Developing Truck Corridor Crash Severity Index

Xiao Qin, Most Afia Sultana, Madhav V. Chitturi, and David A. Noyce

According to NHTSA, more than 400,000 truck accidents occurred in 2009 and approximately 7,800 of those were fatal crashes. Compared with 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 careful planning of the location, design, and operation of driveways, median openings, street connections, and street sections. Through the collection of extensive data on selected arterial corridors that are heavily used by trucks, contributing factors to truck crash frequency and severity were identified with a negative binomial model and multinomial logit model. Corridor truck miles traveled, annual average daily traffic, signal density, shoulder width, and pavement serviceability index and its standard deviation are significant factors for crash frequency prediction. The multinomial logit model identified 12 causal factors for crash severity, such as posted speed limit, lane width, number of lanes, pavement condition index, and undivided roadway portion. Subsequently, a crash severity index for truck arterial corridors was developed. The findings from the study not only will benefit state and local agencies in planning, design, and management of a safer truck arterial corridor, but will also help carriers to optimize their routes from a safety perspective.

Freight transportation is extremely critical to the economic development of a nation. The U.S. economy depends on trucks to deliver nearly 70% of all freight transported annually, accounting for $671 billion worth of manufactured and retail goods, 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*). Although the rapid commercial trucking growth is great news for the country's economy, the increasing truck traffic may negatively affect cars, vans, sport utility vehicles, and other vehicles that share the road. In 2010, large trucks accounted for 4% of all registered vehicles and 10% of the total vehicle miles traveled. Of the fatalities in crashes involving large trucks during 2010, 76% were occupants of other vehicles (*3*). In fact, one person is injured or killed in a truck accident every 16 min and one out of every eight traffic fatalities involves a truck collision (*2*). NHTSA has estimated that more than 400,000 truck accidents occurred in 2009 and approximately 7,800 of those were 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 highway 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% of crashes, whereas 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 because there are more opportunities for conflicts with passenger vehicles at signalized intersections and such analysis is valuable for developing systemwide, corridor-based, and, more important, proactive safety improvement strategies.

Although 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 transport personnel can quantify the hazards associated with particular roadway characteristics, environmental patterns, and driver populations. A quantifiable risk index associated with a roadway segment will help transportation agencies to identify potential safety problems and adopt appropriate remedies before a crash occurrence and thereby reduce the risk exposure to other road users.

Previously, many agencies took 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. AASHTO's newly published *Highway Safety Manual* (HSM) 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.

X. Qin, Room 148, Box 2219, and M. A. Sultana, Room 315, Department of Civil and Environmental Engineering, South Dakota State University, Crothers Engineering Hall, Brookings, SD 57007. M. V. Chitturi, Room B243, and D. A. Noyce, Room 1204, Traffic Operations and Safety Laboratory, Department of Civil and Environmental Engineering, University of Wisconsin–Madison, 2205 Engineering Hall, 1415 Engineering Drive, Madison, WI 53706. Corresponding author: X. Qin, Xiao.qin@sdstate.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2386, Transportation Research Board of the National Academies, Washington, D.C., 2013, pp. 103–111.
DOI: 10.3141/2386-12

LITERATURE REVIEW

Many factors may be involved in truck crashes. The *Large Truck Crash Causation Study* identified human factors (an action or inaction by the driver) and vehicle malfunctions (brake problems) as the two leading causes (*19*). Roadway problems were present in 16% of the two-vehicle cases based on 967 crashes involving 1,127 large trucks and 959 non-truck motor vehicles. Of prime interest to transportation agencies, the impacts of roadway geometric features on truck crashes have 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 multilevel 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 annual average daily traffic (AADT), and lane width significantly affect the number of truck crashes.

Many of the preceding studies were based on either individual intersections or segments, and few studies approached truck safety issues from a corridor perspective (*20–23*). El-Basyouny and Sayed assessed the corridor effects with alternate specifications (*20*). They compared the traditional Poisson lognormal model with two extended Poisson lognormal models by using a data set from 392 urban arterials in the city of Vancouver, British Columbia, Canada, that were clustered into 58 corridors. The results of their study 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 midblock 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 midblock 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 close together (*22*). To account for the correlated data problem they used generalized estimating equations with a negative binomial link function.

Milton et al. used corridor-specific and weather-related variables to predict injury severity proportions with a mixed logistic model (*23*). Within these results, the average daily traffic, 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. In contrast, pavement friction, horizontal curvature per mile, and number of grade breaks per mile have a 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 that 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 roadside environment in an effort to improve the safety of a roadway segment.

