Accident Analysis and Prevention 117 (2018) 106–113

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Effects of truck traffic on crash injury severity on rural highways in Wyoming using Bayesian binary logit models

Mohamed M. Ahmed[a,*], Rebecca Franke[b], Khaled Ksaibati[c], Debbie S. Shinstine[c]

[a] Civil and Architectural Engineering Department, University of Wyoming, 1000 E. University Ave, Dept 3295, Laramie, WY, 82071, United States
[b] Traffic & Safety Bureau, Montana Department of Transportation, 2701 Prospect Avenue, PO Box 201001, Helena, MT, 59620-1001, United States
[c] Wyoming Technology Transfer Center, 1000 E. University Ave, Dept 3295, Laramie, WY, 82071, United States

ARTICLE INFO

Keywords:
Heavy truck safety
Injury severity
Bayesian logit model
Oil and gas
Inclement weather

ABSTRACT

Roadway safety is an integral part of a functioning infrastructure. A major use of the highway system is the transport of goods. The United States has experienced constant growth in the amount of freight transported by truck in the last few years. Wyoming is experiencing a large increase in truck traffic on its local and county roads due to an increase in oil and gas production.

This study explores the involvement of heavy trucks in crashes and their significance as a predictor of crash severity and addresses the effect that large truck traffic is having on the safety of roadways for various road classifications. Studies have been done on the factors involved in and the causation of heavy truck crashes, but none address the causation and effect of roadway classifications on truck crashes.

Binary Logit Models (BLM) with Bayesian inferences were utilized to classify heavy truck involvement in severe and non-severe crashes using ten years (2002–2011) of historical crash data in the State of Wyoming. From the final main effects model, various interactions proved to be significant in predicting the severity of crashes and varied depending on the roadway classification. The results indicated the odds of a severe crash increase to 2.3 and 4.5 times when a heavy truck is involved on state and interstate highways respectively. The severity of crashes is significantly increased when road conditions were not clear, icy, and during snowy weather conditions.

1. Introduction

Large truck crashes tend to have the greatest repercussions of all motor vehicle crashes and is a primary concern in the study of transportation safety. Property damage, loss of goods, and loss of life are all heightened aspects when dealing with a heavy truck crash. The average cost of a heavy-truck related crash in 2005 was estimated to be more than $91,000 (Zaloshnja and Miller, 2007). Since 2009, the numbers of injuries and fatalities due to crashes that involved a heavy truck have been steadily increasing from the previous years (NHTSA, 2014).

The U.S. economy relies heavily upon the trucking industry for the movement of goods. Trucking was responsible for a total of 9.4 million tons of freight in America in 2012, which accounted for 68.5 percent of the total freight weight. This percentage is expected to continue to grow in the next ten years (ATA, 2013).

Not all roads were originally designed to handle the amount of heavy truck traffic that is becoming common on interstates, highways, bridges, and local roads. Large trucks not only contribute to prematurely deteriorated road structures, they also affect crash rates. In 2012, approximately 4000 fatalities and 104,000 injuries occurred nationwide in crashes that involved at least one heavy truck. Of these fatalities, 83 percent were not the operators of the heavy truck (NHTSA, 2014). Roads are not always built wide enough or with the correct geometry needed to accommodate heavy trucks. Driver fatigue from long hours of operating also contributes to the number of crashes involving heavy trucks. Inclement weather can compound this issue.

With oil and gas extraction being a growing industry all around the United States, the amount of heavy trucks operating in rural locations is expected to increase. During a three year period from 2003–2006, a total of 404 fatalities occurred nationwide amongst extraction workers. Of these fatalities, 110 were highway related incidents and multi-vehicle collisions accounted for 36 percent of the highway related fatalities (Bureau of Labor Statistics, 2008). Wyoming and North Dakota are both currently experiencing high oil and gas production. Crash rates for heavy trucks in these states were twice as high as the national average. According to a recent study by the American Transportation Research Institute (ATRI) (Weber and Murray, 2014), Wyoming is ranked the worst in large truck crashes in the US with large truck crash rate of 0.52

* Corresponding author.
E-mail address: mahmed@uwyo.edu (M.M. Ahmed).

https://doi.org/10.1016/j.aap.2018.04.011
Received 7 May 2015; Received in revised form 30 March 2018; Accepted 9 April 2018
Available online 02 May 2018
0001-4575/

crashes/MVMT compared to 0.26 crashes/MVMT national average. The large-truck involvement in fatal crashes in Wyoming has increased from 8.1% of total vehicles involved in fatal crashes in 2009 to 16.8% in 2012 (NHTSA, 2011).

2. Objectives

The objectives of this paper are to explore if the involvement of a heavy truck in a crash is a significant predictor of crash severity, to identify factors combined with heavy truck involvement affecting the severity of a crash, and to discover if a relationship exists between the severity of a crash involving a heavy truck and the classification of the road on which the crash occurred. Together with a descriptive analysis of factors involved in heavy truck crashes, recommendations can be made to improve truck safety.

3. Background

Crash causation can be examined from a variety of angles. Factors are generally broken down into a few different categories including: driver characteristics, external or roadway features, and vehicle factors. In a large truck safety evaluation conducted in Wisconsin (Andrea et al., 2011), driver characteristics and behaviors were shown to be the most significant variables in the outcome of severity of crashes in the state. In the study, many county officials agreed that large truck crash causations are due to driver decisions more so than the infrastructure's design (Andrea et al., 2011).

The critical reasons behind large truck involvement in fatal and injury crashes were explored in a Large Truck Crash Causation Study (LTCCS) (Craft et al., 2007). A critical reason is defined as the immediate reason for an unavoidable collision and is assigned to the vehicle responsible for the crash. Large trucks were assigned the critical reason in 55 percent of the study group. Of those crashes where the truck was the critical reason, 87 percent were driver-related, 10 percent were vehicle-related, and 3 percent were environment-related. A further look at the driver-related crashes revealed that 38 percent were caused by the decision of the driver, such as driving too fast, 28 percent were because the driver was inattentive or distracted, 12 percent were non-performance, such as falling asleep, and 9 percent were based on the performance of the driver, such as a panic or overcompensating (Craft et al., 2007).

