Developing a Truck Corridor Crash Severity Index (CSI)

Xiao Qin, Ph.D., P.E.
(corresponding author)
Assistant Professor
Department of Civil and Environmental Engineering, South Dakota State University
CEH 148, Box 2219, Brookings, SD, 57007
Phone: (605) 688-6355
Fax: (605) 688-6476
Xiao.qin@sdstate.edu

Most Afia Sultana
Research Assistant
Department of Civil & Environmental Engineering
South Dakota State University
Brookings SD 57006
Fax: (605) 655-6474
Email: most.sultana@jacks.sdstate.edu

Madhav Chitturi, Ph.D.
Research Associate
Traffic Operations and Safety (TOPS) Laboratory, University of Wisconsin-Madison
Department of Civil and Environmental Engineering
B243 Engineering Hall
1415 Engineering Drive
Madison, WI 53706
Phone: (608) 890-2439
mchitturi@wisc.edu

David A. Noyce, Ph.D., P.E.
Professor
Director, Traffic Operations and Safety (TOPS) Laboratory, University of Wisconsin-Madison
Department of Civil and Environmental Engineering
1204 Engineering Hall
1415 Engineering Drive
Madison, WI 53706
Phone: (608) 265-1882
noyce@engr.wisc.edu

Word count: 5823+7 (6 Tables and 1 Figure)*250=7573

Submitted for Presentation and Publication to the
92nd Annual meeting of the Transportation Research Board
January 13-17, 2013
Washington, DC

ABSTRACT

According to the National Highway Traffic Safety Administration (NHTSA), over 400,000 truck accidents occurred in 2009 with approximately 7,800 of those are fatal crashes. Compared to extensive studies conducted on freeway truck safety, the research on arterial streets is considerably disproportionate. Making the connections between truck traffic generators, arterial streets are key links in door-to-door deliveries. There is an urgent need to study truck safety on arterial streets because of the strong growth of truck traffic.

Truck related crashes are expected to be reduced through the careful planning of the location, design, and operation of driveways, median openings, street connections and street sections. By collecting extensive data on selected arterial corridors that are heavily used by trucks, truck crash frequency and severity contributing factors have been identified using negative binomial model and multinomial logit (MNL) model, respectively. Corridor truck miles traveled, AADT, signal density, shoulder width, PSI and its standard deviation are significant factors for the crash frequency prediction. MNL identified twelve causal factors for crash severity such as posted speed limit, lane width, number of lanes, pavement condition index, undivided roadway portion and so on. Subsequently, a crash severity index (CSI) for the truck arterial corridors was developed. The findings from the study will not only benefit state and local agencies in planning, design, and manage a safer truck arterial corridor, but also help carriers to optimize their routes from the safety perspective.

INTRODUCTION

Freight transportation is extremely critical to the economic development of a nation. The United States economy depends on trucks to deliver nearly 70 % of all freight transported annually, accounting for $671 billion worth of manufactured and retail goods in the U.S. along with $295 billion in trade with Canada and $195.6 billion in trade with Mexico (1). Trucking revenues totaled $610 billion in 2011, and revenues are estimated to nearly double by 2015 (2). While the rapid commercial trucking growth is great news for the country's economy, the increasing truck traffic may negatively impact cars, vans, SUVs and other vehicles that share the road. In 2010, large trucks accounted for 4 percent of all registered vehicles and 10 percent of the total vehicle miles traveled. Of the fatalities in crashes involving large trucks during 2010, 76 percent were occupants of other vehicles (3). In fact, one person is injured or killed in a truck accident every 16 minutes and one out of every eight traffic fatalities involves a trucking collision (2). The National Highway Traffic Safety Administration (NHTSA) has estimated that over 400,000 truck accidents occurred in 2009 with approximately 7,800 of those are fatal crashes (4). Therefore, it is urgent to improve truck safety and reduce truck-related crashes.

Extensive research has been conducted on site-specific characteristics and their effects on truck crashes, either at intersections or on segments (5-12). Moreover, truck safety on freeways and interstate highways has usually been a focus of research because of the high speed and high truck percentage (8-17). Studies have shown that full access controlled roads have a safer traffic record, accounting for only 24 percent of crashes, while the remainder occurs on arterial or local roadways (7). In contrast, limited research has been conducted on arterial streets, especially from a corridor perspective. Arterial streets connect freeway corridors to the distributors, carriers, vendors, and customers. They are the "last miles" for commercial motor vehicles to deliver the freight to destinations or enter the interstate highway system. Analyzing safety from an arterial corridor perspective is important as there are more opportunities for conflicts with passenger vehicles at signalized intersections and it is valuable for developing system-wide, corridor-based, and more importantly proactive safety improvement strategies.

While emphasizing highway safety, the safety risk index is an effective measure for proactively identifying and analyzing safety issues. More concisely, the safety risk index is a measure by which the transport personnel can quantify the hazards associated with particular roadway characteristics, environmental patterns, and driver population. A quantifiable risk index associated with a roadway segment will help transportation agencies to identify potential safety problems and adopt appropriate remedies prior to a crash occurrence thereby reducing the risk exposure to other road users. Previously, many agencies have taken a reactive approach to safety, only responding to requests for safety improvements or relying heavily on the historic crash statistics. Recently, more agencies have committed to utilizing a more proactive safety management approach that would identify high risk roadway features or high risk locations in the context of a roadway network and implement effective low-cost improvements whenever appropriate. The newly published Highway Safety Manual (HSM) by the American Associations of State Highway and Transportation Officials' (AASHTO) has substantially accelerated the deployment of the proactive safety analysis approach. The HSM recommends the use of the relative severity index (RSI), which is the predicted average crash costs for a site, as the performance measure for the network screen (18). Therefore, the objective of this research is to investigate the relationship between highway and traffic engineering characteristics and truck crashes from a collection of arterial corridors with the purpose of developing a truck arterial corridor crash severity index (CSI) as a holistic measurement of truck crash risk.

