity in a multi-objective optimization and goal programming approach. Analytic Methods in Accident Research 12, 35-47.

Anderson, J. M., Nidhi, K., Stanley, K. D., Sorensen, P., Samaras, C., & Oluwatola, O. A. (2014). Autonomous vehicle technology: A guide for policymakers. Rand Corporation.

Anderson, J., & Hernandez, S. (2017). Roadway classifications and the accident injury severities of heavy-vehicle drivers. Analytic Methods in Accident Research 15, 17-28.

Anderson, J., Kalra, N., Stanley, K., Sorensen, P., Samaras, C., & Oluwatola, O. (2016). Autonomous Vehicle Technology: A Guide for Policymakers. The Rand Corporation.

Andrey, J. (2010). Long-term trends in weather-related crash risks. Journal of Transport Geography, 18(2), 247-258.

Aneziris, O.N., I.A. Papazoglou, H. Baksteen, M. Mud, B.J. Ale, L.J. Bellamy, A.R. Hale, A. Bloemhoff, and J. Post, & J. Oh. (2008). Quantified risk assessment for fall from height, Saf. Sci., 46(2), 198–220.

Assidiq, A.A., Khalifa, O O., Islam, M.R., & Khan, S. (2008). Real time lane detection for autonomous vehicles. In 2008 International Conference on Computer & Communication Engineering, 82-88. IEEE.

Barua, S., El-Basyouny, K., & Islam, M. (2016). Multivariate random parameters collision count data models with spatial heterogeneity. Analytic Methods in Accident Research 9, 1-15.

Behnood, A., & Mannering, F. (2015). The temporal stability of factors affecting driver-injury severities in single-vehicle crashes: some empirical evidence. Analytic Methods in Accident Research 8, 7-32.

Behnood, A., & Mannering, F. (2016). An empirical assessment of the effects of economic recessions on pedestrian-injury crashes using mixed and latent-class models. Analytic Methods in Accident Research 12, 1-17.

Behnood, A., & Mannering, F. (2017a). The effect of passengers on driver-injury severities in single-vehicle crashes: A random parameters heterogeneity-in-means approach. Analytic Methods in Accident Research 14, 41-53.

Behnood, A., & Mannering, F. (2017b). Determinants of bicyclist injury severities in bicycle-vehicle crashes: A random parameters approach with heterogeneity in means and variances. Analytic Methods in Accident Research 16, 35-47.

Behnood, A., Roshandeh, A. M., & Mannering, F. (2014). Latent class analysis of the effects of age, gender, and alcohol consumption on driver-injury severities. Analytic Methods in Accident Research 3, 56-91.

Bhat, C. R., Astroza, S., & Lavieri, P. S. (2017). A new spatial and flexible multivariate random-coefficients model for the analysis of pedestrian injury counts by severity level. Analytic Methods in Accident Research 16, 1-22.

Bijleveld, F. (2005). The covariance between the number of accidents and the number of victims in multivariate analysis of accident-related outcomes. Accident Analysis and Prevention 37(4), 591-600.

Birdsall, M. (2014). Google and ITE: The road ahead for self-driving cars. Institute of Transportation Engineers. ITE Journal, 84(5), 36.

Bischoff, J., & Maciejewski, M. (2016). Simulation of City-wide Replacement of Private Cars with Autonomous Taxis in Berlin, Elsevier Procedia Computer Science 83, 237-244.

Bishop, C.M. (1995). Neural Networks for Pattern Recognition, Oxford: University Press.

Blincoe, L., Miller, T., Zaloshnja, E., & Lawrence, B. (2015). The economic and societal impact of motor vehicle crashes. National Highway Traffic Safety Administration, Washington, DC. Report No.: DOT HS 8 12 013.

Bogue, S., Paleti, R., & Balan, L. (2017). A modified rank ordered logit model to analyze injury severity of occupants in multivehicle crashes. Analytic Methods in Accident Research 14, 22-40.

Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., ... & Zhang, X. (2016). End to end learning for self-driving cars. arXiv preprint arXiv:1604.07316.

Bonneson, J., Lord, D., Zimmerman, K., Fitzpatrick, K., & Pratt, M. (2007). Development of tools for evaluating the safety implications of highway design decisions. Technical Rep. No. FHWA/TX-07/0-4703-4, Texas Transportation Institute, College Station, TX.

Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

Brooke, A., D. Kendrick, A. Meeraus, & R. Raman. (2005). GAMS: A Users Guide, GAMS Development Corp., Washington, D.C.

Brown, D.B. (1980). Use of dynamic programming in optimally allocating funds for highway safety improvements. Transportation Planning and Technology 6, 131–138.

Buda, M., Maki, A., & Mazurowski, M.A. (2017). A systematic study of the class imbalance problem in convolutional neural networks. arXiv preprint arXiv:1710.05381.

Buddhavarapu, P., Banerjee, A., & Prozzi, J. (2013). Influence of pavement condition on horizontal curve safety. Accident Analysis and Prevention 52, 9–18.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. Sociological methods & research, 33(2), 261-304.

Caliendo, C., Guida, M., & Parisi, A. (2007). A crash-prediction model for multilane roads. Accident Analysis & Prevention, 39(4), 657-670.

Cedar, A., & K. Eldar. (2002). Optimal distance between two branches of uncontrolled split intersection, Transportation Research Part A 36, 699–724.

Chang, J., Su, Y., Huang, T., Kang, S., & Hsieh, S. (2009). Measurement of the International Roughness Index (IRI) Using an Autonomous Robot (P3-AT). 26th International Symposium on Automation and Robotics in Construction (ISARC 2009), Austin, Texas, June 24-27, 2009

Chang, L.Y. (2005). Analysis of freeway accident frequencies: negative binomial regression versus artificial neural network. Safety Science, 43(8), 541-557.

Chapra, S. C., & Canale, R. P. 1998. Numerical methods for engineers (Vol. 2). New York: McGraw-Hill.

Chen, C., Seff, A., Kornhauser, A., & Xiao, J. (2015). Deepdriving: Learning affordance for direct perception in autonomous driving. Proceedings of the IEEE International Conference on Computer Vision, 2722-2730.

Chen, C., Zhang, G., Qian, Z., Tarefder, R. A., & Tian, Z. (2016). Investigating driver injury severity patterns in rollover crashes using support vector machine models. Accident Analysis & Prevention, 90, 128-139.

Chen, E., M. Romero, & Tarko, A. (2012). Indiana 2012 five-percent report. Center for Road Safety, Purdue University, W. Lafayette.

Chen, E., & Tarko, A., (2014). Modeling safety of highway work zones with random parameters and random effects models. Analytic Methods in Accident Research 1, 86–95.

Chen, M., Challita, U., Saad, W., Yin, C., & Debbah, M. (2017). Machine learning for wireless networks with artificial intelligence: A tutorial on neural networks. arXiv preprint arXiv:1710.02913.

Chen, M., Tian, Y., Fortino, G., Zhang, J., & Humar, I. (2018). Cognitive internet of vehicles. Computer Communications, 120, 58-70.

Chen, S., Saeed, T. U., Alinizzi, M., Lavrenz, S., & Labi, S. (2019). Safety sensitivity to roadway characteristics: a comparison across highway classes. Accident Analysis & Prevention, 123, 39-50.

Chen, S., Saeed, T. U., Alqadhi, S. D., & Labi, S. (2017b). Safety impacts of pavement surface roughness at two-lane and multi-lane highways: accounting for heterogeneity and seemingly unrelated correlation across crash severities. Transportmetrica A: Transport Science, 16, 75-89.

Chen, S., Saeed, T. U., & Labi, S. (2017a). Impact of road-surface condition on rural highway safety: A multivariate random parameters negative binomial approach. Analytic Methods in Accident Research, 16, 75-89.

Chen, S., Saeed, T.U., Al-Qadhi, S.D., Labi S., & Sinha, K.C. (2016). Safety impacts of pavement rehabilitation at multilane highways. Eighth International Conference on Maintenance and Rehabilitation of Pavements (MAIREPAV8), Singapore.

Cheng, X., Chen, C., Zhang, W., & Yang, Y. (2017). 5G-enabled cooperative intelligent vehicular (5GenCIV) framework: When Benz meets Marconi. IEEE Intelligent Systems, 32(3), 53-59.

Chin, H. C., & Quddus, M. A. (2003). Modeling count data with excess zeroes: an empirical application to traffic accidents. Sociological Methods & Research, 32(1), 90-116.

Chiou, Y. C., & Fu, C. (2015). Modeling crash frequency and severity with spatiotemporal dependence. Analytic Methods in Accident Research 5, 43-58.

Chiou, Y. C., Fu, C., & Chih-Wei, H. (2014). Incorporating spatial dependence in simultaneously modeling crash frequency and severity. Analytic Methods in Accident Research 2, 1-11.

Cho, H., Seo, Y. W., Kumar, B. V., & Rajkumar, R. R. (2014, May). A multi-sensor fusion system for moving object detection and tracking in urban driving environments. In 2014 IEEE International Conference on Robotics and Automation (ICRA) (pp. 1836-1843). IEEE.

Chong, M., Abraham, A., & Paprzycki, M. (2005). Traffic accident analysis using machine learning paradigms. Informatica, 29(1).

Chu, K., Lee, M., & Sunwoo, M. (2012). Local path planning for off-road autonomous driving with avoidance of static obstacles. IEEE Transactions on Intelligent Transportation Systems, 13(4), 1599-1616.