The American Transportation Research Institute (ATRI) examined specific truck driver behavior and infractions as a way of predicting crash involvement (Lueck and Murray, 2011). The study used data that was collected from 2008–2009. It was found that five different truck driver events cause an increase in crash likelihood of 80 percent or more. The greatest increase in crash likelihood occurred when a driver failed to use or improperly used a turn signal at 96 percent. If the driver had been involved in a previous crash or if the driver had been issued an improper passing violation, the likelihood of a future crash occurring increased by 88 percent. The increase in likelihood was 84 percent if a driver had been caught making an improper turn and 80 percent if convicted of an improper or erratic lane change (Lueck and Murray, 2011).

Research presented at the International Driving Symposium (Knipling, 2009) investigated the critical reasons behind single vehicle and multi-vehicle crashes using data from the LTCCS. The conclusions showed that single vehicle truck crashes were the result of driver choices such as driving too fast for conditions or lack of sleep, as well as roadway factors such as curves. Multi-vehicle crashes where fault was placed on the truck driver usually occurred in dense traffic when the driver was not paying adequate attention to the surroundings. Multi-vehicle crashes where the other vehicle was found at fault had similar reasons relating to vehicle interaction errors rather than loss of control (Knipling, 2009).

Determining driver-related crash factors was the subject of a study that focused specifically on two-vehicle fatal collisions (Blower, 1998). This national study focused exclusively on crashes that involved one passenger vehicle and one heavy truck. These types of crashes make up about 60 percent of all truck involved fatal crashes, in which the truck driver survived 98 percent of the time and the passenger vehicle driver was killed in 83 percent. The dataset was separated into two files; one included all information about the truck involved in the crash, and the other file had the non-truck information. It was found that the truck driver was the main contributor to causing the fatal crash 34.1 percent of the time (Blower, 1998).

Another study which used data collected by the LTCCS examined the effect that large trucks had on the severity of crashes (Lemp et al., 2011). A specific focus was placed on long-combination vehicles, which are trucks pulling multiple trailers. The study's model found that the probability increased for injury and fatal crashes if non-bright lighting conditions are present, the road surface is snowy or icy, and/or there is fog.

The road type that a crash occurs on lends important information to the nature and possible causations of the crash. In 1987, a national analysis on large trucks involved in fatal crashes on different road classifications was conducted (Carsten, 1987). This early study considered four variables when predicting crash severity in its model: road class, truck type, truck gross weight, and year of the accident. Of the four types of roadways that were used, the largest amounts of crashes occurred on rural undivided roadways (Carsten, 1987).

More recent national trends in fatal heavy truck crashes on various roadways were observed in 2011 (Jarossi et al., 2011). A descriptive analysis showed that an average of approximately 5000 fatalities per year (2004–2008) occurred as a result of a truck-related crash. In a breakdown of crashes from 2008, 29.4 percent of fatalities occur on state highways, 27.1 percent on interstates, 23.8 percent on US highways, and 8.9 percent on county roads. Almost two-thirds of all the crashes occurred in rural areas. Of the truck drivers involved, 3.1 percent tested for alcohol and 1.2 percent for drugs (Jarossi et al., 2011). These facts do not take into account VMT which may be misleading.

With gas and oil extraction being a growing industry around the United States, the amount of heavy trucks is expected to increase. The National Institute for Occupation Safety and Health-(NIOSH) performs an annual study called the Census of Fatal Occupation Injuries (CFOI) examining the number of fatalities happening among workers in the field. Following a 15% increase in oil and gas extraction fatalities from 2003–2004 (Bureau of Labor Statistics, 2006), a full investigation was done. During a three year period of the 2003–2006 NIOSH study, a total of 404 fatalities happened among extraction workers. 110 of those fatalities were highway- related incidents. Non-collision events accounted for 38 percent of fatalities, while collisions between vehicles accounted for 36% (Mode and Conway, 2008).

When examining all of the CFOIs done by the NIOSH in the past 10 years, highway crashes consistently have been one of the leading causes of worker fatalities (Bureau of Labor Statistics, 2013). Mining had the second highest fatal work injury rate, the majority of which occur in the oil and gas extraction industries.

A safety survey was conducted examining the public view on the impact of increased oil drilling in Western North Dakota (Kubas and Vachal, 2012). Kubas & Vachal's research has shown a significant increase in both the truck traffic and crashes within the most impacted counties which have seen the largest increase in production of oil drilling. The survey revealed that 88.8 percent of the drivers in the affected counties felt less safe driving on the roads than they did 5 years earlier, approximately when drilling started increasing (Kubas and Vachal, 2012).

When analyzing crashes and their severity, statistical models provide valuable insight to the factors that impact crash severity. Several types of statistical models have been used for these purposes.

Savolainen et al. (2011) summarized the evolution of research and statistical analyses of traffic injury severities. From the review it can be seen that binary logit model (BLM), nested logit model (NLM), ordered logit model (OLM) and ordered probit model (OPM), and random parameters (Mixed) logit model (MLM) are among the most popular methods used in analyzing crash injury severities. Other non-traditional statistical models; data mining techniques such as artificial neural networks (ANN) (Abdelwahab and Abdel-Aty, 2001), classification and regression tree (CART) (Chang, 2006), and support vector machine (SVM) (Li et al., 2012), are also used. Recently, Bayesian statistics are gaining momentum in traffic safety analysis, Helai et al. (2008) introduced hierarchical Bayesian binomial logistic models to perform the multi-vehicle crash injury severity analysis. Lemp et al. (2011) analyzed large truck crash severity using heteroskedastic ordered probit models with Bayesian inference approach. Their approach allowed for variations of unobserved components, the authors concluded that the Bayesian approach outperformed the traditional maximum likelihood estimation (MLE). In this study, the BLM models with both the MLE and Bayesian inference techniques were estimated to investigate the effect of heavy vehicles on crash injury severity.