LITERATURE REVIEW

There are many factors that may be involved in truck crashes. The Large Truck Crash Causation Study (LTCCS) identified human factors (an action or inaction by the drivers) and vehicle malfunctions (break problems) as the two leading causes. Roadway problems were present in 16 percent of the two-vehicle cases based on the 967 crashes involving 1,127 large truck and 959 non-truck motor vehicles (19). A prime interest to transportation agencies, the impacts of roadway geometric features on truck crashes has attracted considerable attention from many researchers. Extensive studies have focused on identifying roadway geometric features, traffic operational and pavement characteristics that contribute to truck crashes (5-14, 17). Looking beyond highway geometric data, Wang et al. developed multi-level estimation models by using freeway traffic data (flow, ramp volume, and shoulder width), economic activity data (shipment, county unemployment rate, income) and safety performance data to identify any contributing factors that may increase crash rates (8). They found that factors such as the number of shipments, county unemployment rate, truck and ramp AADT, and lane width significantly affect the number of truck crashes.

Many of the preceding studies were based on either individual intersections or segments, while few studies approached truck safety issues from a corridor perspective (20-23). Sayed and El-Basyouny assessed the corridor effects with alternate specifications (20). They compared the traditional Poisson Log Normal(PLN) model with two extended PLN models using a data set from 392 urban arterials in the city of Vancouver, BC, that were clustered into 58 corridors. The results of their paper provided some strong evidence of the benefit of clustering road segments into rather homogeneous groups (e.g., corridors) and incorporating random corridor parameters in accident prediction models. Research performed by Lee et al. examined factors that affected urban divided arterial road mid-block crashes on a 5.3-km section of urban arterial (21). The authors concluded that the number of access points on urban arterial roadways should be reduced to minimize the number of mid-block crashes. Abdel-Aty and Wang emphasized the fact that signalized intersections within a corridor have a correlated influence on the occurrence of crashes if the intersections are placed closely together (22). To account for the correlated data problem they used generalized estimating equations (GEE) with a negative binomial link function. Milton et al. used corridor specific and weather related variables to predict injury severity proportions using a mixed logistic model (23). Within these results, the average daily traffic (ADT), snowfall, truck average daily traffic, truck percentage, and the number of interchanges per mile were found to be statistically significant random variables for predicting different levels of injury severity. Whereas, the pavement friction, horizontal curvature per mile, and number of grade breaks per mile has fixed effect across all injury levels. These studies demonstrate the importance of corridor effects or corridor-level variables on crash occurrence and injury severities.

The proved relationship between crash frequency, severity and any contributory factors can be applied in a proactive safety analysis. De Leur and Sayed worked on the development of a systematic framework for proactive road safety planning in which they assumed road risk was a function of exposure, collision probability of a vehicle and consequence of a potential collision (24). They also provided some planning recommendations regarding land use shape, road network shape, geometric design elements, roadway functionality and friction, speed at crash prone areas, and road side environment in an effort to improve the safety of a roadway segment. In addition to the planning recommendation for safety improvements, the results of the statistical models of accident frequencies and injury severities can be used to present a road safety risk index. De Leur and Sayed developed two types of road safety risk index, $RSRI_{specific}$ and $RSRI_{combined,}$ based on the risk score of a particular road feature (25). $RSRI_{specific}$ defines the risk associated with each road feature, obtained by combining the scores for the three components of risk, while $RSRI_{combined}$ defines overall risk by combining the $RSRI_{specific}$ scores for all road features. In a recent study, Wu and Zhang proposed a framework for developing a composite Road Risk Index using a logistic function based on exposure, crash rate and crash severity (26). They showed risk index as a function of a predicted number of different crash types multiplied by a relative level of cost due to a particular type of crash using the HSM crash severity distribution and associated crash unit costs. In the HSM network screening process, a site-specific relative severity index (RSI) is calculated by multiplying the observed or predicted average crash frequency for each crash severity with their respective comprehensive crash cost and an average RSI is then obtained by dividing the overall RSI by the total number of observed crashes that occurred at the site (18). Regardless of the differences in the methods examined, they can provide valuable clues for informed decision-making.

METHODOLOGY

This section contains the theoretical concepts and mathematical equations necessary for the development of the truck arterial corridor CSI. Methodologies of predictive methods for crash frequency and crash severity distribution were discussed.

Crash Severity Index (CSI)

Truck corridor CSI was measured by the annual societal economic costs due to truck crashes which occurred along the specific corridor measured by unit length. Expected annual number of truck crashes as well as the proportion of crash by severity can be estimated via corridor geometric characteristics and traffic conditions. Combining annual crash frequency, severities, unit crash cost, and corridor length, the truck arterial corridor CSI is formulated in Equation 1.

$$CSI_i = \frac{\sum_{j=1}^{J} N_i P_j^i U_j}{L_i} \tag{1}$$

where:
CSI_i is the crash severity index for truck corridor i,
N_i is the annual expected number of truck crashes occurred along corridor i,
P_j is the proportion of crash severity j with j=1, J for corridor i,
U_j is the unit crash cost for severity j and
L_i is the length of corridor i.

For any truck corridor under consideration, the CSI value can be estimated using the corridor characteristics and applied either as the ranking tool for the truck safety performance or a proactive method for truck safety planning.

