MID-AMERICA
TRANSPORTATION CENTER

Report # MATC-KSU: 352

Final Report

Nebraska Lincoln

K-STATE Kansas State University

KU THE UNIVERSITY OF KANSAS

MISSOURI S&T University of Science & Technology

THE UNIVERSITY OF IOWA

LINCOLN University

PRAIRIE VIEW A&M UNIVERSITY 1876

Characteristics and Contributory Causes Related to Large Truck Crashes (Phase II) – All Crashes

Sunanda Dissanayake, Ph.D., P.E.
Associate Professor
Department of Civil Engineering
Kansas State University

Siddhartha Kotikalapudi
Graduate Research Assistant

K-STATE Kansas State University

2012

A Cooperative Research Project sponsored by the U.S. Department of Transportation Research and Innovative Technology Administration

MATC

Characteristics and Contributory Causes Related to Large-Truck Crashes (Phase II) – All Crashes

Sunanda Dissanayake, Ph.D., P.E.
Associate Professor
Department of Civil Engineering
Kansas State University

Siddhartha Kotikalapudi
Graduate Research Assistant
Department of Civil Engineering
Kansas State University

A Report on Research Sponsored by

Mid-America Transportation Center

University of Nebraska-Lincoln

March 2012

Technical Report Documentation Page

1. Report No. 25-1121-0001-352	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Characteristics and Contributory Causes Related to Large-Truck Crashes (Phase II) – All Crashes		5. Report Date March 2012 6. Performing Organization Code
7. Author(s) Sunanda Dissanayake and Siddhartha Kotikalapudi		8. Performing Organization Report No. 25-1121-0001-352
9. Performing Organization Name and Address Mid-America Transportation Center 2200 Vine St. PO Box 830851 Lincoln, NE 68583-0851		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Research and Innovative Technology Administration 1200 New Jersey Ave., SE Washington, D.C. 20590		13. Type of Report and Period Final Report Jan. 2010 –March 2012 14. Sponsoring Agency Code TRB RiP No. 24478
15. Supplementary Notes		

16. Abstract

Statistics clearly demonstrate that large-truck crashes contribute to a significant percentage of high-severity crashes. It is therefore important for the highway safety community to identify the characteristics and contributory causes of these types of crashes. The first phase of the current research endeavor examined fatal crash data from the Fatality Analysis Reporting System (FARS) database. In the second phase, presented in the current report, truck-crashes of all severity levels were analyzed with the intention of understanding characteristics and contributory causes, as well as identifying factors contributing to increased severity of truck-crashes. This goal could not be achieved by analyzing fatal crashes alone. Various statistical methodologies such as cross-classification analysis and severity models were developed using Kansas crash data. This study identified the characteristics, contributing causes, and specific factors related to the occurrence and increased severity of large-truck-crashes. By understanding these issues, countermeasures might be developed to mitigate the number and severity of truck crashes.

17. Key Words Truck crashes, crash severity, large trucks, crash contributing causes	18. Distribution Statement		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 99	22. Price

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Akaike Information Criterion (AIC)
Analysis of Variance (ANOVA)
Average Annual Daily Traffic (AADT)
Cochran-Mantel-Haenszel statistic (CMH)
Crash Injury Research and Engineering Network (CIREN)
Crash Severity Index (CSI)
Drowsy Driver Warning System (DDWS)
Fatality Analysis Reporting System (FARS)
Federal Motor Carrier Safety Administration (FMCSA)
Field Operational Test (FOT)
Florida Crash Analysis Reporting (CAR)
General Estimate System of the National Sampling System (NSS GES)
Geographic Information System (GIS)
Heavy Good Vehicle (HGV)
Heteroskedastic Ordered Probit (HOP)
Highway Safety Information System (HSIS)
Indiana Electronic Vehicle-Crash-Record System (EVCRS)
Kansas Accident Reporting System (KARS)
Kansas Department of Transportation (KDOT)
Large-truck Crash Causation Study (LTCCS)
Maximum Likelihood Method (MLM)
Metropolis-Hastings Algorithm (M-H)
National Automotive Sampling Systems Crashworthiness Data System (NASS-CDS)
National Highway Traffic Safety Administration (NHTSA)
National Institute of Scientific Investigation (NISI)
Property Damage Only (PDO)
Run-Off-the-Road (ROR)
Schwarz Criterion (SC)
Standard Ordered Probit (SOP)
Statistical Analysis System (SAS)
Unsafe Driving Act (UDA)
Vehicle Miles Travelled (VMT)
Traffic Accident Surveillance and Analysis System (TASAS)

Acknowledgements

The authors would like to thank Mid-America Transportation Center for funding this project and the Kansas Department of Transportation for providing data from the Kansas Accident Reporting System (KARS) database which has been critical in conducting this study. Authors also wish to acknowledge the support by Mr. Steven Buckley, who served as the Project Monitor.

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Executive Summary

In order to improve the safety of the surface transportation system as a whole, each of several critical areas should be focused upon and addressed separately. Statistics clearly demonstrate that large-truck-crashes contribute to a significant percentage of high-severity crashes. It is therefore important for the highway safety community to identify the characteristics and contributory causes of these types of crashes. Toward this consideration, the first phase of the current research endeavor examined fatal crash data from the Fatality Analysis Reporting System (FARS) database. In the second phase, presented in the current report, truck-crashes of all severity levels were analyzed with the intention of understanding characteristics and contributory causes, as well as identifying factors contributing to increased severity of truck-crashes. This goal could not be achieved by analyzing fatal crashes alone. Various statistical methodologies such as cross-classification analysis and severity models were developed using Kansas crash data. Various driver-, road-, environment-, and vehicle- related characteristics were identified and contributory causes were analyzed.

From the cross-classification analysis, severity of truck-crashes was found to be related to variables such as road surface (type, character and condition), accident class, collision type, driver- and environment-related contributory factors, traffic control type, truck-maneuver, accident location, speed limit, light and weather conditions, time of day, function class, lane class, and Average Annual Daily Traffic (AADT). Furthermore, driver-related contributory causes were found to be more common than any other type of contributory cause for the occurrence of truck-crashes. Failing to give time and attention, being too fast for existing conditions, and failing to yield right of way were the most dominant truck-driver-related contributory causes, among many others.

Through severity modeling, factors such as truck-driver-related contributory cause, accident class, manner of collision, truck-driver under the influence of alcohol, truck maneuver, traffic control device, surface condition, truck-driver being too fast for existing conditions, truck-driver being trapped, damage to the truck, light conditions, etc., were found to be significantly related to increased severity of truck-crashes. Truck-driver being trapped had the highest odds of contributing to a more severe truck crash, with a value of 82.81, followed by the collision resulting in damage to the truck, which had 3.05 times higher odds of increasing the severity of truck-crashes. Truck-driver under the influence of alcohol had 2.66 times higher odds of being involved in a more severe crash.

This study identified the characteristics, contributory causes, and specific factors related to the occurrence and increased severity of large-truck-crashes. By understanding these issues, countermeasures might be developed to mitigate the number and severity of truck crashes.

Chapter 1 Introduction

1.1 Background

The transportation system is one of the most integral factors contributing to the economic progress of any country. In the United States, development of the road network over the past few decades has considerably increased the efficiency of the movement of freight and passengers across the nation. Trucks play a major role in the U.S. transportation system, as they transport a significant portion of the nation's cargo. Many different types of trucks operate in the U.S. The type of truck utilized for a given transportation operation depends on the required duration of travel and the quantity and type of cargo to be transported. Technologies such as the Global Positioning System (GPS) and satellite communications have improved operating conditions for cargo transport by providing drivers with pertinent information on traffic and weather conditions, along with travel routes and directions.

From 1988 to 2008, the number of registered large-trucks in the U.S. increased by 47%, with a corresponding 65% increase in truck vehicle miles traveled (VMT) (1). As the number of large-trucks increases, so does the probability of their being involved in motor vehicle crashes. Table 1.1 displays the numbers and rates of large-trucks involved in crashes in the U.S. occurring between 2000 and 2008. In 2009, one out of every 10 traffic fatalities involved large-trucks (2). Nearly 84% of these fatalities were not truck occupants (3). Also, 7% of all fatal crashes in the United States in 2009 involved a large-truck (4). Aside from impacting the safety of the transportation system, truck crashes represent a substantial economic burden. For example, In Kansas, in 2008, a crash involving a large-truck occurred once every 2.37 hours, resulting in a total financial loss estimated at approximately $ 0.327 billion (6). These figures show that each of the critical areas regarding the large-truck-crashes must be identified and studied for improving overall safety of the transportation system (5).

Table 1.1 Large-truck crashes and involvement rates in the United States

Year	Fatal Crashes			Injury Crashes			PDO Crashes		
	Number of Crashes	Involvement Rate		Number of Crashes	Involvement Rate		Number of Crashes	Involvement Rate	
		per 100 million VMT	per 100,000 Registered Vehicles		per 100 million VMT	per 100,000 Registered Vehicles		per 100 million VMT	per 100,000 Registered Vehicles
2000	4,995	2.43	62.26	101,000	49	1,253	351,000	171	4,377
2001	4,823	2.31	61.38	90,000	43	1,143	335,000	160	4,261
2002	4,587	2.14	57.88	94,000	44	1,189	336,000	156	4,232
2003	4,721	2.17	60.86	89,000	41	1,145	363,000	167	4,681
2004	4,902	2.22	59.99	87,000	39	1,062	324,000	147	3,970
2005	4,951	2.22	58.37	82,000	37	971	354,000	159	4,178
2006	4,766	2.14	54.04	80,000	36	911	300,000	135	3,398
2007	4,633	2.04	51.32	76,000	33	839	333,000	147	3,690
2008	4,089	1.80	45.40	66,000	29	734	309,000	136	3,435

Source: *Traffic Safety Facts 2009*

Large trucks comprised 14.9% of all fatal crashes in the state of Kansas in 2008, in spite of their being involved in only 5.6% of total crashes (6). Figure 1.1 illustrates a comparison between the total number of fatal crashes and the number of fatal crashes involving large trucks in the state of Kansas occurring between 2000-2008.

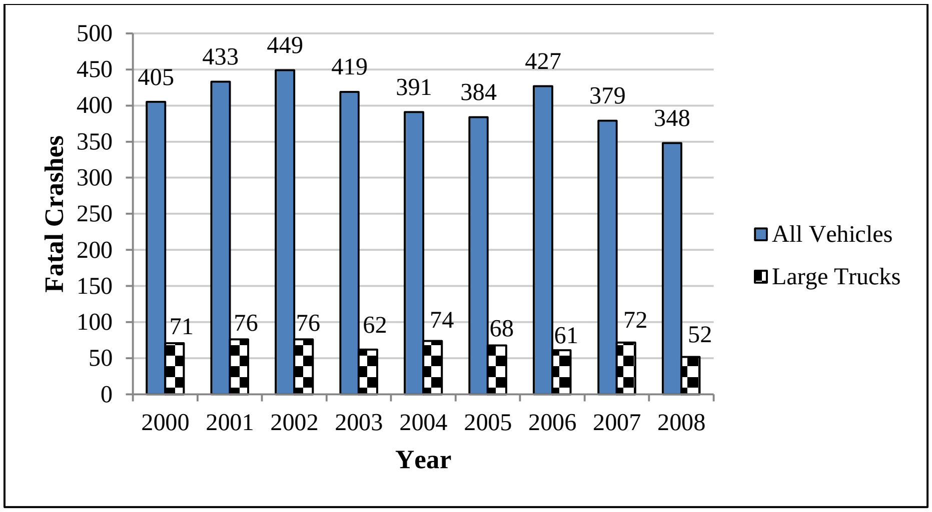

Figure 1.1 Comparisons of total fatal and truck-involved fatal crashes in Kansas

Sheer size and the corresponding spatial requirements for successful maneuvering potentially make the safe operation of a large-truck a more demanding task than operating a smaller, more light-weight vehicle. Moreover, the size of a large truck creates a large blind spot area, resulting in the potential for sideswipe crashes. Figure 1.2 shows the variation of crashes involving large-trucks in Kansas, based on different severity levels.

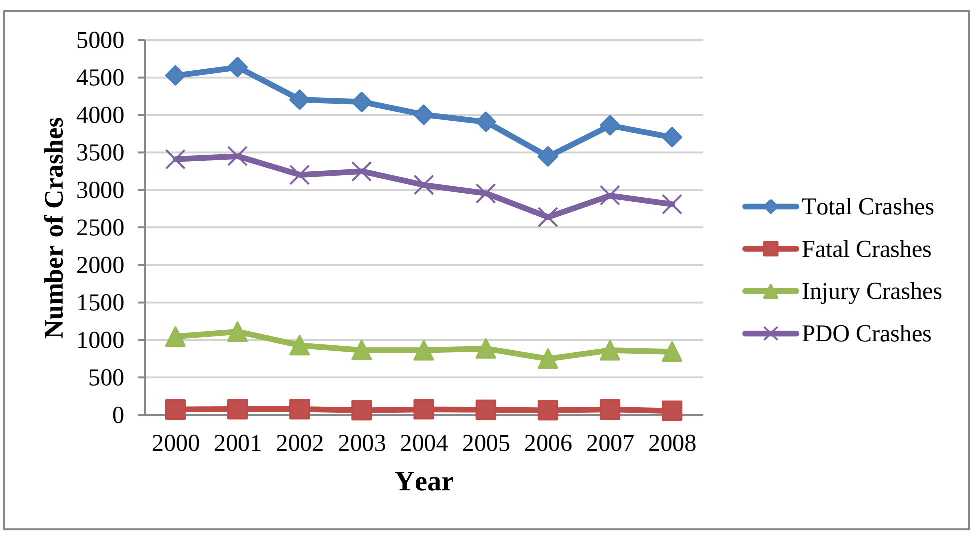

Figure 1.2 Number of large-truck crashes by severity in Kansas

Though the number of large-truck crashes in Kansas decreased in 2008 in comparison to each of the prior ten years, they continued to comprise a uniform percentage (approximately 5.5%) of total crashes. Table 1.2 displays the number of total crashes involving large-trucks in Kansas occurring between 2000-2008, expressed as the percentage of total crashes by severity. As shown, large trucks accounted for a disproportionate share of fatal and injury crashes. These values were deduced from statistics obtained from the Kansas Accident Reporting System (KARS) database.

Table 1.2 Truck-crashes as a percentage of total crashes by severity in Kansas

Year	Fatal Crashes			Injury Crashes			PDO Crashes			Total Crashes		
	Large -truck	All	%	Large -truck	All	%	Large -truck	All	%	Large -truck	All	%
2000	71	405	17.6	1045	19,454	5.4	3409	58,215	5.9	4525	78,074	5.8
2001	76	433	17.6	1110	19,346	5.7	3451	59,028	5.8	4637	78,807	5.9
2002	76	449	16.9	927	18,495	5.0	3201	59,327	5.4	4204	78,271	5.4
2003	62	419	14.8	864	17,037	5.1	3248	57,537	5.7	4174	74,993	5.6
2004	74	391	18.9	862	16,631	5.2	3067	57,080	5.2	4003	74,102	5.4
2005	68	384	17.7	885	16,185	5.5	2954	52,106	5.7	3907	68,675	5.7
2006	61	427	14.3	748	15,792	4.7	2638	49,241	5.4	3447	65,460	5.3
2007	72	379	19.0	862	16,227	5.3	2926	53,983	5.4	3860	70,589	5.5
2008	52	348	14.9	842	14,866	5.7	2808	50,644	5.5	3702	65,858	5.6

Source: *2008 Kansas Traffic Accident Facts*

1.2 Problem Statement

Large-trucks, which are defined in this study as those having a gross vehicle weight rating of 10,000 pounds or more, contribute to a significant proportion of the traffic composition in the U.S. Large-truck crashes represent a major safety concern for the road transportation system. Due to the high severity of crashes involving large-trucks, it is of critical importance to study characteristics of truck crashes closely. In 2009, nearly 296,000 large-trucks were involved in road crashes in the United States, of which 3,215 crashes resulted in at least one fatality (4). In the same year, large trucks accounted for nearly 7% of all vehicles involved in fatal crashes, 2% of all vehicles involved in injury crashes, and 3% of vehicles involved in Property Damage Only (PDO) crashes (4). These figures indicate that large-truck crashes tend to be severe. Truck-crashes are particularly devastating for occupants of the other vehicles involved (e.g., passenger vehicles). In 2009, 98% of all fatalities in two-vehicle, large-truck crashes involving a passenger vehicle occurred among occupants of the passenger vehicle (3).

The current study addresses the need to identify characteristics and contributory causes related to large-truck crashes. In the first phase of the study, the Fatality Analysis Reporting System (FARS) database was used to analyze the characteristics of fatal truck-crashes occurring in the U.S. (7). In phase two, truck crashes occurring in Kansas were analyzed by considering all levels of injury severity. The findings of the current analysis are intended to aid in the identification of countermeasures and areas for future research, in an effort to improve the overall safety of the transportation system.

1.3 Objectives

Mitigation of large-truck crashes can only be accomplished through identification and detailed analysis of their characteristics and contributory causes, as well as the identification of factors associated with their severity. With this in mind the primary objectives of this study are as follows:

1. To identify various characteristics that prevailed at the time of occurrence of large-truck crashes.
2. To identify the vehicle-, road-, driver- and environment-related causes that contributed to the occurrence of large-truck crashes.

3. To identify and evaluate factors contributing to the typically elevated severity of large-truck crashes.

1.4 Outline of the Report

Following this introduction, chapter 2 reviews literature related to the subject of large-truck crashes. Chapter 3 details the methodology adopted by the current study for analysis of the characteristics of large-truck crashes. Utilizing the cross-classification method, this chapter delineates relationships between crash severity and select variables, followed by an overview of the various technical parameters associated with development of the model presented in the current research. In chapter 4, results of the statistical model are discussed. Chapter 5 presents the conclusions of this study. References and appendices are also provided.

Chapter 2 Literature Review

Crashes involving large-trucks are a long-standing and wide-spread issue. Studies conducted in several states have focused on identifying crash severity and contributing characteristics related to truck crashes, utilizing figures from corresponding databases. This chapter summarizes some of the important research conducted in this field that has also contributed to informing current study.

2.1 Characteristics and Contributory Causes of Truck-Crashes

Mulinazzi et al. conducted a study emphasizing high wind and adverse weather conditions as contributory causes for truck-crashes in the U.S. (8). Measures taken by different states to mitigate wind-induced truck-crashes were briefly discussed. Data related to wind-induced truck-crashes on I-70 in Kansas for a six-year time period (2003-2008) were obtained from the Kansas Department of Transportation's KARS database. Data were analyzed to delineate the relationship between crash occurrences and variables such as weather conditions and vehicle and freight characteristics. A multivariate linear regression model was developed using the hourly rate of truck-crashes as the dependent variable, which could predict the possibility of the occurrence of wind-induced truck crashes. Results, however, showed that high wind speed was statistically insignificant toward the prediction of crashes. This study informed the identification of certain corridors in Kansas as potential areas in which to implement a warning system. Further, specific zones on the highways where truck drivers typically do not exhibit altered driving behavior resulting from changing wind speeds were identified. Distributions of wind-induced truck crashes were presented based on variable wind speeds, and suitable recommendations were provided based on research findings.

A study by Golob and Regan was conducted to determine the relationship between truck accidents and traffic flow conditions as well as roadway characteristics on urban freeways (9). Crash data pertaining to accidents, roadways, and traffic flow were obtained from the Traffic Accident Surveillance and Analysis System (TASAS) database for a period of two years for six freeways in Orange County, California. A multinomial logit model was developed to determine the different characteristics of traffic and roadway conditions conducive to weaving, runoff, and rear-end-type truck accidents. The number of truck-involved crashes was found to be inversely proportional to number of lanes and AADT per lane, and directly proportional to the percentage of large-trucks on the road. Other characteristics such as time of day, weather conditions, and days of the week were compared to non-truck crashes, the results of which varied substantially.

Khattak et al. performed a study to determine how single-vehicle truck crashes were influenced by various driver-, vehicle-, environment-, roadway-, and crash-related events (10). In addition to independent explanatory variables, this study considered various interaction terms (e.g., curve*rollover, grade*rollover, seatbelt*rollover, etc.). A comparison was made between rollover and non-rollover truck-involved single vehicle crashes. The study was performed in North Carolina, and utilized data from 1996-1998, obtained through the Highway Safety Information System (HSIS) database. Descriptive statistics, along with cross tabulations, were presented. Binary probit models, with rollover occurrence as the dependent variable, were developed to predict rollover propensity, and ordered-probit models were developed to predict injury severity. Multivariate statistical techniques were used to determine effects and interdependencies among the explanatory variables. Rollovers were found to have occurred in 30% of all truck-crashes and in 43% of truck-crashes occurring on road curvatures. Rollovers were related to increased crash severity.

Dissanayake and Bezwada analyzed characteristics and contributory causes related to fatal crashes involving large-trucks in the U.S. 2003-2007 data was obtained from the FARS database. Various driver-, roadway-, environment-, and vehicle-related factors contributing to the occurrence of large-truck crashes were identified. The likelihood of these factors being present in fatal truck-crashes was compared to their likelihood in fatal non-truck crashes using the Bayesian Statistical Approach (7). A multinomial logistic regression model was developed using the type of crash (truck or non-truck) as the dependent variable. In addition to driver-related factors such as cellular phone usage, failure to give right of way, and inattentiveness, other factors such as inadequate warning signs and poor shoulder conditions were found to be predominant causes that contributed more significantly to truck crashes than non-truck crashes. The model also demonstrated that the majority of single-vehicle fatal truck-crashes occurred on rural roads.

A study carried out by Charbotel et al. assessed the severity of injuries sustained by drivers of trucks involved in crashes (11). The study was performed in the Rhone region of France, using data from Trauma Registry for Road Crash Victims database for the years 1995-1999. Different victim characteristics (e.g., age, place of residence, etc.) and crash characteristics (e.g., place, time, antagonistic driving, and seatbelt wearing) were observed, followed by multivariate analysis using logistic regression. In addition, chi-square tests were performed to compare truck and car crashes. Variables were chosen based on a significance value. Results showed trucks were more dangerous for the safety of other road users. Also, it was concluded that professional driving was a high-risk occupation; risk factors such as driver age, antagonistic driving, and seatbelt usage were identified as relating to the severity of truck crashes.