4. Methodology

The motivation of this study is based on Wyoming's strategic highway safety plan goal to reduce critical crashes. Critical crashes are defined as fatal and serious injury crashes. Wyoming has experienced increased truck traffic over the years. Interstate 80 alone carries over 40 percent truck traffic. This paper presents a statewide study of truck crashes on all roadway classifications in Wyoming.

Once data was collected and organized, a descriptive analysis was conducted. Crash severity is used in this study to assess the effect of trucks on crashes. All crash records and roadway inventories for the state of Wyoming were obtained from the Wyoming Department of Transportation (WYDOT) and a master dataset was created for a 10-year period (2002–2011).

Analysis of the statewide Wyoming dataset focuses on the relationships between different crash factors of vehicle type, road classification, road conditions, and driver influences. Using spreadsheets and various visual tools, relationships between variables were identified.

Upon completion of the descriptive analysis, a statistical analysis was used to verify the findings. First a univariate analysis was conducted to identify if trucks affect crash severity on each of the road types. Then a main effects model was used to test if trucks maintain their significance in the presence of other predictors. Finally, interactions were introduced into the model between trucks and other crash factors to better understand the effect of trucks on crash severity in the presence of these other variables. Binary logit models (BLM) with both the maximum likelihood estimation and Bayesian inference approach were used to classify heavy truck involvement in severe and non-severe crashes, and to identify the effect of driver, roadway, and environmental factors. For all models crash severity is the response variable, y, where $y = 1$ if a crash is severe and $y = 0$ if crash is not severe. Fatal and incapacitating injuries are considered severe, and non-severe crashes include non-incapacitating injury, possible injury, and no injury.

5. Data

The crash records from 2002–2011were investigated. During that time period, a total of 160,613 crashes involving 253,531 vehicles were recorded. Raw crash data along with maintenance inventories were collected from WYDOT and compiled into one database for analysis. The vehicle type and the estimated travel speed at the time of the crash for each vehicle involved was recorded in each crash report. The road surface type (paved/unpaved), vertical grade, horizontal alignment, and the presence of rumble strips data were identified for the specific segment of roadway on which the crash occurred. Information on all drivers that were involved in each crash which included age, gender, and safety equipment use (seat belts) was also included in the crash record. Data detailing the traffic volumes and geometric dimensions of the roadways was retrieved from WYDOT road inventories which included traffic volumes, vehicle miles travelled (VMT), road width, the number of lanes, and shoulder width and material. However, traffic volumes were not available for many of the roadways, particularly the secondary and local roads. The characteristics investigated included: severity, types of vehicles, number of vehicles, first harmful event (FHE), posted speed, vehicle speeds, road condition, road surface type, grade, alignment, weather condition, seat belt usage, drug usage, and alcohol usage. These were analyzed for each highway system.

In order to manage the many categorical predictors for the statistical analysis, all variables were set up as binary (1 or 0) whether the factor was involved in the crash or not (Pei and Fu, 2014). Vehicle type was identified whether trucks were involved or not and was initial separated into three categories for trucks and all other vehicles. Number of vehicles in the crash was handled by assigning the value 1 if more than one vehicle was involved, and 0 if only one vehicle was involved. Road geometrics was condensed to grade (level or not level) and alignment (straight or curve). This was to avoid problems with multicollinearity (Kutner et al., 2004). The first harmful event (FHE) underwent the largest reduction in categories. The crash report contains over 60 variables for FHE. These were reduced to five main categories which include animal, rollover, collision with another vehicle, fixed object, and guardrail. The posted speed limit and the reported mean vehicle speed were used to indicate whether above or below the speed limit. Road condition was identified as dry or not dry, weather as clear or not clear, and road surface type as paved or not paved. The variables used in the descriptive analysis can be observed in Table 1.

The severity level of a crash is dictated by the worst injury that was incurred by all passengers in the crash. Only one severity level is assigned to each crash, so the percentage is taken from the total number of crashes (160,613). Of the 160,613 crashes, heavy trucks were involved in 13,273 crashes, which is 8.3 percent of the total crashes, but they only make up 6.0 percent of the total vehicle count on all roadways types in Wyoming. Truck involvement refers to when there is at least one heavy truck weighing more than 26,000 pounds involved in a crash. Cars are involved in 58.5 percent of the crashes, and make up 47.3 percent of the vehicles. Over 50 percent of crashes are multiple vehicle crashes. Single vehicle crashes account for 46.2 percent of the crashes reported.

Table 2 shows the percentage of crashes broken down by each severity level, and separated by whether or not one or more of the vehicles involved in the crash was a heavy truck. The percentages are slightly higher for more severe crashes when trucks are involved than when no truck is involved. The percentage of fatal crashes is almost twice as high when a truck is involved. The usage of safety restraints was evaluated and found that truck drivers were using their safety restraints 83.0 percent of the time. As seen in Table 2, alcohol was involved 2.1 percent of the time, which is less than the percentage of crashes where no heavy truck was involved at seven percent. Drugs were involved in 0.3 percent of the crashes, which is less than half of that for crashing not involving a truck at 0.7 percent.

The road classifications were grouped into categories based on their ownership and maintenance. Most crashes took place on primary federal highways and interstates. This could be due to large traffic volumes and a high number of road miles, which is typical for these types of roads. Sixty-eight percent of heavy trucks were involved in crashes on interstates. This is approximately double any of the other vehicle types on any of the other road types.

When comparing the posted speeds where the crashes occurred, 39.1 percent of heavy truck crashes occurred where the speed was posted at 75 mph. The second most occurred on roads where the posted speed was 30 mph at 19.3 percent of the truck crashes. Weather and road conditions at the time of the crashes were also analyzed. Weather can affect visibility and roadway condition. Most crashes occurred

Table 1
Summary of Variables.