Modeling Methods for Crash Frequency

Count-data modeling (Poisson, negative binomial) techniques are widely using for crash frequency as the number of accidents n_i on roadway segment per unit of time is a non-negative integer. When the variance is larger than the mean, the data are said to be over dispersed. Over dispersed count data are usually modeled with a negative binomial distribution because the Poisson distribution has a restrictive assumption of equal variance and mean. In a Poisson model, the probability of the number of truck crashes for corridor i, n_i is as follows:

$$P(n_i) = \frac{\exp(-\lambda_i)\lambda_i^{n_i}}{n_i!} \tag{2}$$

where $P(n_i)$ is the probability of a corridor i having n_i crashes and λ_i is the expected number of crashes in corridor i. The negative binomial model is an extension of the Poisson where the Poisson parameter λ follows a gamma probability distribution. The standard log link function for the negative binomial model can be expressed as a linear model of the covariates in Equation 3.

$$\lambda_i = \exp(\beta_{0i} + \beta_1 x_{1i} + \cdots + \beta_k x_{ki})\exp(\varepsilon_i) \tag{3}$$

where βs are coefficients of explanatory variables and $\exp(\varepsilon_i)$ is the term adjusting for over-dispersion and is gamma distributed. The models were estimated by using generalized linear modeling. For this modeling, the SAS GEMOD procedure was used (27).

Modeling Methods for Crash Severity

Ordered Probit (OP) Model

The consequence of a crash can be modeled as a discrete outcome. An extensive and detailed review of the discrete choice probabilistic models and their applications in predicting crash severities is discussed by Savolainen et al. (28). It has been accepted by many researchers that there is an ordinal nature to crash severities, i.e. injury severity can be ranked from high to low as fatal injury (K), incapacitating injury (A), non-incapacitating injury (B), possible injury (C), and property-damage-only (O). To model injury severities as the ordinal response, researchers most frequently used discrete choice models such as ordered Probit (OP) models (28). An OP model is a special case of the Probit model where more than two outcomes of an ordinal dependent variable is modeled, usually estimated using maximum likelihood. The underlying relationship to be characterized is as Equation 4.

$$y^* = \mathbf{X}'\boldsymbol{\beta} + \varepsilon \tag{4}$$

where y* is the exact but unobserved dependent variable; **X** is the vector of independent variables, and **β** is the vector of regression coefficients which needs to be estimated. The ε is a random error term and assumed to follow a standard normal distribution. Furthermore y* cannot be observed, instead the categories of response can only be observed, as expressed in Equation 5.

$$y = \begin{cases} 1 & if\, y^* \le 0 \\ 2 & if\, 0 < y^* \le \mu \\ 3 & if\, \mu < y^* \end{cases} \tag{5}$$

μ represents thresholds to be estimated along with the parameter vector **β**.

Multinomial Logistic (MNL) Model

When modeling crash severities as an ordinal dependent variable, some restrictions can potentially affect the estimated results (28). The primary concern is the manner in which the explanatory variables affect the probabilities of the discrete outcome, i.e. the shift in the cutoff thresholds is constrained to move in the same direction. On the other hand, non-ordinal probabilistic models, such as multinomial logit (MNL) models, allow variables to have opposite effects regardless of the order of the injury severities. MNL model is a regression model which generalizes logistic regression by allowing more than two discrete outcomes. MNL relies on the assumption of independence of irrelevant alternatives (IIA), i.e. the odds of preferring one class over another do not depend on the presence or absence of other "irrelevant" alternatives. The mathematical model underlying MNL is to construct a linear predictor function that constructs the relationship between outcomes from a set of weights that are linearly combined with the explanatory variables of a given observation:

$$\mathrm{U_{ij}} = \mathbf{X_i'}\boldsymbol{\beta}_j + \varepsilon_{\mathrm{ij}} \tag{6}$$

where $\mathbf{X_i}$ is the vector of explanatory variables describing observation i, $\boldsymbol{\beta}_\mathbf{j}$ is a vector of weights (or regression coefficients) corresponding to outcome j, and U_{ij} is the utility associated with assigning observation i to get category j. The ε_{ij} is an error term that accounts for the random noise and assumed to be independently and identically distributed with a Gumbel extreme value distribution, and its logistic formulation is given by:

$$P_i(j) = \frac{EXP\left[\boldsymbol{\beta}_j' \boldsymbol{X}_i\right]}{1+\sum_{j=1}^{K-1} EXP\left[\boldsymbol{\beta}_j' X_i\right]} \quad \text{for j} = 1,\ldots,\mathrm{K}-1 \tag{7}$$

In a multinomial logit model, for K possible outcomes, running (K-1) independent binary logistic regression models, in which one outcome is chosen as a "pivot" and then the other (K-1) outcomes are separately regressed against the pivot outcome. If the last outcome K is chosen as the pivot, the estimated coefficients are usually presented as a log odds ratio between the probability of a given category and the reference one, resulting in (K-1) estimates for each independent variable if the response variable has K levels, as specified in Equation 8.

$$\log\left[\frac{P_i(j)}{P_i(K)}\right] = \boldsymbol{\beta}_\mathbf{j}\mathbf{X_i} \quad \text{for j} = 1,..\mathrm{K}-1 \tag{8}$$

Note that $\boldsymbol{\beta}_\mathbf{j}$ is a vector of estimable parameters representing the log odds ratio between the probabilities of two alternatives.

In a similar attempt, Geedipally et al. applied MNL models for estimating the proportion of crashes by collision type and then multiplied by the total number of crashes estimated with a total crash model to obtain the crash counts for each crash type at a site (29). They concluded that it is a promising method based on comparisons with the fixed proportion method and the method of developing respective collision type models.