In addition to planning recommendations 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 (RSRI). De Leur and Sayed developed two types of RSRI ($\text{RSRI}_{\text{specific}}$ and $\text{RSRI}_{\text{combined}}$) based on the risk score of a particular road feature (*25*). $\text{RSRI}_{\text{specific}}$ defines the risk associated with each road feature, obtained by combining the scores for the three components of risk, and $\text{RSRI}_{\text{combined}}$ defines overall risk by combining the $\text{RSRI}_{\text{specific}}$ scores for all road features.

In a recent study, Wu and Zhang proposed a framework for developing a composite road risk index by using a logistic function based on exposure, crash rate, and crash severity (*26*). They showed the 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 with the HSM crash severity distribution and associated crash unit costs. In the HSM network screening process, a site-specific RSI is calculated by multiplying the observed or predicted average crash frequency for each crash severity with their respective comprehensive crash cost; 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 are discussed.

Crash Severity Index

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

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

where

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

For any truck corridor under consideration, the CSI value can be estimated by using the corridor characteristics and applied either as the ranking tool for 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 used for crash frequency because the number of accidents n_i on a roadway segment per unit of time is a nonnegative integer. When the variance is larger than the mean, the data are said to be overdispersed. Overdispersed 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 corridor i's 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 in which 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 the β's are coefficients of explanatory variables and $\exp(\varepsilon_i)$ is the term adjusting for overdispersion and is gamma distributed. The models were estimated by using generalized linear modeling. For this modeling, the SAS GENMOD procedure was used (*27*).

Modeling Methods for Crash Severity

Ordered Probit 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 given by Savolainen et al. (*28*). It has been accepted by many researchers that there is an ordinal nature to crash severities; that is, injury severity can be ranked from high to low as fatal injury (K), incapacitating injury (A), nonincapacitating injury (B), possible injury (C), and property-damage-only (O). To model injury severities as the ordinal response, researchers most frequently use discrete choice models such as ordered probit (OP) models (*28*). An OP model is a special case of the probit model in which more than two outcomes of an ordinal dependent variable is modeled, usually estimated by using maximum likelihood. The underlying relationship to be characterized is

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

where

- y^* = exact but unobserved dependent variable,
- $\boldsymbol{X}$ = vector of independent variables, and
- $\boldsymbol{\beta}$ = vector of regression coefficients, which needs to be estimated.

The ε is a random error term that is assumed to follow a standard normal distribution. Furthermore y^* cannot be observed; instead, only the categories of response can be observed, expressed as follows:

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

where μ represents thresholds to be estimated along with the parameter vector $\boldsymbol{\beta}$.

Multinomial Logistic Model

When crash severities are modeled 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; that is, the shift in the cutoff thresholds is constrained to move in the same direction. However, nonordinal probabilistic models, such as multinomial logit (MNL) models, allow variables to have opposite effects regardless of the order of the injury severities. The MNL model is a regression model that generalizes logistic regression by allowing more than two discrete outcomes. MNL relies on the assumption of independence of irrelevant alternatives; that is, 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:

$$U_{ij} = \boldsymbol{X}_i'\boldsymbol{\beta}_j + \varepsilon_{ij} \tag{6}$$

where

- $\boldsymbol{X}_i$ = vector of explanatory variables describing observation i,
- $\boldsymbol{\beta}_j$ = vector of weights (or regression coefficients) corresponding to outcome j, and
- U_{ij} = utility associated with assigning observation i to get category j.

The ε_{ij} is an error term that accounts for the random noise and is assumed to be independently and identically distributed with a Gumbel extreme value distribution; its logistic formulation is

$$P_i(j) = \frac{\exp[\boldsymbol{X}_i'\boldsymbol{\beta}_j]}{1 + \sum_{j=1}^{K-1}\exp[\boldsymbol{X}_i'\boldsymbol{\beta}_j]} \qquad \text{for } j = 1, \ldots, K-1 \tag{7}$$

In an MNL model, for K possible outcomes, running $(K-1)$ independent binary logistic regression models, 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:

$$\log\left[\frac{P_i(j)}{P_i(K)}\right] = \boldsymbol{X}_i'\boldsymbol{\beta}_j \qquad \text{for } j = 1, \ldots, K-1 \tag{8}$$

Note that $\boldsymbol{\beta}_j$ 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 5 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. Because of the challenge of short (less than 1 mi) or very short segments (less than 0.1 mi) in the data set, it was necessary to collapse short segments into longer ones so that they could be treated as a corridor. This operation was done by using collapsing criteria to dissolve adjacent roadway segments with similar or the same annual average daily truck traffic (AADTT).