Variables	Description	Coded Response	Interstate		State Highways	
			Count	% of Total Crashes	Count	% of Total Crashes
Animal (FHE)	First Harmful Event (FHE) was striking an Animal	0 = No Animal 1 = Animal	5406	16.34%	15415	23.61%
Rollover (FHE)	First Harmful Event (FHE) was a Rollover	0 = No Rollover 1 = Rollover	7663	23.16%	8405	12.87%
Guardrail (FHE)	First Harmful Event (FHE) was striking a Guardrail	0 = No Guardrail 1 = Guardrail	3792	11.46%	1287	1.97%
Fixed Object (FHE)	First Harmful Event (FHE) was striking a Fixed Object	0 = No Fixed Object 1 = Fixed Object	8474	25.61%	9180	14.06%
Location (FHE)	First Harmful Event (FHE) Location	0 = On Roadway 1 = Off Roadway	18600	56.20%	45906	70.32%
Multi-Vehicles	Number of Vehicles involved in a crash	0 = Single Vehicle 1 = Multiple Vehicles	7619	23.03%	30012	44.78%
Impaired	Driver impairment	0 = Not Impaired 1 = Impaired	24	0.07%	92	0.14%
Road Conditions	Road Surface Condition	0 = Dry 1 = Wet, snow, etc.	17832	53.90%	17988	26.84%
Weather	Weather Conditions	0 = Clear 1 = Fog, rain, etc.	13951	42.17%	11956	17.84%
Speeding	Speed for a crash is higher or lower than the Posted Speed	0 ≤ Posted Speed Limit 1 > Posted Speed Limit	1961	5.93%	5948	9.11%
Grade	Longitudinal Grade	0 = Level 1 = Not Level	8794	26.58%	31667	47.25%
Alignment	Horizontal Alignment	0 = Straight 1 = Curve	5131	15.51%	22432	33.47%
All Trucks	Truck involvement	0 = No Trucks 1 = Truck(s)	9040	27.32%	4601	7.05%
Motorcycle	Motorcycle involvement	0 = No Motorcycle 1 = Motorcycle(s)	455	1.37%	2121	3.25%
Median	Presence of Median	0 = Median 1 = No Median	6626	20.02%	237	0.36%

Table 2
Crashes and Truck Involvement.

Severity	Truck Involved		No Truck Involved	
	Crashes	Percent	Crashes	Percent
No Injury	9818	74.0%	107726	73.1%
Possible Injury	1161	8.8%	14269	9.7%
Non-Incapacitating Injury	1314	9.9%	14565	9.9%
Incapacitating Injury	709	5.3%	6093	4.1%
Fatal	205	1.5%	1191	0.8%
Unknown	66	0.5%	3496	2.4%
Total Crashes	13273	100%	147340	100%
Alcohol Involved				
Yes	280	2.1%	10387	7.0%
No	12992	97.9%	127185	86.3%
No Answer	1	0.0%	9768	6.6%
Total Crashes	13273	100%	147340	100%
Drugs Involved				
Yes	46	0.3%	1069	0.7%
No	4961	37.4%	46890	31.8%
No Answer	8266	62.3%	99381	67.5%
Total Crashes	13273	100%	147340	100%

when the weather was clear. However, snow was most common in crashes that occurred during inclement weather. Inclement weather can lead to less-than-optimal road conditions, which affect handling of the vehicle. In Fig. 1, the analysis shows that the next most common road condition to dry roads was ice and frost. Twenty-eight percent of truck involved crashes occurred on roads with ice and frost, compared to fifteen percent of non-truck crashes occurring on ice and frost covered roads.

6. Statistical analysis

The descriptive data analysis suggested certain trends when trucks are involved in a crash compared to when they are not. Statistical analysis was performed on the data to learn about the effects of trucks on crash severity. Logistic regression models with both maximum likelihood estimation and Bayesian inference were used to examine all aspects of the data and the effect of trucks on crash severity. Models were constructed for each type of road classification. Various models were estimated for single-vehicle crashes, multi-vehicles crashes, and all crashes. The models that combined both single and multiple vehicles crashes provided the best fit.

6.1. Bayesian logistic regression

The study utilized a Bayesian logistic regression approach to estimate the probability of a crash resulting in a severe injury conditional upon that a crash has occurred. Bayesian logistic regression has the formulation of a logistic equation and can handle both continuous and categorical explanatory variables. The classical logistic regression treats the parameters of the models as fixed, unknown constants and the data is used solely to best estimate the unknown values of the parameters. In the Bayesian approach, the parameters are treated as random variables, and the data is used to update beliefs about the behavior of the parameters to assess their distributional properties. The interpretation of Bayesian inference is slightly different than the classical statistics; the Bayesian derives updated posterior probability of the parameters and construct credibility intervals that have a natural interpretation in terms of probabilities. Moreover, Bayesian inference can effectively avoid the problem of over fitting that occurs when the number of observations is limited and the number of variables is large.

The Bayesian logistic regression models the relationship between the dichotomy response variable (severe/non-severe) and the