Torre and Rossi identified potentially dangerous locations for safety regarding heavy good vehicles (HGVs). Data was obtained for four countries (Italy, France, the United Kingdom, and Finland) from a common database. Crashes were grouped according to road section, type of heavy vehicle, and type of accident (12). Analysis was conducted by investigating the distribution of different explanatory variables obtained from the database, or by using the equation for accident rate, a measure of crash occurrence. The findings were used to identify situations in which trucks had a higher probability of being involved in a crash. The study identified tractor semitrailer as the truck type most likely to be involved in a severe crash. Rural highways, urban highways, primary roads, and secondary roads were identified, in that order, as the most accident-prone locations. Work zone location characteristics such as narrow roads, traffic signs, barriers, and barricades increased the probability of a crash occurrence as compared to other roadways, especially as the size of the vehicle increased.

A (2000) study by Khattak and Darga examined the occurrence of truck crashes in North Carolina in the year 2000. The research involved a comparison between truck and non-truck vehicles, at both work zone and non-work zone areas (13). The HSIS database, along with police reports, were used to obtain statistical data such as type of work zone, presence of warning signs and cones, type of activity in the work zone, crash location, construction impact of the work zone on the roadway, etc. Severity measures of various crashes were presented as ‘most seriously injured occupant,’ or ‘total harm.’ The study combined crash frequency and injury severity. An ordered-probit model was developed for injury severity. The study showed that multi-vehicle crashes involving trucks were the most harmful type of collisions among all crash types.

Data pertaining to the state of Michigan from 1987-1988 was used in a study by Blower et al. Accident counts were taken from police reports and classified based on configuration, time of day, road type, and area type. Accident rates (a measure of exposure being vehicle miles traveled) were used as the dependent variable (14). Contingency tables were prepared and accident rates of heavy truck-tractors were modeled using the log-linear method. Two models were developed, one each for fatal crashes and property-damage-only crashes, respectively. Chi-square statistics and deviance were used to obtain goodness-of-fit statistics. The study showed that for all truck types, with the exception of bobtails, the probability of being involved in an accident depended more on the operating environment than on the configuration of the truck. Further, characteristics such as time of day, road type, and area type were more likely to cause a crash, as compared to whether the vehicle was a single or double truck.

A study by Mannila (15) analyzed all two-vehicle crashes involving two cars or a car and a truck, and various contributory causes. Crash data for a five-year period from 2000 to 2004 were obtained from the General Estimate System of the National Sampling System (NSS GES) and FARS databases. Crashes were classified into different categories based on the type of collision (e.g., angled, rear-end, head-on, etc.). Statistical analysis was conducted using logistic regression. Binary-logit models and multinomial logistic-regression models were used to identify significant contributing factors. Results obtained for car-truck crashes were compared with car-car crashes. The study showed that various environmental causes, driver-related causes, and speeding significantly increased the risk of car-truck crashes. Angled collisions were found to constitute the highest percentage of car-truck crashes. Also, speeding and alcohol involvement were found to increase the risk of crash involvement for both cars and trucks.

Duncan et al. modeled injury severities of occupants involved in rear-end collisions between trucks and passenger cars. The HSIS was used to obtain data for the state of North Carolina, which, according to 1993-1995 HSIS data, has long truck routes and a high number of rear-end collisions involving trucks (16). Factors influencing injury severity in truck-involved, rear-end collisions were initially presented, and then modeled using the ordered-probit model.

Interactions among independent variables were also taken into consideration while modeling. Variables such as light conditions, speed, speed limits, gender of the driver, influence of alcohol, and grade were found to increase injury severity of occupants of passenger cars involved in truck crashes.

2.2 Logistic Regression

Moghaddam et al. performed a study to identify the main factors responsible for increasing crash severity on urban highways (17). Highways of Tehran, Iran, were selected for analysis. Data from 2004-2008 relating to various factors prevailing during the occurrence of crashes were considered for analysis. Binary-logit models were developed to determine the simultaneous influence of human factors; road, vehicle, and weather conditions; and traffic features on crash severity. Selection of significant variables was carried out using the backward-regression method. Developed models showed that crash severity varied with the influence of multiple factors interacting simultaneously, rather than the action of any single factor. Factors such as age and gender of the driver, light conditions, behavior of the driver, defective vehicular components, manner of collisions, multi-vehicle crashes, etc., were found increase crash severity.

Liu et al. illustrated patterns of injury severity and contact sources by age. The National Highway Traffic Safety Administration's (NHTSA) National Automotive Sampling Systems Crashworthiness Data System (NASS-CDS) was used to obtain data for the years 1993-2004. Data was analyzed based on rollovers and seat belt usage (18). Frequency tables were presented, and chi-square analysis was performed to determine the dependence of injury severity on age. A logistic-regression model was developed in order to predict the severity of injury based on age. Odds ratios were used as supportive information. The study showed that males sustained more severe injuries than did females among young-driver crashes, while females sustained more severe injuries than did males among older-driver crashes. A majority of the severe crashes resulted in injuries to the head or chest. Further, seat belt usage was found to significantly reduce injury severity.

Dissanayake compared factors affecting the severity of injury to young and older drivers involved in single-vehicle crashes (19). Binary-logistic-regression models for both driver groups were developed using crash severity as the dependent variable. Variables related to roadway, environment, driver, and vehicle characteristics were used as explanatory variables. Five different models were developed, representing five different levels of severity. Data for this study was obtained from the Florida traffic-crash database, which was obtained from the state data program. The models were checked for goodness of fit. The driver being under the influence of drugs/alcohol was found to reduce the severity of crashes involving older drivers. Speeding and the driver not using a restraint device were important factors contributing to higher crash severity. Curved highways and driver ejection increased the severity of young-driver crashes. Crashes with frontal-impact points increased the severity of older-driver crashes.

A study performed by Conroy et al. illustrated differences in injury patterns, severity, and sources of injury as they related to the type of damage sustained by the vehicle in head-on collisions (20). Field investigations were conducted at multiple centers, and crash data for 1997-2006 were obtained from the Crash Injury Research and Engineering Network (CIREN) program. Different variables related to occupants, vehicles, and crashes were identified, and their relation to injury severity was identified using chi-square or Fisher exact statistics and odds ratios. Logistic-regression models were developed and analyzed. Hoshmer-Lemeshow goodness-of-fit statistics were used to check the fit of the developed logistic-regression model. The study showed that distribution of damage across the frontal plane, intrusion, and vehicle body type were important factors to consider in the study of occupant injuries in motor vehicle crashes.

Malyshkina and Mannering studied the effects of increasing speed limits of rural interstate and multilane non-interstate routes in the state of Indiana from 2004 to 2006, since speed limits were increased there in 2005 (21). Data was obtained from the Indiana Electronic Vehicle-Crash-Record System (EVCRS) database, where data were available in three different categories, namely, roadway and environmental data, vehicle data, and occupant data. The study was performed with consideration of the social and economic burden of truck crashes. A multinomial-logit model was developed using accident severity as the dependent variable. The study showed that speed limits did not significantly affect injury severity on interstates, unlike non-interstates, where higher speeds were associated with greater injury severity.

Gabauer and Gabler studied the effects of airbags and seatbelts on the injury severity of occupants involved in longitudinal-barrier crashes (22). Data from 1997-2007 were obtained from the National Automotive Sampling System/Crashworthiness Data System. Binary-logistic-regression models were developed to predict the risk of occupant injury, and a comparison was made based on the type of restraint used. The study showed that concrete barriers were associated with a higher rate of airbag deployment than were metal barriers. In single-event, longitudinal-barrier crashes, seatbelts and airbags were found to reduce the severity of injuries sustained by occupants.

2.3 Severity Modeling

A study performed by Eboli and Mazzulla explored the relationships between road accident severity and number of people injured, number of vehicles involved, and other accident characteristics (23). Data pertaining to Cosenza province, Italy, for the year 2003 was considered. Accident severity was related to factors such as road characteristics, environmental context, and driver characteristics. A developed structural equation model contained latent variables which were unobserved road accident aspects that cannot be explained by observed variables. The parameter-estimated standard error, critical ratio, level of statistical significance of each variable, and various goodness-of-fit indices were calculated, along with indirect effects of observed variables on latent variables.

Wang studied the characteristics of the crashes occurring in work-zone areas, and factors contributing to different injury severity levels (24). Crash data was obtained for the state of Florida for a period of five years from 2002-2006 using the Florida Crash Analysis Reporting (CAR) system database. A descriptive statistical analysis for work-zone crashes for different age groups was performed, along with a comparison between crashes occurring in work-zone and non-work-zone areas. An ordered probit model was developed to model injury severity. The study showed that middle-age drivers were involved in a higher percentage of work-zone crashes and no-injury crashes. Careless driving and failing to yield the right of way were important driver-related contributory factors in work-zone crashes. Also, heavy vehicles were found to have greater involvement in work-zone crashes.

Liu and Dissanayake examined issues relating to speed limits on gravel roads in Kansas. The study was performed in three facets, included field studies, questionnaire surveys, and statistical analysis of crash data (25). The field study was performed in Riley County, and included on-site data collection. Questionnaire surveys included a collection of opinions and comments from local county engineers. Data from the KARS database was extracted for the years 2003-2005. A contingency table test method was performed as part of the statistical analysis. Data obtained from the three methods were analyzed. The study showed that 55 mph was the most acceptable speed limit on gravel roads in Kansas under existing road conditions. Lower speed limits were found to characterize less severe crashes.

Lemp et al. examined various factors affecting crash severity of occupants involved in heavy-duty truck crashes by analyzing records contained in the recent Large-truck Crash Causation Study (LTCCS) data, provided by the United States Federal Motor Carrier Safety Administration (FMCSA) and NHTSA. Data was also obtained through interviews with drivers, passengers, and witnesses. Standard Ordered Probit (SOP) models and Heteroskedastic Ordered Probit (HOP) models were used to illustrate the impact of various vehicle, environmental, and occupant characteristics on injury outcomes (26). The same set of variables was used in both SOP and HOP models. HOP models offered greater model flexibility than SOP models, since they captured the effect of crash characteristics on the variance or uncertainty in crash severity. Crash severity and injury severity were used as response variables, and all independent variables were broadly classified as crash-level variables, largest-truck attributes, and vehicle- and driver-related variables. SOP and HOP models developed were compared using log likelihood values, and then analyzed. Analysis showed the probability of the occurrence of a fatal crash increases with the number of vehicles involved and the number of truck occupants. Also, fatality likelihood was observed to increase with the number of truck trailers and decrease with length and gross vehicular weight rating of the truck.

Kockelman developed an ordered probit model to examine the risk of different levels of injury severity under the categories of all crashes, single-vehicle crashes, and two-vehicle crashes, respectively (27). Data related to crash, vehicle, and persons was obtained from the National Automotive Sampling System's General Estimate System (NASS GES) for the year 1998. The data was a sample of police-reported crash records. These explanatory variables were used to model driver injury severity, both with and without the speed variable. The study showed that rollovers and head-on collisions increased the severity of the crash. Late-night driving on weekends and in daylight conditions had negligibly small influential effects. Light-duty trucks were observed to provide relatively better safety to their occupants.

A study performed by Ma and Kockelman used data from state highways in Washington for the year 1996, using the Highway Safety Information System (HSIS) database (28). In this study, a multivariate Poisson specification, as well as a Bayesian technique, was used to perform a joint study of crash frequency and severity. In addition, Gibb's sampler, as well as the Metropolis-Hastings (M-H) algorithm, was established to estimate parameters of interest for Bayesian statistical inference. For the purpose of comparison, a series of univariate Poisson models for injury counts were estimated. Tables were developed for all injury-severity levels, showing the frequency of a condition under different injury-severity levels. Expected percentage changes in injury rates corresponding to changes in variables were calculated, and a cost analysis was conducted using NHTSA's estimate-of-injury costs. The study showed that the travel time saved by increasing the speed limit by 10mi/hr was not worth the economic loss generated by resultant crashes.

2.4 Countermeasure Ideas

The I-80 corridor in Iowa was considered in a study by Burke, as it is a major connector of many areas of the country (29). Prior to the study there had been an increase in that location in number of trucks on the interstate, which in turn resulted in greater congestion, greater pavement deterioration, and a spike in auto-truck accidents. Burke discussed advantages and disadvantages of providing an exclusive travel lane for trucks, and discussed the design of a truck lane by taking both passing lanes and the breakdown lane into consideration. Respective costs were predicted based on factors such as cumulative mileage, right-of-way costs, terrain costs, etc. The study demonstrated that dedicated truck lanes help in getting long-term benefits.

Rau considered the detection of driver drowsiness and the effects of employing a warning system for commercial vehicle drivers (30). The NHTSA identified drowsiness as the most important factor relating to safety concern among commercial vehicle drivers. NHTSA's five years of data from 1989 to 1993 were utilized. A field operational test (FOT) was performed during 2004-2005, in which three main participating research partners had analyzed and predicted the effectiveness of employing warning systems, such as the drowsy driver warning system (DDWS). Analyzing results from the FOT, it was concluded that further research was needed pertaining to highway safety benefits, fleet acceptance, operational utility, and fatigue management practices, in order to minimize fatigue-related crashes.

Council et al. examined driver being at fault in non-fatal crashes, a provision of crash-based validation for unsafe driving acts (UDAs), and identified critical combinations of crash types at specific roadway locations through an analysis of the total harm resulting from the combination of crash and type of site (31). Analysis was performed for the state of North Carolina and findings were compared to earlier standard findings. Findings obtained were observed to differ slightly from standard findings. Truck drivers were found to be more at fault in collisions occurring due to backing, right turn, left turn, rear-end and sideswipe crashes. The driver of the car was found to be more at fault during collisions due to maneuvers, such as head-on and angled collisions.

Montella and Pernetti studied a 127.5 km section of motorway in Italy (32). Data were considered for the years 2001-2005, and were obtained from a number of sources, including police reports, hospital reports, and site investigations. The main aim of this study was to point out risk factors associated with the motorway that could be considered by highway agencies and designers as suggestions for suitable safety countermeasures to aid in reducing run-off-the road (ROR) crash frequency and severity. The chi-square test with Yate’s correction was performed to determine whether the parameter was significant. Number of ROR crashes for both trucks and cars were obtained and compared. Crash severities in relation to various significant parameters were analyzed. Results showed that the severity of crashes involving motor vehicles was significantly higher than those involving other vehicles. Also, severities of crashes on the roadways involving blunt-end terminals were higher than those on roadways with longitudinal barriers (e.g., guardrails).

Wang et al. considered loss-of-life and financial burden as they relate to traffic accidents (33). The study examined causes of crashes on two-lane, rural highways in Washington. Six study routes were chosen based on length, location, and geometric characteristics for a period of six years between 1999-2004. Corresponding data were obtained from the HSIS Roadway Video Image Data, and Geographic Information System (GIS), retrieved from the Washington Department of Transportation. Segments of roads and intersections were considered in two different categories. T-tests and analysis of variance (ANOVA) were performed to identify significant contributory causes in the occurrence of a crash. The data was used to develop the Poisson regression model, negative binomial regression, zero-inflated Poisson, and negative binomial models. Effects of factors such as speed limit, degree of curvature, shoulder width, grade percentage, etc., on risks involved in all types of crashes and rear-end crashes were summarized. Cost-effective methods of mitigating risk on roadway segments, such as avoiding frequent speed-limit changes, widening road surface and shoulder widths, etc., were discussed.

Li and Bai developed a variable—the crash severity index (CSI)—which was modeled as a measure of risk levels associated with work-zone crashes (34). Data relating to fatal crashes occurring between 1998-2004 and injury crashes between 2003-2004 was obtained from a Kansas Department of Transportation (KDOT) database. Four CSI models were developed using the logistic regression method, and were analyzed using crash data. The chi-square statistic and the Cochran-Mantel-Haenszel (CMH) statistic were used to ensure accuracy of the factors associated with the risk involved in crashes. CSI values for most work zone crashes were found to be consistent with actual crash severity outcomes. Benefits of implementing the method of using CSI values were presented, along with countermeasures to mitigate risk involved in work zone crashes.

Oh et al. analyzed pedestrian-vehicle crashes in Korea, with the aim of mitigating fatalities and injury severity among pedestrians. Considering pedestrians as the most vulnerable elements in the transportation system (35), this study focused on developing a probabilistic pedestrian-fatality model. Relevant data was collected for a period of one year using accident report forms. This data was analyzed by the National Institute of Scientific Investigation (NISI) and Center for Accident Analysis of Hanyang University. A binary logistic regression model was developed using pedestrian fatality as the dependent variable. Out of all available data provided for explanatory variables, collision speed, vehicle type, and pedestrian age were selected for modeling. Collision speed was the most significant variable. The model was developed primarily with the aim of providing countermeasures, in the realm of both transportation safety and automobile operations. The study showed that the probability of a fatality decreased as the age of the pedestrian increased. Heavy vehicles had a greater probability of causing more severe crashes, as compared to lighter vehicles. Findings of the study were summarized, and areas for future research were discussed.

Dissanayake and Lu analyzed differences between domestic and international drivers in the U.S., considering crashes that may have occurred as a result of unfamiliarity with road rules among international tourists (36). A comparison was made between domestic and international drivers in regards to the comprehension of traffic-control devices. The study was performed at the departing areas of two international airports in Florida (Tampa and Orlando). Survey forms and a questionnaire were supplied to passengers; using cross classification, these were later analyzed and checked for relationships among variables. The study showed that international respondents were satisfied with the transportation system in the U.S., but less satisfied with traffic-control devices. Both domestic and international respondents were less satisfied with the availability and accuracy of information associated with the transportation system.

Dissanayake and Ratnayake performed a study to explore the reduction of crash severity on rural highways in Kansas, and to identify suitable countermeasures to enhance the safety of the rural highways (37). Data was obtained from the KARS database for the years 1998-2002. Modeling approaches comprised of ordered choice which included ordered-probit and ordered-logit models along with log-linear models. The study found that crashes involving drivers with no safety equipment involved more severe injuries. Further, injury severity was high when the driver ejected out of the vehicle after the crash. Single-vehicle crashes and head-on collisions were found to be relatively more severe than other crash types. A list of possible countermeasures to mitigate crashes in rural areas was a provided and discussed.

Chapter 3 Methodology

3.1 Data

The first phase of this study utilized data from the FARS database to identify characteristics and contributory causes related to large-truck crashes in the U.S. (7). However, this database contained information pertaining only to fatal crashes, and could not be used to study crashes of different severity levels. Data for the second phase was obtained from KDOT's KARS database, which contains details of police-reported crashes at all severity levels occurring in the state of Kansas. The database consists of a complete dataset containing information related to every truck crash in Kansas, as well as a limited dataset consisting of data pertaining to truck-crashes occurring only on the state highway system, which is comprised of Kansas highways, interstate highways and U.S. routes. This database is an integration of various driver-, vehicle-, environment- , and road-related characteristics that prevailed at the time of a crash. The database includes some inaccurate or missing values, either because of lack of complete information or due to human error during electronic data entry. Details such as name, address, contact number, and other such personal information related to individuals involved in crashes are restricted to the public in order to maintain privacy. Data obtained from this database were redefined by codes to simplify the data entry process. These codes are explained in KDOT's Kansas Motor Vehicle Accident Report Coding Manual (39).

Injury severity was determined and categorized as fatal, disabling, non-incapacitating, possible, or Property Damage Only (PDO), based on the level of injury sustained by vehicle occupants. Crash severity was the dependent variable used for analysis in this study, and was identified based on the highest level of injury severity sustained by the occupants involved in a crash. Severity type was recorded as fatal if the death of an occupant occurred within 30 days of the occurrence of the crash (39). A disabling injury was defined as one which resulted in preventing an occupant from performing normal activities, such as walking and driving, after the occurrence of the crash. Conversely, non-incapacitating injuries were those that occurred during a crash but did not result in disability. All other injury types were categorized as possible injuries. A severity type classified as PDO was one which involved neither fatality nor notable injury to a vehicle occupant.

For the purposes of this study, a truck having a gross vehicular weight rating of 10,000 pounds or more was considered to be a large-truck. Based on vehicle body type, large-trucks included single heavy trucks, truck and trailer(s), and tractor-trailer(s), as obtained from the Kansas Motor Vehicle Accident Report Coding Manual (39). Data pertaining to crashes involving large-trucks in Kansas occurring between 2004-2008 was considered for this study. For crashes involving more than one truck, information pertaining only to one truck was considered, as the number of such crashes was negligible.

Various truck crash characteristics were available from different files in the database. These files were combined for this study using the accident key variable, which is unique for each crash. Once combined, data were filtered using Microsoft Access and Microsoft Excel to avoid redundancy. The filtered dataset resulted in a total of 18,919 separate truck crash records. The finalized dataset was exported to Statistical Analysis System (SAS) version 9.2 (40) for further analysis.

3.2 Cross-Classification Analysis

Cross-classification analysis, also known as contingency table analysis, can be performed to verify the dependency of various factors on the severity of truck-crashes. This test is used to identify the relationship between a pair of variables, one of them being crash severity. This analysis is associated with the hypothesis testing procedure, where the null hypothesis (H_0) and alternate hypothesis (H_A) for the study are defined as follows: H_0: Variable considered is independent of the crash severity. H_A: Null hypothesis is not true

If the null hypothesis is supported, there exists no relationship between the examined variable and truck crash severity. The level of confidence was considered to be 95%. In the cross-classification procedure, variables are subdivided into suitable categories and arranged in rows and columns. In this study, columns contained the five levels of crash severity, and rows contained the combined subcategories of the variables under consideration. For example, the variable ‘Light Condition’ can be categorized as Daylight, Dark with Lights, Dark without Lights, Dusk, Dawn, etc. These categories of variables are combined to obtain reasonably large values in the cells for cross-classification analysis. This is because smaller values of sample variables create smaller values for expected frequencies, which could influence inaccurate results (41).