DATA COLLECTION AND PROCESSING

The data used in this research consisted of five years (2005 to 2009) of crash counts, and geometric, pavement, and traffic volume data. Truck crashes were retrieved from the online Wisconsin crash database through the WisTransportal System (30). In order to undertake the investigation of truck crashes from a corridor perspective based on arterial roads, the truck corridor selection was confined to principal arterials and minor arterials. Recognizing the challenge of short (less than 1 mile) or very short segments (less than 0.1 mile) in the dataset, it was necessary to collapse short segments into longer ones so that it can be treated as a corridor. This was done by using collapsing criteria to dissolve adjacent roadway segments with similar or same annual average daily truck traffic (AATT). After a sensitivity analysis to specify a reasonable corridor length, it was determined to collapse adjacent segments having AATT differences within the range of 100 trucks per day. Next, three more criteria were applied to identify the beginning and end of the study corridors: 1) threshold of the corridor length is no less than one mile, 2) threshold value of truck annual average daily traffic 800 or more, and 3) study segment must be within five miles of an Interstate highway or a freeway. This resulted in 100 corridors containing 720 smaller segments. The descriptive statistics for key variables used in the crash frequency and severity models can be seen in Table 1.

During this five year period, 8,196 truck related crashes occurred in selected corridors, notably more than 50% of the crashes occurred in the South-East region and near the Milwaukee area where most truck activities occur. There was a decreasing trend of crashes over the five year period with 2009 showing the lowest number of crashes. Among these truck crashes 66% were property damage only (O); 21% were possible injuries (C); 9% were non-incapacitating injuries (B); 3% were incapacitating injuries (A); and 1% were fatal injuries (K). From the results of single and multiple vehicle crashes that were studied, 88% of the crashes were multi-vehicle crashes.

Corridor-level variables were created for each of the 100 corridors. As shown in Table 1, the total annual crash frequency had a mean of 82 and a standard deviation of 71, with a maximum of 407 crashes. The percentage of observations with more than 50 crashes within a corridor was found to be over 50%. Corridor lengths vary from relatively short (1.03 mi) to very long (16.94 mi) with an average segment length of 4.88 mi. The mean corridor AADT was 16,256 with a standard deviation of 6,107. Signal density and Access point density were calculated by the ratio of the number of signalized intersections and corridor lengths and the number of un-signalized intersections and corridor lengths. The maximum access point density of 30.47 exists in a 2.56 mile corridor where a total of 78 access points were counted, including 60 residential and commercial driveways and 18 other types of access points. The maximum speed of 60 mph identifies the corridor that contains a portion of a principal arterial with the 65mph posted speed limit. Similarly, the maximum lane width of 18 feet reflects a portion of a principal arterial corridor that has very wide lane width i.e. 22 feet. In addition, the proportion of corridor by the number of lanes, median presence, and speed limited were calculated. In particular, the corridor data was analyzed carefully for the good, fair, poor condition of roadways with less than or greater than 40mph horizontal curvature speed.

TABLE 1 Summary Statistics of Crash, Geometric and Traffic Variables for 100 Corridors

Variable	Description	Mean	STDV	Min	Max
Crash count	5 Year crash count for each corridor	82	71	14	407
Crash Severity					
	O	54	49	9	276
	C	17	16	0	84
	B	8	7	0	41
	A	3	3	0	11
	K	1	2	0	6
L	Length of the corridor (miles)	4.88	3.42	1.03	16.94
AADT	Annual average daily traffic	16256	6107	8172	39435
AATT	Annual average daily truck traffic	1077	211	800	1892
TRKPT	Truck percentage (%)	7.1	1.4	4.8	10.2
N_br	Number of Bridges	1.01	1.38	0	8
Sigden	Signal density (signals/mile)	0.51	0.87	0	4.33
Accden	Access point density (access points/mile)	5.29	4.81	0	30.47
SPD	Posted speed limited in mph	45	9	30	60
Lnwd	Lane width in feet	12.3	0.8	10	18
Mednwd	Median width in feet	14	12.9	0	47.3
Lshwd	Left shoulder width in feet	3.8	3.4	0	10.9
Rshwd	Right shoulder width in feet	5.6	4.2	0	15
Divund_U	Portion of undivided segments within a corridor	0.48	0.4	0	1
Divund_D	Portion of divided segments within a corridor	0.52	0.4	0	1
NL_1	Portion of segment with one lane	0.01	0.06	0	0.47
NL_2	Portion of segment with two lane	0.81	0.3	0	1
NL_3	Portion of segment with three lane	0.06	0.2	0	1
NL_4	Portion of segment with four lane	0.12	0.25	0	1
Hcl_g	Portion of segment with Horizontal curve speed less than 40mph_Good	0.95	0.19	0	1
Hcl_f	Portion of segment with Horizontal curve speed less than 40mph _Fair	0.03	0.17	0	1
Hcl_p	Portion of segment with Horizontal curve speed less than 40mph _Poor	0.01	0.07	0	0.43
Hcg_g	Portion of segment with Horizontal curve speed greater than 40mph_Good	0.89	0.29	0	1
Hcg_f	Portion of segment with Horizontal curve speed greater than 40mph _Fair	0.09	0.26	0	1
Hcg_p	Portion of segment with Horizontal curve speed greater than 40mph _Poor	0.02	0.09	0	0.59
PSI	Pavement Serviceability Index(0-5)	3.05	0.92	0.88	4.75
STD(PSI)	Standard deviation of PSI	0.58	0.42	0	1.98
IRI	International Roughness Index in mm	0.08	0.08	0	0.427
PCI	Pavement Condition Index (0-100)	77.09	24.35	0	100

RESULTS ANALYSIS & DISCUSSION

When traveling along an arterial corridor, truck drivers must adjust to design inconsistencies such as posted speed limits, signal timing, and geometric variations as well as heed the drivers of other motor vehicles to avoid any potential collisions. The expected number of truck crashes can be modeled as the product of traffic exposure and the truck crash rate, which may be a function of truck volume, AADT, and other factors. There is no fixed formula for measuring traffic exposure; different methods can be applicable depending on the way that segment length and traffic volume were specified (10, 31, 32). For example, Miaou (10) used AATT as an exposure variable and AADT as a surrogate variable to indicate traffic condition while modeling truck crashes. Whereas, Venkataraman (31) used AADT and the length of a segment as exposure variables in modeling Interstate crash occurrences. Using vehicle miles traveled (VMT), which is the product of segment length, AADT, and the number of days a year in the unit of million or 100 million, as the traffic exposure measurement is also common. Therefore, a variety of model specifications have been tested before the selection was narrowed down to the three representative ones.