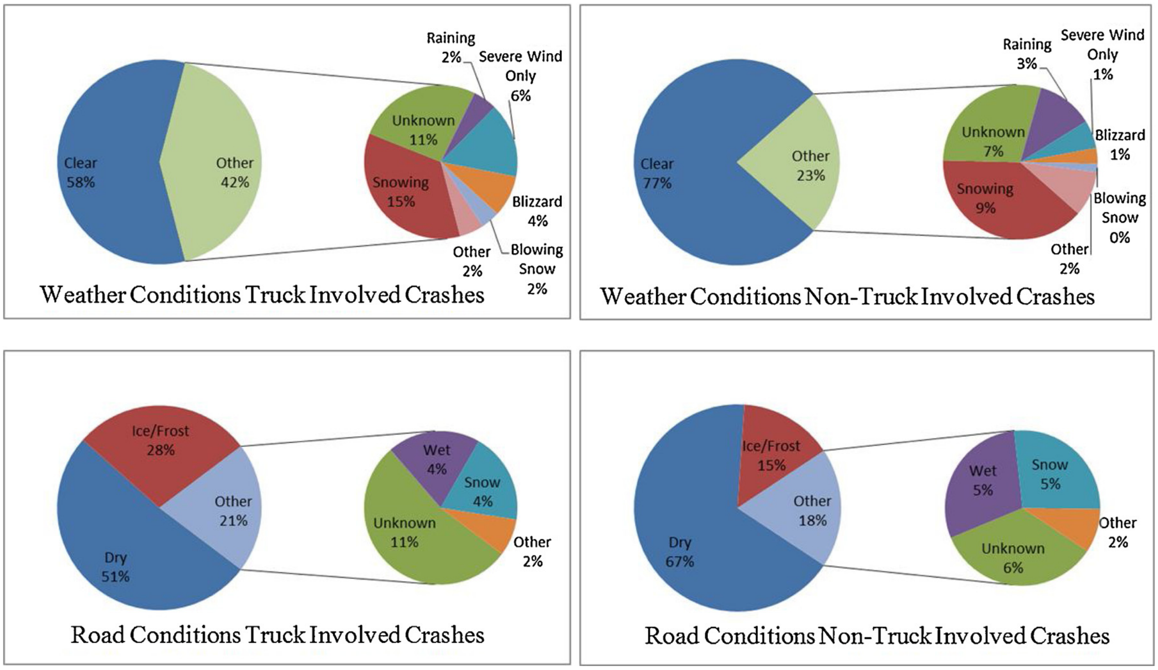

Fig. 1. Weather and Road Conditions during Crashes.

explanatory variables of roadway geometry, traffic, and weather. Suppose that the response variable y has the outcomes y = 1 or y = 0 with respective probability *p* and 1-*p*. The logistic regression equation can be expressed as:

$$log\left(\frac{p}{1-p}\right) = \beta_0 + \beta X \tag{1}$$

Where β_0 is the intercept, β is the vector of coefficients for the explanatory variables, and X is the vector of the explanatory variables. The logit function relates the explanatory variables to the probability of an outcome y = 1. The expected probability that y = 1 for a given value of the vector of explanatory variables X can be theoretically calculated as:

$$p(y = 1) = \frac{exp(\beta_0 + \beta X)}{1 + exp(\beta_0 + \beta X)} = \frac{e^{\beta_0+\beta X}}{1 + e^{\beta_0+\beta X}} \tag{2}$$

One advantage of the Bayesian approach over the classical model is the applicability of choosing the parametric family for prior probability distributions. There are three different priors that can be used; 1) informative prior distributions based on the literature, experts' knowledge or explicitly from an earlier data analysis, 2) weak informative priors that do not supply any controversial information but are strong enough to pull the data away from inappropriate inferences, or 3) uniform priors or non-informative priors that basically allow the information from the likelihood to be interpreted probabilistically. In this study, uniform priors following normal distribution with initial values for the estimation of each parameter from the maximum likelihood method was used. Different types of prior distributions using the results from this study as prior could be considered for further research once more data become available to update the estimated models.

All models were estimated by Bayesian inference using the freeware Winbugs (Lunn et al., 2000). For each model, three Monte Carlo Markov Chains (MCMC) that contains the approximate posterior distribution of samples by use of Gibbs sampling and the adaptive rejection of 10,000 iterations were set up in Winbugs based on the convergence speed and the magnitude of the dataset. After ensuring the convergence, first 1000 samples were discarded as adaptation and burn-in. The 95% Bayesian credible interval (95% BCI) was utilized to assess the significance of covariates. BCI provides probability interpretations with normality assumption on unknowns and confidence interval estimations (Gelman et al., 2003). The Deviance Information Criterion DIC, a Bayesian generalization of Akaike Information Criterion AIC, is used to measure the model complexity and fit. DIC is a combination of the deviance for the model and a penalty for the complexity of the model. The deviance is defined as $-2\log(likelihood)$. The effective number of parameters, pD, is used as a measure of the complexity of the model, $pD = Dbar{-}Dhat$, where $Dbar$ is the posterior mean of the deviance, and $Dhat$ is a point estimate of the deviance for the posterior mean of the parameters. DIC is given by DIC = Dhat + 2 pD (Spiegelhalter et al., 2002). Moreover, receiver operating characteristic (ROC) curve analysis was used to compare across models.

Specific crash variables were chosen for the model. First and foremost, truck involvement was tested. Light truck, medium truck, and heavy truck were each measured. A light truck is considered to weigh less than 10,000 pounds, a medium truck between 10,000 pounds and 26,000 pounds, and a heavy truck is more than 26,000 pounds (WYDOT, 2013). A stepwise variable selection method was utilized to find significant predictors of crash severity. The factors identified from that model were the variables used for this analysis, in addition to the truck involvement variables.

A Bayesian binary logit regression approach was used to model the response that the crash is severe or not severe. As mentioned earlier, a crash is considered severe if its outcome is a fatal or incapacitating injury. Severe crashes are coded with the value one (1). Not severe crashes consist of non-incapacitating injury, possible injury, and no injury and are given the value zero (0). Each value of a crash variable was also converted into a binary or dichotomous response of either one or zero (1 or 0). The All Trucks variable takes into account if any light, medium, or heavy truck was involved in the crash or not.