If there are ‘n’ rows and ‘m’ columns in the matrix, then the degrees of freedom are given by the following expression (42):

$$\text{Degrees of Freedom} = (n-1)*(m-1) \tag{3.1}$$

Entries in the contingency table are recorded as the observed frequencies ‘O_{ij}’, where i and j denote the corresponding row and column. Expected values for any cell in the matrix ‘E_{ij}’ are calculated by multiplying the sum of the observations in the corresponding row and the column and dividing it by the sample size of the matrix (42). In other words,

$$E_{ij} = \frac{\text{(Row Total)*(Column Total)}}{\text{Sample Size}} \tag{3.2}$$

Having found this value, the chi-square (χ^2) statistic is computed as follows (42) :

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_{ij}-E_{ij})^2}{E_{ij}} \tag{3.3}$$

where,

k is the number of cells in the contingency table.

Using the value of the obtained degrees of freedom from equation 3.1., the rejection region for a confidence interval of 95% can be determined from the standardized chi-square distribution table, which gives the tabular chi-square value. This value is compared with the calculated chi-square value obtained using equation 3.3. If the calculated value is greater than the tabular value, then the null hypothesis is rejected, which means a relationship exists between the variable under consideration and the crash severity. On the other hand, if the calculated value is less than the tabular value, then the null hypothesis is not rejected, which means the two variables are independent of each other. Though this test is not very accurate or perfect, it gives a rough idea about the relationship between the variables. SAS version 9.2 (40) was used to perform the cross-classification analysis.

3.3 Multicollinearity

The filtered data was imported into SAS Version 9.2 (40) for further analysis. All candidate variables considered in the modeling procedure were recoded as binary values of 0 or 1. Independent candidate variables were first checked for linear dependencies using the correlation matrix. Presence of two correlated variables in the model at the same time reduces the accuracy of the impact of one variable on the crash severity. . The PROC CORR statement available in SAS Version 9.2 (40) was used to generate the matrix. Each of the values generated in the matrix are Pearson's correlation coefficients. Their magnitudes determine the strength of the relationship between the corresponding variables. According to Oh et al., a Pearson's correlation coefficient of 0.5 indicates high multicollinearity between the corresponding pair of explanatory variables (35). Hence, a correlation coefficient of 0.5 was chosen as the cutoff value. Pairs of variables having a coefficient of 0.5 or higher were considered one by one to minimize the effect of multicollinearity. The pair of explanatory variables with the highest correlation coefficient was considered first. Each of the two variables was alternately placed in the model and model strength was calculated using model-fit statistics. The variable that resulted in a weaker model was discarded, while the other variable was retained in the model. The procedure was repeated for the pair of variables having the next highest magnitude of the correlation coefficient. The process was continued until no remaining pair of variables had a correlation coefficient of 0.5 or higher. This substantially mitigated the effect of multicollinearity among the explanatory variables.

3.4 Binary Logistic Regression

The odds ratio, which is defined as the ratio of the probability of the occurrence of an event to that of its non-occurrence (38), was used to describe the influence of each of the candidate variables on crash severity. In this study, the "event" refers to cases in which the crash-severity variable was given a value of 1. The odds ratio (O) was provided by the following expression:

$$O = \frac{p}{1-p} \tag{3.4}$$

where,

p = probability that the crash severity takes a value of 1

Probabilities are generally bounded, while linear functions are unbounded. Transforming the probability to odds and taking its logarithm removes the bounded nature of the dependent variable. A logistic model is obtained by setting the logarithm of odds of the dependent variable to a linear function of the explanatory variables (38). A logistic regression model with k explanatory variables and i = 1, 2 … n individuals has a general form as follows (38):

$$\log \left[\frac{P_i}{(1-P_i)}\right] = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \ldots\ldots\ldots\ldots\ldots \beta_k x_{ik} \tag{3.5}$$

where,

α = value of the intercept,

β = estimates of different independent variables in the model, and

$x_{i1}, x_{i2} \ldots x_{ik}$ = interval-level or indicator variables associated with crash i.

The expression for p_i can be obtained by solving the logistic equation 3.4 as follows:

$$p_i = \frac{1}{1+\exp(-\alpha-\beta_1 x_{i1}-\beta_2 x_{i2}- \beta_3 x_{i3}-\ldots\ldots\ldots-\beta_k x_{ik})} \tag{3.6}$$

Since p_i is the probability of the variable crash severity displaying a value of 1, the value of p_i ranges between 0 and 1 for all values of x’s and β’s. A logistic regression model predicts the probability that the dependent variable displays a given value for a particular set of explanatory variables (19). In the case of a binary logistic regression model, the dependent variable displays values of either 0 or 1.

The binary logistic regression model is an efficient tool for modeling crash severity, which has been considered as a dichotomous dependent variable (38). Crash severity, denoted as 'Y' in this case, is redefined as follows: Y = 1, if the occupants involved in the truck crash sustained injury of any severity level. Y = 0, if the occupants involved in the truck crash did not sustain any injury.

A total of 46 independent variables related to vehicle, driver, road, and environmental characteristics, such as alcohol, light conditions, speed limit, etc., were considered for the model. The PROC LOGISTIC statement, available in SAS Version 9.2 (40), was used to develop models using the three variable selection methods, which include forward selection, stepwise selection, and backward elimination methods. In the forward selection method, a model initially contains no variables. Variables enter one by one until all the variables in the model have significant p-values (40). A p-value of 0.05 was chosen as the level of significance in the current study. Any variable with a p-value greater than 0.05 did not remain in the model (27). In the forward selection procedure, a variable once entered into the model will never leave the model (40). In the backward elimination method, a model initially contains all variables, and each variable is removed one by one until all remaining variables have a significant p-value of 0.05 (40); disregarded variables are not re-entered into the model. The stepwise selection procedure is a combination of forward and backward selection methods, where variables rotate in and out of the model until the best possible fit is obtained (40). These methods are used to identify the significant variables that are to remain in the model.

The maximum likelihood method (MLM) was used to estimate the coefficients of the explanatory variables in the model. Maximum likelihood is a general approach toward estimation that is used widely in many different methods of statistical modeling. According to P. D. Allison, "The basic principle of this method is to choose those parameter values as the estimates which if true, would maximize the probability of observing what we have, in fact, observed (38)."

The value of the R^2 statistic, which represents the amount of variability in the model explained by the independent variables, was used in selecting the best model, with greater R^2 values corresponding to a better model. MLM generates important model fit statistics, such as the Akaike Information Criterion (AIC), Schwarz Criterion (SC), and the value of twice the negative of log likelihood (-2 log L), both for the intercept-only and the fitted model. AIC and SC values are calculated as follows (38):

$$AIC = -2 * \text{log-likelihood} + 2k \tag{3.7}$$

$$SC = -2 * \text{log-likelihood} + k \log (n) \tag{3.8}$$

where,

k = number of estimated parameters, and

n = sample size.

These statistics can be used to make comparisons among a set of models obtained by different variable selection methods, with smaller values representing a better model (38).

Other goodness-of-fit statistics include the percentage concordant, percentage discordant, percent tied, pairs, Somer' s D, Goodman – Kruskal Gamma, Tau-a, and C values, which can evaluate the strength of the developed model. Descriptions of these parameters are as follows (7):

- Percent concordant – A pair of observations with different observed responses is said to be concordant if the observation with the lower ordered response value has a lower predicted mean score than the observation with the higher ordered response value.
- Percent discordant – If the observation with the lower ordered response value has a higher predicted mean score than the observation with the higher ordered response value, then the pair is discordant.
- Percent tied – If a pair of observations with different responses is neither concordant nor discordant, it is a tie.
- Pairs – This is a number of distinct ways of pairing up different observations. The concordant pairs, discordant pairs, and tied pairs aggregate to give the total number of pairs. Each of the percent concordant, percent discordant and percent tied is calculated with respect to the total number of pairs.
- Somer's D – Somer's D is used to determine the strength and direction of the relationship between pairs of variables. Its values range from -1.0 (all pairs disagree) to 1.0 (all pairs agree). It is defined as $(n_c-n_d)/t$, where n_c is the number of pairs that are concordant, n_d the number of pairs that are discordant, and t is the total number of pairs with different responses (38).
- Gamma – A Goodman-Kruskal Gamma value closer to 1 indicates good association among model variables. This method does not penalize for ties on either variable. Its values range from -1.0 (no association) to 1.0 (perfect association). It is defined as (n_c -

n_d)/ ($n_c + n_d$), where n_c is the number of pairs that are concordant and n_d is the number of pairs that are discordant (38).

- Tau-a – Kendall's Tau-a is a modification of Somers' D that takes into account the difference between the number of possible paired observations and the number of paired observations with different responses. It is defined as (n_c-n_d)/n, where n_c is the number of pairs that are concordant, n_d the number of pairs that are discordant, and n the total number of pairs (38).
- c – Another measure of rank correlation of ordinal variables is 'c'. It ranges from 0 (no association) to 1 (perfect association). It is a variant of Somers' D index. The value of c is given as (38):

$$c = 0.5 * (1 + \text{Somer's D}) \tag{3.9}$$

Chapter 4 Results and Discussion

This chapter summarizes the characteristics and contributory causes of crashes involving large-trucks in Kansas, utilizing five years of combined data from 2004-2008. Data obtained and analyzed from both the complete and limited datasets of the KARS database are presented.

A total of 18,919 truck-crashes were recorded on all Kansas roads, out of which 11,762 truck-crashes occurred on the state highway system. Analysis of the KARS database showed that large-trucks in Kansas resulted in more fatalities in the other vehicle as compared to the truck occupant itself. Greater than 81% of fatalities that occurred in crashes involving trucks occurred among occupants of the other vehicle.

4.1 Characteristics of Large-Truck Crashes on All Roads

4.1.1 Road-Related Features

The roadway where a truck crash occurs is an important consideration in understanding the characteristics of large-truck crashes. Figures 4.1-4.3 show the distribution of truck-crashes in Kansas based on the type, condition, and character of the road. Blacktop surface type, dry surface conditions, and straight and level surface characteristics have, respectively, constituted the majority of crashes under each category. One possible explanation is that more trucks travel under such conditions, resulting in a greater probability of those conditions characterizing a crash.

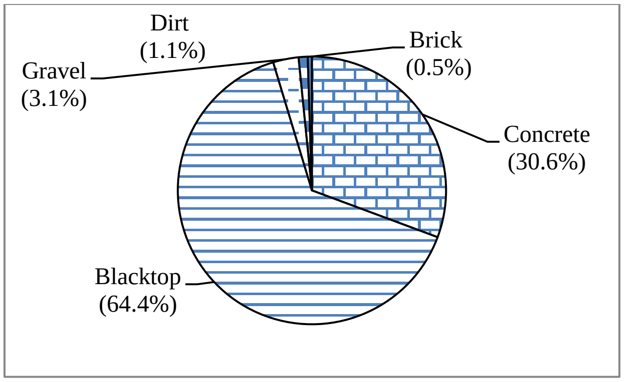

Figure 4.1 Distribution of truck-crashes in Kansas based on road surface type

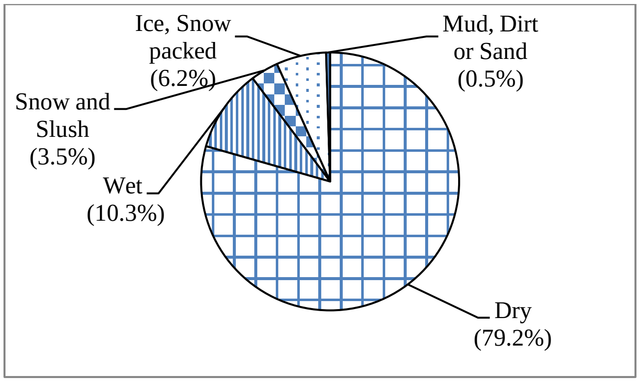

Figure 4.2 Distribution of truck-crashes in Kansas based on road surface condition

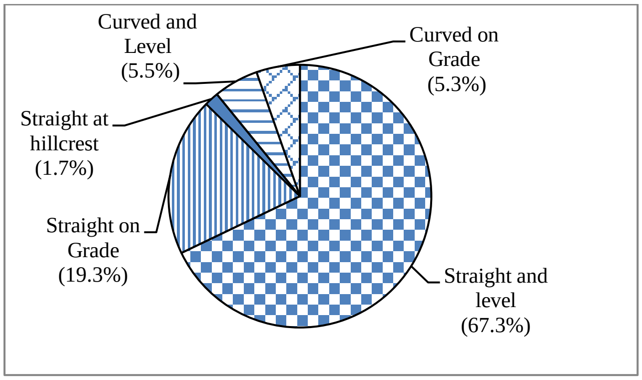

Figure 4.3 Distribution of truck-crashes in Kansas based on road surface character

4.1.2 Light Conditions

Large-truck crashes under different light conditions were categorized. Figure 4.4 shows the distribution of truck crashes based on different light conditions. A majority of truck-crashes have occurred in daylight conditions. One possible reason for this finding could be that more trucks are on the road during the day. Percentages of crashes under other light conditions were considerably low when compared to the daylight condition.

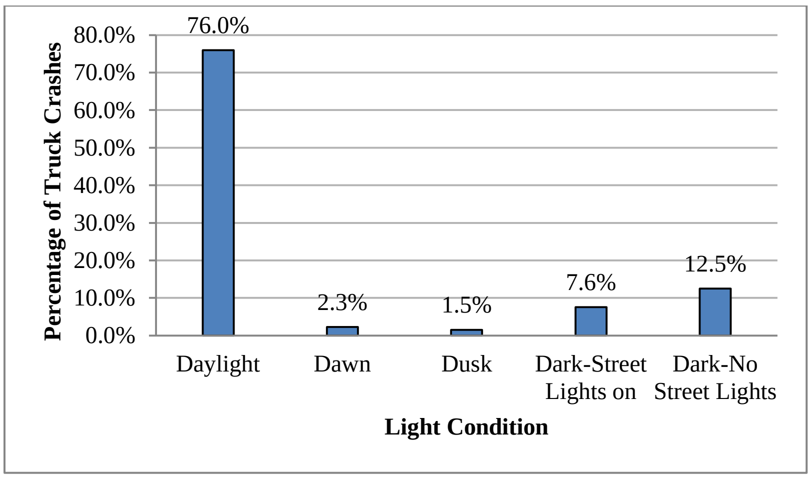

Figure 4.4 Distribution of truck-crashes based on light conditions

4.1.3 Weather Conditions

Large-truck crashes in Kansas were categorized based on weather conditions that prevailed during crash occurrences. The distribution of crashes is presented in figure 4.5. Analysis reveals that a majority of truck-crashes occurred under no adverse weather conditions. Rain, mist, and drizzle accounted for the greatest number of truck-crashes among adverse weather conditions, perhaps because those conditions are much more common than other adverse weather conditions.

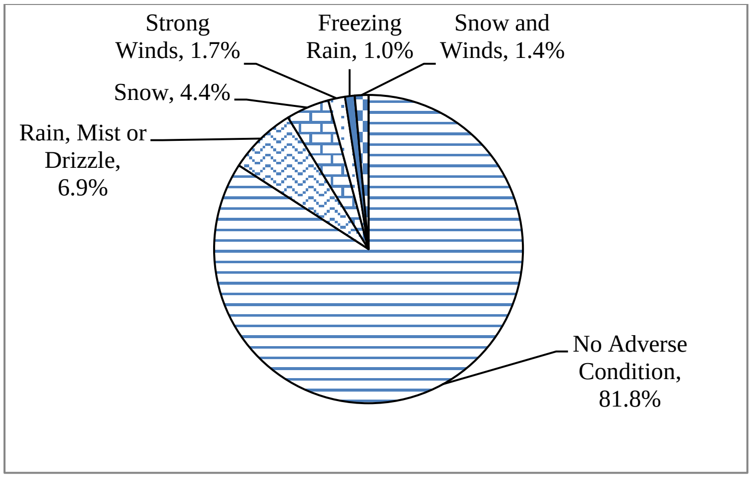

Figure 4.5 Distribution of truck-crashes based on weather conditions

4.1.4 Time of Day

Traffic conditions vary at different times of day, creating variable driving conditions based on time. Figure 4.6 shows the distribution of crashes based on time of day. Analysis of the data revealed that a majority of truck-crashes occurred in the afternoon between noon and 3:00 p.m., closely followed by the hours between 9:00 a.m. and 12:00. Overwhelming majority (77.6%) of truck-crashes occurred from 6 a.m. to 6 p.m. This statistic may be influenced by the fact that these hours comprise the entire business day, potentially placing more vehicles on the road during that time period. On the other hand, very few crashes occur at midnight, for example, due to relatively low traffic.

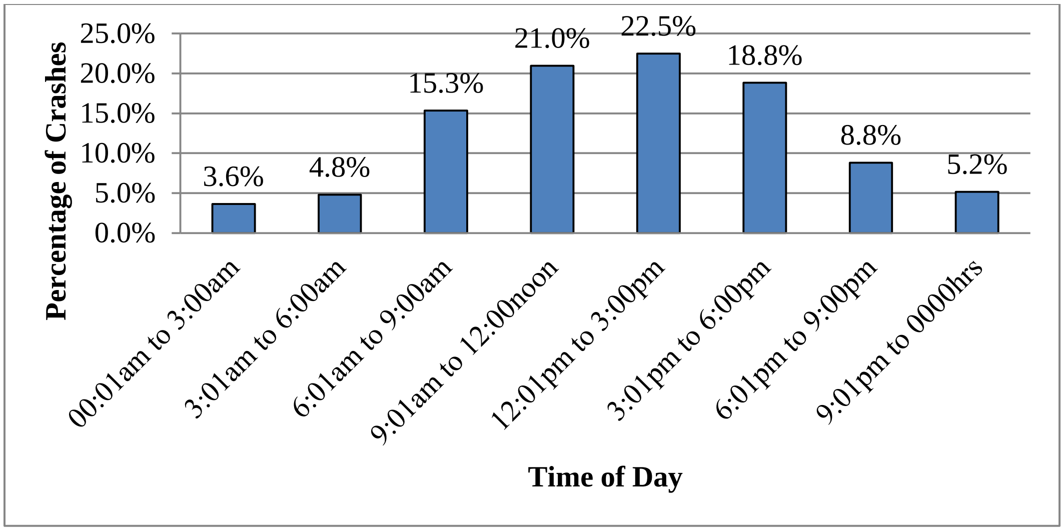

Figure 4.6 Distribution of truck-crashes based on time of day

4.1.5 Age of Truck Driver

Age of the truck driver is one of the factors useful for understanding the characteristics of crashes involving large trucks. Figure 4.7 shows the distribution of crashes involving large-trucks based on the age of the truck driver. Data analysis reveals that a majority of truck drivers involved in crashes were 21-40 years of age, followed by those who were between 41-60 years old. While there were some young and older drivers, 80.4% of truck drivers involved in crashes were between 20 and 60 years old.

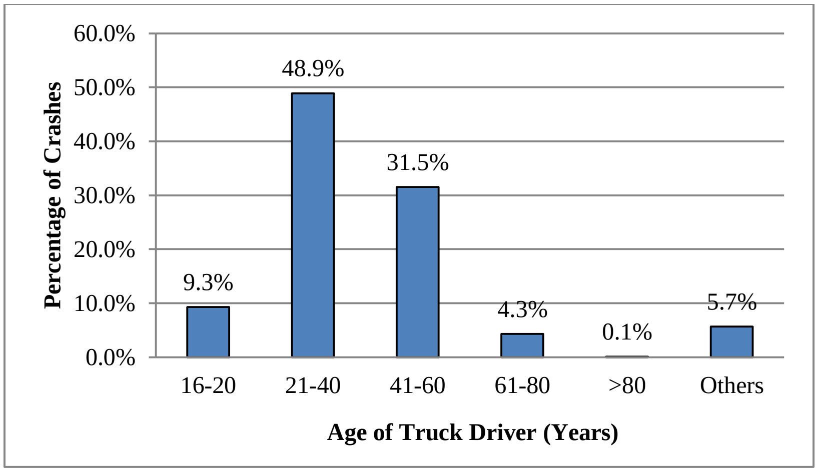

Figure 4.7 Distribution of truck-crashes based on age of truck driver

4.1.6 Gender of Truck Driver

Analysis of the KARS data showed that among truck drivers involved in crashes, nearly 79% were males. Figure 4.8 shows the distribution of large-truck crashes based on gender of the truck driver. This could be due to there being more male truck drivers than female drivers.

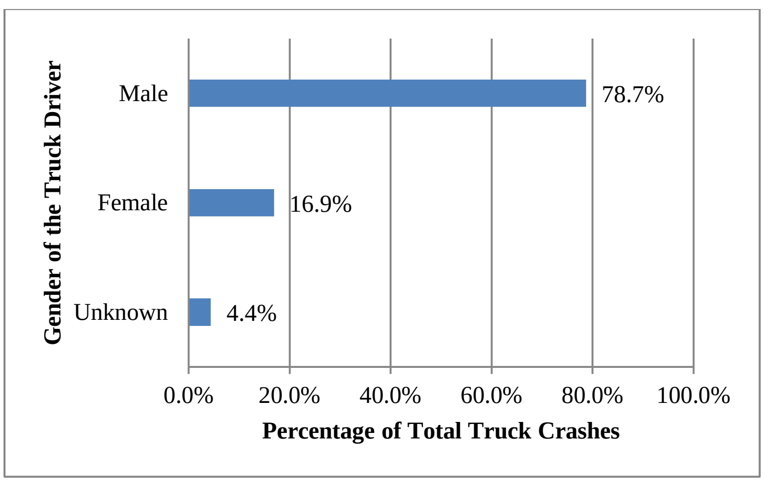

Figure 4.8 Distribution of truck-crashes based on gender of truck driver

4.1.7 Vehicle Maneuvers

Vehicle-related factors are important considerations in the analysis of truck-crashes and the development of solutions to mitigate them. Truck maneuverability is one such feature. Due to their large size, the maneuverability of large-trucks is significantly limited in comparison to other vehicles. Figure 4.9 shows the distribution of large-truck crashes based on truck maneuvers at the time of crash occurrence. Analysis of the data shows that more than half of all crashes occurred when the truck was going straight and following the road. Right and left turns were the other maneuvers that resulted in a significant number of crashes, followed by backing and changing lanes. Other truck maneuvers include merging, parking, backing, avoiding maneuver, stopping or slowing, and illegal parking. These maneuvers contribute to a small percentage of the total large-truck crashes in Kansas, respectively.