After a sensitivity analysis to specify a reasonable corridor length, it was determined to collapse adjacent segments having AADTT 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: (*a*) threshold of the corridor length is no less than 1 mi, (*b*) threshold value of truck AADT is 800 or more, and (*c*) study segment is within 5 mi of an Interstate highway or a freeway. This operation 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.

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

Variable	Description	Mean	SD	Min.	Max.
Crash count	5-year crash count for each corridor	82	71	14	407
Crash severity					
O	Property damage only	54	49	9	276
C	Possible injury	17	16	0	84
B	Nonincapacitating injury	8	7	0	41
A	Incapacitating injury	3	3	0	11
K	Fatal injury	1	2	0	6
L	Length of corridor (mi)	4.88	3.42	1.03	16.94
AADT	Annual average daily traffic	16,256	6,107	8,172	39,435
AADTT	Annual average daily truck traffic	1,077	211	800	1,892
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/mi)	0.51	0.87	0	4.33
Accden	Access point density (access points/mi)	5.29	4.81	0	30.47
SPD	Posted speed limit (mph)	45	9	30	60
Lnwd	Lane width (ft)	12.3	0.8	10	18
Mednwd	Median width (ft)	14	12.9	0	47.3
Lshwd	Left shoulder width (ft)	3.8	3.4	0	10.9
Rshwd	Right shoulder width (ft)	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 lanes	0.81	0.3	0	1
NL_3	Portion of segment with three lanes	0.06	0.2	0	1
NL_4	Portion of segment with four lanes	0.12	0.25	0	1
Hcl_g	Portion of segment with horizontal curve speed less than 40 mph_good	0.95	0.19	0	1
Hcl_f	Portion of segment with horizontal curve speed less than 40 mph_fair	0.03	0.17	0	1
Hcl_p	Portion of segment with horizontal curve speed less than 40 mph _poor	0.01	0.07	0	0.43
Hcg_g	Portion of segment with horizontal curve speed greater than 40 mph_good	0.89	0.29	0	1
Hcg_f	Portion of segment with horizontal curve speed greater than 40 mph_fair	0.09	0.26	0	1
Hcg_p	Portion of segment with horizontal curve speed greater than 40 mph_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 (mm)	0.08	0.08	0	0.427
PCI	Pavement condition index (0–100)	77.09	24.35	0	100

NOTE: SD = standard deviation; min. = minimum; max. = maximum.

During this 5-year period, 8,196 truck-related crashes occurred in selected corridors; notably more than 50% of the crashes occurred in the southeast region and near the Milwaukee area, where most truck activities occur. There was a decreasing trend of crashes over the 5-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 nonincapacitating injuries (B), 3% were incapacitating injuries (A), and 1% were fatal injuries (K). From the results of single- and multiple-vehicle crashes studied, 88% of the crashes involved more than one vehicle.

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 unsignalized intersections and corridor lengths. The maximum access point density of 30.47 exists in a 2.56-mi 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, which contains a portion of a principal arterial with a 65-mph posted speed limit. Similarly, the maximum lane width of 18 ft reflects a portion of a principal arterial corridor that has very wide lanes (22 ft). In addition, the proportion of corridor by the number of lanes, median presence, and speed limited was calculated. In particular, the corridor data were analyzed carefully for the good, fair, and poor condition of roadways with less than or greater than 40-mph horizontal curvature speed.

ANALYSIS AND DISCUSSION OF RESULTS

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 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 et al. used AADTT as an exposure variable and AADT as a surrogate variable to indicate traffic condition when they modeled truck crashes (*10*). In contrast, Venkataraman et al. used AADT and the length of a segment as exposure variables in modeling Interstate crash occurrences (*31*). Use of vehicle miles traveled, which is the product of segment length, AADT, and the number of days a year in units of millions or 100 millions, as the traffic exposure measurement is also common (*31*). Therefore, a variety of model specifications were tested before the selection was narrowed down to the three representative ones.