Cigna, C. (2008). Safety Costs: A real issue or opinion? A critical analysis based on case history, Chemical Engineering Transactions 13, 17–22.

Cleveland, D. (1987). Effect of resurfacing on highway safety, in relationship between safety and key highway features: A synthesis of prior research, State of the Art Report 6, Transportation Research Board, Washington, DC.

Coruh, E., Bilgic, A., & Tortum, A. (2015). Accident analysis with aggregated data: The random parameters negative binomial panel count data model. Analytic Methods in Accident Research 7, 37-49.

Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., ... & Tsirigos, A. (2018). Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. Nature medicine, 24(10), 1559.

Council, F., & Stewart, J. (1999). Safety effects of the conversion of rural two-lane to four-lane roadways based on cross-sectional models. Paper presented at the 78th Annual Meeting of the Transportation Research Board, Washington, DC.

Cox, M. (1975). Reserved bus lanes in Dallas, Texas. Journal of Transportation engineering, 101(ASCE# 11720 Proceeding).

Cozzani V., & E. Salzano. (2004). The quantitative assessment of domino effect caused by overpressure. Part II: case-studies, Journal of Hazardous Materials 107, 81–94.

Dai, H., & Cao, Z. (2017). A Wavelet Support Vector Machine-Based Neural Network Metamodel for Structural Reliability Assessment. Computer-Aided Civil and Infrastructure Engineering, 32(4), 344-357.

Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. Paper presented at the Proceedings of the 23rd international conference on Machine learning.

Dikmen, M., & Burns, C. (2017). Trust in autonomous vehicles: The case of Tesla Autopilot and Summon, IEEE International Conference on Systems, Man, and Cybernetics, Banff, Canada 0.1109/SMC.2017.8122757

Dinu, R., & Veeraragavan, A. (2011). Random parameter models for accident prediction on two-lane undivided highways in India. Journal of Safety Research 42 (1), 39-42.

Dong, C., S.S. Nambisan, S.H. Richards, & Ma, Z. (2015). Assessment of the effects of highway geometric design features on the frequency of truck involved crashes using bivariate regression, Transportation Research Part A 75, 30–41.

Donnell, E. T., & Mason Jr, J. M. (2006). Predicting the frequency of median barrier crashes on Pennsylvania interstate highways. Accident Analysis & Prevention, 38(3), 590-599.

Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., & Koltun, V. (2017). CARLA: An open urban driving simulator. arXiv preprint arXiv:1711.03938.

Duda, R., Hart, P., & Stork, D. (2001). Pattern Classification. 2nd Ed., Wiley-Interscience, New York, 55.

El-Assaly, A., & Hempsey, L. (2000). The relationship between pavement width and highway safety on two-lane paved highway in Alberta, Canada. ASCE Construction Congress VI: Building Together for a Better Tomorrow in an Increasingly Complex World, 671-679).

El-Basyouny, K., Barua, S., Islam, M., & Li, R. (2014b). Assessing the effect of weather states on crash severity and type by use of full Bayesian multivariate safety models. Transportation Research Record 2432, 65-73.

El-Basyouny, K., Barua, S., & Islam, T. (2014a). Investigation of time and weather effects on crash types using full Bayesian multivariate Poisson lognormal models. Accident Analysis and Prevention 73, 91-99.

El-Basyouny, K., & Sayed, T. (2009). Accident prediction models with random corridor parameters. Accident Analysis and Prevention 41(5), 1118-1123.

Eluru, N., Paleti, R., Pendyala, R., & Bhat, C. (2010). Modeling injury severity of multiple occupants of vehicles: Copula-based multivariate approach. Transportation Research Record 2165, 1-11.

Everett, S. R., Athigakunagorn, N., Woldermariam, W., Varadarajan, V., Arman, M., Roshandeh, A. M., ... & Sinha, K. C. (2014). Impact of HB-1481 on Indiana's highway revenue generation, asset degradation, modal distribution, and economic development and competitiveness (No. FHWA/IN/JTRP-2014/14). Purdue University. Joint Transportation Research Program.

Fang, Q., Li, H., Luo, X., Ding, L., Luo, H., & Li, C. (2018). Computer vision aided inspection on falling prevention measures for steeplejacks in an aerial environment. Automation in Construction 93, 148-164.

Farah, H., S. Bekhor, A. Polus, & Toldeo, T. (2009). A passing gap acceptance model for two-lane rural highways. Transportmetrica 5(3), 159–172.

FHWA. (1995). Interactive highway safety design model: Accident predictive module. Public Roads 59(2).

FHWA. (2006). Low-cost treatments for horizontal curve safety, Report No. FHWA-SA-07-002, prepared by Vanasse Hangen Brustlin, Inc and Transportation Research Corporation for the Federal Highway Administration, Washington, DC.

FHWA. (2014). Facts and statistics, office of safety. Federal Highway Administration, Washington, DC.

FHWA. (2015). Highway Statistics Table FI-30. FHWA Office of Highway Policy Information. https://www.fhwa.dot.gov/policyinformation/statistics/2014/fi30.cfm

FHWA. (2013). FHWA Functional Classification Guidelines: Section II - Concepts, Definitions, and System Characteristics. Retrieved from U.S. Department of Transportation, http://www.fhwa.dot.gov/planning/processes/statewide/related/functional_classification

FHWA. (2016). Local and Rural Road Safety Program. Retrieved from FHWA Office of Safety Programs, Federal Highway Administration, U.S. Department of Transportation, http://safety.fhwa.dot.gov/local_rural/

Fildes, B. N., & Lee, S.J. (1993). The speed review: Road environment, behavior, speed limits, enforcement and crashes. Report Nos. CR 127 (FORS), CR 3/93 (RSB). Federal Office of Road Safety, Roads and Traffic Authority, Sydney, Canberra, Australia.

Fitzpatrick, K., Harwood, D., Anderson, & I., Balke, K. (2000). Accident mitigation guide for congested rural two-lane highways, NCHRP Report 440. Transportation Research Board, Washington DC.

Fitzpatrick, K., W.H. Schneider IV, & Park, E. (2005). Comparisons of crashes on rural two-lane and four-lane highways in Texas. Texas Transportation Inst., College Station, TX.

Friedman, J., Hastie, T., & Tibshirani, R. (2001). The elements of statistical learning (Vol. 1): Springer series in statistics New York, NY, USA:.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software, 33(1), 1.

Gårder, P. (2006). Segment characteristics and severity of head-on crashes on two-lane rural highways in Maine. Accident Analysis & Prevention, 38(4), 652–661.

Garnowski, M., & Manner, H. (2011). On factors related to car accidents on German Autobahn connectors. Accident Analysis and Prevention 43(5), 1864-1871.

Geni, B., & Peng, B. (2008). Roadway departure crashes: how can they be reduced? Institute of Transportation Engineers ITE Journal 78(12), 44–48.

Ghosh-Dastidar, S., & Adeli, H. (2003). Wavelet clustering – neural network model for freeway incident detection. Computer-Aided Civil and Infrastructure Engineering, 18(5), 325-338.

Giarratana, C. (2016). How AI is driving the future of autonomous cars. https://readwrite.com/2016/12/20/ai-driving-future-autonomous-cars-tl4/. Retrieved May,10, 2019.

Gitelman, V., Doveh, E. & Hakkert, S. (2010). Designing a composite indicator for road safety, Safety Science, 48(9), 1212–1224.

Gkritza, K., & Mannering, F. (2008). Mixed logit analysis of safety-belt use in single-and multi-occupant vehicles. Accident Analysis and Prevention 40(2), 443-451.

Gomes, M. J., Cunto, T.L., & Silva, A. R. (2017). Geographically weighted negative binomial regression applied to zonal level safety performance models. Accident Analysis and Prevention 106, 254-261.

Goodall, N. J. (2014). Ethical decision making during automated vehicle crashes. Transportation Research Record, 2424(1), 58-65.

Google Cloud. (2018). Advanced Guide to Inception v3 on Cloud TPU. https://cloud.google.com/tpu/docs/inception-v3-advanced Accessed: 08/25/2018

Gordon, T.J., & Lidberg, M. (2015). Automated driving and autonomous functions on road vehicles, Vehicle System Dynamics 53(7), 958-994.

Grazioli, F., Kusmenko, E., Roth, A., Rumpe, B., & von Wenckstern, M. (2017). Simulation framework for executing component and connector models of self-driving vehicles. In MODELS (Satellite Events) (pp. 109-115).

Greaves, S. P., & Ellison, A. B. (2011). Personality, risk aversion and speeding: An empirical investigation. Accident Analysis & Prevention, 43(5), 1828-1836.

Greenblatt, J.B. & Saxena, S. (2015). Autonomous taxis could greatly reduce greenhouse-gas emissions of US light-duty vehicles. Nature Climate Change 5, 860–863.

Greenblatt, N. A. (2016). Self-driving cars and the law. IEEE spectrum, 53(2), 46-51.

Greene, W. (2008). Econometric Analysis, 6th Ed., Upper Saddle River, NJ: Prentice-Hall.

Gross, F., & Donnell, E.T. (2011). Case-control and cross-sectional methods for estimating crash modification factors: Comparisons from roadway lighting and lane and shoulder width safety effect studies. Journal of Safety Research, 117–129.

Gross, F., & Jovanis, P.P. (2007). Estimation of the safety effectiveness of lane and shoulder width: Case-control approach. ASCE Journal of Transportation Engineering, 362–369.