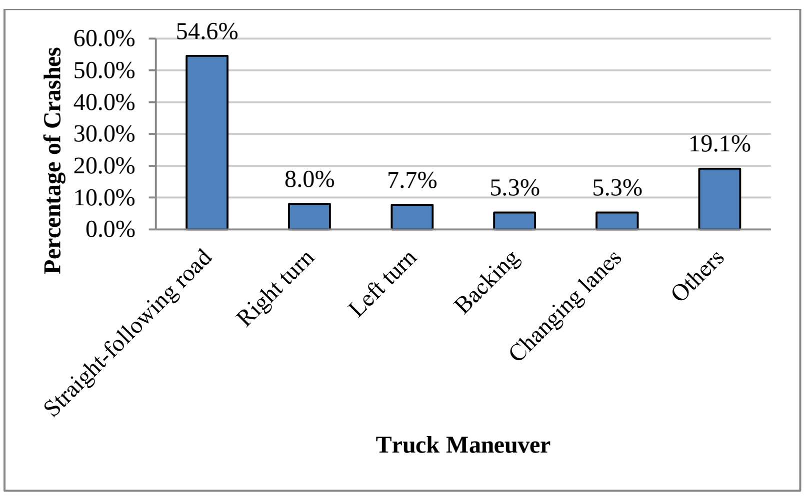

Figure 4.9 Distribution of truck-crashes based on truck maneuvers

4.1.8 Manner of Collision

The majority of truck crashes involved two vehicles, a truck and another non-truck vehicle, followed by a significant percentage of single-vehicle crashes. Figure 4.10 shows the distribution of truck-crashes based on the number of vehicles involved.

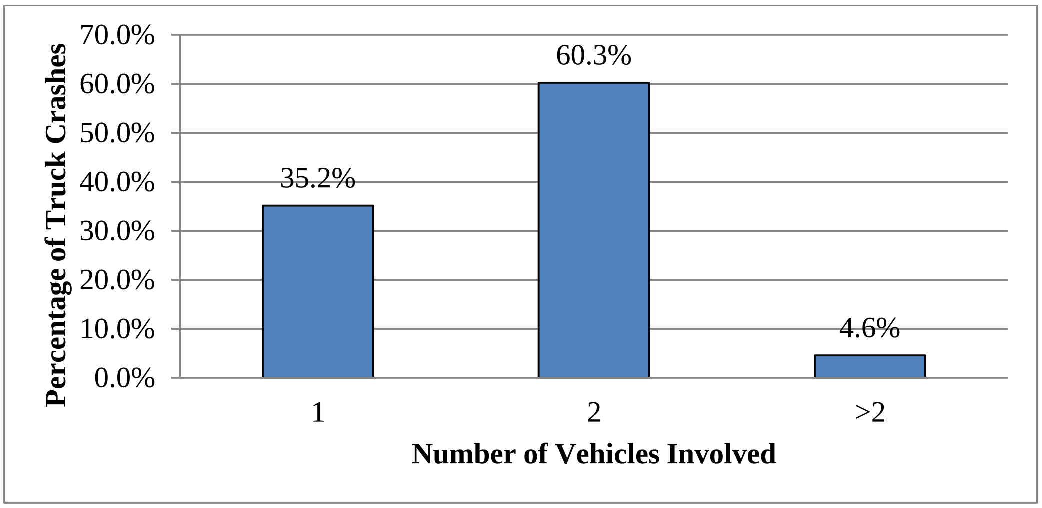

Figure 4.10 Distribution of truck-crashes based on number of vehicles involved

Truck-crashes involving more than one vehicle were further classified on the basis of their manner of collision, as shown in figure 4.11. Data revealed that the majority (30.4%) of multi-vehicle truck-crashes were angled collisions. Rear-end and sideswipe collisions also characterized a significant proportion of multi-vehicle truck-crashes.

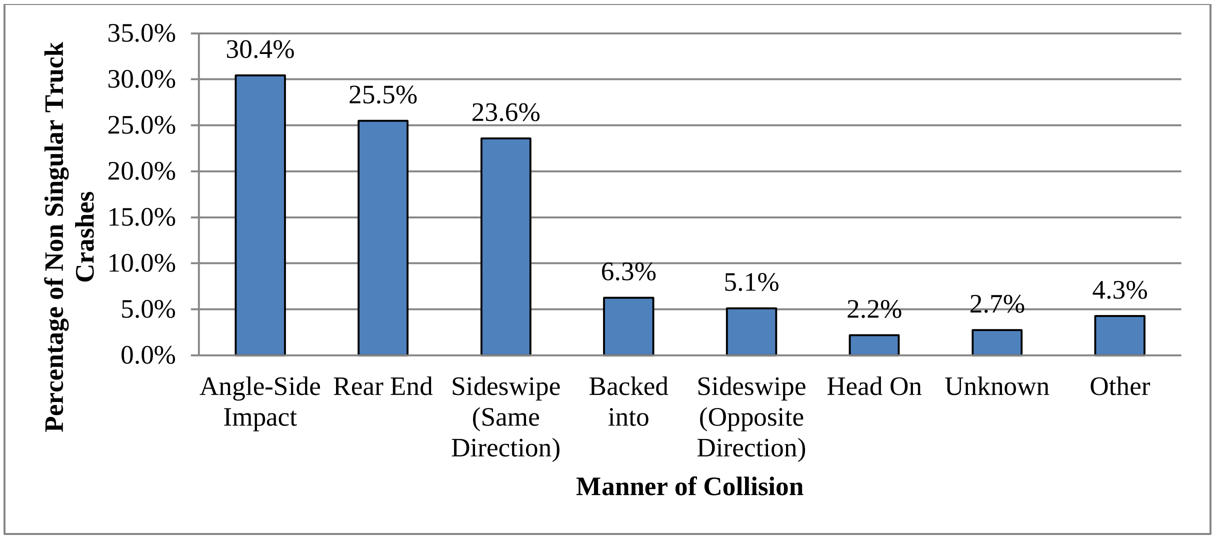

Figure 4.11 Distribution of multi-vehicle truck-crashes based on manner of collision

4.1.9 Vehicle Body Type

Figure 4.12 shows the distribution of two-vehicle crashes involving one truck and one non-truck vehicle based on body type of the other vehicle. Analysis of data showed that a majority of large-truck, two-vehicle crashes involved a car, followed by pickup trucks and sport utility vehicles. Other vehicles include trains, buses, farm equipment, and camper-rvs.

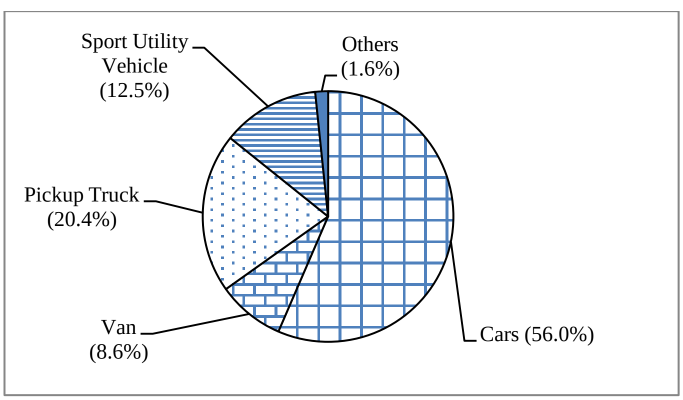

Figure 4.12 Distribution of two-vehicle truck-crashes based on body type

4.1.10 Day of the Week

The number of truck-crashes that occurred during weekends was relatively fewer than those on weekdays. Figure 4.13 shows the distribution of truck-crashes based on day of the week. Analysis of the data showed that the percentage of crashes occurring on each of the weekdays was rather uniform without much variation, with slightly more crashes being recorded on Wednesdays.

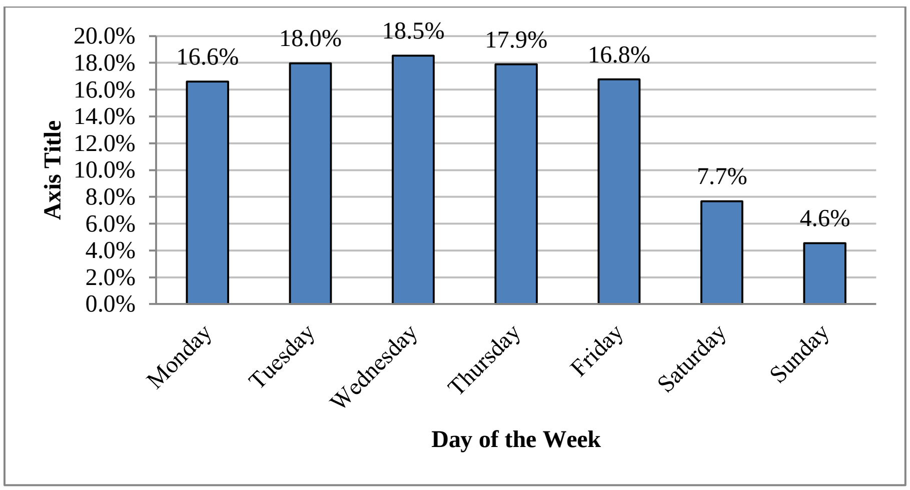

Figure 4.13 Distribution of truck-crashes based on day of the week

4.1.11 Accident Location

Location of the crash is another important parameter for understanding the characteristics of truck-crashes. Figure 4.14 shows the distribution of truck-crashes based on crash location. Data analysis revealed that the majority of truck crashes occurred on non-intersection areas, followed by intersection areas.

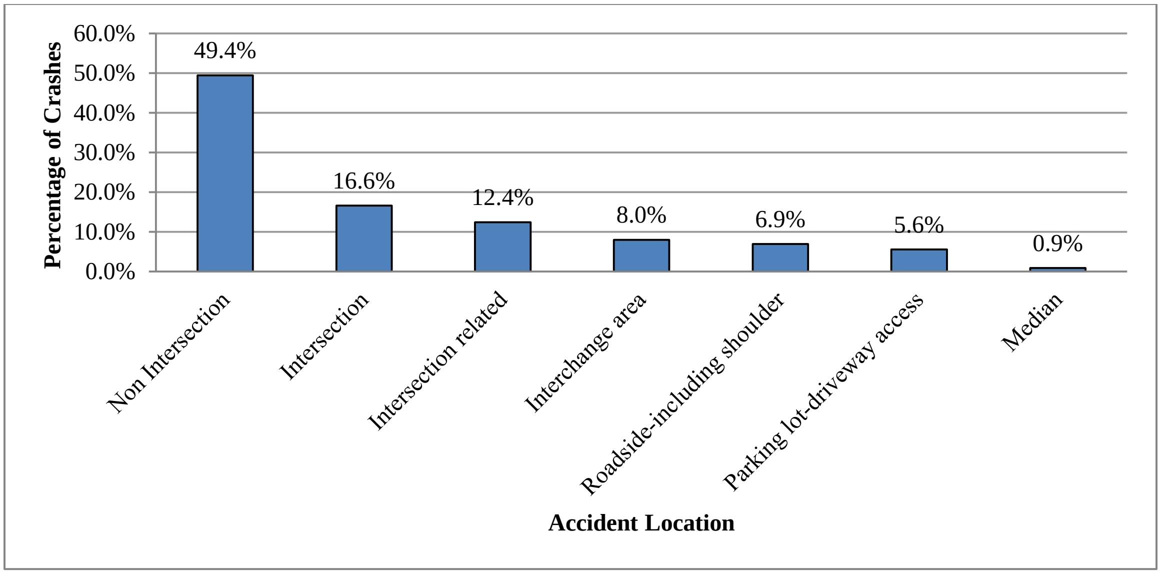

Figure 4.14 Distribution of truck-crashes based on accident location

4.1.12 Speed Limit

Speed is an important factor influencing the severity of the crashes. Control of the vehicle becomes difficult as the vehicle attains higher speeds. Figure 4.15 shows the distribution of truck-crashes based on the speed limit at the location where the crash occurred. The speed limit of the roadway on which the truck was traveling at the occurrence of the crash can be considered its approximate crash speed, even though this may not be an accurate assumption, depending on whether and by how fast the truck was speeding.

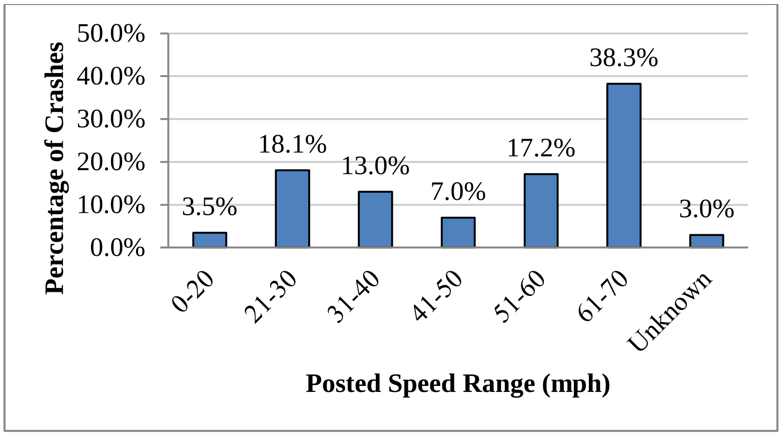

Figure 4.15 Distribution of truck-crashes based on posted speed limit

4.1.13 Pedestrian-Involved, Large-Truck Crashes

Truck crashes involving pedestrians contribute to a very small percentage of total truck crashes in Kansas, amounting to only 80 crashes in the five-year data period. 85 pedestrians were involved in truck crashes. Among all pedestrian-involved truck crashes, 80% of all pedestrians were males. Figures 4.16 and 4.17 show the distribution of pedestrian-involved truck-crashes based on gender and pedestrian age, respectively.

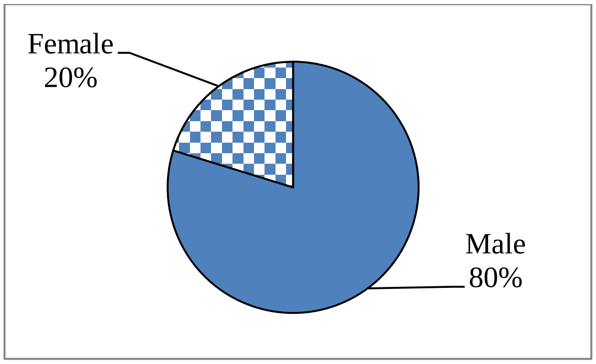

Figure 4.16 Distribution of truck-crashes based on gender of pedestrian involved

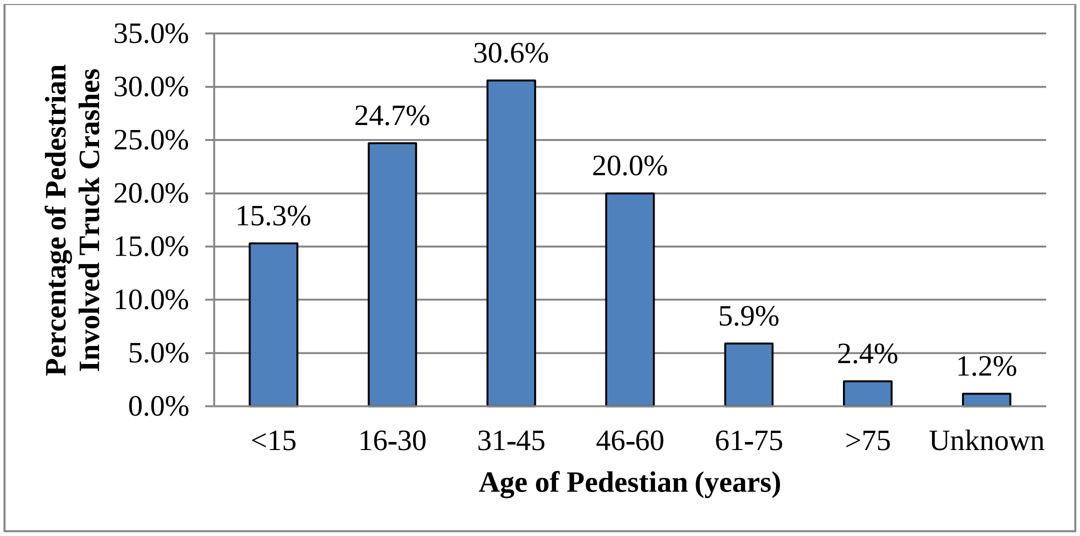

Figure 4.17 Distribution of truck-crashes based on age of pedestrian involved

A majority of the crashes occurred when the pedestrian was either entering or crossing the roadway. Figure 4.18 shows the distribution of pedestrian-involved truck crashes based on pedestrian action.

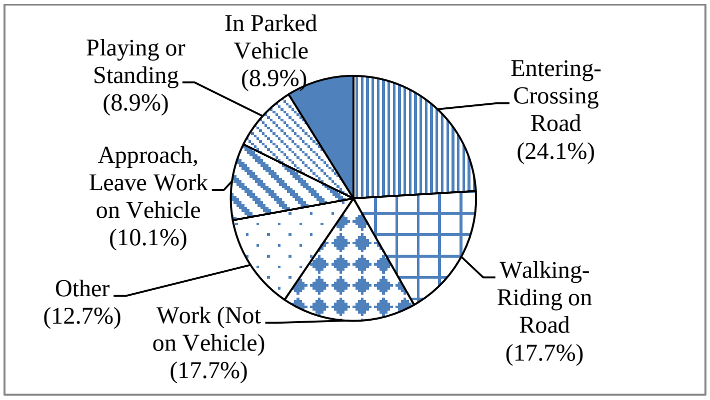

Figure 4.18 Distribution of truck-crashes based on action of pedestrian involved

Another factor in pedestrian-involved large-truck-crashes is the type of pedestrian. Figure 4.19 shows the distribution of pedestrian-involved, large-truck crashes based on type of pedestrian.

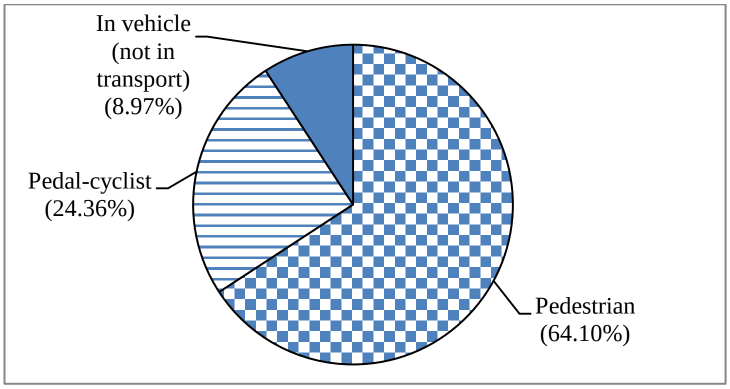

Figure 4.19 Distribution of truck-crashes based on type of pedestrian involved

<u>4.2 Characteristics of Large-Truck Crashes on the State Highway System</u>

The following variables correspond to 2004-2008 data for truck-crashes that occurred on the state highway system in Kansas, which includes Kansas highways, interstate highways, and U.S. Routes. A total of 11,762 truck-crashes were recorded on the state highway system, constituting 62.2% of all truck-crashes occurring in Kansas during that five year period.

4.2.1 Accident Class

Figure 4.20 shows the distribution of large truck crashes based on accident class (i.e. the type of collision). The majority of truck crashes involved a collision with another motor vehicle, followed by collisions with fixed objects.

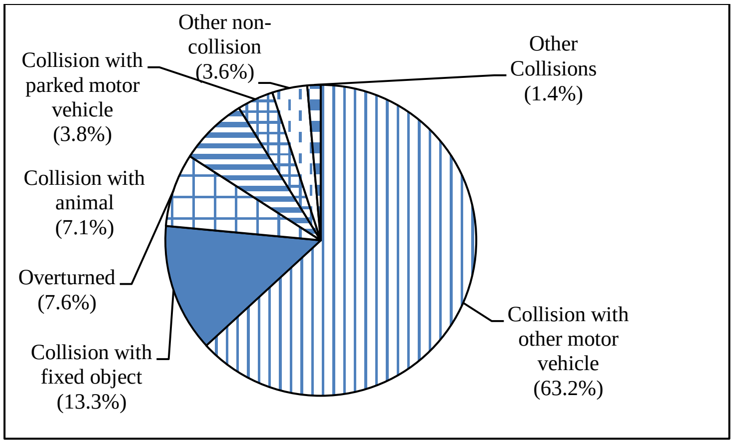

Figure 4.20 Distribution of truck-crashes based on accident class

4.2.2 Lane Class

Analysis of lane class in large truck-crashes occurring on the state highway system is presented in figure 4.21, which shows the distribution of highway truck crashes based on the lane class. The analysis showed that a majority of truck-crashes occurred on two-lane, undivided roadways, followed closely by four-lane, divided roadways. Small percentages of truck-crashes were recorded on two lane divided and eight lane divided highways.