The purpose of the modeling is to identify variables which are significant predictors of crash severity. The variables that were used are listed in Table 1. The response variable is the function of a set of predictor variables (Kutner et al., 2004). A simple logistic regression analysis was performed first to examine the significance of truck involvement as a predictor by itself. This univariate analysis involved running the logistic regression model with only one predictor variable. The analysis was done separately for each roadway model with each weight of truck. A univariate analysis demonstrates if a variable is important by itself and its individual relationship with the response. Next, an analysis was conducted to assess the main effects. The main effects model determined predictor variables when interaction terms are not present. The main effects model was calibrated separately for each roadway model with each weight of truck. Once a main effects model was run, pairwise interaction terms were added to the model. Interaction terms occur when a variable has a different effect on the outcome, severity in this case, depending on the value of another variable. Each significant factor was paired with the truck variable and tested for significance. Insignificant interaction terms were removed from the model based on an iterative process to arrive at a final model.

Odds ratios were estimated based on variables that showed up consistently in the models. These ratios explain the probability of a severe crash happening when certain conditions are present versus when they are not present. The odds of a single variable are calculated as odds = $\frac{\pi}{1-\pi}$. For interaction terms, the odds of the two variables that make up the interaction must be taken into account. Odds ratios are computed using Eq. (3) shown below.

$$\hat{OR} = \left(\frac{\pi_1}{1-\pi_1}\right)\left(\frac{1-\pi_2}{\pi_2}\right) \tag{3}$$

From the odds ratios that are calculated from interactions, the odds of a severe crash in the presence of two predictors can be estimated.

7. Modeling results and discussions

Many approaches were attempted to model injury severity of truck crashes in Wyoming. Multi-vehicle that involve a truck, single truck crash models, and all crashes for interstate and state highways were estimated. It is worth mentioning that the estimated models for multi-vehicle crashes that involve a truck, and single truck crashes provided the least areas under the ROC curve and very few significant explanatory variables. Since the all vehicle crashes models for Interstate and State classifications provided better fit and more explanatory variables, this study focused on the main hypothetical question of determining the odds that a crash will be more severe having a truck involved in the crash.

Since the light trucks and medium trucks were shown to be insignificant in the preliminary analysis, only *all truck*s variable was retained in the final model. Logistic regression models with the Bayesian inference approach were used for each specific data and the estimated coefficients, credible intervals and model fits are shown in Table 2. The two final models show similar classification ability since the ROC areas were almost the same.

Some of the interactions that were significant on all road types were, *all trucks* with *multi-vehicle crashes,* and *all trucks* with *undesirable road conditions*. Since these terms showed to be significant, it prompts further investigation into their effects on crash severity.

Two final models have been presented in this paper; state, and interstate models. The first model (state) included the road classifications of state highway, primary federal highway, and secondary federal highway, while the second model is for only interstates. For the state model, heavy trucks and all trucks variables were shown to be significant in the univariate analysis. In the main effects model, no weight of truck was shown to be significant in the presence of other variables. This lack of significance did not rule out the chance that trucks would be involved in significant interaction terms. The final model showed that the presence of a truck in a crash is significant when combined with any of the following terms: multiple vehicles, inclement road conditions, and curvy alignment. For the interstate model, the all trucks variable (light, medium, or heavy truck) was significant in the univariate analysis, but no trucks showed up in the main effects model. In the final interaction model the significant factors were shown to be interactions of truck and guardrail (FHE), truck and fixed object (FHE), truck and number of vehicles, truck and impairment, truck and road condition, and truck and grade.

The multi-vehicle crashes has a positive sign for the Interstate roads and a negative sign for state roads, which may imply that the injury severity on interstate roads for multi-vehicle crashes is more than on the state roads. This can be explained by the fact that lane departure single-vehicle critical crashes are more common on state roads compared to interstate roads. Single-vehicle critical crashes occurred on state roads comprised about 72% of total crashes according to the Wyoming Strategic Highway Safety Plan (WSHSP), 2012. In Wyoming, interstate roads carry significant volume of truck traffic while state roads have drastically low volume traffic in general with less truck traffic. Moreover, interstate roads have higher design standards than state roads of having guardrail and better rumble strips. Previous studies (Holdridge et al., 2005) concluded that crashing into the leading end of guardrail increases the propensity toward fatal injury, also another study by (Shankar et al., 2000) indicated that some types of guardrail, i.e., thrie-beam is associated with an increase in the probabilities of non-severe crashes. A recent study by (Johnson and Gabler, 2014) concluded that the odds of injury in guardrail frontal crashes with the end terminals is 5.1 times greater than frontal crashes to the guardrail face. Moreover, they found that rollover crashes with guardrail end terminal were 6.7 times higher than if the collision was with the guardrail face. The interaction term for state roads was not significant due to the fact that heavy trucks are much less on state roads and guardrail is not widely used on state roads in Wyoming. Moreover, there is no distinction in Wyoming crash database between the different types of guardrail used in the state. The negative sign of the interaction between guardrail and all trucks variables in the *Interstate* model could be explained by the fact that based on the type of the guardrail, heavy trucks can survive striking a guardrail, while small cars might not. There is also no specific information in the crash database whether vehicles involved in crashes have strike the end terminal, or the face of the guardrail.

Crashes involving animals were not as harmful as other types of crashes, most likely since the animals receive most of the impact, this result is consistent with a study by Khattak et al. (Khattak et al., 2002). Similar to a study by the National Highway Traffic Safety Administration (NHTSA) (Liu et al., 2004), rollover crashes and crashes occurred off road were more injurious on the state roads. It is worth mentioning that run-off-road crashes is ranked the second crash type according to the LTCCS (Craft et al., 2007).

The odds ratios can be used to interpret the presence of a truck in a crash on the severity of a crash when multiple vehicles are involved. The odds ratio is interpreted as the odds for when one or both interaction terms are present and comparing it to the odds when they are not present. When a multi-vehicle crash occurs but there is no truck involved, the estimated odds that the crash will be severe is 1.8 times as likely as when the crash is single vehicle with no truck. However, when the crash involves multiple vehicles and there is a truck involved, the estimated odds of a severe crash increase to 2.3 times as likely. When an impaired driver gets into a collision where no truck is involved, the estimated odds of a severe crash are 2.7 times as likely as a non-impaired driver where no truck is involved. When there is a truck involved in a crash with an impaired driver, the odds almost double to being 5 times higher than with a non-impaired driver and no truck. And finally, on the interstate where a truck is involved in a multi-vehicle collision, the estimated odds are 4.5 times greater of a severe crash occurring than when a truck is not involved in a single vehicle crash (Table 3).