Gross, F., Jovanis, P.P., & Eccles, K. (2009). Safety effectiveness of lane and shoulder width combinations on rural, two-lane, undivided roads. Transp. Res. Record 2103, 42–49.

Hadi, M.A., Aruldhas J., Chow, L.F., & Wattleworth, J.A. (1995). Estimating safety effects of cross-section design for various highway types using negative binomial regression. Transportation Research Record 1500, 169–177.

Hadley, G. (1964). Nonlinear and Dynamic Programming, Addison-Wesley, Reading, Mass.

Hale, A. (2009). Why safety performance indicators? Saf. Sci., 47(4), 479–480.

Hall, J., Rutman, E., & Brogan, J. (2003). Highway safety challenges on low-volume rural Roads. Proceedings of Institute of Transportation Engineers Annual Meeting, Seattle, WA.

Hancock, P. A., Nourbakhsh, I., & Stewart, J. (2019). On the future of transportation in an era of automated and autonomous vehicles. Proceedings of the National Academy of Sciences, 201805770.

Hankey, J.M., Wielwille, W.W., Cannell, W.J., Kieliszewski, C. A., Medina, A., Dingus, T., & Cooper, L.M. (1999). Identification and evaluation of driver errors: Driver error taxonomy development, Tech. Rep. prepared by the Virginia Tech. Transp. Institute for the Federal Highway Administration, Washington, DC.

Hartmann, S. (1996). The World as a Process: Simulations in the Natural and Social Sciences. In R. Hegselmann, et al. Eds. Modeling and Simulation in the Social Sciences from the Philosophy of Science Point of View. Theory and Decision Library. Kluwer, Dordrecht, 77–100.

Harwood, D., Kohlman Rabbani, E., Richard, K., McGee, H., & Gittings, G. (2003). System-wide impact of safety and traffic operations design decisions of resurfacing, restoration, or rehabilitation Projects. NCHRP Report 486, Transportation Research Board, Washington, D.C.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York: Springer-Verlag.

Hauer, E. (2010). On prediction in road safety, Safety Science, 48(9), 1111–1122.

He, K., Girshick, R., & Dollár, P. (2018). Rethinking ImageNet Pre-training. ArXiv:1811.08883

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (770-778).

Hecht, J. (2018). Lidar for self-driving cars. Optics and Photonics News, 29(1), 26-33

Hee Lee, G., Faundorfer, F., & Pollefeys, M. (2013). Motion estimation for self-driving cars with a generalized camera. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2746-2753).

Heron, M. (2016). Deaths: Leading causes for 2014. National Vital Statistics Reports 65(5), Centers for Disease Control and Prevention, National Center for Health Statistics.

Heydari, S., Fu, L., Joseph, L., & Miranda-Moreno, L. F. (2016). Bayesian nonparametric modeling in transportation safety studies: Applications in univariate and multivariate settings. Analytic Methods in Accident Research 12, 18-34.

Heydari, S., Fu, L., Miranda-Moreno, L. F., & Joseph, L. (2017). Using a flexible multivariate latent class approach to model correlated outcomes: A joint analysis of pedestrian and cyclist injuries. Analytic Methods in Accident Research 13, 16-27.

Hinton, G.E., Osindero, S., & Teh, Y.W. (2006). A fast learning algorithm for deep belief nets. Neural Computation 18(7), 1527-1554.

Hinze, J., Thurman, S., & Wehle, A. (2013). Leading indicators of construction safety performance, Saf. Sci., 51, 23–28.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

Hong, J., Shankar, V. N., & Venkataraman, N. (2016). A spatially autoregressive and heteroskedastic space-time pedestrian exposure modeling framework with spatial lags and endogenous network topologies. Analytic Methods in Accident Research 10, 26-46. http://www.publications.usace.army.mil/Portals/76/Publications/EngineerManuals

Huang, H., Zhou, H., Wang, J., Chang, F., & Ma, M. (2017). A multivariate spatial model of crash frequency by transportation modes for urban intersections. Analytic Methods in Accident Research 14, 10-21.

Huang, S. S., Chen, C. J., Hsiao, P. Y., & Fu, L. C. (2004). On-board vision system for lane recognition and front-vehicle detection to enhance driver's awareness. In IEEE International Conference on Robotics and Automation. Proceedings. ICRA'04. 2004 (Vol. 3, 2456-2461). IEEE.

INDOT. (2000). Pavement Surface Report, December, 2000. INDOT Pavement Management Section, Program Development Division.

INDOT. (2003). Indiana Safety Management System Database, Safety Management Unit, Program Development Division, Indiana Department of Transportation. Indianapolis, IN.

INDOT. (2013). Chapter 52: Pavement and Underdrain Design Elements. Retrieved from Indiana Design Manual: http://www.in.gov/indot/design_manual/files/Ch52_2013.pdf

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167.

Irfan, M., Khurshid, M.B., Bai, Q., Morin, T.L., & Labi, S. (2011). Establishing optimal project-level strategies for pavement maintenance and rehabilitation. Journal of Engineering Optimization 44(5), 565-589.

Ivers, R., Senserrick, T., Boufous, S., Stevenson, M., Chen, H. Y., Woodward, M., & Norton, R. (2009). Novice drivers' risky driving behavior, risk perception, and crash risk: findings from the DRIVE study. American journal of public health, 99(9), 1638-1644.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112): Springer.

Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. Intelligent Data Analysis, 6(5), 429-449.

Jenn, D. C. (2019). Radar and laser cross section engineering. American Institute of Aeronautics and Astronautics, Inc..

Jha, M.K., and P. Schonfeld. (2004). highway alignment optimization model using geographic information systems. Transportation Research Part A 38, 455–481.

Jiang, X., & Adeli, H. (2007). Pseudospectra, MUSIC, and dynamic wavelet neural network for damage detection of highrise buildings. International Journal for Numerical Methods in Engineering, 71(5), 606-629.

Jiang, Y., & Sinha, K. C. (1989). Bridge service life prediction model using the Markov chain. Transportation Research Record, 1223, 24-30.

Jonah, B. A., Thiessen, R., & Au-Yeung, E. (2001). Sensation seeking, risky driving and behavioral adaptation. Accident Analysis & Prevention, 33(5), 679-684.

Jovanis, P. & Gross, P. (2007). Estimation of safety effectiveness of changes in shoulder width with case control and cohort methods. Transportation Research Record 237–245.

Kalra, N., & Paddock, S. M. (2016). Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?. Transportation Research Part A: Policy and Practice, 94, 182-193.

Kandhai, K., Smith, M., & Kannah, A. (2011). Immersive driving simulation for driver education and analysis. 16th International Conference on Computer Games, IEEE, Louisville, KY. 10.1109/CGAMES.2011.6000355

Kerner, B. S. (2016). Failure of classical traffic flow theories: Stochastic highway capacity and automatic driving. Physica A: Statistical Mechanics and its Applications, 450, 700-747.

Keskar, N. S., Mudigere, D., Nocedal, J., Smelyanskiy, M., & Tang, P. T. P. (2016). On large-batch training for deep learning: Generalization gap and sharp minima. arXiv preprint arXiv:1609.04836.

Kim, J.-K., Ulfarsson, G. F., Shankar, V. N., & Mannering, F. L. (2010). A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. Accident Analysis & Prevention, 42(6), 1751-1758.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kjellen, U. (2009). The safety measurement problem revisited, Saf. Sci., 47(4), 486–489.

Knight, S. (2018). Accidents have eroded public trust in self-driving vehicle technology, Techspot, www.techspot.com/news/74752.

Knuiman, M., Council, F., & Reinfurt, D. (1993). Association of median width and highway accident rates. Transportation Research Record 1401, 70-82.

Konduri, S., Labi, S., & Sinha, K. (2003). Incident occurrence models for freeway incident management. Transportation Research Record 1856, 125-130.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems (1097-1105).

Labi, S. (2014). Introduction to civil engineering systems: A systems perspective to the development of civil engineering facilities. John Wiley & Sons.

Labi, S. (2016). The tri-zone concept: exploring the relationships between the intensity of safety features and safety performance. ASCE International Conference on Transportation & Development, Houston, TX.

Labi, S., (2006). Effects of geometric characteristics of rural two-lane roads on safety. Technical Report FHWA/IN/JTRP-2005/2, Joint Transportation Research Program, W. Lafayette, IN.

Labi, S., (2011). Efficacies of roadway safety improvements across functional subclasses of rural two-lane highways, Journal of Safety Research 42, 231-239.

Labi, S., Chen, S., Preckel, P. V., Qiao, Y., & Woldemariam, W. (2017). Rural two-lane highway shoulder and lane width policy evaluation using multiobjective optimization. Transportmetrica A: Transport Science, 13(7), 631-656.

Labi, S., Rodriguez, M. M., & Sinha, K.C. (2006). Assessing preservation needs for a bridge network: A comparison of alternative approaches. Structure & Infrastructure Engineering 4(3), 221-235.

Labi, S., Saeed, T.U., Volovski, M., & Alqadhi, S.D. (2015). An exploratory discussion of the impacts of driverless vehicle operations on the man-made environment. 1st International Conference on Mechanical and Transportation Engineering. Kuala Lumpur, Malaysia.

Lamptey, G. (2004). Development of analytical and software tools for highway safety management, M.S. Thesis, Purdue University, W. Lafayette, IN.

Lamptey, G., L. Singh, S. Labi, & Sinha, K. (2010). Systematic framework for incorporating safety in network-level transportation planning. J. Transp. Eng. 136, 436–444.