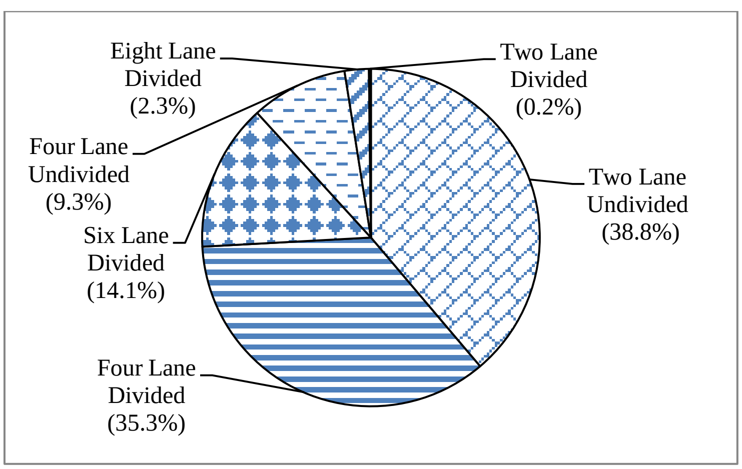

Figure 4.21 Proportion of truck-crashes based on lane class

4.2.3 Road-Function Class

Among truck-crashes that have been recorded on the state highway system, more than a quarter have occurred on rural principle arterials. Figure 4.22 shows the distribution of large-truck-related crashes based on road-function class. Arterials and Interstates combined comprised nearly 78% of truck-crashes.

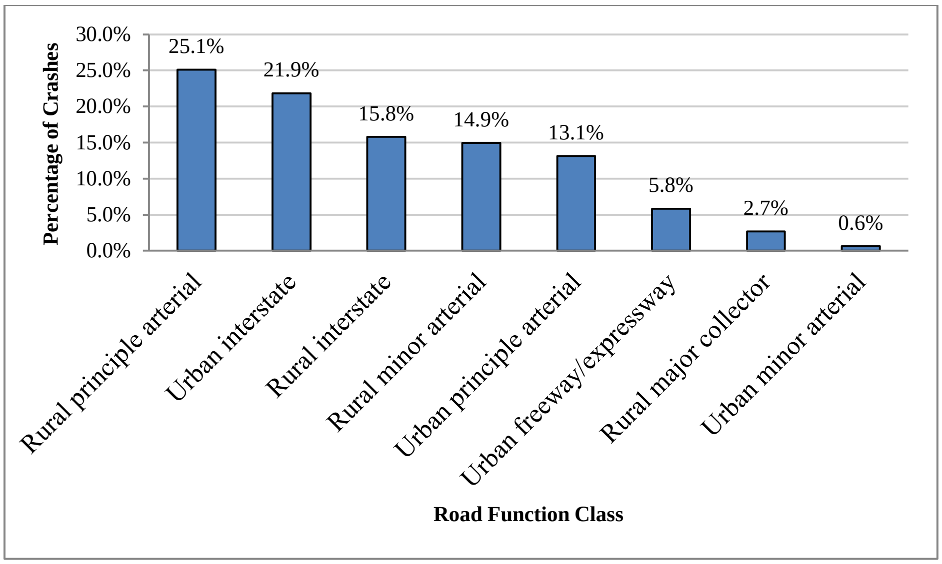

Figure 4.22 Distribution of truck-crashes based on function class

4.2.4 Average Annual Daily Traffic

AADT is defined as the average of 24-hour traffic counts collected daily over the course of a year (43). Figure 4.23 shows the distribution of truck-crashes that occurred on the state highway system, based on AADT. The percentage of truck-crashes generally decreased with increasing AADT, and a majority of truck-crashes were on roadways where AADT was less than 10,000 vehicles per day (vpd).

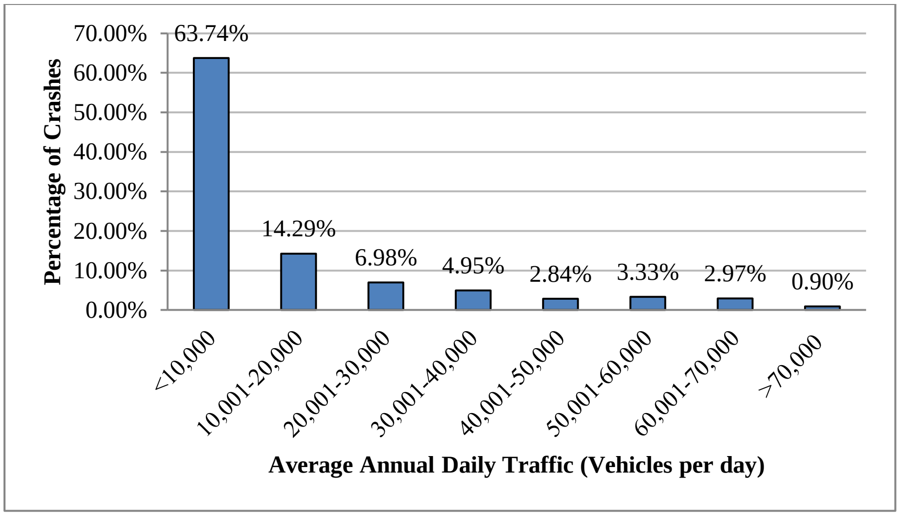

Figure 4.23 Distribution of truck-crashes based on average annual daily traffic

4.3 Contributory Causes of Large-Truck Crashes

Study of the factors contributing to truck-crashes is important in order to improve the overall safety of the transportation system. Contributory causes of large-truck crashes can be broadly classified as driver-related, vehicle-related, environment-related, and road-related. Table 4.1 shows the number of crashes based on the contributory-cause category involved. Though some crashes may have more than one contributory cause reported, all crashes may not necessarily have a contributory cause identified. Analysis of KARS data revealed that some crashes were influenced by two or more contributory causes.

Table 4.1 Number of truck-crashes based on type of contributory cause

Type of Contributory Cause	Number of Truck-Crashes	Percentage of Crashes Involving a Contributory Cause
Driver-related	13,260	73.00%
Environment-related	2,360	13.00%
Road-related	1,409	7.80%
Vehicle-related	1,112	6.10%

Based on the data presented in table 4.1., truck-driver-related factors were the most common type of contributory cause involved in truck-crashes. Table 4.2 provides details of the truck-driver-related causes contributing to truck-crashes. Among all truck-driver-related contributory causes, the majority of truck-crashes occurred when the truck driver failed to give enough time and attention to the activity in hand. Other causes, such as the truck driver going too fast for driving conditions, failing to yield the right of way, changing lanes improperly, following too closely, and making improper turns also contributed significantly to truck-crashes.

Table 4.2 Number of truck-crashes based on truck-driver-related contributory causes

Truck-Driver-Related Contributory Cause	Number of Crashes	Percentage of Crashes Involving Driver-Related Causes
Failed to give time and attention	6,458	35.4%
Speeding	2,063	11.3%
Failed to yield right of way	1,644	9.0%
Improper lane change	1,196	6.6%
Followed too closely	1,178	6.5%
Made improper turn	1,016	5.6%
Disregarded traffic signs, signal	770	4.2%
Avoidance or evasive action	742	4.1%
Improper backing	726	4.0%
Improper passing	487	2.7%
Wrong side or wrong way	337	1.9%
Fell asleep	307	1.7%
Under influence of alcohol	250	1.4%
Other distraction in or on vehicle	216	1.2%
Reckless/careless driving	197	1.1%
Ill or medical condition	105	0.6%
Did not comply with license restriction	91	0.5%
Improper or no signal	77	0.4%
Impeding traffic, too slow	76	0.4%
Distraction-mobile(cell) phone	71	0.4%
Under influence of drugs	66	0.4%
Aggressive/antagonistic driving	46	0.3%
Improper parking	46	0.3%
Distraction- other electronic devices	40	0.2%
Unknown	24	0.1%
Others	18	0.1%
Total	18,247	100.0%

Truck-related factors were the next most important contributory causes of large-truck crashes. Table 4.3 shows the number of truck-crashes in Kansas over the period of 2004-2008 based on the specific truck-related contributory cause involved. Analysis of the data showed that a majority of truck-crashes involving a truck-related contributory cause had occurred due to falling cargo, followed by defective tires, brakes, and wheels, respectively. These statistics were obtained as part of police reports and may not represent completely accuracy, as officers are not professional vehicle inspectors.

Table 4.3 Number of truck-crashes based on truck-related contributory causes

Truck-Related Contributory Cause	Number of Crashes	Percentage of Crashes Involving Vehicle-Related Causes
Falling cargo	389	33.73%
Defective tires	220	19.08%
Defective brakes	175	15.18%
Defective wheel(s)	128	11.10%
Trailer-coupling related	85	7.37%
Other lights	48	4.16%
Unattended or driverless (not in motion)	41	3.56%
Unattended or driverless (in motion)	22	1.91%
Defective windows-windshield	18	1.56%
Defective exhaust system	12	1.04%
Headlights related	5	0.43%
Other	5	0.43%
Unknown	5	0.43%
Total	1,153	100%

After truck-driver and truck-related causes, environmental factors were the most important type of contributory cause related to large-truck crashes. Table 4.4 shows the number of truck-crashes in Kansas from 2004-2008 based on the environment-related contributory causes involved. Animals contributed to a majority of environment-related truck-crashes. Rain, mist or drizzle, falling snow, strong winds, etc. are other important contributory causes.

Table 4.4 Number of truck-crashes based on environment-related contributory causes

Environment-Related Contributory Cause	Number	Percentage of Total
Animal-related	966	37.80%
Rain, mist, or drizzle	388	15.17%
Falling snow	352	13.77%
Strong winds	336	13.14%
Sleet, hail, freezing rain	185	7.23%
Vision obstruct – glare	93	3.64%
Vision obstruct – cultural	77	3.01%
Fog, smoke, or smog	75	2.93%
Blowing sand, soil, dirt	39	1.53%
Vision obstruct – vegetation	26	1.02%
Reduced visibility due to cloud cover	17	0.67%
Unknown	2	0.08%
Total	2,556	100%

Road features are an important safety consideration pertaining not only to trucks, but all road vehicle types. Table 4.5 shows road-related contributory causes involved in large-truck crashes. Analysis showed that icy or slushy conditions contributed to the majority of truck-crashes that involved road-related contributory causes. Other factors like wet, snow-packed, and debris conditions also contributed to a significant number of environment-related truck-crashes.

Table 4.5 Number of truck-crashes based on road-related contributory causes

Road-Related Contributory Cause	Number of Crashes	Percentage of Crashes Involving Road Related Factor
Icy or slushy road	686	45.70%
Wet surface	281	18.70%
Snow-packed condition	239	15.90%
Debris or obstruction	113	7.50%
Road under construction/maintenance	79	5.30%
Shoulders-related	69	4.60%
Ruts, holes ,or bumps on road	20	1.30%
Inoperative traffic control device	14	0.90%
Others	1	0.10%
Total	1,502	100.00%

4.4 Cross-Classification Analysis

Cross-classification analysis was performed to test for a relationship between select factors and truck-crash severity. Twenty-three variables were considered. Table 4.6 shows the results of the cross-classification analysis. The null hypothesis was supported for the variables of day of week, truck-related contributory causes, pedestrian-related contributory causes, gender of the truck driver, and age of truck driver, signifying that these variables did not affect the severity of truck-crashes. A sample calculation for obtaining the values of table 4.6 is provided in appendix A. These variables, along with others, were further analyzed using binary logistic-regression modeling, which is discussed in subsequent sections.

Table 4.6 Cross-classification analysis

Parameter	Degrees of Freedom	Chi-Square (χ^2) Value		Reject/Not Reject Null Hypothesis	Related to Crash Severity Yes/No
		Calculated Value	Tabular Value		
Accident class	8	159.2	15.5	Reject	Yes
Accident location	8	189.1	15.5	Reject	Yes
Age of the truck driver	12	9.8	21	Not Reject	No
Average annual daily traffic (AADT)	12	196.3	21	Reject	Yes
Manner of collision	12	1413.5	21	Reject	Yes
Contributory causes	12	106.6	21	Reject	Yes
Day of the week	24	29.9	36.4	Not Reject	No
Truck-driver-related contributory cause	24	598.7	36.4	Reject	Yes
Environment-related contributory cause	12	197.8	21	Reject	Yes
Function class	12	291.9	21	Reject	Yes
Gender of truck driver	4	3.1	9.5	Not Reject	No
Lane class	8	288.6	15.5	Reject	Yes
Light conditions	8	42.4	15.5	Reject	Yes
Pedestrian-related contributory cause	6	5.7	12.6	Not Reject	No
Road surface character	8	86.5	15.5	Reject	Yes
Road surface condition	8	23.8	15.5	Reject	Yes
Road surface type	8	29.6	15.5	Reject	Yes
Speed limit	8	653	15.5	Reject	Yes
Time of day	28	44.2	32.6	Reject	Yes
Traffic control type	20	571.7	31.4	Reject	Yes
Truck maneuver	20	568	31.4	Reject	Yes
Truck-related contributory cause	4	7.8	9.5	Not Reject	No
Weather conditions	12	22.8	21	Reject	Yes

4.5 Binary Logistic-Regression Analysis of Truck-Crashes

The binary-logistic regression technique was used to model the severity of truck-crashes in Kansas during the five-year period from 2004 to 2008. Crash severity, which is the dependent variable in this model, is dichotomous, taking a value of 0 for a crash with no injury (Property Damage Only) and a value of 1 for an injury of any severity level.

A total of 46 variables were considered in the model development using SAS Version 9.2 (40). Table 4.7 shows the description of all variables initially considered in the analysis, along with their corresponding means and variances. These variables were checked for multicollinearity using Pearson's correlation matrix to identify the significantly independent candidate variables.

Table 4.7 Description of variables considered in the model

Variable	Mean	Standard Deviation	Description
ALCOHOL	0.0159	0.1249	=1 if the truck driver was under the influence of alcohol; =0 otherwise
BRAKES	0.0355	0.185	=1 if the crash occurred due to defective brakes, exhaust, headlights, windows-windshield, tires, or falling cargo; =0 otherwise
CARELESS	0.0181	0.1334	=1 if the truck driver was distracted or was too aggressive; =0 otherwise
CC_DR	0.699	0.4587	=1 if the crash occurred had a driver-related contributory cause; =0 otherwise
CC_ENV	0.1246	0.3303	=1 if the crash occurred had environment-related contributory cause; =0 otherwise
CC_RD	0.0745	0.2626	=1 if the crash occurred had road-related contributory cause; =0 otherwise
CC_VEH	0.0583	0.2343	=1 if the crash occurred had truck-related contributory cause; =0 otherwise
CLASS	0.6317	0.4824	=1 if the crash involved collision with a motor vehicle in transport; =0 otherwise

Table 4.7 Description of variables considered in the model (cont.)

Variable	Mean	Standard Deviation	Description
COLLISION	0.1793	0.3836	=1 if the crash involved a head-on collision; =0 otherwise
CONSTR_MAINT	0.0587	0.2351	=1 if crash occurred in construction, maintenance or utility zone; =0 otherwise
CONTROL	0.8108	0.3917	=1 if the crash site had a traffic control device; =0 otherwise
DAMAGE	0.8643	0.3425	=1 if the truck had damage, =0 otherwise
DAY	0.8777	0.3276	=1 if crash occurred during weekdays; =0 otherwise
DRUGS_ALCOHOL	0.0162	0.1262	=1 if the truck driver was under the influence of drugs or alcohol; =0 otherwise
EVASIVE	0.0481	0.2140	=1 if the truck driver took evasive action or was too slow; =0 otherwise
GENDR	0.7870	0.4095	=1 if the driver of the truck was a male; =0 otherwise
IMP_MAN	0.1313	0.3377	=1 if the truck driver made improper maneuver; =0 otherwise
INOPERATIVE	0.0048	0.0688	=1 if the crash occurred at construction site or had inoperative traffic control device;
			=0 otherwise
LIGHT	0.7596	0.4273	=1 if the light condition was daylight; =0 otherwise
LOCATION	0.2907	0.4541	=1 if the crash occurred at an intersection or intersection-related; =0 otherwise
MANEUVER	0.5456	0.4979	=1 if the truck was straight following road during crash; =0 otherwise
MIDDLE_AGED	0.6877	0.4635	=1 if the driver of the truck was between 26 and 64 years; =0 otherwise
OLD	0.022	0.1467	=1 if the driver of the truck was 65 years or more; =0 otherwise
ONAT_TC	0.8324	0.3735	=1 if the traffic-control device was on the road on which the crash had occurred;
			=0 otherwise
RAIN	0.0205	0.1417	=1 if the crash occurred during rain, mist, or drizzle; =0 otherwise
RUTS	0.0106	0.1025	=1 if the roadway had ruts, holes, or bumps; =0 otherwise
S_CHAR	0.6733	0.4690	=1 if surface character was straight and level; =0 otherwise
S_COND	0.7915	0.4062	=1 if the surface condition was dry;
			=0 otherwise

Table 4.7 Description of variables considered in the model (cont.)

Variable	Mean	Standard Deviation	Description
S_TYPE	0.6439	0.4789	=1 if the surface type was blacktop;
			=0 otherwise
SAFETY_EQUIPT	0.9456	0.2269	=1 if safety equipment was used;
			=0 otherwise
SMOG_SAND	0.0060	0.0774	=1 if smog, smoke, fog, dirt, or blowing sand were prevailing during the crash occurrence; =0 otherwise
SNOW	0.0418	0.2000	=1 if the crash occurred during snow, sleet, hail, freezing rain conditions; =0 otherwise
SPEED	0.1433	0.3504	=1 if the truck driver exceeded posted speed limit or was too fast for conditions;
			=0 otherwise
SPEED_LIMIT_1	0.3457	0.4756	=1 if speed limit was less than 40 mi/h;
			=0 otherwise
SPEED_LIMIT_2	0.0701	0.2550	=1 if speed limit was between 40 and 50 mi/h; =0 otherwise
SPEED_LIMIT_3	0.1718	0.3773	=1 if speed limit was between 50 and 60 mi/h; =0 otherwise
SPEED_LIMIT_4	0.3825	0.486	=1 if speed limit was between 60 and 70 mi/h; =0 otherwise
TIME_ATTN	0.4145	0.4927	=1 if the truck driver fell asleep, failed to yield right of way, or failed to give time and attention; =0 otherwise
TIME_DAY	0.8438	0.3631	=1 if crash occurred between 6 am and 8 pm; =0 otherwise
TRAPPED	0.0195	0.1383	=1 if truck driver was trapped; =0 otherwise
UNATTND	0.0033	0.0576	=1 if the crash occurred during unattended driver condition; =0 otherwise
VSN_OBSTRUCT	0.0573	0.2324	=1 if the crash occurred during a vision obstruction; =0 otherwise
WEATHER	0.1818	0.3857	=1 if the weather conditions were adverse;
			=0 otherwise
WET	0.0605	0.2385	=1 if the crash occurred in wet or icy conditions; =0 otherwise
WRONG	0.1327	0.3393	=1 if the truck driver made improper turn, was on wrong side or wrong way, or followed too closely; =0 otherwise
YOUNG	0.2320	0.4221	=1 if driver of the truck was between 16 and 25 years; =0 otherwise

Pearson's correlation matrix was developed using SAS Version 9.2 (40). The correlation matrix is presented in appendix B. Among the independent variables, a total of 12 correlated pairs achieved a significance level of $p \leq 0.5$, which was the cutoff criteria selected for the current analysis (38). One variable from each pair was discarded, so that the variable providing the stronger model (the variable with the higher-magnitude Pearson's statistic) remained. Using this method, the following variables were excluded: wet or icy road conditions, obstructions to the truck driver's vision, truck driver under the influence of drugs/alcohol, younger truck drivers (less than 25 years old), defective brakes, exhaust system, headlights windows/ windshield, tires, or falling cargo, weather conditions, time of day, accident location, environment-related contributory causes, speed limit between 60 and 70 mi/hr, truck driver falling asleep, failing to give right of way or failing to give time and attention were all discarded by this method. Table 4.8 shows the variables retained after checking multicollinearity.

Table 4.8 Variables retained among correlated pairs

Correlated Variable-Pair	Pearson's Correlation Coefficient	Variable Retained
CC_RD, WET	0.895	CC_RD
DAMAGE, VSN_OBSTRUCT	0.831	DAMAGE
ALCOHOL, DRUGS_ALCOHOL	0.822	ALCOHOL
YOUNG, MIDDLE_AGED	-0.816	MIDDLE_AGED
CC_VEH, BRAKES	0.771	CC_VEH
WEATHER, S_COND	-0.750	S_COND
TIME_DAY, LIGHT	0.729	LIGHT
ONAT_TC, LOCATION	-0.689	ONAT_TC
CC_ENV, VSN_OBSTRUCT	0.653	none
SPEED_LIMIT_1, SPEED_LIMIT_4	-0.572	SPEED_LIMIT_1
CC_ENV, SNOW	0.553	SNOW
CC_DR, TIME_ATTN	0.552	CC_DR

After eliminating the correlated variables, the model development was left with a set of 35 variables. Three variable selection methods, which included the Forward Selection method, Backward Elimination method, and Stepwise Selection method, were performed to select variables which were significant enough to remain in the model. A p-value of 0.05 was chosen as the significance criteria, and any variable having a p-value greater than 0.05 was considered to be insignificant, and was excluded from the model (27). Table 4.9 shows the comparison of the model-fit statistics obtained from the three variable selection methods.

Table 4.9 Comparison of model-fit statistics from the three variable selection methods

Criterion	Forward Selection Method		Stepwise Selection Method		Backward Elimination Method	
	Intercept Only	Intercept and Covariates	Intercept Only	Intercept and Covariates	Intercept Only	Intercept and Covariates
AIC	20820.1	17391.8	20820.1	17390.9	20820.1	17390.3
SC	20828	17613.7	20828.0	17610.6	20828.0	17605.7
-2logL	20818.1	17337.8	20818.1	17334.9	20818.1	17330.3
R^2	0.1680		0.1682		0.1684	

Based on these statistics, the model obtained by the Backward Elimination method was found to be the slightly better model, because of relatively lower AIC, SC and -2logL values, and higher R^2 values. Table 4.10 shows further goodness-of-fit parameters obtained by using the LOGISTIC procedure in SAS version 9.2 (40) for the three variable selection methods. From table 4.10, the relatively lower percent discordant value and values of Somer’s D and Gamma closer to 1 reinforces the finding that the Backward Elimination method produced the better model among the three variable selection methods.