As shown in Table 2, Model 1 uses million vehicle miles traveled as the traffic exposure and truck percentage (TRKPT) as one of the explanatory variables in the crash rate function. Model 2 uses truck miles traveled (TMT) as the traffic exposure, assuming that 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 AADTT and AADT in the traffic exposure, and segment length is treated as an offset. This model structure emphasizes the interaction between trucks and nontruck motor vehicles.

TABLE 2 Negative Binomial Model Structures

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

NOTE: TRKPT = truck percentage; VMT = vehicle miles traveled; AIC = Akaike information criterion.

The statistically significant variables vary across the three models because of different model specifications. For brevity, they are represented as $\boldsymbol{X'\beta}$ in the model. The final model was selected on the basis of the model statistical goodness of fit and the number of meaningful and statistically significant variables. The AIC is a measure of the statistical goodness of fit. The general formula is $\text{AIC} = 2k - 2\ln(L)$, 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 3 summarizes the parameter estimates, standard deviation, t-statistics, and variables that are significant at the 95% confidence limit. Along with the intercept, million TMT, AADT, signal density, and standard deviation of the 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 on the basis of the slope variance, rut depth, cracking, and patching. A PSI value of 5 means a perfect riding condition of a road surface and vice versa. The model results imply that corridor-based safety performance could be improved by better pavement condition, wider shoulder width, and more consistent signal timing design (e.g., protected phases, longer clearance interval).

Following the crash frequency prediction, the crash severity distribution was also estimated on the basis of corridor-level variables.

TABLE 3 Negative Binomial Estimates for Accident Frequency Prediction

Effect	Estimate	SE	t-Statistic	p-Value
Constant	2.7523	0.255	11	.0001
TMT	0.8404	0.08	10.2	.0001
AADT (thousands)	0.023	0.009	2.54	.0366
Shoulder width	–0.042	0.02	–2.24	.0283
Signal density	0.186	0.042	2.95	.0036
PSI	–0.2115	0.061	–3.53	.0009
STD(PSI)	0.26	0.112	2.27	.0278
Dispersion	0.180	0.027	6.67	.0001

NOTE: AIC = 966; Pearson chi-square/degrees of freedom = 1.07; SE = standard error.

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

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 by using the sum of absolute difference (SAD) as follows:

$$\text{SAD}(j) = \sum_{i=1}^{100} \left| P_i(j) - O_i(j) \right| \tag{9}$$

where

$\text{SAD}(j)$ = sum of absolute difference for all 100 corridors for injury severity type j,
$P_i(j)$ = predicted probability for injury severity type j on corridor i, and
$O_i(j)$ = observed probability for injury severity type j on corridor i.

Table 4 shows the SAD of injury severity proportions for 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 SAD in the MNL model was smaller than in the OP models for all levels.

In the MNL model results shown in Table 5, the posted speed limit, shoulder width, PSI, standard deviation of PSI, pavement condition index, number of lanes, lane width, AADTT, AADT, and undivided portion of roadway were all determined 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 and 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's being 1.42 ($e^{0.348}$) times that of O. Similarly, a one-lane corridor increases the probabilities of Level B's being 3.97 ($e^{1.38}$) times that of O and the injury severity due to the effect of PSI for Level B 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 with 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 the HSM for the estimation of total crash costs of each corridor (*18*). A worksheet designed to facilitate the calculation is illustrated in Figure 1.

The observed truck corridor CSIs were calculated and compared with the predicted ones. Figure 2 shows that both predicted CSI and observed CSI skewed to the left; this distribution suggests that 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 $200,000~$300,000 than in other intervals. For those overestimated corridors, some common characteristics such as narrower shoulder width, higher standard deviation of AADTT, lower PSI, and narrower lane width were observed; these characteristics seem to contribute considerably to the predicted crash frequency and severity. Nevertheless, the overestimated corridors are the ones with low CSI; this finding suggests very few serious injury crashes.