Lamptey, G., Labi, S., & Li, Z. (2008). Decision support for optimal scheduling of highway pavement preventive maintenance within resurfacing cycle. Elsevier Decision Support Systems 46(1), 376-387.

Lamptey, G., Labi, S., & Sinha, K. (2005). Investigating the sensitivity of optimal network safety needs to key safety management inputs. Transportation Research Record 1922, 52-61.

Lamptey, G., Samdariya, A. J., Labi, S., & Sinha, K. C. (2011). Indiana’s safety management system software package 1 (SMSS) A review. In 3rd International Conference on Road Safety and Simulation.

Law, W.K., Chan, A.H. & Pun, K. (2006). Prioritizing the safety management elements. A hierarchical analysis for manufacturing enterprises, Ind. Mng. Data Syst., 106(6), 778–792.

LeBlanc, D. J., Johnson, G. E., Venhovens, P. J. T., Gerber, G., DeSonia, R., Ervin, R. D., ... & Pilutti, T. E. (1996). CAPC: A road-departure prevention system. IEEE Control Systems Magazine, 16(6), 61-71.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature 521(7553), 436.

LeCun, Y., Bottou, L., Orr, G.B., & Müller, K.R. (2012). Efficient backprop. In Neural networks: Tricks of the trade (9-48). Springer, Berlin, Heidelberg.

Lee, J., & Mannering, F. (2002). Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. Accident Analysis & Prevention, 34(2), 149-161.

Lee, J., Abdel-Aty, M., & Jiang, X. (2015). Multivariate crash modeling for motor vehicle and non-motorized modes at the macroscopic level. Accident Analysis and Prevention 78, 146-154.

Less, T., & Adeli, H. (2010). Computational earthquake engineering of bridges. Scientia Iranica. Transaction A, Civil Engineering, 17(5), 325.

Li, X., Lord, D., Zhang, Y., & Xie, Y. (2008). Predicting motor vehicle crashes using support vector machine models. Accident Analysis & Prevention, 40(4), 1611-1618.

Li, Y., Liu, C., & Ding, L. (2013). Impact of pavement conditions on crash severity. Accident Analysis and Prevention 59, 399-406.

Li, Z., Liu, P., Wang, W., & Xu, C. (2012). Using support vector machine models for crash injury severity analysis. Accident Analysis & Prevention, 45, 478-486.

Lin, Y.Z., Nie, Z. H., & Ma, H.W. (2017). Structural damage detection with automatic feature‐extraction through deep learning. Computer‐Aided Civil and Infrastructure Engineering, 32(12), 1025-1046.

Litman, T. (2017). Autonomous vehicle implementation predictions (p. 28). Victoria, Canada: Victoria Transport Policy Institute.

Lord, D., Washington, S., & Ivan, J. M. (2004). Statistical challenges with modeling motor vehicle crashes: understanding the implications of alternative approaches: Texas A & M University System, Texas Transportation Institute.

Lu, Z., Li, J., & Minghui, Z. (2010, March). Intelligent control research based on the smart car. In 2010 2nd International Conference on Advanced Computer Control (Vol. 2, pp. 292-297). IEEE.

Luo, H., Xiong, C., Fang, W., Love, P.E., Zhang, B., & Ouyang, X. (2018). Convolutional neural networks: Computer vision-based workforce activity assessment in construction. Automation in Construction 94, 282-289.

Ma, J., Hu, J., Hale, D. K., & Bared, J. (2016). Dynamic hard shoulder running for traffic incident management. Transportation Research Record: Journal of the Transportation Research Board(2554), 120-128.

Ma, J., Kockelman, K. M., & Damien, P. (2008). A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. Accident Analysis & Prevention, 40(3), 964-975.

Ma, J., & Kockelman, K. (2006). Bayesian multivariate Poisson regression for models of injury count by severity. Transportation Research Record 1950, 24-34.

Ma, X., Chen, S., & Chen, F. (2017). Multivariate space-time modeling of crash frequencies by injury severity levels. Analytic Methods in Accident Research 15, 29-40.

Macadam, C.C. (2003). Understanding and modeling the human driver. Vehicle System Dynamics, 40(1-3), 101-134.

Machin, M. A., & Sankey, K. S. (2008). Relationships between young drivers' personality characteristics, risk perceptions, and driving behaviour. Accident analysis & prevention, 40(2), 541-547.

Mahoney, K, Julian, F., & Taylor, H. (2006). Good practices: Incorporating safety into resurfacing and restoration projects. FHWA-SA-07-001, Federal Highway Administration, Washington, DC.

Malyshkina, N. V., & Mannering, F. L. (2009). Markov switching multinomial logit model: an application to accident-injury severities. Accident Analysis & Prevention, 41(4), 829-838.

Mannering, F. (2009). An empirical analysis of driver perceptions of the relationship between speed limits and safety. Transportation Research Part F: Traffic Psychology and Behavior, 12(2), 99–106.

Mannering, F. L., & Bhat, C. R. (2014). Analytic methods in accident research: Methodological frontier and future directions. Analytic methods in accident research, 1, 1-22.

Mannering, F. L., Shankar, V., & Bhat, C. R. (2016). Unobserved heterogeneity and the statistical analysis of highway accident data. Analytic methods in accident research, 11, 1-16.

Martinez, M., Sitawarin, C., Finch, K., Meincke, L., Yablonski, A., & Kornhauser, A. (2017). Beyond grand theft auto v for training, testing and enhancing deep learning in self driving cars. arXiv preprint arXiv:1712.01397.

Mayora, J., & Piña, R. (2009). An assessment of the skid resistance effect on traffic safety under wet-pavement conditions. Accident Analysis and Prevention 41(4), 881-886.

Melachrinoudis, E., & Kozanidis, G. (2002). A mixed integer knapsack model for allocating funds to highway safety improvements. Transportation Research Part A 36, 789–803.

Mikusova, M. (2017). Crash avoidance systems and collision safety devices for vehicle occupants. MATEC Web of Conferences 107, 00024 (2017) DOI: 10.1051/matecconf/2017107000

Miller, J. (2018). Autonomous vehicle makers want to set their own standards. Autonomous Vehicle Technology, Axios, www.axios.com.

Milton, J. C., Shankar, V. N., & Mannering, F. L. (2008). Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accident Analysis & Prevention, 40(1), 260-266.

Milton, J., & Mannering, F. (1998). The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies. Transportation, 25(4), 395-413.

Minnesota Department of Transportation (MNDOT). (1997). Highway spending: A program evaluation report. Office of the Legislative Auditor of the State of Minnesota, St. Paul, MN.

Mishra, S., Golias, M.M., Sharma, S., & Boyles, S. (2015). Optimal funding allocation strategies for safety improvements on urban intersections. Transp. Research Part A 75, 113–133.

Mothafer, G. I., Yamamoto, T., & Shankar, V. N. (2017). A negative binomial crash sum model for time invariant heterogeneity in panel crash data: some insights. Analytic Methods in Accident Research 14, 1-9.

Murillo-Hoyos, J., Athigakunagorn, N., & Labi, S. (2015). Methodology for safety improvement programming using constrained network-level optimization. Transportation Research Part C 50, 106-116.

Mussone, L., Ferrari, A., & Oneta, M. (1999). An analysis of urban collisions using an artificial intelligence model. Accident Analysis & Prevention, 31(6), 705-718.

Nabian, M. A., & Meidani, H. (2018). Deep learning for accelerated seismic reliability analysis of transportation networks. Computer-Aided Civil and Infrastructure Engineering 33(6), 443-458.

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th international conference on machine learning (ICML-10) (807-814).

Naranjo, J. E., Gonzalez, C., Garcia, R., & De Pedro, T. (2008). Lane-change fuzzy control in autonomous vehicles for the overtaking maneuver. IEEE Transactions on Intelligent Transportation Systems, 9(3), 438-450.

National Center for Statistics and Analysis. (2015). Rural/urban comparison: 2013 data. (Traffic Safety Facts, Report No. DOT HS812181). National Highway Traffic Safety Administration, Washington, DC.

NHTSA. (2017). 2016 Fatal Motor Vehicle Crashes: Overview. Traffic safety facts research note, October 2017. https://crashstats.nhtsa.dot.gov/Api/Public/Publication/812456/ Accessed July 2018.

NHTSA. (2016). Traffic Safety Facts 2013 Data: Rural/Urban Comparison. Retrieved from U.S. Department of Transportation.

NHTSA. (2013a). Preliminary Statement of Policy Concerning Automated Vehicles National Highway Traffic Safety Administration, Washington, DC.

NHTSA. (2013b). Traffic Safety Facts 2012, National Highway Traffic Safety Administration, Washington, DC.

NVIDIA. (2018). Self-driving safety report. https://www.nvidia.com/content/dam/en-zz/Solutions/self-driving-cars/safety-report/NVIDIA-Self-Driving-Safety-Report-2018.pdf Retrieved May,12, 2019.

Øien, K., Utne, I.B., Tinmannsvik, R.K., & Massaiu, S. (2011). Building safety indicators: part 2 – application, practices and results, Saf. Sci., 49(2), 162–171.

Örnek, E. & Drakopoulos, A. (2007). Analysis of run-off-road crashes in relation to roadway features and driver behavior. In Proceedings of the 2007 Mid-Continent Transportation Research Symposium, Ames, IA.