As shown in Table 2, Model 1 uses million VMT as the traffic exposure and truck percentage (TRKPT) as one of the explanatory variables in the crash rate function. Model 2 used truck mile traveled (TMT) as the traffic exposure, assuming truck crashes are proportional to the truck volume and segment length. AADT is treated as one of the explanatory variables, representing the traffic density. Model 3 uses both AATT and AADT in the traffic exposure and segment length is treated as an offset. This model structure emphasizes the interaction between trucks and non-truck motor vehicles. Note that the statistically significant variables vary across three models due to different model specification. For brevity, they are represented as $\mathbf{X\beta}$ in the model. The final model was selected based on the model statistical goodness-of-fit and the number of meaningful and statistically significant variables. The Akaike information criterion (AIC) is a measure of the statistical goodness-of-fit. The general formula is AIC= 2k - 2ln(L) where k is the number of parameters in the statistical model and L is the maximized value of the likelihood function for the estimated model. The preferred model is the one with the minimum AIC value, which is Model 2.

TABLE 2 NB Model Structures

Model	Equation	AIC value
Model 1	$\mu = (VMT)^{\alpha}\ EXP\ (\beta_0+\beta_1 TRKPT+\mathbf{X\beta})$ where VMT is million VMT	968
Model 2	$\mu = (TMT)^{\alpha}\ EXP\ (\beta_0+\beta_1 AADT+\mathbf{X\beta})$ where TMT is million truck miles traveled	966
Model 3	$\mu = length * AATT^{\alpha 1} AADT^{\alpha 2} EXP\ (\beta_0+\mathbf{X\beta})$	982

Table 3 summarizes the parameter estimates, standard deviation, t-statistics and variables that are significant at 95% confidence limit. Along with the intercept, million truck miles traveled (TMT), AADT, signal density and standard deviation of Pavement Serviceability Index (PSI) are positively associated with the number of truck crashes. The closely spaced signalized intersections along corridors could influence each other in operation as well as in safety (22). The shoulder width and PSI are negatively associated with the number of truck crashes. Among these crash contributing factors, the PSI value was calculated based on slope variance, rut depth, cracking and patching. A PSI value of 5 means the perfect riding condition of a road surface and vice versa. The model results imply that the corridor-based safety performance could be improved by better pavement conditions, wider shoulder widths, and more consistent signal timing designs (e.g. protected phases, longer clearance interval, etc.).

TABLE 3 NB Estimates for Accident Frequency Prediction

Effect	Estimate	Std. Err.	t- Statistics	p-value
Constant	2.7523	0.255	11	0.0001
TMT	0.8404	0.08	10.2	0.0001
AADT in thousands	0.023	0.009	2.54	0.0366
Shoulder width	-0.042	0.02	-2.24	0.0283
Signal density	0.186	0.042	2.95	0.0036
PSI	-0.2115	0.061	-3.53	0.0009
STD(PSI)	0.26	0.112	2.27	0.0278
Dispersion	0.180	0.027	6.67	0.0001
AIC = 966; Pearson Chi-Square / DF=1.07				

Following the crash frequency prediction, the crash severity distribution was also estimated based on corridor-level variables. Both MNL and OP models were used for the prediction of probabilities for crash injury severity proportions for each corridor. The predicted probabilities were compared with the observed proportion using the sum of absolute difference (SAD) as follows:

$$SAD^j = \sum_{i=1}^{100} \left| P_i^j - O_i^j \right| \qquad (10)$$

Where:
SAD^j is the sum of absolute difference for all 100 corridors for injury severity type j;
P_i^j is the predicted probability for injury severity type j on corridor i; and
O_i^j is the observed probability for injury severity type j on corridor i;

Table 4 shows the sum of absolute difference of injury severity proportions of MNL and OP models. The MNL model was chosen to calculate the predicted number of crashes for the five levels within a corridor because the sum of the absolute difference in MNL was smaller than OPM for all levels.

TABLE 4 Sum of Absolute Difference of Injury Severity Proportions

Model	O	C	B	A	K
OP	6.29	6.02	3.81	2.16	1.50
MNL	6.16	5.06	3.70	1.82	1.27

In the MNL model results shown in Table 5, the posted speed limit, shoulder width, pavement serviceability index, standard deviation of PSI, pavement condition index, number of lanes, lane width, AATT, AADT and undivided portion of roadway segment were all determined

TABLE 5 Coefficient Estimates for MNL

Variable	C		B		A		K	
	Coef. (Std. Err.)	Z (p-value)	Coef. (Std. Err.)	Z (p-value)	Coef. (Std. Err.)	Z (p-value)	Coef. (Std. Err.)	Z (p-value)
Intercept	-	-	-2.44 (1.08)	-2.24 (.02)	-7.13 (2.0)	-3.40 (.001)	-12.51 (4.0)	-3.11 (.002)
AADT	-	-	-.043 (.024)	-1.83 (.06)	-	-	-	-
AATT	-	-	.001 (.000)	1.99 (0.04)	-	-	-	-
SPD	-	-	-	-	.052 (.01)	3.22 (.001)	.059 (.03)	1.85 (.06)
Ln width	-.096 (.04)	-1.94 (.053)	-	-	-	-	.393 (.22)	1.76 (.07)
NL_1			1.38 (0.61)	2.25 (.02)	-	-	-	-
NL_2	-.378 (.17)	-2.21 (.02)	-	-	-	-	-	-
NL_3	-.480 (.19)	-2.41 (.01)	-	-	-	-	-	-
Shoulder width	-	-	-	-	.111 (.03)	2.87 (.004)	-	-
Divund_U	-	-	.348 (.18)	1.93 (.053)	-	-	-	-
PCI	-.003 (.001)	-1.69 (.09)	-.004 (.002)	-2.06 (.03)	-	-	-	-
PSI	-	-	.173 (.08)	2.15 (0.03)	-	-	-	-
STD(PSI)	-	-			-.735 (.20)	-3.61 (.000)	-1.25 (.42)	-2.89 (.003)