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

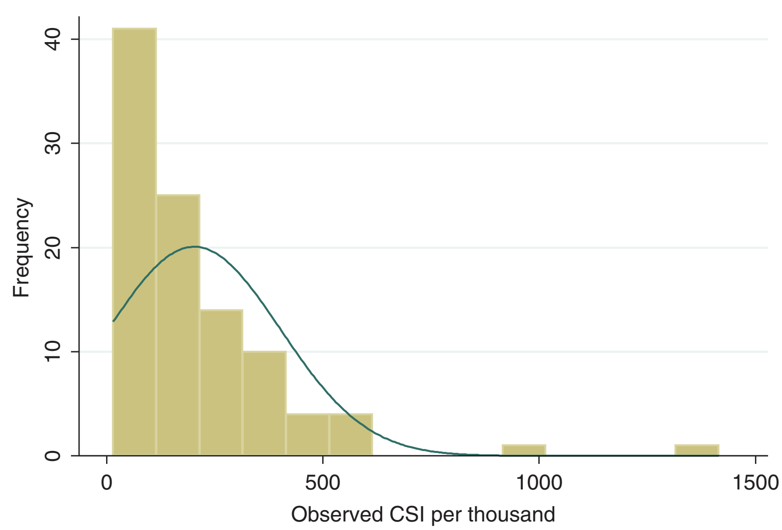
TABLE 5 Coefficient Estimates for MNL

	C		B		A		K	
Variable	Coeff. (SE)	*Z* (*p*-value)	Coeff. (SE)	*Z* (*p*-value)	Coeff. (SE)	*Z* (*p*-value)	Coeff. (SE)	*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)	—	—	—	—
AADTT	—	—	.001 (.000)	1.99 (0.04)	—	—	—	—
SPD	—	—	—	—	.052 (.01)	3.22 (.001)	.059 (.03)	1.85 (.06)
Lnwd	−.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)	—	—	—	—
SD(PSI)	—	—	—	—	−.735 (.20)	−3.61 (.000)	−1.25 (.42)	−2.89 (.003)

NOTE: Number of observations = 1,986; probability > chi-square = 0; log likelihood = −7,755.43; coeff. = coefficient; — = variable not statistically significant at 10% level of significance.

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

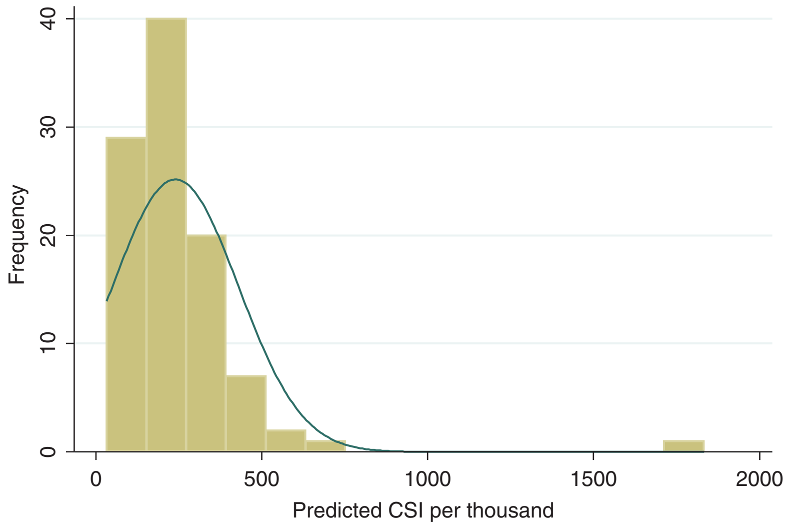
FIGURE 1 CSI estimation worksheet. For definition of variables, see Table 1. PDO = property damage only. Unit crash cost from *Highway Safety Manual* (18).



(a)

FIGURE 2 Histogram of (*a*) observed CSI per thousand.

(continued on next page)



(b)

FIGURE 2 *(continued)* **Histogram of (*b*) predicted CSI per thousand.**

components of arterials. By taking adequate measures based on the CSI, road agencies can direct trucks to arterial roadways with adequate geometries and pavement conditions. The CSI can also be employed in a truck route network analysis so that highway safety can be incorporated into the route choice. Motor carriers can make informed decisions based on not only logistics but also safety.

CONCLUSIONS

Because of rapid truck travel growth in the United States, concern among transportation agencies about truck-related safety issues has increased. Although numerous studies have been conducted for truck safety on the Interstate Highway System, 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 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, a rigorous effort was 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 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, and the PCI 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, such as signal density, affect only truck crash frequency, and other factors, such as posted speed limit, lane width, number of lanes, pavement condition index, and undivided roadway portion, only affect crash severity. The common factors that affect both are AADT, AADTT, shoulder width, and PSI and its standard deviation. Therefore, when different safety improvement strategies are compared, any change to the value of the factors related to crash frequency, severity, and especially to both should be comprehensively and carefully evaluated.