Table 4.10 Associations of predicted probabilities and observed responses

Statistic	Forward Selection Method	Stepwise Selection Method	Backward Elimination Method
Percent Concordant	76	76	76
Percent Discordant	23.7	23.7	23.6
Percent Tied	0.4	0.4	0.4
Pairs	65,142,718	65,142,718	65,142,718
Somers' D	0.523	0.523	0.524
Gamma	0.525	0.525	0.526
Tau-a	0.19	0.191	0.191
c	0.762	0.762	0.762

Following are descriptions of the variables in table 4.9 for the Backward Elimination method (7):

- Percent concordant**:** A pair of observations with different observed responses is concordant if the observation with the lower ordered response value has a lower predicted mean score than the observation with the higher ordered response value. Of the examined pairs, 76% were found to be concordant.
- Percent discordant**:** If the observation with the lower ordered response value has a higher predicted mean score than the observation with the higher ordered response value, then the pair is discordant. Of the pairs under examination, 23.6% were found to be discordant.
- Percent tied: 0.4% of observations were found to be neither concordant nor discordant.
- Pairs: The concordant pairs, discordant pairs and tied pairs totaled of 65,142,718 distinct pairs.

- Somer’s D: The value of Somer’s D was found to be 0.524, which is closer to 1, which indicates that more pairs agreed than disagreed. Somer’s D is used to determine the strength and direction of relation between pairs of variables. Its values range from -1.0 (all pairs disagree) to 1.0 (all pairs agree).
- Gamma: The Goodman-Kruskal Gamma has a value of 0.526, which indicates good association among the variables in the model. Its values range from -1.0 (no association) to 1.0 (perfect association).
- Tau-a**:** This value was found to be 0.191 for the model obtained. Kendall's Tau-a takes into the account the difference between the number of possible paired observations and the number of paired observations with different responses.
- c**:** This value was found to be 0.762 for the model obtained. It ranges from 0 (no association) to 1 (perfect association).

A total of 26 variables were found to be significant and remained in the model. Table 4.11 shows the parameter estimates and odds ratios, as obtained using the Backward Elimination method. The models obtained by the other two methods are presented in appendix C. These parameter estimates and odds-ratio values are used to understand the relationship of the variable under consideration with crash severity.

Table 4.11 Parameter estimates and odds ratios of large-truck crash severity model

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Sq	Odds Ratio	95% Wald Confidence Limits For Odds Ratio
Intercept*	-1.522	0.163	87.15	<0.0001	NA**	NA**
ALCOHOL*	0.979	0.135	52.5	<0.0001	2.66	2.04,3.47
CARELESS*	0.334	0.126	7.08	0.0078	1.40	1.09, 1.79
CC_DR*	0.6	0.054	126.08	<0.0001	1.82	1.64, 2.02
CC_RD*	-0.332	0.084	15.49	<0.0001	0.72	0.61, 0.85
CC_VEH	-0.09	0.093	0.94	0.3329	0.91	0.76, 1.10
CLASS	0.102	0.052	3.81	0.0509	1.11	1.00, 1.23
COLLISION*	0.471	0.052	82.71	<0.0001	1.60	1.45, 1.77
CONSTR_MAINT*	-0.267	0.083	10.33	0.0013	0.77	0.65, 0.90
CONTROL*	0.308	0.057	29.58	<0.0001	1.36	1.22, 1.52
DAMAGE*	1.116	0.083	181	<0.0001	3.05	2.60, 3.59
DAY	-0.003	0.058	0.00	0.9661	1.00	0.89, 1.12
EVASIVE*	0.427	0.079	29.37	<0.0001	1.53	1.31, 1.79
GENDR*	-0.129	0.049	7.06	0.0079	0.88	0.80, 0.97
IMP_MAN*	-0.453	0.068	44.48	<0.0001	0.64	0.56, 0.73
INOPERATIVE	-0.247	0.328	0.57	0.4508	0.78	0.41, 1.48
LIGHT	0.06	0.049	1.50	0.2209	1.06	0.96,1.17
MANEUVER*	0.321	0.041	61.54	<0.0001	1.38	1.27, 1.49
MIDDLE_AGED*	0.102	0.043	5.74	0.0166	1.11	1.02, 1.20
OLD	0.092	0.14	0.43	0.5141	1.10	0.83, 1.44
ONAT_TC*	-0.521	0.054	93.75	<0.0001	0.60	0.53, 0.66
RAIN*	0.33	0.132	6.25	0.0124	1.39	1.07, 1.80
RUTS	-0.148	0.224	0.44	0.5091	0.86	0.56, 1.34
S_CHAR*	-0.114	0.041	7.86	0.0051	0.89	0.82, 0.97
S_COND*	0.256	0.056	20.68	<0.0001	1.29	1.16, 1.44
S_TYPE*	0.132	0.04	10.62	0.0011	1.14	1.05, 1.24
SAFETY_EQUIPT*	-1.378	0.075	337.60	<0.0001	0.25	0.22, 0.29
SMOG_SAND	0.355	0.218	2.65	0.1037	1.43	0.93, 2.19
SNOW	0.151	0.099	2.34	0.1261	1.16	0.96, 1.41
SPEED*	0.442	0.054	66.12	<0.0001	1.56	1.40, 1.73
SPEED_LIMIT_1*	-0.801	0.051	248.48	<0.0001	0.45	0.41, 0.50
SPEED_LIMIT_2*	-0.39	0.077	25.92	<0.0001	0.68	0.58, 0.79
SPEED_LIMIT_3*	0.116	0.052	5.01	0.0252	1.12	1.01, 1.24
TRAPPED*	4.417	0.344	165.04	<0.0001	82.81	42.21, 162.44
UNATTND	0.483	0.329	2.16	0.142	1.62	0.85, 3.09
WRONG	0.014	0.058	0.06	0.8034	1.01	0.91, 1.14

*- Significant at 0.05 level
NA**- Not Applicable

The following sections explain the variables that are significant in the model at a p-value of 0.05, with regard to parameter estimates and odds ratios:

4.5.1 Roadway Characteristics

The estimated coefficient for the variable S_TYPE was positive (i.e., 1.14), indicating that blacktop-surface type had 1.14 times greater odds of causing more severe truck-crashes as compared to concrete and other surface types. Similarly, the variable S_COND had a positive coefficient estimate; the dry-surface condition had 1.29 times greater odds of causing a more severe crash as compared to wet and other surface conditions. However, a negative coefficient for the variable S_CHAR indicated that straight- and leveled-surface characteristics had 0.89 times lesser odds of causing a more severe crash as compared to other surface characteristics.

The variable CC_RD had a negative coefficient, which indicates the road-related contributory cause had 0.72 times lesser odds of causing a more severe truck crash as compared to other factors.

4.5.2 Crash Characteristics

As variables SPEED_LIMIT_1 and SPEED_LIMIT_2 have negative coefficients for the parameter estimates, vehicles speeds lower than 50 mph resulted in a decreasing probability of severe truck-crashes. On the other hand, the variable SPEED_LIMIT_3 had a positive coefficient; speed limits ranging from 60-70 mph increased the odds of a more severe crash by 1.12 times. This figure demonstrates that the severity of the crash increased with increasing vehicle speeds. Further, a positive coefficient estimate for the variable COLLISION reveals that head-on collisions had 1.60 times greater odds of causing a more severe crash as compared to other collision types, such as angled and sideswipe collisions.

A positive coefficient estimate for the variable CONTROL shows that large-trucks had 1.36 times greater odds of being involved in a more severe crash when there was a traffic-control device at the location of the crash, as compared to locations where there was no traffic-control device. In addition, a negative coefficient estimate for the ONAT_TC indicates that large trucks have 0.59 times lesser odds of being involved in a more severe -crash lower when a traffic-control device is on the road along which the truck is travelling as compared to being on the road perpendicular to it.

A positive coefficient estimate for the MANEUVER variable shows that the odds of a severe large-truck crash were 1.38 times greater when the driver of the truck was going straight and following the road, as compared to when he/she made a maneuver such as left turn, right turn, U-turn, etc. Also, the variable DAMAGE had a positive coefficient estimate, indicating that any damage to the vehicle involved in the crash resulted in 3.05 times greater odds of increasing the severity of the crash, as compared to the case when minimal damage occurs to the involved truck.

A positive coefficient for the variable RAIN shows that large-trucks were 1.39 times more likely to be involved in a more severe crash under rain, mist, or drizzle conditions as compared to other conditions.

4.5.3 Driver Characteristics

A positive coefficient of the variable ALCOHOL shows that large-trucks were 2.66 times more likely to be involved in more severe crashes when the driver was under the influence of alcohol. Further, a positive coefficient estimate for the MIDDLE_AGED variable shows that large-trucks were 1.11 times more likely to be involved in a more severe crash when the driver was middle-aged, as compared to old and young drivers. The negative coefficient of the GENDR variable shows that large-trucks with male drivers were 0.88 times less likely to be involved in a more severe crash, compared to female drivers. The TRAPPED variable, which has the highest magnitude of odds ratio among all the variables, had a positive coefficient estimate indicating that large-truck-involved crashes were 82.81 times more likely to be more severe when the driver was trapped, as compared to other conditions such as being ejected, not being ejected, etc. Similarly, a negative coefficient estimate for the SAFETY_EQUIPT variable shows that large-trucks had 0.25 times lesser odds of being involved in a more severe crash when the driver wore safety equipment, compared to when he/she did not wear safety equipment. This further supports the fact that the use of safety belts reduces crash severity.

The variable CC_DR had a positive coefficient, indicating that large-trucks were 1.82 times more likely to be in a more severe crash when there was a driver-related cause contributing to the occurrence of the crash, as compared to other conditions. A positive coefficient estimate for the variable SPEED shows that large-trucks had 1.56 times greater odds of being in a more severe crash when the driver was speeding. This shows that speeding increased crash severity. A positive coefficient estimate for the variable EVASIVE shows that large-trucks had 1.53 times greater odds of being in a more severe crash when the driver took an evasive action or was slow for the existing conditions. Similarly, a positive coefficient estimate for CARELESS shows that large-trucks were 1.40 times more likely to be involved in a more severe crash when the driver was aggressive, reckless, or antagonistic while driving. However, the variable IMP_MAN has a negative coefficient, which indicates that large-trucks had 0.64 times lower odds of being involved in a more severe crash when the driver took an improper action such as improper backing, improper passing, improper turning, improper or no signal, etc. as compared to other conditions.

The binary logistic-regression method provided a good measure to identify factors contributing to increasing severities of crashes involving large-trucks. The model developed shows that 10 out of 26 candidate variables, which included those related to the use of safety equipment, obstruction of vision, speed limits between 0 and 40 mi/hr, location of the traffic-control device, making an improper maneuver, speed limits between 40 and 50 mph, road-related contributory causes, construction, maintenance or utility zones, gender of the truck driver, and surface character, had a negative coefficient for the parameter estimates in the decreasing order of the magnitude, and the rest of the variables had positive coefficients.

Chapter 5 Conclusions and Summary

5.1 Conclusions

This study identified characteristics of truck-crashes, factors contributing to their occurrence, and factors associated with increased crash severity relating to vehicle, driver, environment, road, and other related factors. Crash data, obtained from Kansas Department of Transportation's Kansas Accident Reporting System (KARS) database for the five-year period of 2004-2008 was utilized for this study. This database was a compilation of police-reported crash data in the state of Kansas.

The majority of truck-crashes were found to have occurred during daylight conditions and under no adverse weather conditions. Of all truck-crashes, 35.2% were single-vehicle truck-crashes, and the majority of multi-vehicle truck crashes were characterized by angular collisions. Most of the non-truck-vehicles involved in two-vehicle truck crashes were cars. More than 75% of all truck-crashes occurring in the study period happened on weekdays. Of all truck-crashes, 54.6% occurred when the truck was moving straight and following the road—the most common among all truck maneuvers. The majority of truck-crashes occurred when the truck was driven by a male truck-driver between 20 and 60 years of age. Most of the pedestrians involved in truck crashes were males between the ages of 16 and 60. Non-intersection locations predominantly characterized truck-crashes. Most truck-crashes occurred between noon and 3:00 p.m. Blacktop surface type, dry surface conditions, and straight and level surface characteristics were dominant in their respective truck-crashes categories. Further, more truck-crashes were recorded in high speed-limit locations. Among all truck-crashes on the state highway system, 63.2% involved collision with another motor-vehicle, and majority occurred on arterials and interstates under low AADT conditions.

Cross-classification analysis was performed over a subset of variables to identify the relationship of truck-crash severity with select independent variables. Among the factors considered, variables such as type, character, and condition of the road surface; accident class; type of collision; driver- and environment-related contributory causes; traffic-control type; vehicle maneuver; accident location; speed limit; light and weather conditions; time of day; road function class; lane class; and Average Annual Daily Traffic (AADT) were found to be related to the severity of truck-crashes.

Analysis of the factors contributing to the occurrences of truck crashes showed that driver-related factors were the most dominant type of contributory cause. The most significant factor involved in the majority of truck-crashes in cases where a driver-related contributory cause was recorded was when truck drivers failed to give time and attention. Moreover, other driver-related factors such as speeding, drivers failing to yield right of way, and improper lane change also contributed to the occurrence of truck-crashes. Falling cargo comprised 33.73% of truck-related causes, while animal-related factors comprised 37.80% of environment-related causes contributing to the occurrence of truck crashes. Among all truck-crashes caused by road-related factors, icy and slushy road condition was the most dominant factor, contributing to 45.70% of truck-crashes.

Severity modeling was performed using binary a logistic-regression model in order to identify and evaluate the factors contributing to increased crash severity. Severity was considered as a dichotomous dependent variable in order to develop the model. The goodness-of-fit statistics and overall percentage concordant value of 76% were evidence of good model fit. Based on the developed model, important factors were identified.

A truck-driver being trapped resulted in an 82.81 times greater likelihood of increased crash severity—the highest odds ratio of all variables examined. Damage to the truck, with an odds ratio of 3.05, was another important factor associated with increased severity of truck-crashes. Further, truck-crashes had 2.66 times higher odds of being more severe when the truck-driver was under the influence of alcohol. Truck driver-related causes had 1.82 times higher odds of increasing the severity of truck-crashes. Over-speeding, aggressiveness, and evasive driving by the truck-driver were among the truck-driver-related factors likely to increase the severity of truck-crashes. Head-on collisions were 1.60 times more likely to contribute to more severe truck crashes. Traffic control devices resulted in a 1.36 times higher odds ratio of increased crash severity. Dry-surface conditions, having an odds ratio of 1.29, and blacktop-surface type, with an odds ratio of 1.14, were likely to cause more severe truck-crashes. Speed limits of 50-60 mph resulted in 1.12 times higher odds, and middle-aged drivers had 1.11 times higher odds of contributing to higher crash severity.

On the other hand, certain variables were found to present lower odds of severity of truck-crashes. Straight- and level-surface characteristics had 0.89 times lower odds of contributing to increasing severity of truck-crashes. Further, construction/maintenance zones had 0.77 times lower odds, and road-related contributory cause had 0.72 times lower odds of contributing to more severe truck crashes. Male truck drivers and improper truck maneuvers, with odds ratios of 0.88 and 0.64, respectively, were found to have lower odds of contributing to more severe truck crashes.

These findings can potentially aid researchers in understanding the various characteristics and causes contributing to the occurrences and increasing severity of truck crashes. Various conditions have been elaborated upon. By addressing these issues and developing suitable countermeasures, both the number and severity of truck crashes could potentially be mitigated, which would improve the overall safety of the surface transportation system.

References

1. National Center for Statistics and Analysis, National Highway Traffic Safety Administration. 2010. “Traffic safety facts - 2008 data: Large trucks.” *Publication DOT HS 811 158*, U.S. Department of Transportation, Washington, D.C.

2. Federal Motor Carrier Safety Administration. 2010. “Large truck and bus crash facts 2008.” *Publication FMCSA-RRA-10-043*, U.S. Department of Transportation, Washington, D.C.

3. Insurance Institute of Highway Saftey and Highway Loss Data Institute, Highways Safety Research and Communications. 2010. “Fatality facts 2009: Large trucks.” Retrieved from www.iihs.org

4. National Center for Statistics and Analysis, National Highway Traffic Safety Administration. 2011. “Traffic safetyf facts 2009 data.” *Publication DOT HS 811 402*, U.S. Department of Transportation, Washington, D.C.

5. National Center for Statistics and Analysis, National Highway Traffic Safety Administration. 2003. “An Analysis of Fatal Large-Truck Crashes.” *Publication DOT HS 809 569*, U.S. Department of Transportation, Washington, D.C.

6. Kansas Department of Transportation, Geometric and Accident Data Unit. 2009. “2008 Kansas traffic accident facts book.” Kansas Department of Transportation, Topeka, Kansas.

7. Dissanayake, S., and N. Bezwada. 2010. "Characteristics and contributory causes related to large-truck crashes (Phase I) - fatal crashes." *Report no. 25-1121-0001-11*, Mid-America Transportation Center, University of Nebraska, Lincoln, NE.

8. Mullinazi, T. E., S. D. Schrock., R. Rescot., R. Jasrotia., K. Hovey., Y. Li., and J. Pohlman. 2010. "Mitigating wind-induced truck-crashes." Mid-America Transportation Center, University of Nebraska, Lincoln, NE.

9. Golob, T. F., and A. C. Regan. 2004. "Traffic conditions and truck accidents on urban freeways." *Publication UCI-ITS-WP-04-3*, Institute of Transportation Studies, University of California, Irvine, CA.

10. Khattak, A. J., R. J. Schneider., and F. Targa. 2002. "A comparison of rollovers with non-rollovers and analysis of injury severity in large-truck crashes." Southeastern Transportation Center, University of Tennessee, Knoxville, TN.

11. Charbotel, B., J. L. Martin., B. Gadegbeku., and M. Chiron. 2003. "Severity factors for truck drivers' injuries." *American Journal of Epidemiology,* 158: 753-759.

12. Torre, F. L., and L. Rossi. 2009. "Heavy good vehicle accident analysis for the identification of potentially dangerous situations." University of Florence, Italy.

13. Khattak, A. J., and F. Targa. 2004. *Paper No. 04-2458. Injury Severity and Total Harm in Truck-Involved Work-Zone Crashes*. CD-ROM. Transportation Research Board of the National Academics, Washington, D.C.

14. Blower, D., K. L. Campbell., and P. E. Green. 1993. "Accident rates for heavy truck-tractors in Michigan." *Accident Analysis and Prevention,* 25: 307-321.

15. Mannila, K. K. "Analysis of various car-truck crash types based on GES and FARS crash databases using multinomial and binary logit model." MA thesis, Department of Civil and Environmental Engineering, University of Central Florida, Orlando, FL.

16. Duncan, C. S., A. J. Khattak., and M. F. Council. 1998. "Applying the ordered-probit model to injury severity in truck passenger car rear-end collisions." *Transportation Research Record,* 1635: 63-71.

17. Moghaddam, R. F., R. T. Moghaddam., P. M. Khiavi., and A. M. Ghorbani. 2009. "Crash severity modeling in urban highways using backward-regression method." *World Academy of Science, Engineering and Technology*, 60: 223-228.

18. Liu , C., D. Utter., and C. L. Chen. 2007. "Characteristics of crash injuries among young, middle-aged, and older drivers." *NHTSA Report No. DOT HS 810 857,* National Highway Traffic Safety Administration, Washington D.C.

19. Dissanayake, S. "Comparison of severity affecting factors between young and older drivers involved in single-vehicle crashes." *IATSS Research,* 28: 48-54.

20. Conroy, C., G. T. Tominaga., S. Erwin., S. Pacyna., T. Velky., F. Kennedy., M. Sise., and R. Coimbra. 2008. "The influence of vehicle damage on injury severity of drivers in head-on motor vehicle crashes." *Accident Analysis and Prevention,* 40: 1589-1594.

21. Malyshkina, N. V., and F. Mannering. 2008. "Effects of increases in speed limits on severities of injuries in accidents." *Transportation Research Record: Journal of the Transportation Research Board,* 2083: 122-127.

22. Gabauer, D. J., and H. C. Gabler. 2010. "The effects of airbags and seatbelts on occupant injury in longitudinal-barrier crashes." *Journal of Safety Research,* 41: 9–15.

23. Eboli, L., and G. Mazzulla. 2007. "A structural equation model for road-accident analysis." *Proceedings of the 4th International SIIV Congress,* Palermo, Italy.

24. Wang, Q. "Study on crash characteristics and injury severity at roadway work zones." MA thesis, Department of Civil and Environmental Engineering, University of South Florida, Tampa, FL.

25. Liu, L., and S. Dissanayake. 2007. "Speed-limit-related issues on gravel roads." *Proceedings of the Mid-Continent Transportation Research Symposium,* Ames, Iowa, IA.

26. Lemp, J. D., K. M. Kockelman., and A. Unnikrishnan. 2011. "Analysis of large-truck crash severity using heteroskedastic-ordered probit models." *Accident Analysis and Prevention,* 43: 370-380.

27. Kockelman, K. M. 2001. "Driver injury severity: An application of ordered probit models. " *Accident Analysis and Prevention*, 34: 313-321.

28. Ma, J., and K. M. Kockelman. 2006. "Bayesian multivariate poisson regression for models of injury count, by severity." *Transportation Research Record: Journal of the Transportation Board*, 1950: 24-34.

29. Burke, N. "Dedicated truck lanes as a solution to capacity and safety issues on interstate highway corridors." *Proceedings of the Midwest Transportation Consortium*, Ames, IA.

30. Rau. P. S. 2005. "Drowsy driver detection and warning system for commercial vehicle drivers: Field operational test design, data analysis, and progress." *Paper No. 05-0192,* National Highway Traffic Safety Administration. U.S. Department of Transportation, Washington, D.C.