Note: Number of observation = 1986, Prob>chi-square=0; LL= -7755.43
"-" represents the variables that are not statistically significant at 10% level of significance.

to be statistically significant variables for predicting different levels of injury severity at the 10% significance level. In the MNL model, the coefficient estimates are explained as the comparison between injury level *i* with the base level O. For example, if a road is undivided, a driver's chance of getting injured increases significantly, with respective probabilities of level B being 1.42 ($e^{0.348}$) times that of O. Similarly, a one lane corridor increases the probabilities of level B being 3.97 ($e^{1.38}$) times that of O and injury severity due to the effect of PSI for level B is 1.2 ($e.^{173}$) times that of the base level.

In the final phase of the research, the predicted crash frequency and the predicted severity proportions for each corridor were employed to develop the truck corridor CSI using Equation 1. The total number of predicted crashes for a corridor was multiplied by the corresponding injury severity proportions in order to get the crash frequency for each severity type. Then those predicted injury severity frequencies were multiplied by the respective comprehensive crash cost provided in HSM for the estimation of total crash costs of each corridor (18). A worksheet was designed to facilitate the calculation as illustrated in Table 6.

TABLE 6 CSI Estimation Worksheet

<table>
<tr><td colspan="2">Corridor Location Information
Highway name:
From / To:
Nearby Interstate Highway:
Region:</td></tr>
<tr><td rowspan="2">Variables
AADT
AATT
L
Shoulder width
Signal density
Ln width
NL_1
NL_2
NL_3
Divund_U
SPD
PCI
PSI
STD(PSI)</td><td>Calculation of expected number of crashes
$TMT = \frac{365 \times AATT \times L}{1000000}$
$N = \mathrm{TMT}^{0.084} \exp^{(2.75 + 0.02 * \mathrm{AADT} - 0.042 * \text{Shoulder width} + 0.186 * \text{Signal density} - 0.212 * \mathrm{PSI} + .258 * \mathrm{STD(PSI)})}$</td></tr>
<tr><td>Calculation of predicted injury severity proportion
(coefficients refer to Table 5)
$\log\left[\frac{P_i(k-1)}{P_i(k)}\right] = \alpha_{k-1} X_{(k-1)}$
$P^k = P(O) * e^{\alpha_k X_k}$
$P^A = P(O) * e^{\alpha_A X_A}$
$P^B = P(O) * e^{\alpha_B X_B}$
$P^C = P(O) * e^{\alpha_C X_C}$
$P^O = \frac{1}{1 + \sum_{j=1}^{4} e^{\alpha_j * x}}$</td></tr>
<tr><td>Unit crash cost
($) (18)
U_{PDO} = 7,400
U_C = 44,900
U_B = 79,000
U_A = 216,000
U_K = 4,008,900</td><td>Calculation of corridor crash severity index (CSI)
$CSI = \frac{\sum_{j=1}^{J} N P^j U_j}{L}$</td></tr>
<tr><td colspan="2">Glossary: Refer to Table 1</td></tr>
</table>

The observed truck corridor CSIs were calculated and compared with the predicted ones. Figure 1 shows that both predicted CSI and observed CSI skewed to the left, suggesting the CSI is not symmetrically distributed. The average annual predicted CSI was found to be $ 239,830 per mile with a standard deviation of $190, 269, which was higher than the actual average annual CSI of $202, 850 per mile with a standard deviation of $198, 751. The overestimation was more apparent in the range of $200K~$300K than in other intervals. For those overestimated corridors, some common characteristics such as narrower shoulder width, higher standard deviation of AATT, lower pavement serviceability index, narrower lane width were observed, which seem to contribute considerably to the predicted crash frequency and severity. Nevertheless, the overestimated corridors are the ones with low CSI, suggesting very few serious injury crashes.