Pal, R., & Sinha, K.C. (1998). Optimization approach to highway safety improvement programming. Transportation Research Record 1640, 1–9.

Pan, X., You, Y., Wang, Z., & Lu, C. (2017). Virtual to real reinforcement learning for autonomous driving. arXiv preprint arXiv:1704.03952.

Panakkat, A., & Adeli, H. (2009). Recurrent neural network for approximate earthquake time and location prediction using multiple seismicity indicators. Computer-Aided Civil and Infrastructure Engineering, 24(4), 280-292.

Pande, A., & Abdel-Aty, M. (2008). A computing approach using probabilistic neural networks for instantaneous appraisal of rear-end crash risk. Computer-Aided Civil and Infrastructure Engineering 23(7), 549-559.

Park, E., & Lord, D. (2007). Multivariate Poisson-lognormal models for jointly modeling crash frequency by severity. Transportation Research Record: Journal of the Transportation Research Board (2019), 1-6.

Park, J., (2016). Safety effects for widening shoulder width on rural multi-lane roadways 4 in developing crash modification functions using multivariate adaptive regression splines. Safety 2, 3–12.

Park, J., Abdel-Aty, M., Wang, J.H., & Lee, C. (2015). Assessment of safety effects for widening urban roadways in developing crash modification functions using nonlinearizing link functions. Accident Analysis & Prevention 79, 80–87.

Paterson, W. D. (1987). Road deterioration and maintenance effects: Models for planning and management.

Patole, S. M., Torlak, M., Wang, D., & Ali, M. (2017). Automotive radars: A review of signal processing techniques. IEEE Signal Processing Magazine, 34(2), 22-35.

Pham, M.-H., Bhaskar, A., Chung, E., & Dumont, A.G. (2010). Random forest models for identifying motorway rear-end crash risks using disaggregate data. Paper presented at the Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on.

Podgorski, D. (2015). Measuring operational performance of OSH management system – A demonstration of AHP-based selection of leading key performance indicators, Safety Science 73, 146–166.

Powers, D. M. (2011). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.

Prato, C. G., Rasmussen, T. K., & Kaplan, S. (2014). Risk factors associated with crash severity on low-volume rural roads in Denmark. Journal of Transportation Safety & Security, 6(1), 1-20.

Provost, F. J., Fawcett, T., & Kohavi, R. (1998). The case against accuracy estimation for comparing induction algorithms. Paper presented at the ICML.

Pulugurtha, S. S., Ogunro, V., Pando, M. A., Patel, K. J., & Bonsu, A. (2013). Preliminary Results towards Developing Thresholds for Pavement Condition Maintenance: Safety Perspective. Procedia - Social and Behavioral Sciences, 104, 302-311.

Qiao, Y., Chen, S., Alinizzi, M., & Labi, S. (2018, July). Modeling the relationships between pavement distress and performance. In Advances in Materials and Pavement Prediction: Papers from the International Conference on Advances in Materials and Pavement Performance Prediction (AM3P 2018), April 16-18, 2018, Doha, Qatar (p. 11). CRC Press.

Qiao Y, Saeed TU, Chen S, Nateghi R, & Labi S (2018) Acquiring insights into infrastructure repair policy using discrete choice models. Transportation Research Part A: Policy and Practice. 113, 491–508.

Rajasekaran, S., Gayathri, S., & Lee, T.L. (2008). Support vector regression methodology for storm surge predictions. Ocean Engineering, 35(16), 1578-1587.

Ramos, S., Gehrig, S., Pinggera, P., Franke, U., & Rother, C. (2017, June). Detecting unexpected obstacles for self-driving cars: Fusing deep learning and geometric modeling. In 2017 IEEE Intelligent Vehicles Symposium (IV) (pp. 1025-1032). IEEE.

Reiman, T. & Pietikäinen, E. (2012). Leading indicators of system safety – monitoring and driving the organizational safety potential, Saf. Sci., 50(10), 1993–2000.

Reimer, B. (2014). Driver assistance systems and the transition to automated vehicles: A path to increase older adult safety and mobility?. Public Policy & Aging Report, 24(1), 27-31.

Riviere, C., Lauret, P., Ramsamy, J. M., & Page, Y. (2006). A Bayesian neural network approach to estimating the energy equivalent speed. Accident Analysis & Prevention, 38(2), 248-259.

Robson, L.S., Clarke, J.A., Cullen, K., Bielecky, A., Severin, C., Bigelow, P.L., & Mahood. Q. (2007). The effectiveness of occupational health and safety management system interventions: a systematic review, Saf. Sci., 45(3), 329–353.

Rumar, K. (1985). The Role of Perceptual and Cognitive Filters in Observed Behavior, Human Behavior and Traffic Safety, Eds.: L. Evans and R. Schwing, Plenum Press, 164–165.

Russo, B., Savolainen, P., Schneider, W., & Anastasopoulos, P.Ch. (2014). Comparison of factors affecting injury severity in angle collisions by fault status using a random parameters bivariate ordered probit model. Analytic Methods in Accident Research 2, 21-29.

Saeed, TU., Qiao, Y., Chen, S., Gkritza, K., & Labi, S. (2017). Methodology for probabilistic modeling of highway bridge infrastructure condition: accounting for improvement effectiveness and incorporating random effects. Journal of Infrastructure Systems. 23(4)

Saeed, TU., Nateghi, R., Hall, T., & Waldorf, B. (2019a). Statistical analysis of area-wide alcohol-related driving crashes: a spatial econometric approach. Geographical Analysis (accepted, in press)

Saeed, TU., Burris, M., Labi, S., & Sinha, KC. (2019b). Forecasting autonomous vehicle use using consumers' preferences: An empirical discourse. Working paper. Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA.

Saeed, TU., Labi, S., Sinha, K.C. (2019c). Impacts of autonomous vehicle operations on highway infrastructure and roadway design. Working paper. Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA.

Saeed, T.U. (2019). Road infrastructure readiness for autonomous vehicles. PhD Dissertation. Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA.

Salmon, P.M., Regan, M.A., & Johnston, I. (2005). Human error and road transport: Phase one– Literature review. No. 256. 2005.

Sarwar, M. T., Anastasopoulos, P. C., Golshani, N., & Hulme, K. F. (2017). Grouped random parameters bivariate probit analysis of perceived and observed aggressive driving behavior: a driving simulation study. Analytic methods in accident research, 13, 52-64.

Sarwar, M. T., Anastasopoulos, P.Ch., Golshani, N., & Hulme, K. F. (2017). Grouped random parameters bivariate probit analysis of perceived and observed aggressive driving behavior: a driving simulation study. Analytic Methods in Accident Research 13, 52-64.

Sarwar, M. T., Fountas, G., & Anastasopoulos, P.Ch. (2017). Simultaneous estimation of discrete outcome and continuous dependent variable equations: A bivariate random effects modeling approach with unrestricted instruments. Analytic Methods in Accident Research 16, 23-34.

Sawalha, Z., & Sayed, T. (2001). Evaluating safety of urban arterial roadways. Journal of Transportation Engineering 127(2), 151–158.

Sayers, M., Gillespie, T., & Paterson, W. (1986). Guidelines for conducting and calibrating road roughness measurements, World Bank Technical Paper 46, The World Bank, Washington, D.C.

Sayers, M., & Steven, M. (1998). The little book of profiling. University of Michigan Transportation Research Institute (UMTRI), Ann Arbor, MI.

Scholkopf, B., & Smola, A. J. (2001). Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press.

Schrock, S., Parsons, R., & Zeng, H. (2011). Estimation of safety effectiveness of widening shoulders and adding passing lanes on rural two-lane roads. Transp. Res. Rec. 2203, 57–63.

Seraneeprakarn, P., Huang, S., Shankar, V., Mannering, F., Venkataraman, N., & Milton, J. (2017). Occupant injury severities in hybrid-vehicle involved crashes: A random parameters approach with heterogeneity in means and variances. Analytic Methods in Accident Research 15, 41-55.

Serhiyenko, V., Mamun, S. A., Ivan, J. N., & Ravishanker, N. (2016). Fast Bayesian inference for modeling multivariate crash counts. Analytic Methods in Accident Research 9, 44-53.

Sgouru, E., Katsakiori, P., Goutsos, S., & Manatakis, E. (2010). Assessment of selected safety performance evaluation methods in regards to their conceptual, methodological and practical characteristics, Saf. Sci., 48(8), 1019–1025.

Shankar, V., Milton, J., & Mannering, F. (1997). Modeling accident frequencies as zero-altered probability processes: an empirical inquiry. Accident Analysis and Prevention 29(6), 829-837.

Shi, P., & Valdez, E. (2014). Multivariate negative binomial models for insurance claim counts. Insurance 55, 18-29.

Siddique, N., & Adeli, H. (2013). Computational intelligence: synergies of fuzzy logic, neural networks and evolutionary computing. John Wiley & Sons.

Silva, M. M., & Liautaud, G. (2011). Performance-based Road Rehabilitation and Maintenance Contracts (CREMA) in Argentina: A Review of Fifteen Years of Experience (1996-2010) (No. TP36).

Sinha, K. C., & Labi, S. (2007). Transportation Decision Making. Principles of Project Evaluation and Programming. John Wile & Sons.

Sinha, K., Labi, S., Hodge, S., Tine, G., & Shah, H. (2005). An assessment of highway financing needs in Indiana. Joint Transportation Research Program, 14.

Sinha, K.C., Kaji, T., & Liu, C.C. (1981). Optimal allocation of funds for highway safety improvement projects. Transportation Research Record 808, 24–30.