8. Conclusions

Although Wyoming is considered to be primarily a rural state, crashes have different causes and effects on the various road classifications. Of particular importance in this research is the effect of truck involvement on the severity of crashes.

A descriptive analysis was completed which investigated crashes for the entire state of Wyoming. In the 10 year study period (2002–2011),

Table 3
Parameters Estimates.

Model	State			Interstate		
		Credible interval			Credible interval	
	Mean	2.5%	97.5%	Mean	2.5%	97.5%
Intercept	−3.094	−3.406	−2.739	−4.153	−4.622	−3.830
Animal (FHE)	−1.676	−1.878	−1.483	−1.295	−1.719	−0.955
Rollover (FHE)	0.4651	0.324	0.605	1.497	1.305	1.679
Guardrail (FHE)	0.191*	−0.023	0.442	0.515	0.368	0.702
Fixed Object (FHE)	−0.512	−0.671	−0.355	0.204*	−0.030	0.427
Multi-vehicles	−0.103*	−0.256	0.043	0.467	0.226	0.706
On/ off road (FHE)	0.249	0.133	0.359	0.462	0.343	0.585
Impaired	1.316	1.199	1.432	1.000	0.796	1.205
Road Condition (wet/snow)	−0.824	−0.962	−0.699	−0.887	−1.018	−0.758
Weather (inclement)	0.054	−0.074	0.195	−0.176	−0.295	−0.049
Speeding	0.969	0.726	1.148	0.741	0.540	0.984
Grade (steep)	−0.424	−0.615	−0.274	−0.076	−0.201	0.055
Curved Alignment	0.708	0.514	0.899	0.256	0.148	0.363
All Trucks	−0.594	−0.860	−0.341	−0.231	−0.778	0.574
Motorcycle-involved crash	2.376	2.243	2.502	2.877	2.601	3.143
Median	–	–	–	−0.496	−0.845	−0.177
Speeding*Grade	0.398	0.213	0.621	–	–	–
Speeding*Alignment	−0.534	−0.750	−0.314	–	–	–
Guardrail*All Trucks	–	–	–	−0.620	−1.284	−0.130
Fixed Object*All Trucks	–	–	–	0.660	0.239	1.040
Vehicles*All Trucks	0.870	0.602	1.151	1.042	0.774	1.295
Impaired*All Trucks	–	–	–	0.646	0.150	1.086
Road Condition*All Trucks	0.440	0.096	0.743	0.650	0.437	0.859
Grade*All Trucks	–	–	–	−0.21*	−0.420	0.021
Alignment*All Trucks	0.326	0.034	0.608	–	–	–
pD: no of effective variables	19.770	–	–	22.478	–	–
DIC	19018	–	–	13,088	–	–
ROC	0.785	–	–	0.789	–	–

* Variables significant at 90% Credible Interval.

fatal crashes made up less than 1 percent of the total crashes. However, truck involved crashes had 1.54 percent fatalities. At 68 percent, heavy trucks had a higher observed percentage of vehicles that were involved in crashes that occurred on interstates. Truck involved crashes occurred 19 percent more often when there is inclement weather such as severe wind or snow than no truck-involved crashes. In addition, when the road conditions were not dry, truck involved crashes occurred more often when there was ice or frost. Safety belt use in trucks was higher than that of drivers of other vehicles. Alcohol and drug involvement in truck drivers were both lower than that of drivers of other vehicles as well.

An extensive analysis has been conducted in this study to explore if the involvement of a heavy truck in a crash is a significant predictor of crash severity, and to identify roadway and weather factors affecting crash severity. The analysis was also performed at various roadway classifications. A Bayesian Binary Logit Models were calibrated to verify results from the statewide data analysis and to further study the relationship between truck involvement and crash severity. Bayesian approach accounts for the uncertainty associated with parameter estimates and provide exact measures of uncertainty on the posterior distributions of these parameters and hence overcome the maximum likelihood methods' problem of overestimating precision because of ignoring this uncertainty. From the final main effects model, where truck interactions were introduced into the model, various interactions proved to be significant in predicting the severity of crashes. The interstate system model showed that trucks involved with multi-vehicle crashes were more significant in estimating severity. In interstate models, the effect of truck crashes where striking the guardrail, impairment, or multiple vehicles was higher. Moreover, the involvement of trucks in crashes results in higher severity. The odds of a severe crash increases to 2.3 and 4.5 times when a heavy truck is involved on state and interstate highways respectively.

9. Recommendations

Based on the analysis, when weather and road conditions are not clear and dry, trucks have a higher percentage of crashes than other vehicles (28% truck involved crashes occurred during adverse weather condition compared to only 15% of non-truck crashes occurring in the same conditions). Improvements such as enhanced advanced warning systems could provide vital information to the truckers about the existing conditions. WYDOT is currently in the process of implementing such a program where monitors will be installed at truck stops, restaurants and gas stations that will provide real time visual information on the existing weather and road conditions.

The analysis showed that there is high compliance with seatbelt use and low drug and alcohol use. This could be due to the policies and training programs conducted by the trucking industry. Their training programs could be expanded to include the risks involved in driving in inclement weather. Providing advanced warning systems and targeted training, roadway safety could potentially be realized and fatal crashes involving trucks could be reduced. The higher serious and fatal crash rates of truck related crashes are not unique to Wyoming and thus these improvements could be implemented to reduce these types of crashes throughout the country.

Acknowlegements

The WYT^2/LTAP center at the University of Wyoming provided extensive resources to assist in the compilation of the data sets used and the development of the models. WYDOT was extremely helpful and responsive to provide the needed bulk crash data, inventories and traffic data. Special acknowledgement goes to WYDOT and FHWA that provided the resources to make this research possible.