ACKNOWLEDGMENTS

The work was sponsored by the National Center for Freight and Infrastructure Research and Education at the University of Wisconsin–Madison. The authors thank Chase Cutler for editorial assistance.

REFERENCES

1. *When Trucks Stop America Stops.* American Trucking Associations, Arlington, Va. http://www.trucking.org/Newsroom/Trucks%20Are/When%20Trucks%20Stop%20America%20Stops.pdf. Accessed July 5, 2012.
2. Leonard, V. Trucking Industry Needs Drivers—But They Need To Be Safe? *Motor Vehicle Accidents, Trucking Accidents.* Searcy Blog, West Palm Beach, Fla. http://www.searcylaw.com/blog/category/motor-vehicle-accidents. Accessed July 5, 2012.
3. *Traffic Safety Facts: 2010 Data—Large Trucks.* NHTSA, U.S. Department of Transportation, 2012. http://www-nrd.nhtsa.dot.gov/Pubs/811628.pdf. Accessed July 20, 2012.
4. *Traffic Safety Facts 2009.* NHTSA, U.S. Department of Transportation, 2009. http://www-nrd.nhtsa.dot.gov/Pubs/811402.pdf. Accessed Oct. 20, 2012.
5. Schneider, W. H., IV, K. H. 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,* Transportation Research Board of the National Academies, 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,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 54–59.

7. Daniel, J., I. Steven, and J. Chien. Truck Safety Factors on Urban Arterials. *Journal of Transportation Engineering,* ASCE, Vol. 130, No. 6, 2004, p. 742–752.
8. Wang, H., K. Jang, C.-Y. Chan, and Z. Shi. Multi-Level Mixed-Effects Regression Model for Truck Collisions. Presented at 90th Annual Meeting of the Transportation Research Board, 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 and Prevention,* Vol. 26, No. 4, 1994, pp. 471–482.
10. Miaou, S.-P., P. S. Hu, T. Wright, A. K. Rathi, and S. C. Davis. Relationship Between Truck Accidents and Highway Geometric Design: A Poisson Regression Approach. In *Transportation Research Record 1376,* TRB, National Research Council, Washington, D.C., 1992, pp. 10–18.
11. Miaou, S.-P., and H. Lum. Statistical Evaluation of the Effects of Highway Geometric Design on Truck Accident Involvements. In *Transportation Research Record 1407,* TRB, 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 and Prevention,* Vol. 41, No. 1, 2009, pp. 153–159.
14. Mohamedshah, Y. M., J. F. Paniati, and A. G. Hobeika. Truck Accident Models for Interstates and Two-Lane Rural Roads. In *Transportation Research Record 1407,* TRB, 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,* Transportation Research Board of the National Academies, 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,* Vol. 39, 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, Vol. II.* Virginia Transportation Research Council, Charlottesville, Va., Nov. 1990.
18. *Highway Safety Manual,* 1st ed. AASHTO, Washington, D.C., 2010.
19. *The Large Truck Crash Causation Study.* FMCSA, U.S. Department of Transportation, March 2006. http://www.fmcsa.dot.gov/facts-research/research-technology/analysis/fmcsa-rra-07-017.pdf. Accessed Oct. 20, 2012.
20. El-Basyouny, K., and T. Sayed. Accident Prediction Models with Random Corridor Parameters. *Accident Analysis and 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. Presented at 90th Annual Meeting of the Transportation Research Board, Washington, D.C., 2011.
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,* Transportation Research Board of the National Academies, 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 and 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. *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,* Transportation Research Board of the National Academies, Washington, D.C., 2002, pp. 33–42.
26. Wu, H., and Z. Zhang. Framework for Developing Road Risk Indices Using Quantile Regression Based Crash Prediction Model. Presented at 91st Annual Meeting of the Transportation Research Board, Washington, D.C., 2012.
27. GENMOD Procedure. SAS Institute, Cary, N.C. http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#genmod_toc.htm. Accessed Jan. 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,* Transportation Research Board of the National Academies, Washington, D.C., 2010, pp. 12–20.
30. WisTransPortal. http://transportal.cee.wisc.edu. Accessed Sept. 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: Exploratory Insights from Random Parameter Negative Binomial Approach. In *Transportation Research Record: Journal of the Transportation Research Board, No. 2236,* Transportation Research Board of the National Academies, 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 and Prevention,* Vol. 36, No. 2, 2004, pp. 183–191.

The Safety Data, Analysis, and Evaluation Committee peer-reviewed this paper.