Sivak, M., & Schoettle, B. (2015). Road safety with self-driving vehicles: General limitations and road sharing with conventional vehicles.

Sivinski, R. (2011). Crash prevention effectiveness of light-vehicle electronic stability control: an update of the 2007 NHTSA evaluation. Tech. Rep. Nr. DOT HS 811 634, prepared for the National Highway Traffic Safety Administration.

Smith, S., Bellone, J., Bransfield, S., Ingles, A., Noel, G., Reed, E., & Yanagisawa, M. (2015). Benefits estimation framework for automated vehicle operations. Tech. Rep. DOT-VNTSC-FHWA-15-12, The Volpe Center, Cambridge, MA.

Sokolowski, J.A., & Banks, C.M. (2009). Principles of modeling and simulation. Wiley & Sons.

Song, J., Ghosh, M., Miaou, S., & Mallick, B. (2006). Bayesian multivariate spatial models for roadway traffic crash mapping. Journal of Multivariate Analysis 97(1), 246-273.

Sunley, W.T. (1990). What's That Extra Two Feet of Lane or Shoulder Width Worth? Illinois Municipal Review, April 1990 Issue, 19-20.

Sweden, G. O. o. The Vision Zero Initiative. Retrieved from http://www.visionzeroinitiative.com/ Accessed July 2018.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. Procs., IEEE conference on computer vision and pattern recognition, 2818-2826.

Tang, Z., Chen, S., Cheng, J., Ghahari, S. A., & Labi, S. (2018). Highway Design and Safety Consequences: A Case Study of Interstate Highway Vertical Grades. Journal of Advanced Transportation, 2018.

Tay, R. S., Champness, P. G., & Watson, B. C. (2003). Personality and speeding: Some policy implications. IATSS research, 27(1), 68-74.

Tensorflow. https://www.tensorflow.org/ Accessed: 08/25/2018.

Tingvall, C. (1997). The zero vision. In Transportation, Traffic Safety and Health (pp. 37-57): Springer.

Torrão, G. A., Coelho, M. C., & Rouphail, N. M. (2012). Effect of vehicle characteristics on crash severity: Portuguese experience. Inj. Prev, 18, 310-316.

Treat, J.R., Tumbas, N.S., McDonald, S.T., Shinar, D., Hume, R.D., Mayer, R.E., Stansifer, R.l., & Castellan, N.J. (1979). Tri-level study of the causes of traffic accidents, Tech. Rep. prepared by the Institute for Research in Public Safety for the Federal Highway Administration, Washington, DC.

Tsoukalas, L.H., & Uhrig, R.E. (1996). Fuzzy and neural approaches in engineering, Wiley & Sons, Inc.

U.S. Department of Health and Human Services, 2016. Centers for Disease Control and Prevention, National Center for Health Statistics. 1600 Clifton Road Atlanta, GA. http://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm, Accessed July 2018.

United Nations High Commissioner for Refugees. (2016). Safe Road Use in UNHCR: 2016 Survey Findings.

Urmson, C., Anhalt, J., Bagnell, D., Baker, C., Bittner, R., Clark, M. N., ... & Gittleman, M. (2008). Autonomous driving in urban environments: Boss and the urban challenge. Journal of Field Robotics, 25(8), 425-466.

USACE. 2016. Civil Works Construction Cost Index System (CWCCIS), Manual Nr. 1110-2-1304, Department of the Army, Washington, DC.

USDOT (2004). Transportation Statistics Annual Report, 2004. Bureau of Transportation Statistics, U.S. Department of Transportation, Washington, DC.

USDOT (2014). Beyond Traffic 2045 – Trends and Choices, U.S. Department of Transportation.

Van Schalkwyk, I., & Washington, S. (2008). Cost-effective safety improvements on two-lane rural state roads in Washington State. Tech. Rep. WA-RD 695.1, Washington State Department of Transportation, Olympia, WA.

Venkataraman, N., Shankar, V., Ulfarsson, G. F., & Deptuch, D. (2014). A heterogeneity-in-means count model for evaluating the effects of interchange type on heterogeneous influences of interstate geometrics on crash frequencies. Analytic Methods in Accident Research 2, 12-20.

Venkataraman, N., Ulfarsson, G., & Shankar, V. (2013). Random parameter models of interstate crash frequencies by severity, number of vehicles involved, collision, and location type. Accident Analysis and Prevention 59, 309-318.

Venkataraman, N., Ulfarsson, G., Shankar, V., Oh, J., & Park, M. (2011). Model of relationship between interstate crash occurrence and geometrics: Exploratory insights from random parameter negative binomial approach. Transportation Research Record 2236, 41-48.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica: Journal of the Econometric Society, 307-333.

Wang, C., Quddus, M., & Ison, S. (2011). Predicting accident frequency at its severity levels and its application in site ranking using a two-stage mixed multivariate model. Accident Analysis and Prevention 43(6), 1979-1990.

Wang, Y., Teoh, E. K., & Shen, D. (2004). Lane detection and tracking using B-Snake. Image and Vision computing, 22(4), 269-280.

Waseem, M., Ahmed, A., Saeed, TU. (2019). Factors affecting motorcyclists' injury severities: An empirical assessment using random parameters logit model with heterogeneity in means and variances. Accident Analysis and Prevention 123, 12-19.

Washington, S., Haque, M. M., Oh, J., & Lee, D. (2014). Applying quantile regression for modeling equivalent property damage only crashes to identify accident blackspots. Accident Analysis & Prevention, 66, 136-146.

Washington, S., Karlaftis, M., & Mannering, F. (2011). Statistical and Econometric Methods for Transportation Data Analysis. Boca Raton, FL: Chapman and Hall/CRC.

Wegman, F., & Hagenzieker, M. (2010). Scientific research on road safety management. Safety Science 48(9), 1081–1084.

Wierwille, W.W., Hanowski, R.J., Hankey, J.M., Kieliszewski, C.A., Lee, S.E., Medina, A., Keisler, A.S., & Dingus, T.A. (2002). Identification and evaluation of driver errors: Overview and recommendations. Tech. Rept. FHWA-RD-02-003, Prepared for the Federal Highway Administration, Washington, DC.

Williams, A. F. (2007). Contribution of the components of graduated licensing to crash reductions. Journal of Safety Research, 38(2), 177-184.

Winkelmann, R. (2000). Seemingly unrelated negative binomial regression. Oxford Bulletin of Economics and Statistics 62(4), 553-560.

Winkelmann, R. (2008). Econometric Analysis of Count Data, 5th Ed., Germany: Springer.

Winston, C., Maheshri, V., & Mannering, F. (2006). An exploration of the offset hypothesis using disaggregate data: The case of airbags and antilock brakes. AEI-Brookings Joint Center Working Paper 06-10. DOI.org/10.2139/ssrn.980069.

World Health Organization. (2015). Global Health Estimates (GHE) 2013: Deaths by Age, Sex and Cause. Author at http://www.who.int/mediacentre/factsheets/fs310/en/ Accessed July 2018.

Wu, H., Han, Z., Murphy, M.R., & Zhang, Z. (2015). Empirical bayes before–after study on safety effect of narrow pavement widening projects in Texas. Transportation Research Record 2515, 63–69.

Wu, J. (2018). An automatic procedure for vehicle tracking with a roadside LiDAR sensor. Institute of Transportation Engineers. ITE Journal, 88(11), 32-37.

Wu, Z., Sharma, A., Mannering, F., & Wang, S. (2013). Safety impacts of signal-warning flashers and speed control at high-speed signalized intersections. Accident Analysis and Prevention 54, 90-98.

Wymann, B., Espié, E., Guionneau, C., Dimitrakakis, C., Coulom, R., & Sumner, A. (2000). Torcs, the open racing car simulator. Software available at http://torcs. sourceforge. net, 4, 6.

Xie, Y., Lord, D., & Zhang, Y. (2007). Predicting motor vehicle collisions using Bayesian neural network models: An empirical analysis. Accident Analysis & Prevention, 39(5), 922-933.

Xiong, Y., & Mannering, F. (2013). The heterogeneous effects of guardian supervision on adolescent driver-injury severities: A finite-mixture random-parameters approach, Transportation Research Part B 49, 39-54.

Yan, C., Xu, W., & Liu, J. (2016). Can you trust autonomous vehicles: Contactless attacks against sensors of self-driving vehicle. DEF CON, 24.

Yang, J., & Coughlin, J. F. (2014). In-vehicle technology for self-driving cars: Advantages and challenges for aging drivers. International Journal of Automotive Technology, 15(2), 333-340.

Yang, X. J., Pradhan, A., Tilbury, D., & Robert, L. (2018). Human autonomous vehicles interactions: An interdisciplinary approach.

Yao, B., Chen, C., Cao, Q., Jin, L., Zhang, M., Zhu, H., & Yu, B. (2017). Short - term traffic speed prediction for an urban corridor. Computer-Aided Civil and Infrastructure Engineering, 32(2), 154-169.

Yasmin, S., Eluru, N., Pinjari, A.R., & Tay, R. (2014). Examining driver injury severity in two vehicle crashes – A copula based approach. Accident Analysis and Prevention 66, 120-135.

Yu, R., & Abdel-Aty, M. (2013). Utilizing support vector machine in real-time crash risk evaluation. Accident Analysis & Prevention, 51, 252-259.