References

Abdelwahab, H., Abdel-Aty, M., 2001. Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections. Transp. Res. Rec. (1746), 6–13.

Andrea, B., Francisco, S., Noyce, D., 2011. Wisconsin Large Truck Safety and Enforcement Study. University of Wisconsin, Traffic Operations and Safety (TOPS) Laboratory. Wisconsin Department of Transportation, Madison.

ATA, 2013. U.S. Freight Transportation Forecast to 2024. American Trucking Associations, Arlington.

Blower, D., 1998. The Relative Contribution of Truck Drivers and Passenger Vehicle Drivers to Truck-Passenger Vehicle Traffic Crashes. University of Michigan. Ann Arbor: Federal Highway Administration.

Bureau of Labor Statistics, 2006. Census of Fatal Occupational Injuries. US Department of Labor.

Bureau of Labor Statistics, 2008. Census of Fatal Occupational Injuries (2003–2006). US Department of Labor.

Bureau of Labor Statistics, 2013. Census of Fatal Occupational Injuries Charts, 1992–2012. US Department of Labor.

Carsten, O., 1987. Road Class and Large Truck Involvements in Fatal Accidents. The University of Michigan Transportation Research Institute, Ann Arbor.

Chang, L.W., 2006. Analysis of traffic injury severity: an application of nonparametric classification tree techniques. Accid. Anal. Prev. 38 (5), 1019–1027.

Craft, R., 2007. The Large Truck Crash Causation Study, Analysis Brief. Federal Motor Carrier Safety Administration, Washington, DC.

Helai, H., Chor, C., Haque, M., 2008. Severity of driver injury and vehicle damage in traffic crashes at intersections: a bayesian hierarchical analysis. Accid. Anal. Prev. 40 (1), 45–54.

Holdridge, J.M., Shankar, V.N., Ulfarsson, G.F., 2005. The crash severity impacts of fixed roadside objects. J. Saf. Res. 139–147.

Gelman, A., Carlin, J.B., Stern, H.S., 2003. Bayesian Data Analysis, 2nd ed. Chapman & Hall, New York.

Jarossi, L., Matteson, A., Woodrooffe, J., 2011. Trucks Involved in Fatal Accidents Factbook 2008. Center for National Truck and Bus Statistics, Ann Arbor.

Johnson, N.S., Gabler, H.C., 2014. Injury risk in frontal crashes with guardrail and guardrail end terminals, Washington D.C. Transportation Research Board Annual Meeting.

Khattak, A., Pawlovich, M., Souleyrette, R., Hallmark, S., 2002. Factors related to more severe older driver traffic crash injuries. J. Transp. Eng 243-149.

Knipling, R., 2009. Three large truck crash causation categories: what they tell us about crash causation. Driving symposium on human factors in driver assessment, training and vehicle design. Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.

Kubas, A., Vachal, K., 2012. Oil County Traffic Safety Survey, 2012. Upper Great Plains Transportation Institute. Fargo: North Dakota State University.

Kutner, M., Nachtsheim, C., Neter, J., 2004. Applied Linear Regression Models, 4th ed. New York, McGraw-Hill.

Lemp, J., Kockelman, K., Unnikrishnan, A., 2011. Analysis of large truck crash severity using heteroskedastic ordered probit models. Accid. Anal. Prev. 43 (1), 370–380.

Li, Z., Liu, P.W., Xu, C., 2012. Using support vector machine models for crash injury severity analysis. Accid. Anal. Prev. 45, 478–486.

Liu, C., Utter, D., Chen, C.-L., 2004. Characteristics of Crash Injuries Among Young, Middle-Aged, and Older Drivers. National Highway Traffic Safety Administration, Washington, DC.

Lueck, M., Murray, D., 2011. Predicting Truck Crash Involvement: A 2011 Update. Arlington, VA, ATRI.

Lunn, D., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS – a bayesian modelling framework: concepts, structure, and extensibility. Stat. Comput. 10, 325–337.

Mode, N., Conway, G., 2008. Fatalities Among Oil and Gas Extraction Workers - United States, 2003–2006. National Institute for Occupational Safety and Health, Atlanta.

NHTSA, 2011. Traffic Safety Facts, 2009 Data- Large Trucks. U.S. Department of Transportation, Washington, DC.

NHTSA, 2014. Traffic Safety Facts, 2012 Data- Large Trucks. U.S. Department of Transportation, Washington, DC.

Savolainen, P., Mannering, F., Lord, D., Quddus, M., 2011. The statistical analysis of highway crash-injury severities. Accid. Anal. Prev. 43 (5), 370–380.

Shankar, V.N., Albin, R.B., Milton, J.C., Nebergall, M., 2000. In-Service, performance-based roadside design policy: preliminary insights from Washington state's bridge rail study. Transp. Res. Rec. 1720, 72–79.

Spiegelhalter, D., Best, N., Carlin, B., van der Linde, A., 2002. Bayesian measures of model complexity and fit (with discussion). J. R. Stat. Soc. Ser. B (Stat. Methodol.) 64 (4), 583–639.

Weber, A., Murray, D., 2014. Evaluating the Impact of Commercial Motor Vehicle Enforcement Disparities on Carrier Safety Performance. American Transportation Research Institute.

WYDOT, 2013. Wyoming's 2012 Report on Traffic Crashes. WYDOT, Cheyenne.

Pei, Yulong, Fu, Chuanyun, 2014. Investigating crash injury severity at unsignalized intersections in Heilongjiang Province, China. J. Traffic Transp. Eng. (English Edition) 1 (4), 272–279. http://dx.doi.org/10.1016/S2095-7564(15)30272-5. ISSN 2095-7564.

Zaloshnja, E., Miller, T., 2007. Unit Costs of Medium and Heavy Truck Crashes. U.S. Department of Transportation, Washington, DC.