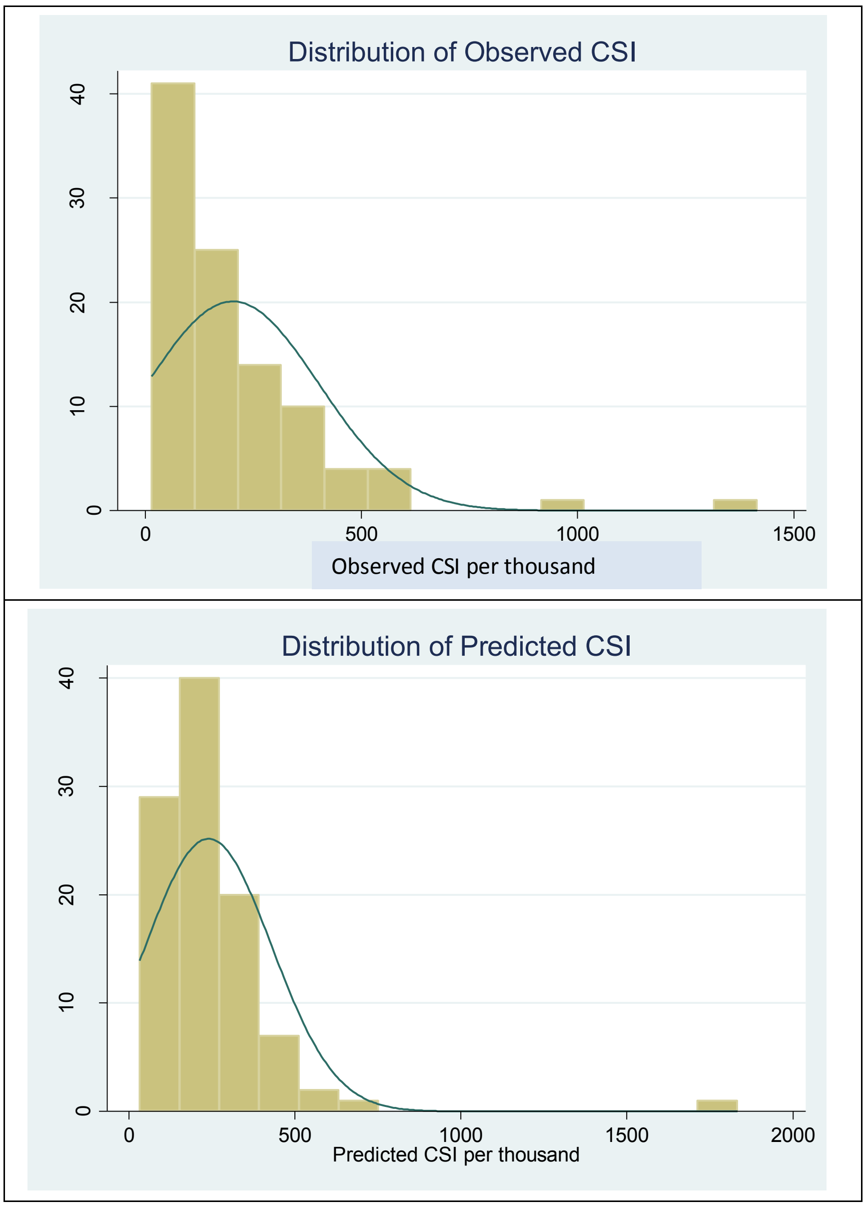

FIGURE 1 Histogram of observed and predicted CSI per thousand.

The developed CSI can play a vital role in quantifying the overall risk to the traveling public posed by each truck corridor. The CSI is designed to alert motor carriers and transportation agencies of potential safety issues so that preventive measures can be taken. The index could assist transportation agencies in allocating safety improvement funding and enhancing the identified geometric design components of arterials. By taking adequate measures based on the CSI, the road agencies can direct trucks to arterial roadways with adequate geometries and pavement conditions. The CSI can also be employed to a truck route network analysis so that highway safety can be incorporated into the route choice. The motor carriers can make informed decision based on not only logistics but also safety.

CONCLUSIONS

Due to rapid truck travel growth in the county, concern amongst transportation agencies about truck related safety issues have increased. Although numerous studies have been conducted for truck safety on the Interstate highway system, the research on truck crashes on arterial streets, especially from the arterial corridor perspective, is relatively limited. Arterial streets are the "last miles" for trucks to deliver the freight to destinations or enter the Interstate highway system. Improving truck safety from an arterial corridor standpoint is crucial for developing more proactive, corridor-based safety strategies. In this study, rigorous effort has been made in the selection of the truck corridors based on corridor length, truck volume and their proximity to interstate highways. Based on the selected truck corridors, a quantifiable crash severity index (CSI) was developed to provide a holistic measurement of the truck crash risk.

The truck corridor based CSI is defined as the annual societal economic costs due to truck crashes per unit length. It is a composite average of the truck crashes by severity with the weights determined by the crash unit cost. The truck crash count by severity for each corridor can be estimated by combining a crash severity model and a crash frequency model through a set of corridor-level variables. The negative binomial model was used to predict the total number of truck crashes, where million truck miles traveled, AADT, signal density, shoulder width, the pavement serviceability index and its standard deviation were identified as statistically significant variables. The MNL model was employed to estimate the injury severity proportion. The model results showed that some factors only affect truck crash frequency such as signal density and other factors only affect crash severities such as posted speed limit, lane width, number of lanes, pavement condition index and undivided roadway portion. The common factors that affect both are AADT, AATT, shoulder width, PSI and its standard deviation. Therefore, when comparing different safety improvements strategies, any change to the value of the factors related to crash frequency, severity, and especially both should be comprehensively and carefully evaluated.

ACKNOWLEDGEMENT

The work was sponsored by the National Center for Freight & Infrastructure Research & Education (CFIRE) at the University of Wisconsin-Madison. The authors thank Chase Cutler for editorial assistance.