Zaremba, W., Sutskever, I., & Vinyals, O. (2014). Recurrent neural network regularization. arXiv preprint arXiv:1409.2329.

Zegeer, C.V., Stewart, R., Council, F., & Neuman, T.R. (1994). Accident relationships of roadway width on low-volume roads. Transportation Research Record 1445, 160–168.

Zegeer, C.V., Deen, R.C., & Mayes, J.G. (1980). The effect of lane and shoulder widths on accident reductions on rural, two-lane roads. Transportation Research Record 806, 33–43.

Zegeer, C.V., Stewart, R., & Council, F. (1994). Roadway widths for low traffic-volume roads, NCHRP Report 362. Transportation Research Board, Washington, DC.

Zeng, Q., Huang, H., Pei, X., & Wong, S. C. (2016). Modeling nonlinear relationship between crash frequency by severity and contributing factors by neural networks. Analytic Methods in Accident Research 10, 12-25.

Zhang, A., Wang, K. C., Fei, Y., Liu, Y., Chen, C., Yang, G., & Qiu, S. (2018). Automated pixel−level pavement crack detection on 3d asphalt surfaces with a recurrent neural network. Computer-Aided Civil and Infrastructure Engineering.

Zimmerman, K. A., & Larson, R. (2005). Improving safety as part of a pavement preservation program. In First National Conference on Pavement Preservation. Kansas City, Missouri, Transportation Research Board, 159-172.

VITA

RESEARCH INTERESTS

- **Applications:** Autonomous and connected transportation; Highway safety analysis; Transportation infrastructure systems management; Human-computer interaction; Automation; Control; Smart cities; Intelligent transport systems; Urban computing.
- **Operations Research Tools:** Machine learning and data mining; Deep learning; Computer vision; Optimization; Simulation; Econometrics and statistics; Mathematical programming; Numerical methods; Big data analytics.

EDUCATION

Purdue University, Lyles School of Civil Engineering — Aug 2014 – Aug 2019
Ph.D. Candidate in Civil Engineering **-** Bilsland Fellowship
Major Concentration: Computational Science & Engineering (CS&E)
Mini-MBA program: Applied Management Principles
GPA: 3.96/4.0
Advisors: Purdue University **-** Dr. Samuel Labi (main advisor), Dr. Kumares C. Sinha, Dr. Paul V. Preckel, Dr. Lefteri H. Tsoukalas; MIT **-** Dr. Richard de Neufville

Purdue University, Next Generation Transportation Systems (NEXTRANS) — July 2018 – Aug 2019
Purdue University, Center for Connected and Automated. Transportation (CCAT) — July 2018 – Aug 2019
Bilsland Research Fellow**:** Machine Learning and Artificial Intelligence

University of Southern California, Viterbi School of Engineering — Aug 2012 – May 2014
M.S. in Transportation Engineering
Advisor: Dr. James E. Moore, II

Shandong University (China), School of Civil Engineering — Sept 2007 – Jun 2011
B.E. in Engineering Mechanics

HONORS / AWARDS

- **Nomination for Schmidt Science Fellows**, 1 of the 5 nominees from Purdue University, 2018 https://engineering.purdue.edu/Engr/AboutUs/News/Spotlights/2018/engineering-students-nominated-for-elite-Schmidt-Fellows-program
- **Bilsland Fellowship Award**, 2018 – Office of Interdisciplinary Graduate Programs (OIGP), Purdue University https://engineering.purdue.edu/CE/AboutUs/News/Transportation_Features/sikai-chen-receives-2018-bilsland-dissertation-fellowship-award
- **Pai Tao Yeh Fellowship**, 2018
- **STV Civil Engineering Award**, 2018 – Purdue University
- **Summer Research Grant** (for excellence in teaching and research), 2018 – Purdue University (PRF)
- **Estus H. & Vashti L. Magoon Award for Excellence in Teaching**, 2016 – College of Engineering, Purdue University.
- **Summer Research Grant** (for excellence in teaching and research), 2016 – Purdue University (PRF)

TEACHING EXPERIENCE / INTERESTS

Teaching Experience (4 years as a graduate student)

Teaching Assistant, Purdue University — Fall, 2014 – Spring, 2018
CE 398: Introduction to Civil Engineering Systems Design
(Estus H. & Vashti L. Magoon Award for Excellence in Teaching, 2016)

Teaching Assistant, Purdue University — Spring, 2016
CE 568: Highway Infrastructure Management Systems

Instructor, Purdue University — Summer, 2016
Introduction to Civil Engineering Systems Design. Short-term course designed for exchange students from Brazil.

Teaching Interests

Undergraduate level: Engineering Systems Design; Statistics; Mathematical Programming; Introduction to Machine Learning and Data Mining; Introduction to Numerical Method.
Graduate level: Highway Infrastructure Management Systems; Statistical and Econometric Method in Transportation Data Analysis; Transportation Systems Evaluation; Machine Learning and Data Mining; Numerical Analysis.

Programming languages

- General: Python, Matlab, C/C++, R
- Deep learning: TensorFlow, Keras
- Statistical analysis: SAS, Stata, SQL
- Big data analysis: Hadoop, Spark, RHIPE
- Optimization: GAMS, Gurobi

PROFESSIONAL SERVICE

Young member – ASCE Committee for Planning, Economics and Finance (PEF)
Member – ASCE Connected and Autonomous Vehicle (CAV) Impacts Committee
Member – ASCE Transportation and Development Institute (T&DI) Student Chapter @ Purdue
Member – Transportation Research Forum (TRF)
Member – international Society for Maintenance And Rehabilitation of Transport infrastructures (iSMARTi)
Treasurer Chair – 13th Annual Inter-University Symposium on Infrastructure Management (AISIM)
Co-Chair of Planning Committee – 1st Annual Conference on Next-Generation Transport Systems (NGTS)
Reviewer – ICTAI 2018: The Annual IEEE International Conference on Tools with Artificial Intelligence (ICTAI)
Main Contributor – Solution Manual of Introduction to Civil Engineering Systems: A Systems Perspective to the Development of Civil Engineering Facilities

OTHER ACTIVITIES

Instruments: Saxophone, Pipa (a traditional Chinese instrument)
Sports: Soccer, Swimming, Skiing, Skating, Ping-Pong

PUBLICATIONS

Peer-Reviewed Journal Articles

Chen, S., Saeed, T.U., & Labi, S. (2017). Impact of road-surface condition on rural highway safety: A multivariate random parameters negative binomial approach. Analytic Methods in Accident Research, 16, 75-89. https://doi.org/10.1016/j.amar.2017.09.001

Bhargava, A., Labi, S., Chen, S., Saeed, T.U., & Sinha, K.C. (2017). Predicting cost escalation pathways and deviation severities of infrastructure projects using risk-based econometric models and Monte Carlo simulation. Computer-Aided Civil and Infrastructure Engineering, 32(8), 620-640. http://onlinelibrary.wiley.com/doi/10.1111/mice.12279/full

Chen, S., Saeed, T. U., Alqadhi, S.D., & Labi, S. (2017). Safety impacts of pavement surface roughness at two-lane and multi-lane highways: accounting for heterogeneity and seemingly unrelated correlation across crash severities. Transportmetrica A: Transport Science, 15(1), 18-33. http://dx.doi.org/10.1080/23249935.2017.1378281

Saeed, T.U., Qiao, Y., Chen, S., Gkritza, K., & Labi, S. (2017). Methodology for probabilistic modeling of highway bridge infrastructure condition: Accounting for improvement effectiveness and incorporating random effects. Journal of Infrastructure Systems, 23(4), 04017030. http://ascelibrary.org/doi/full/10.1061/(ASCE)IS.1943-555X.0000389

Labi, S., Chen, S., Preckel, P.V., Qiao, Y., & Woldemariam, W. (2017). Rural two-lane highway shoulder and lane width policy evaluation using multiobjective optimization. Transportmetrica A: Transport Science, 13(7), 631-656. http://dx.doi.org/10.1080/23249935.2017.1315841

Ghahari, S. A., Ha, P., Chen, S., Alinizzi, M., Agbelie, B., & Labi, S. (2018). Inputs for bridge painting decision support: a synthesis. Infrastructure Asset Management, 5(2), 56-74. https://doi.org/10.1680/jinam.17.00029

Alqadhi, S., Ghahari, S. A., Chen, S., Volovski, M., & Woldemariam, W. (2018). Costs and benefits of highway resurfacing: a case study of Interstate 465 in Indiana, USA. Infrastructure Asset Management, 5(2), 45-55. https://doi.org/10.1680/jinam.17.00036

Qiao, Y., Saeed, T. U., Chen, S., Nateghi, R., & Labi, S. (2018). Acquiring insights into infrastructure repair policy using discrete choice models. Transportation Research Part A: Policy and Practice, 113, 491-508. https://doi.org/10.1016/j.tra.2018.04.020

Tang, Z., Chen, S., Cheng, J., Ghahari, S. A., & Labi, S. (2018). Highway design and safety consequences: A case study of interstate highway vertical grades. Journal of Advanced Transportation, 2018. Article ID 1492614 https://doi.org/10.1155/2018/1492614

Alinizzi, M., Chen, S., Labi, S., & Kandil, A. (2018). A methodology to account for one−way infrastructure interdependency in preservation activity scheduling. Computer−Aided Civil and Infrastructure Engineering, 33 (2018), 905–925. https://doi.org/10.1111/mice.12380