31. Council. M. F., L. D. Harkey., D. T. Nabors., A. J. Khattak., and Y. M. Mohamedshah. 2003. "Examination of fault, unsafe driving acts, and total harm in car-truck collisions." *Transportation Research Record: Journal of the Transportation Research Board*, 1830: 63-71.

32. Montella. A., and M. Pernetti. 2010. "In-depth investigation of run-off-the-road crashes on the motorway Naples-Candela." *Proceedings of the 4th International Symposium on Highway Geometric Design*, Valencia, Spain.

33. Y. Wang., N. H. Nguyen., A. B. E. Levy., and Y. J. Wu. 2008. "Cost-effective safety improvements for two-lane rural roads." *Report TNW2008-04*, Transportation Northwest (TransNow), University of Washington, Seattle, WA.

34. Y. Li., and Y. Bai. 2008. "Development of crash-severity-index models for the measurement of work-zone risk levels." *Accident Analysis and Prevention*, 40: 1724-1731.

35. Oh. C., Y. Kang., B. Kim., and W. Kim. 2005. "Analysis of pedestrian-vehicle crashes in Korea: Focused on developing probabilistic pedestrian fatality model." *Proceedings of the 19th International Technical Conference on the Enhanced Safety of the Vehicles*, Washington, D.C.

36. Dissanayake, S., and J. J. Lu. 2001. "Traffic-control device comprehension - differences between domestic and international drivers in USA." *IATSS Research*, 25: 80-87.

37. Dissanayake, S., and I. Ratnayake. 2004. "Identification of countermeasures to reduce severity of rural highway crashes." *Proceedings of the Mack-Blackwell Transportation Center*, 2051, University of Arkansas, Fayetteville, AR.

38. Allision, D. P. 1999. *Logistic regression ssing SAS system*. Cary, NC: John Wiley & Sons, Inc.

39. Kansas Department of Transportation. 2009. *Kansas Motor Vehicle Accident Report Coding Manual*. Retrieved from www.ksdot.org

40. SAS Institute Inc. 2008. *SAS Online Doc 9.2*. Cary, NC: SAS Institute Inc.

41. Howell. D. C. 2009. “Chi-square test - analysis of contingency tables.” University of Vermont, Burlington, VT. Retrieved from www.uvm.edu

42. Dixon, W. J., and F. J. Massey, Jr. 1983. *Introduction to statistical analysis (4th ed.)*. New York, NY: McGraw-Hill, 319-322.

43. Garber. N. J., and L. A. Hoel. 2009. *Traffic and highway engineering*. St. Paul, MN: West Publishing Company, 115-116.

Appendix A Cross-Classification Analysis

Table A.1 shows the number of truck-crashes in Kansas based on the speed limit. This variable was used for cross-classification analysis. A sample calculation is presented following table A.1.

Table A.1 Number of truck-crashes in Kansas based on speed limit

Speed Limit (mi/h)	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total Number of Crashes
		Disabled	Non Incapacitating	Possible		
50+	287	537	1,395	949	7,507	10,675
30-49	32	77	512	522	5,440	6,583
0-29	2	7	36	43	1,011	1,099
Unknown	6	16	63	42	435	562
Total	327	637	2,006	1,556	14,393	18,919

<u>Sample Calculation</u>

Null hypothesis (H_0): Speed limit and crash severity are independent of each other.

Alternate hypothesis (H_A): Null hypothesis is not true.

Values shown in Table A.1 are observed frequencies (O).

Expected frequencies (E) are given as:

$$E_{ij} = \frac{\text{(Row Total)*(Column Total)}}{\text{Sample Size}}$$

i.e., the expected frequency of fatal crashes at the speed limit of 30-49 mi/h is given as:

$$E_{21} = \frac{(6{,}583)*(327)}{18{,}919}$$
$$= 113.8$$

Similarly, the expected frequencies of all cells are calculated. Table A.2 shows the expected frequencies of truck crashes.

Table A.2 Expected frequencies of truck-crashes in Kansas based on speed limit

Speed Limit	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total Number of Crashes
		Disabled	Non Incapacitating	Possible		
50+	184.509	359.426	1,132	877.969	8,121	10,675
30-49	113.782	221.649	698.002	541.421	5,008	6,583
0-29	18.9953	37.0032	116.528	90.3877	836	1,099
Unknown	9.71373	18.9225	59.5894	46.2219	427.553	562
Total	327	637	2006	1556	14393	18,919

Now, the statistic chi-square (χ^2) is calculated using the formula:

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}}$$

Using the formula, the calculated chi-square value obtained is 653.03.

- Degrees of freedom= (3-1)* (5-1)

  = 8

- Chi-square value from the chi-square distribution table for 8 degrees of freedom and 95% confidence is 15.51.

Since the calculated chi-square value (653.03) was greater than the chi-square value from the table (15.51), the null hypothesis is rejected. Hence, there exists a relationship between speed limit and crash severity.

Following are some of the other tables used for analyzing the relationship of the corresponding variables with crash severity, using cross-classification analysis. In all of the following tables, the “unknown” and “others” categories were ignored, as they constituted a negligible percentage of the total crashes.

Table A.3 Number of truck-crashes in Kansas based on accident location

Accident Location	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total Crashes
		Disabled	Non Incapacitating	Possible		
Non-Intersection-On Roadway	185	296	921	688	7,258	9,348
Intersection-On Roadway	97	159	426	308	2,154	3,144
Intersection-Related-On Roadway	15	48	179	199	1,914	2,355
Interchange Area-On Roadway	17	49	165	122	1,162	1,515
Roadside-Including Shoulder-Off Roadway	12	56	209	134	898	1,309
Pklot-Drvway Access-On Roadway	0	20	83	90	861	1,054
Median-Off Roadway	1	9	21	14	117	162
Total	327	637	2,006	1,556	14,393	18,919

Table A.4 Number of truck-crashes in Kansas based on light conditions

Light Condition	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total Number of Crashes
		Disabled	Non Incapacitating	Possible		
Daylight	229	482	1,513	1,265	10,882	14,371
Dark-No Street Lights	61	91	268	144	1798	2,362
Dark-Street Lights On	23	34	150	89	1138	1,434
Dawn	10	20	40	33	331	434
Dusk	4	9	34	23	223	293
Total	327	637	2,006	1,556	14,393	18,919

Table A.5 Number of truck-crashes in Kansas based on weather conditions

Weather Condition	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total Number of Crashes
		Disabled	Non Incapacitating	Possible		
No Adverse Conditions	272	526	1,636	1,231	11,814	15,479
Rain, Mist or Drizzle	17	34	135	136	991	1,313
Snow	4	29	79	73	647	832
Strong Winds	9	12	57	25	222	325
Snow and Winds	6	7	21	22	207	263
Freezing Rain	7	7	21	21	129	185
Fog	6	9	19	13	109	156
Sleet	1	3	5	16	109	134
Rain and Winds	1	5	17	8	92	123
Blowing Dust/Sand	3	4	8	2	19	36
Total	327	637	2,006	1,556	14,393	18,919

Table A.6 Number of truck-crashes in Kansas based on time of day

Time of the Day	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total Number of Crashes
		Disabled	Non Incapacitating	Possible		
0000 hrs-3:00 am	14	26	83	40	521	684
3:01 am-6:00 am	21	33	110	60	683	907
6:01 am-9:00 am	53	94	314	238	2202	2,901
9:01am-12:00 noon	51	139	387	365	3,022	3,964
12:01pm -3:00 pm	75	147	451	354	3,226	4,253
3:01pm-6:00 pm	55	115	379	319	2,693	3,561
6:01 pm-9:00pm	33	50	179	124	1,280	1,666
9:01 pm-11:59pm	25	33	103	56	758	975
Total	327	637	2,006	1,556	14,385	18,919

Table A.7 Number of truck-crashes in Kansas based on road function class

Road Function Class	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total Number of Crashes
		Disabled	Non Incapacitating	Possible		
Rural Other Principal Arterial	153	189	370	219	2,020	2,951
Urban Interstate	9	71	268	256	1,966	2,570
Rural Interstate	22	73	236	151	1,377	1,859
Rural Minor Arterial	56	96	262	132	1,211	1,757
Urban Other Principal Arterial	19	33	131	132	1,229	1,544
Urban Freeway/Expressway	6	19	68	67	528	688
Rural Major Collector	3	21	54	34	204	316
Total	271	505	1,394	997	8,595	11,762

Table A.8 Number of Truck-Crashes in Kansas Based on AADT*

AADT*	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total Number of Crashes
		Disabled	Non Incapacitating	Possible		
0-10,000	249	402	954	594	5,298	7,497
10,001-20,000	8	42	200	140	1,291	1,681
20,001-30,000	7	16	75	81	642	821
30,001-40,000	4	19	53	62	444	582
50,001-60,000	1	10	39	41	301	392
60,001-70,000	0	6	32	35	276	349
40,001-50,000	2	4	33	30	265	334
80,001 and above	0	2	3	7	43	55
70,001-80,000	0	4	5	7	35	51
Total	271	505	1394	997	8,595	11,762

*AADT is the average annual daily traffic.

Table A.9 Number of truck-crashes in Kansas based on lane class

Lane Class	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total Number of Crashes
		Disabled	Non Incapacitating	Possible		
Two Lane Undivided	200	292	614	352	3,107	4,565
Four Lane Divided	53	148	492	355	3,108	4,156
Six Lane Divided	6	48	169	184	1,250	1,657
Four Lane Undivided	10	7	90	81	901	1,089
Eight Lane Divided	1	8	28	25	211	273
Total	271	505	1,394	997	8,595	11,762

Table A.10 Number of truck-crashes in Kansas based on road-surface type

Road Surface Type	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total Crashes
		Disabled	Non Incapacitating	Possible		
Concrete	79	175	587	541	4,399	5,781
Blacktop	229	433	1,330	948	9,242	12,182
Gravel, Dirt and Brick	18	27	80	59	695	879
Total	327	637	2,006	1,556	14,393	18,919

Table A.11 Number of truck-crashes in Kansas based on road-surface conditions

Surface Condition	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total
		Disabled	Non Incapacitating	Possible		
Dry	280	520	1,619	1199	11,357	14,975
Wet	27	58	213	180	1,472	1,950
Ice or Snow Packed, Snow or Slush, Mud, Dirt or Sand and Debris	20	58	168	168	1,522	1,936
Total	327	637	2,006	1,556	14,393	18,919

Table A.12 Number of truck-crashes in Kansas based on the road-surface character

Road Surface Character	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total
		Disabled	Non Incapacitating	Possible		
Straight and Level	215	407	1263	986	9,868	12739
Straight on Grade and Straight at Hill Crest	67	149	415	360	2,995	3986
Curved and Level, Curved on Grade and Curved at Hillcrest	45	81	322	197	1439	2084
Total	327	637	2,006	1,556	14,393	18,919

Table A.13 Number of truck-crashes in Kansas based on day of week

Day of the Week	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total Number of Crashes
		Disabled	Non Incapacitating	Possible		
Monday	59	111	359	277	2,335	3,141
Tuesday	50	112	334	295	2,609	3,400
Wednesday	58	137	357	280	2,676	3,508
Thursday	56	106	338	289	2,594	3,383
Friday	58	103	334	234	2,441	3,170
Saturday	28	41	173	110	1100	1,452
Sunday	18	27	111	71	634	861
Total	327	637	2,006	1,556	14,393	18,919

Table A.14 Number of truck-crashes in Kansas based on accident class

Accident Class	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total
		Disabled	Non Incapacitating	Possible		
Collision with Other Motor Vehicle	278	444	1,266	1,125	8,838	11,951
Collision with Fixed Object	7	74	255	158	2,023	2,517
All others	42	119	485	273	3,530	4,449
Total	327	637	2,006	1,556	14,393	18,919

Table A.15 Number of truck-crashes in Kansas based on contributory cause involved

Contributory Cause	Fatal Crashes	Injury Crashes			Property Damage Only Crashes	Total Number of Crashes
		Disabled	Non Incapacitating	Possible		
Driver Related	289	558	1,644	1,211	9,558	13,260
Environment related	30	57	226	146	1,901	2,360
Road Condition Related	19	43	152	121	1,150	1,485
Vehicle and Pedestrian Related	20	34	122	73	893	1,142
Total	358	692	2,144	1,551	13,502	18,247

Appendix B Correlation Matrix

Table B.1 shows the Pearson's correlation matrix used in the study. The Pearson's correlation coefficient greater than 0.5 for the pairs of variables which were interdependent has been highlighted:

Table B.1 Correlation matrix

Variable	ALCOHOL	LOCATION	SPEED_LIMIT_1	SPEED_LIMIT_2	SPEED_LIMIT_3	SPEED_LIMIT_4	WEATHER	S_TYPE
ALCOHOL	1.000	0.006	-0.033	0.026	-0.004	0.022	-0.012	0.002
LOCATION	0.006	1.000	0.296	0.117	-0.067	-0.287	-0.081	0.102
SPEED_LIMIT_1	-0.033	0.296	1.000	-0.200	-0.331	**-0.572**	-0.094	0.072
SPEED_LIMIT_2	0.026	0.117	-0.200	1.000	-0.125	-0.216	-0.025	0.056
SPEED_LIMIT_3	-0.004	-0.067	-0.331	-0.125	1.000	-0.359	-0.029	-0.115
SPEED_LIMIT_4	0.022	-0.287	**-0.572**	-0.216	-0.359	1.000	0.119	0.004
WEATHER	-0.012	-0.081	-0.094	-0.025	-0.029	0.119	1.000	-0.004
S_TYPE	0.002	0.102	0.072	0.056	-0.115	0.004	-0.004	1.000
S_COND	0.015	0.058	0.047	0.023	0.034	-0.072	**-0.750**	0.019
S_CHAR	0.014	0.136	0.125	0.021	-0.059	-0.074	-0.061	0.004
CONSTR_MAINT	-0.008	-0.061	-0.025	0.002	0.094	-0.047	-0.059	-0.072
LIGHT	-0.146	0.157	0.185	0.059	0.035	-0.245	-0.089	-0.039
ONAT_TC	-0.018	**-0.689**	-0.182	-0.066	0.039	0.173	0.058	-0.068
TIME_DAY	-0.169	0.138	0.161	0.048	0.039	-0.215	-0.048	-0.034
DAY	-0.077	0.038	0.071	0.024	-0.002	-0.077	-0.050	0.003
CLASS	0.050	0.281	0.119	0.098	-0.050	-0.120	-0.016	-0.085
MANEUVER	0.023	-0.177	-0.211	-0.033	0.054	0.192	0.070	0.023
DAMAGE	0.042	-0.033	-0.204	-0.007	0.071	0.148	0.062	-0.020
YOUNG	0.036	0.091	0.034	0.032	0.014	-0.056	0.001	-0.016
MIDDLE_AGED	-0.026	-0.058	-0.036	-0.022	-0.012	0.055	-0.003	0.020
OLD	-0.010	-0.041	-0.023	-0.012	0.023	0.011	-0.014	0.012
GENDR	0.019	-0.045	-0.049	-0.028	0.030	0.036	-0.021	0.045
SAFETY_EQUIPT	-0.087	-0.004	0.045	0.002	-0.075	0.013	0.041	-0.008

Table B.1 Correlation matrix (cont.)

Variable	ALCOHOL	LOCATION	SPEED_LIMIT_1	SPEED_LIMIT_2	SPEED_LIMIT_3	SPEED_LIMIT_4	WEATHER	S_TYPE
TRAPPED	0.074	0.002	-0.078	-0.021	0.031	0.064	0.012	0.017
CONTROL	0.009	0.122	-0.155	0.020	-0.050	0.190	0.045	0.003
COLLISION	0.040	0.059	-0.025	0.064	0.007	-0.014	0.008	-0.062
CC_RD	-0.028	-0.100	-0.105	-0.014	0.003	0.097	0.358	-0.035
CC_DR	0.071	0.153	0.169	0.053	-0.022	-0.173	-0.034	-0.042
CC_VEH	-0.026	-0.050	-0.063	-0.002	0.017	0.055	-0.054	-0.010
CC_ENV	-0.034	-0.154	-0.191	-0.055	-0.015	0.226	0.346	0.038
DRUGS_ALCOHOL	0.822	0.008	-0.030	0.022	0.000	0.019	-0.017	-0.002
SPEED	0.031	0.017	-0.075	0.000	0.014	0.047	0.226	-0.050
WRONG	0.044	0.087	0.040	0.026	-0.021	-0.037	-0.064	-0.020
IMP_MAN	-0.011	-0.015	0.074	-0.013	-0.012	-0.048	-0.069	-0.051
TIME_ATTN	0.020	0.157	0.157	0.056	-0.026	-0.155	-0.119	0.037
EVASIVE	0.007	-0.060	-0.095	-0.001	0.027	0.076	-0.004	-0.005
CARELESS	0.084	-0.007	-0.016	0.000	0.001	0.015	-0.017	0.004
SMOG_SAND	-0.004	-0.008	-0.042	-0.013	0.017	0.030	0.130	-0.002
RAIN	-0.003	-0.047	-0.058	-0.006	-0.007	0.060	0.295	-0.011
SNOW	-0.022	-0.099	-0.109	-0.037	-0.029	0.140	0.424	0.000
VSN_OBSTRUCT	-0.028	-0.126	-0.150	-0.047	-0.003	0.181	-0.039	0.060
WET	-0.027	-0.087	-0.093	-0.017	-0.022	0.103	0.403	-0.027
BRAKES	-0.017	-0.019	-0.056	-0.007	0.018	0.050	-0.047	-0.009
UNATTND	-0.007	-0.019	0.027	0.009	-0.012	-0.228	-0.006	0.001
RUTS	-0.013	-0.050	-0.051	-0.004	0.046	0.021	0.001	-0.016
INOPERATIVE	0.004	-0.019	-0.010	0.008	0.028	-0.016	-0.007	-0.010

Table B.1 Correlation matrix (cont.)

Variable	S_COND	S_CHAR	CONSTR_MAINT	LIGHT	ONAT_TC	TIME_DAY	DAY	CLASS	MANEUVER	DAMAGE
ALCOHOL	0.015	0.014	-0.008	-0.146	-0.018	-0.169	-0.077	0.050	0.023	0.042
LOCATION	0.058	0.136	-0.061	0.157	**-0.689**	0.138	0.038	0.281	-0.177	-0.033
SPEED_LIMIT_1	0.047	0.125	-0.025	0.185	-0.182	0.161	0.071	0.119	-0.211	-0.204
SPEED_LIMIT_2	0.023	0.021	0.002	0.059	-0.066	0.048	0.024	0.098	-0.033	-0.007
SPEED_LIMIT_3	0.034	-0.059	0.094	0.035	0.039	0.039	-0.002	-0.050	0.054	0.071
SPEED_LIMIT_4	-0.072	-0.074	-0.047	-0.245	0.173	-0.215	-0.077	-0.120	0.192	0.148
WEATHER	**-0.750**	-0.061	-0.059	-0.089	0.058	-0.048	-0.050	-0.016	0.070	0.062
S_TYPE	0.019	0.004	-0.072	-0.039	-0.068	-0.034	0.003	-0.085	0.023	-0.020
S_COND	1.000	0.065	0.053	0.085	-0.040	0.032	0.044	-0.009	-0.040	-0.051
S_CHAR	0.065	1.000	-0.019	-0.007	-0.107	-0.012	0.002	0.076	-0.040	-0.041
CONSTR_MAINT	0.053	-0.019	1.000	0.041	0.043	0.021	0.003	0.051	-0.025	-0.005
LIGHT	0.085	-0.007	0.041	1.000	-0.108	**0.729**	0.114	0.236	-0.147	-0.098
ONAT_TC	-0.040	-0.107	0.043	-0.108	1.000	-0.098	-0.022	-0.241	0.056	-0.036
TIME_DAY	0.032	-0.012	0.021	**0.729**	-0.098	1.000	0.115	0.222	-0.129	-0.085
DAY	0.044	0.002	0.003	0.114	-0.022	0.115	1.000	0.054	-0.034	-0.035
CLASS	-0.009	0.076	0.051	0.236	-0.241	0.222	0.054	1.000	-0.213	0.015
MANEUVER	-0.040	-0.040	-0.025	-0.147	0.056	-0.129	-0.034	-0.213	1.000	0.133
DAMAGE	-0.051	-0.041	-0.005	-0.098	-0.036	-0.085	-0.035	0.015	0.133	1.000
YOUNG	-0.015	0.028	-0.001	0.056	-0.077	0.055	0.006	0.232	-0.048	0.030
MIDDLE_AGED	0.014	-0.026	-0.004	-0.033	0.047	-0.042	0.006	-0.211	0.046	-0.026
OLD	0.018	-0.012	-0.004	-0.027	0.026	-0.020	-0.011	-0.139	0.030	-0.007
GENDR	0.034	-0.020	-0.022	-0.052	0.024	-0.055	0.001	-0.266	0.069	-0.005
SAFETY_EQUIPT	-0.051	0.016	0.007	-0.013	0.013	-0.007	0.010	0.048	-0.028	-0.029

Table B.1 Correlation matrix (cont.)