REFERENCES

1. American Trucking Associations (ATA). http://www.trucking.org/Newsroom/Trucks%20Are/When%20Trucks%20Stop%20America%20Stops.pdf. Accessed on July 5, 2012.
2. Trucking Industry Needs Drivers – But They Need To Be Safe? *Published by Vincent Leonard in Motor Vehicle Accidents, Trucking Accidents*. http://www.searcylaw.com/blog/category/motor-vehicle-accidents/. Accessed on July 5, 2012.
3. United States. NHTSA National Center for Statistics and Analysis. *Traffic Safety Facts, 2010 Data - Large Trucks*. Washington DC, 2012. Web. http://www-nrd.nhtsa.dot.gov/Pubs/811628.pdf, Accessed on July 20, 2012.
4. National Highway Traffic Safety Administration (NHTSA). http://www-nrd.nhtsa.dot.gov/Pubs/811402.pdf. Accessed on October 20, 2012.
5. Schneider, W. H., K. Zimmerman, D. Van Boxel, and S. Vavilikolanu. Bayesian Analysis of the Effect of Horizontal Curvature on Truck Crashes Using Training and Validation Data Sets. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2096, National Research Council, Washington, D.C., 2009, pp. 41-46.
6. Daniel, J., C. Tsai, and S. Chien. Factors in Truck Crashes on Roadways with Intersections. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1818, National Research Council, Washington, D.C., 2002, pp. 54-59.
7. Daniel, J., I. Steven, and J. Chien. Truck Safety Factors on Urban Arterials. In *Journal of Transportation Engineering*, No. 130, 2004, pp. 742.
8. Wang, H., K. Jang, Y. C. Chan, and Z. Shi. Multi-Level Mixed-Effects Regression Model for Truck Collisions. In Transportation Research Board Annual Meeting, CD-ROM. Washington, D.C., 2011.
9. Miaou, S. P. The Relationship between Truck Accidents and Geometric Design Of Road Sections: Poisson Versus Negative Binomial Regressions. *Accident Analysis & Prevention*, Vol. 26, No. 4, 1994, pp. 471-482.
10. Miaou, S-P., Hu, P.S., Wright, T., Rathi, A.K., and Davis, S.C. Relationship between Truck Accidents and Geometric Design: A Poisson Regression Approach. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1376, National Research Council, Washington, D.C., 1992, pp. 10-18.
11. Miaou, S-P., and Lum, H. A Statistical Evaluation of the Effects of Highway Geometric Design on Truck Accident Involvements. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1407, National Research Council, Washington, D.C., 1993, pp. 11-23.
12. Miaou, S-P., and H. Lum. Modeling Vehicle Accidents and Highway Geometric Design Relationships. *Accident Analysis and Prevention*, Vol. 25, No. 6, 1993, pp. 689-709.
13. Anastasopoulos, P. C., and F. L. Mannering. A Note on Modeling Vehicle Accident Frequencies with Random-Parameters Count Models. *Accident Analysis & Prevention*, Vol. 41, No. 1, 2009, pp. 153-159.
14. Mohamedshah, Y., J.F. Paniati, and A.G. Hobeika. Truck Accident Models for Interstates and Two-lane Rural Roads. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1407, National Research Council, Washington, D.C., 1993, pp. 35-41.

15. Peeta, S., W. Zhou, and P. Zhang. Modeling and Mitigation of Car-Truck Interactions on Freeways. In *Transportation Research Record: Journal of the Transportation Research Board*, No.1899, National Research Council, Washington, D.C., 2004, pp. 117-126.
16. Peeta, S., P. Zhang, and W. Zhou. Behavior-based Analysis of Freeway Car-truck Interactions and Related Mitigation Strategies. *Transportation Research Part B*; 39B, No. 5, 2005, pp. 417-451.
17. Garber, N. J. and S. C. Joshua. Traffic and Geometric Characteristics Affecting the Involvement of Large Trucks in Accidents Volume II, Virginia Transportation Research Council, Charlottesville, VA, November 1990.
18. Highway Safety Manual (1st edition), AASHTO, 2010.
19. The Large Truck crash causation study (LTCCS). http://www.fmcsa.dot.gov/facts-research/research-technology/analysis/fmcsa-rra-07-017.pdf. Accessed on October 20, 2012.
20. El-Basyouny, K., and T. Sayed. Accident Prediction Models With Random Corridor Parameters. *Accident Analysis & Prevention*, Vol. 41, No. 5, 2009, pp. 1118-1123.
21. Lee, C., X. Xu, and V. Nguyen. Non-Intersection-Related Crashes at Midblock In Urban Divided Arterial Road With High Truck Volume. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2448, National Research Council, Washington, D.C., 2011, pp. 20p.
22. Abdel-Aty, M., and X. Wang. Crash Estimation at Signalized Intersections along Corridors: Analyzing Spatial Effect and Identifying Significant Factors. In *Transportation Research Record: Journal of the Transportation Research Board*, No 1953, National Research Council, Washington, D.C., 2006, pp. 98-111.
23. Milton, J. C., V. N. Shankar, and F. L. Mannering. Highway Accident Severities and the Mixed Logit Model: An Exploratory Empirical Analysis. *Accident Analysis & Prevention*, Vol 40, No. 1, 2008, pp. 260-266.
24. De Leur, P., and T. Sayed. A Framework for Proactively Consider Road Safety within the Road Planning Process. In *Canadian Journal of Civil Engineering*, Vol. 30, No. 4, 2003, pp. 711-719.
25. De Leur, P., and T. Sayed. Development of A Road Safety Risk Index. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1784, National Research Council, Washington, D.C., 2002, pp. 33-42.
26. Wu, H., and Z. Zhang. A Framework For Developing Road Risk Indices Using Quantile Regression Based Crash Prediction Model. In Transportation Research Board 91st Annual Meeting, CD-ROM. Washington, D.C., 2012.
27. SAS GENMOD Procedure. http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#genmod_toc.htm. Accessed January, 2012.
28. Savolainen, P. T., F. L. Mannering, D. Lord, and M. A. Quddus. The Statistical Analysis of Highway Crash-Injury Severities: A Review and Assessment of Methodological Alternatives. *Accident Analysis and Prevention*, Vol. 43, No. 5, 2011, pp. 1666–1676.
29. Geedipally, S. R., S. Patil, and D. Lord. Examination of Methods to Estimate Crash Counts by Collision Type. In *Transportation Research Record: Journal of the Transportation Research Board,* No. 2165, National Research Council, Washington, D.C., 2010, pp. 12–20.
30. WisTransportal. http://transportal.cee.wisc.edu/, accessed on September 5, 2011.
31. Venkataraman, N. S., G. F. Ulfarsson, V. Shankar, J. Oh, and M. Park. Model of Relationship between Interstate Crash Occurrence and Geometrics. In *Transportation*

Research Record: Journal of the Transportation Research Board, No. 2236, National Research Council, Washington, D.C., 2011, pp. 41-48.

32. Qin, X., J. N. Ivan, and N. Ravishanker. Selecting exposure measures in crash rate prediction for two-lane highway segments. *Accident Analysis & Prevention*, Vol. 36, No. 2, 2004, pp. 183–191.