Chen, S., Saeed, T.U., Alinizzi, M., Lavrenz, S., & Labi, S. (2019). Safety sensitivity to roadway characteristics: A comparison across highway functional classes. Accident Analysis & Prevention, 123, 39-50. https://doi.org/10.1016/j.aap.2018.10.020

Chen, S., Tang, Z., Zhou, H., & Cheng, J. (2019). Extracting topographic data from online sources to generate a digital elevation model for highway preliminary geometric design. Journal of Transportation Engineering, Part A: Systems, 145(4), 04019003. https://ascelibrary.org/doi/full/10.1061/JTEPBS.0000212

Ghahari, S., Alabi, B., Alinizzi, M., Alqadhi, S., Chen, S., & Labi, S. (2019). Examining the relationship between infrastructure investment and performance using state-level data. Journal of Infrastructure Systems (Accepted, in press) https://doi.org/10.1061/(ASCE)IS.1943-555X.0000505

Peer-Reviewed Conference Papers

Chen, S., Saeed, T.U., Al-Qadhi, S.D., Labi S., & Sinha, K.C. (2016). Safety impacts of pavement rehabilitation at multilane highways. Eighth International Conference on Maintenance and Rehabilitation of Pavements (MAIREPAV8), Singapore. doi: 10.3850/978-981-11-0449-7-212-cd http://rpsonline.com.sg/proceedings/9789811104497/html/212.xml

Labi, S., Alqadhi, S.D., Murillo-Hoyos, J., & Chen, S. (2016). Mathematical constructs for assessing the effectiveness of environmental remediation actions. CECAR 7. Waikiki, Oahu, Hawaii http://www.eventscribe.com/2016/CECAR7/presinfo.asp?pres=148689

Chen, S., Qian, C., & Chen, Y. (2011). Study on the taxicab industry planning restricted to the policy of the government in china. In ICCTP 2011: Towards Sustainable Transportation Systems (pp. 485-496). http://ascelibrary.org/doi/10.1061/41186(421)48

Qiao, Y., Chen, S., Alinizzi, M., & Labi, S. (2018). Modeling the relationships between pavement distress and performance. Advances in Materials and Pavement Prediction (pp. 11-13) (AM3P 2018, Doha, Qatar). CRC Press. https://www.taylorfrancis.com/books/e/9780429855801/chapters/10.1201%2F9780429457791-2

Chen, S., Saeed, T.U., Alqadhi, S., & Labi, S. (2018). Comparative analysis of safety impacts of pavement surface roughness at two-lane and multi-lane highways: accounting for heterogeneity and seemingly unrelated correlation across crash severities. Transportation Research Board 97th Annual Meeting, Washington D.C., USA https://trid.trb.org/view/1495200

Chen, S., Ghahari, S., Miralinaghi, M., & Labi, S. (2019). Assessing performance outcomes and ranking of jurisdictions–A nonparametric efficiency approach for asset management (No. 19-05065). Transportation Research Board 98th Annual Meeting, Washington D.C., USA https://trid.trb.org/view/1572709

Miralinaghi, M., Woldemariam, W., Abraham, D. M., Chen, S., & Labi, S. (2019). Network-level scheduling of construction projects considering user and business impacts (No. 19-03829). Transportation Research Board 98th Annual Meeting, Washington D.C., USA https://trid.trb.org/view/1572257

Ghahari, S., Alabi, B. N. T., Alinizzi, M., Alqadhi, S., Chen, S., & Labi, S. (2019). The bridge investment-performance nexus at an aggregate level–Accounting for situational and measurement biases (No. 19-04878). Transportation Research Board 98th Annual Meeting, Washington D.C., USA https://trid.trb.org/view/1573231

Miralinaghi, M., Seilabi, S. E., Chen, S., Hsu, Y. T., & Labi, S. (2019). Time-dependent traffic network design with multi-class projects (No. 19-03832). Transportation Research Board 98th Annual Meeting, Washington D.C., USA https://trid.trb.org/view/1572805

Conference Presentations

Chen, S., Labi, S., Qiao, Y., Preckel, P. V., Bai, Q., & Woldemariam, W. (2017). Multiobjective optimization of lane and shoulder widths at rural two-lane highways. Transportation Research Board 96th Annual Meeting. Washington, D.C., USA http://amonline.trb.org/63532-trb-1.3393340/t006-1.3408722/391-1.3408916?qr=1

Chen, S., Qiao, Y., Labi, S., & Preckel, P.V. (2017). Lane and shoulder widths at rural two-lane highways-design policy evaluation. Presenting at the First International Conference on Maintenance and Rehabilitation of Constructed Infrastructure Facilities (MAIREINFRA1). J.W. Marriott Dongdaemun Square Hotel, Seoul, Korea. July 19-21, 2017 http://maireinfra.org/wp-content/uploads/2017/03/MAIREINFRA1-Conference_Program_Full_Version_Final.pdf

Bhargava, A., Chen, S., Qiao, Y., Labi, S., & Sinha, K.C. (2016). Addressing uncertainty in bridge projects – estimating the cost escalation pathway likelihood and overrun severities for infrastructure expansion contracts. ASCE International Conference on Transportation & Development. Houston, Texas, USA. June 26-29, 2016 https://www.asce-ictd.org/sites/asce-ictd.org/2018/files/inline-files/ictd-2016-program-final.pdf

Chen, S., Labi, S., & Sinha, K. C. (2017). Overweight-vehicle research that led to indiana's new trucking law (Act 1481). 13th Annual Inter-University Symposium on Infrastructure Management (AISIM). Purdue University, USA. June 23, 2017 https://aisimpurdue.wixsite.com/aisim13purdue/presentations

Nafakh, A., Chen, S., Saeed, T.U., & Labi, S. (2018). Evaluation of an automated people mover in downtown Indianapolis. ASCE International Conference on Automated People Movers & Automated Transit Systems. Tampa, Florida, USA. April 29 - May 2, 2018 http://www.apmconference.org/sites/apmconference.org/files/u4/apm-2018-final-program.pdf

Labi, S., Saeed, T.U., Chen, S., & Nafakh, A. (2018). Evaluation of hyperloop transportation systems. ASCE International Conference on Automated People Movers & Automated Transit Systems. Tampa, Florida, USA. April 29 - May 2, 2018 http://www.apmconference.org/sites/apmconference.org/files/u4/apm-2018-final-program.pdf

Chen, S., Alinizzi, M., Alqadhi, S.D., Miralinaghi, M., Volovski, M., Agbelie, B., & Labi, S. (2018). Changes in highway agency expenditures and revenue in an era of cavs. 2018 global symposium on connected and automated vehicles and infrastructure. University of Michigan, Ann Arbor, Michigan, USA. March 7-8, 2018 http://ccat.umtri.umich.edu/wp-content/uploads/2018/03/2018-SymposiumAgenda.pdf

Chen, S., Leng, Y., & Labi, S. (2019). Evaluating CAV safety impacts with simulation and artificial intelligence. ASCE International Conference on Transportation & Development (ICTD 2019). Alexandria, Virginia. June 09-12, 2019. https://www.asce-ictd.org/sites/asce-ictd.org/2019/files/inline-files/ictd-2019-final-program.pdf

Chen, S., Ha, Y., & Labi, S. (2019). Developing a sustainable transportation system under electric and autonomous vehicles dedicated lane deployment scheme. ASCE International Conference on Transportation & Development (ICTD 2019). Alexandria, Virginia. June 09-12, 2019. https://www.asce-ictd.org/sites/asce-ictd.org/2019/files/inline-files/ictd-2019-final-program.pdf

Technical Report

Saeed, T. U., Qiao, Y., Chen, S., Alqadhi, S., Zhang, Z., Labi, S., & Sinha, K. C. (2017). Effects of bridge surface and pavement maintenance activities on asset rating (Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2017/19). West Lafayette, IN: Purdue University. https://doi.org/10.5703/1288284316573

Peer-Reviewed Journal Articles Submitted and Currently Under Review

Chen, S., Leng, Y., & Labi, S. (2019). A deep learning algorithm for simulating autonomous driving: Considering prior knowledge and temporal information. (Under review: Computer-Aided Civil and Infrastructure Engineering)

Miralinaghi, M., Woldemariam, W., Abraham, D. M., Chen, S., & Labi, S. (2019). Network-level scheduling of construction projects considering user and business impacts. (Under review: Computer-Aided Civil and Infrastructure Engineering)

Qiao, Y., Chen, S., Alinizzi, M., Alamaniotis, M., & Labi, S. (2019). Investigating relationships between IRI and pavement distresses using machine learning techniques. (Under review: Transportation Research Part C: Emerging Technologies)

Feng, J., Chen, S., Ye, Z., & Labi, S. (2019). A novel method to solve the repositioning problem in personal transport unit (PTU)-sharing system. (Under review: Transportmetrica A: Transport Science)

Miralinaghi, M., Seilabi, S., Chen, S., Hsu, Y., & Labi, S. (2019). Optimizing the selection and scheduling of multi-class projects using a Stackelberg framework. (Under review: European Journal of Operational Research)

Chen, S., Ghahari, S., Miralinaghi, M., & Labi, S. (2019). A nonparametric efficiency methodology for comparative assessment of infrastructure agency performance. (Under review: Transportmetrica A: Transport Science)

Chen, S., Alinizzi, M., Leng, Y., Miralinaghi, M., & Labi, S. (2019). Efficacy of machine learning in highway safety analysis: predicting the fatality status of highway segments. (Under review: Computer-Aided Civil and Infrastructure Engineering)