Variable	S_COND	S_CHAR	CONSTR_MAINT	LIGHT	ONAT_TC	TIME_DAY	DAY	CLASS	MANEUVER	DAMAGE
TRAPPED	0.007	-0.022	-0.012	-0.021	-0.019	-0.018	0.005	-0.010	0.040	0.055
CONTROL	0.028	-0.036	0.033	-0.001	-0.122	0.002	-0.006	0.163	0.161	0.095
COLLISION	-0.020	0.008	0.049	0.098	0.098	0.092	0.026	0.357	0.035	0.066
CC_RD	-0.430	-0.081	0.005	-0.037	0.072	-0.004	-0.038	-0.024	0.057	0.058
CC_DR	0.007	-0.014	0.052	0.181	-0.106	0.155	0.036	0.315	-0.205	0.037
CC_VEH	0.072	-0.025	-0.013	0.069	0.051	0.056	0.015	-0.055	0.085	-0.086
CC_ENV	-0.248	-0.033	-0.057	-0.249	0.105	-0.204	-0.072	-0.208	0.161	0.106
DRUGS_ALCOHOL	0.021	0.006	-0.004	-0.125	-0.019	-0.141	-0.065	0.046	0.016	0.043
SPEED	-0.259	-0.102	-0.005	0.021	-0.058	0.031	-0.021	0.029	0.068	0.111
WRONG	0.066	0.036	0.028	0.074	-0.006	0.063	0.025	0.184	-0.114	0.024
IMP_MAN	0.062	0.029	0.040	0.068	0.047	0.057	0.015	0.214	-0.285	-0.076
TIME_ATTN	0.111	0.024	0.024	0.099	-0.137	0.075	0.035	0.164	-0.035	0.012
EVASIVE	-0.001	-0.031	0.006	0.008	0.045	0.003	-0.019	-0.019	-0.056	0.046
CARELESS	0.029	0.009	0.003	-0.024	-0.006	-0.028	-0.009	0.036	0.000	0.035
SMOG_SAND	-0.029	0.006	-0.011	-0.022	0.008	-0.021	-0.015	0.017	0.009	0.013
RAIN	-0.274	-0.027	-0.017	-0.049	0.040	-0.036	-0.014	0.000	0.005	0.029
SNOW	-0.325	-0.037	-0.031	-0.041	0.066	-0.025	-0.056	-0.055	0.075	0.057
VSN_OBSTRUCT	0.059	0.004	-0.047	-0.303	0.082	-0.260	-0.051	-0.275	0.163	**0.831**
WET	-0.490	-0.078	-0.040	-0.041	0.063	0.000	-0.037	0.002	0.047	0.061
BRAKES	0.053	-0.015	-0.013	-0.049	0.028	0.038	0.004	-0.057	0.058	0.004
UNATTND	0.007	-0.016	-0.007	-0.008	0.016	-0.003	-0.001	0.010	-0.008	-0.004
RUTS	0.007	-0.025	0.007	-0.010	0.040	-0.016	-0.016	-0.077	0.043	0.003
INOPERATIVE	0.017	-0.009	0.166	0.017	0.006	0.006	0.009	0.007	-0.002	0.009

Table B.1 Correlation matrix (cont.)

Variable	YOUNG	MIDDLE_AGED	OLD	GENDR	SAFETY_EQUIPT	TRAPPED	CONTROL	COLLISION
ALCOHOL	0.036	-0.026	-0.010	0.019	-0.087	0.074	0.009	0.040
LOCATION	0.091	-0.058	-0.041	-0.045	-0.004	0.002	0.122	0.059
SPEED_LIMIT_1	0.034	-0.036	-0.023	-0.049	0.045	-0.078	-0.155	-0.025
SPEED_LIMIT_2	0.032	-0.022	-0.012	-0.028	0.002	-0.021	0.020	0.064
SPEED_LIMIT_3	0.014	-0.012	0.023	0.030	-0.075	0.031	-0.050	0.007
SPEED_LIMIT_4	-0.056	0.055	0.011	0.036	0.013	0.064	0.190	-0.014
WEATHER	0.001	-0.003	-0.014	-0.021	0.041	0.012	0.045	0.008
S_TYPE	-0.016	0.020	0.012	0.045	-0.008	0.017	0.003	-0.062
S_COND	-0.015	0.014	0.018	0.034	-0.051	0.007	0.028	-0.020
S_CHAR	0.028	-0.026	-0.012	-0.020	0.016	-0.022	-0.036	0.008
CONSTR_MAINT	-0.001	-0.004	-0.004	-0.022	0.007	-0.012	0.033	0.049
LIGHT	0.056	-0.033	-0.027	-0.052	-0.013	-0.021	-0.001	0.098
ONAT_TC	-0.077	0.047	0.026	0.024	0.013	-0.019	-0.122	0.098
TIME_DAY	0.055	-0.042	-0.020	-0.055	-0.007	-0.018	0.002	0.092
DAY	0.006	0.006	-0.011	0.001	0.010	0.005	-0.006	0.026
CLASS	0.232	-0.211	-0.139	-0.266	0.048	-0.010	0.163	0.357
MANEUVER	-0.048	0.046	0.030	0.069	-0.028	0.040	0.161	0.035
DAMAGE	0.030	-0.026	-0.007	-0.005	-0.029	0.055	0.095	0.066
YOUNG	1.000	**-0.816**	-0.082	-0.115	-0.051	0.020	0.021	0.112
MIDDLE_AGED	**-0.816**	1.000	-0.222	0.265	0.033	-0.006	-0.013	-0.078
OLD	-0.082	-0.222	1.000	0.064	-0.032	0.005	-0.022	-0.048
GENDR	-0.115	0.265	0.064	1.000	-0.039	-0.016	-0.055	-0.070
SAFETY_EQUIPT	-0.051	0.033	-0.032	-0.039	1.000	-0.114	0.027	-0.004

Table B.1 Correlation matrix (cont.)

Variable	YOUNG	MIDDLE_AGED	OLD	GENDR	SAFETY_EQUIPT	TRAPPED	CONTROL	COLLISION
TRAPPED	0.020	-0.006	0.005	-0.016	-0.114	1.000	0.038	0.020
CONTROL	0.021	-0.013	-0.022	-0.055	0.027	0.038	1.000	0.093
COLLISION	0.112	-0.078	-0.048	-0.070	-0.004	0.020	0.093	1.000
CC_RD	0.014	-0.002	-0.011	-0.010	0.031	0.001	0.048	0.012
CC_DR	0.108	-0.068	-0.046	-0.063	-0.023	0.043	0.092	0.172
CC_VEH	-0.026	0.008	0.018	0.014	-0.019	-0.007	0.037	-0.052
CC_ENV	-0.055	0.068	0.022	0.061	0.039	0.012	0.046	-0.057
DRUGS_ALCOHOL	0.031	-0.019	-0.107	0.019	-0.091	0.070	0.015	0.041
SPEED	0.045	-0.020	-0.021	0.001	-0.030	0.066	0.094	0.045
WRONG	0.064	-0.048	-0.019	-0.028	-0.005	0.014	0.067	0.283
IMP_MAN	0.045	-0.044	-0.031	-0.077	0.021	-0.032	0.012	-0.102
TIME_ATTN	0.051	-0.023	-0.019	-0.017	-0.035	0.047	0.022	0.105
EVASIVE	-0.011	0.016	0.010	0.010	-0.010	0.025	0.028	0.017
CARELESS	0.029	-0.040	-0.012	-0.022	-0.049	0.035	0.012	0.009
SMOG_SAND	-0.002	-0.001	0.016	-0.003	-0.011	0.014	-0.004	0.035
RAIN	0.011	-0.009	-0.004	-0.004	0.015	0.017	0.018	0.012
SNOW	-0.013	0.017	-0.002	0.003	0.023	0.030	0.516	0.000
VSN_OBSTRUCT	-0.078	0.092	0.031	0.089	0.034	-0.013	0.014	-0.104
WET	0.023	-0.010	-0.008	-0.016	0.037	0.001	0.060	0.024
BRAKES	-0.005	0.008	0.020	0.020	-0.030	-0.004	0.037	-0.022
UNATTND	-0.021	-0.005	-0.002	-0.021	-0.010	0.005	-0.035	0.002
RUTS	-0.017	0.016	-0.005	0.014	-0.011	0.004	-0.021	-0.035
INOPERATIVE	-0.003	0.007	-0.010	-0.005	0.010	-0.010	0.002	0.010

Table B.1 Correlation matrix (cont.)

Variable	CC_RD	CC_DR	CC_VEH	CC_ENV	DRUGS_ALCOHOL	SPEED	WRONG	IMP_MAN	TIME_ATTN	EVASIVE	CARELESS
TRAPPED	0.001	0.043	-0.007	0.012	0.070	0.066	0.014	-0.032	0.047	0.025	0.035
CONTROL	0.048	0.092	0.037	0.046	0.015	0.094	0.067	0.012	0.022	0.028	0.012
COLLISION	0.012	0.172	-0.052	-0.057	0.041	0.045	0.283	-0.102	0.105	0.017	0.009
CC_RD	1.000	-0.036	-0.036	0.263	-0.025	0.232	-0.064	-0.078	-0.128	0.023	-0.029
CC_DR	-0.036	1.000	-0.244	0.260	0.084	0.268	0.257	0.255	**0.552**	0.148	0.089
CC_VEH	-0.036	-0.244	1.000	-0.062	-0.028	-0.054	-0.070	-0.075	-0.143	-0.025	-0.030
CC_ENV	0.263	0.260	-0.062	1.000	-0.034	0.068	-0.102	0.109	-0.213	-0.003	0.039
DRUGS_ALCOHOL	-0.025	0.084	-0.028	-0.034	1.000	0.038	0.041	-0.003	0.023	0.004	0.086
SPEED	0.232	0.268	-0.054	0.068	0.038	1.000	-0.043	0.105	-0.065	-0.016	0.052
WRONG	-0.064	0.257	-0.070	-0.102	0.041	-0.043	1.000	-0.082	-0.019	0.001	0.004
IMP_MAN	-0.078	0.255	-0.075	0.109	-0.003	0.105	-0.082	1.000	-0.072	-0.027	0.007
TIME_ATTN	-0.128	**0.552**	-0.143	-0.213	0.023	-0.065	-0.019	-0.072	1.000	-0.065	0.015
EVASIVE	0.023	0.148	-0.025	-0.003	0.004	-0.016	0.001	-0.027	-0.065	1.000	0.001
CARELESS	-0.029	0.089	-0.030	0.039	0.086	0.052	0.004	0.007	0.015	0.001	1.000
SMOG_SAND	0.017	-0.017	-0.008	0.206	0.001	0.029	0.002	-0.018	-0.023	0.008	0.005
RAIN	0.263	-0.012	-0.015	0.383	-0.016	0.089	-0.023	-0.022	-0.053	0.023	-0.011
SNOW	0.329	-0.072	-0.025	**0.553**	-0.016	0.150	-0.056	-0.067	-0.120	0.005	-0.022
VSN_OBSTRUCT	-0.051	-0.317	-0.058	**0.653**	-0.028	-0.090	-0.089	-0.087	-0.179	-0.020	-0.032
WET	**0.895**	-0.003	-0.042	0.299	-0.027	0.261	-0.056	-0.070	-0.119	0.017	-0.026
BRAKES	-0.032	-0.190	**0.771**	-0.050	-0.020	-0.041	-0.048	-0.053	-0.119	-0.019	-0.026
UNATTND	-0.002	-0.034	0.232	-0.002	-0.007	-0.018	-0.023	-0.009	-0.015	0.004	-0.008
RUTS	0.365	-0.086	0.014	-0.002	-0.009	-0.014	-0.033	-0.040	-0.054	0.015	-0.010
INOPERATIVE	0.244	-0.013	-0.011	-0.005	0.009	0.002	-0.007	-0.004	-0.010	0.017	-0.009

Table B.1 Correlation matrix (cont.)

Variable	SMOG_SAND	RAIN	SNOW	VSN_OBSTRUCT	WET	BRAKES	UNATTND	RUTS	INOPERATIVE
TRAPPED	0.014	0.017	0.030	-0.013	0.001	-0.004	0.005	0.004	-0.010
CONTROL	-0.004	0.018	0.516	0.014	0.060	0.037	-0.035	-0.021	0.002
COLLISION	0.035	0.012	0.000	-0.104	0.024	-0.022	0.002	-0.035	0.010
CC_RD	0.017	0.263	0.329	-0.051	**0.895**	-0.032	-0.002	0.365	0.244
CC_DR	-0.017	-0.012	-0.072	-0.317	-0.003	-0.190	-0.034	-0.086	-0.013
CC_VEH	-0.008	-0.015	-0.025	-0.058	-0.042	**0.771**	0.232	0.014	-0.011
CC_ENV	0.206	0.383	**0.553**	**0.653**	0.299	-0.050	-0.002	-0.002	-0.005
DRUGS_ALCOHOL	0.001	-0.016	-0.016	-0.028	-0.027	-0.020	-0.007	-0.009	0.009
SPEED	0.029	0.089	0.150	-0.090	0.261	-0.041	-0.018	-0.014	0.002
WRONG	0.002	-0.023	-0.056	-0.089	-0.056	-0.048	-0.023	-0.033	-0.007
IMP_MAN	-0.018	-0.022	-0.067	-0.087	-0.070	-0.053	-0.009	-0.040	-0.004
TIME_ATTN	-0.023	-0.053	-0.120	-0.179	-0.119	-0.119	-0.015	-0.054	-0.010
EVASIVE	0.008	0.023	0.005	-0.020	0.017	-0.019	0.004	0.015	0.017
CARELESS	0.005	-0.011	-0.022	-0.032	-0.026	-0.026	-0.008	-0.010	-0.009
SMOG_SAND	1.000	-0.002	0.032	0.007	0.003	-0.004	-0.005	0.025	0.014
RAIN	-0.002	1.000	0.082	-0.024	0.292	-0.008	0.005	0.025	0.006
SNOW	0.032	0.082	1.000	-0.032	0.369	-0.243	-0.003	-0.004	0.005
VSN_OBSTRUCT	0.007	-0.024	-0.032	1.000	-0.045	-0.047	-0.006	-0.167	-0.014
WET	0.003	0.292	0.369	-0.045	1.000	-0.033	0.001	0.000	0.011
BRAKES	-0.004	-0.008	-0.243	-0.047	-0.033	1.000	0.014	0.000	-0.009
UNATTND	-0.005	0.005	-0.003	-0.006	0.001	0.014	1.000	-0.006	-0.004
RUTS	0.025	0.025	-0.004	-0.167	0.000	0.000	-0.006	1.000	0.045
INOPERATIVE	0.014	0.006	0.005	-0.014	0.011	-0.009	-0.004	0.045	1.000

Appendix C Variable Selection Methods

Following are models and goodness-of-fit statistics for forward selection and stepwise selection methods of variable selection procedures, respectively:

Forward Selection Method

Table C.1 shows parameter estimates and odds-ratio values of the variables in the model obtained by the forward selection method.

Table C.1 Model obtained by forward selection method

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Sq	Odds Ratio	95% Wald Confidence Limits For Odds Ratio
Intercept*	-1.494	0.163	84.47	<0.0001		
ALCOHOL*	0.973	0.135	51.9	<0.0001	2.65	2.03,3.45
CARELESS*	0.331	0.125	6.98	0.0083	1.39	1.09,1.78
CC_DR*	0.589	0.053	122.43	<0.0001	1.8	1.62,2.00
CC_RD*	-0.303	0.082	13.51	0.0002	0.74	0.63,0.87
CC_VEH	-0.09	0.093	0.94	0.3329	0.91	0.76, 1.10
CLASS	0.103	0.052	3.92	0.0477	1.11	1.00,1.23
COLLISION*	0.473	0.052	83.78	<0.0001	1.61	1.45,1.78
CONSTR_MAINT*	-0.271	0.083	10.68	0.0011	0.76	0.65,0.90
CONTROL*	0.307	0.057	29.47	<0.0001	1.36	1.22,1.52
DAMAGE*	1.12	0.083	182.14	<0.0001	3.06	2.60,3.60
DAY	-0.003	0.058	0	0.9661	1	0.89, 1.12
EVASIVE*	0.43	0.079	29.83	<0.0001	1.54	1.32,1.80
GENDR*	-0.129	0.049	7.08	0.0078	0.88	0.80,0.97
IMP_MAN*	-0.455	0.068	44.85	<0.0001	0.64	0.56,0.73
INOPERATIVE	-0.247	0.328	0.57	0.4508	0.78	0.41, 1.48
LIGHT	0.06	0.049	1.5	0.2209	1.06	0.96,1.17
MANEUVER*	0.321	0.041	61.66	<0.0001	1.38	1.27,1.49
MIDDLE_AGED*	0.104	0.043	5.95	0.0147	1.11	1.021,1.21
OLD	0.092	0.14	0.43	0.5141	1.1	0.83, 1.44
ONAT_TC*	-0.517	0.054	92.35	<0.0001	0.6	0.54,0.66
RAIN*	0.312	0.132	5.64	0.0176	1.37	1.06,1.77
RUTS	-0.148	0.224	0.44	0.5091	0.86	0.56, 1.34
S_CHAR*	-0.113	0.041	7.72	<0.0001	0.89	0.83,0.97
S_COND*	0.234	0.055	18.32	<0.0001	1.26	1.14,1.41
S_TYPE*	0.133	0.04	10.87	0.001	1.14	1.06,1.24
SAFETY_EQUIPT*	-1.379	0.075	338.08	<0.0001	0.25	0.217, 0.292
SMOG_SAND	0.355	0.218	2.65	0.1037	1.43	0.93, 2.19
SNOW	0.17	0.098	3	0.0831	1.19	0.978, 1.437
SPEED*	0.449	0.054	68.62	<0.0001	1.57	1.41, 1.74
SPEED_LIMIT_1*	-0.807	0.051	253.93	<0.0001	0.45	0.40, 0.49
SPEED_LIMIT_2*	-0.396	0.076	26.95	<0.0001	0.67	0.58, 0.78
SPEED_LIMIT_3*	0.11	0.052	4.6	0.032	1.12	1.01, 1.24
TRAPPED*	4.43	0.344	166.15	<0.0001	83.95	42.80, 164.66
UNATTND	0.483	0.329	2.16	0.142	1.62	0.85, 3.09
WRONG	0.014	0.058	0.06	0.8034	1.01	0.91, 1.14

*significant at 0.05 level

Table C.2 Model fit statistics of the binary logistic-regression analysis

Criterion	Intercept Only	Intercept and Covariates
AIC	20820.1	17391.8
SC	20828	17613.7
-2logL	20818.1	17337.8

Table C.3 Associations of predicted probabilities and observed responses

Statistic	Value
Percent Concordant	76
Percent Discordant	23.7
Percent Tied	0.4
Pairs	65,142,718
Somers' D	0.523
Gamma	0.525
Tau-a	0.19
C	0.762

- R^2 = 0.1680

Stepwise Selection Method

Table C.4 shows parameter estimates and odds-ratio values of the variables in the model obtained by the stepwise selection method:

Table C.4 Model obtained by stepwise selection method

Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Sq	Odds Ratio	95% Wald Confidence Limits For Odds Ratio
Intercept*	-1.513	0.163	86.21	<0.0001		
ALCOHOL*	0.976	0.135	52.24	<0.0001	2.65	2.04,3.46
CARELESS*	0.333	0.125	7.06	0.0079	1.4	1.09,1.79
CC_DR*	0.595	0.053	124.20	<0.0001	1.81	1.63,2.01
CC_RD*	-0.333	0.084	15.54	<0.0001	0.72	0.61,0.85
CC_VEH	-0.09	0.093	0.94	0.3329	0.91	0.76, 1.10
CLASS	0.106	0.052	4.10	0.0429	1.11	1.00,1.23
COLLISION*	0.473	0.052	83.56	<0.0001	1.6	1.45,1.78
CONSTR_MAINT*	-0.269	0.083	10.49	0.0012	0.76	0.65,0.90
CONTROL*	0.304	0.057	28.87	<0.0001	1.36	1.23,1.51
DAMAGE*	1.117	0.083	181.40	<0.0001	3.06	2.6,3.6
DAY	-0.003	0.058	0.00	0.9661	1	0.90, 1.12
EVASIVE*	0.43	0.079	29.80	<0.0001	1.54	1.32,1.80
GENDR*	-0.129	0.049	7.07	0.0078	0.88	0.80,0.97
IMP_MAN*	-0.455	0.068	44.79	<0.0001	0.64	0.56,0.73
INOPERATIVE	-0.247	0.328	0.57	0.4508	0.78	0.41, 1.48
LIGHT	0.06	0.049	1.50	0.2209	1.06	0.96,1.17
MANEUVER*	0.32	0.041	61.06	<0.0001	1.38	1.27,1.49
MIDDLE_AGED*	0.103	0.043	5.87	0.0154	1.11	1.02,1.21
OLD	0.092	0.14	0.43	0.5141	1.1	0.83, 1.44
ONAT_TC*	-0.52	0.054	93.26	<0.0001	0.6	0.54,0.66
RAIN*	0.329	0.132	6.23	0.0125	1.39	1.073,1.80
RUTS	-0.148	0.224	0.44	0.5091	0.86	0.56, 1.34
S_CHAR*	-0.114	0.041	7.88	0.005	0.89	0.82,0.97
S_COND*	0.255	0.056	20.57	<0.0001	1.29	1.16,1.44
S_TYPE*	0.132	0.04	10.69	0.0011	1.14	1.05,1.24
SAFETY_EQUIPT*	-1.38	0.075	338.74	<0.0001	0.25	0.22,0.29
SMOG_SAND	0.355	0.218	2.65	0.1037	1.43	0.93, 2.19
SNOW	0.17	0.098	3.00	0.0831	1.19	0.98,1.44
SPEED*	0.444	0.054	66.83	<0.0001	1.56	1.40,1.733
SPEED_LIMIT_1*	-0.801	0.051	249.34	<0.0001	0.45	0.41,0.50
SPEED_LIMIT_2*	-0.39	0.077	26.07	<0.0001	0.68	0.58,0.79
SPEED_LIMIT_3*	0.115	0.052	5.00	0.0254	1.12	1.01,1.24
TRAPPED*	4.419	0.344	165.23	<0.0001	83.01	42.32,162.84
UNATTND	0.483	0.329	2.16	0.142	1.62	0.85, 3.09
WRONG	0.014	0.058	0.06	0.8034	1.01	0.91, 1.14

*significant at 0.05 level

Table C.5 Model fit statistics of the binary logistic-regression analysis

Criterion	Intercept Only	Intercept and Covariates
AIC	20820.1	17390.9
SC	20828	17610.6
-2logL	20818.1	17334.9

Table C.6 Associations of predicted probabilities and observed responses

Statistic	Value
Percent Concordant	76
Percent Discordant	23.7
Percent Tied	0.4
Pairs	65,142,718
Somers' D	0.523
Gamma	0.525
Tau-a	0.191
C	0.762

- R^2 = 0.1682