Analysis of Crash Severity for Hazard Material Transportation Using Highway Safety Information System Data

Xiuguang Song1, Jianqing Wu1,2, Hongbo Zhang1,2, and Rendong Pi1,2

Abstract
Crash severity, as a major concern in the routing and scheduling of hazardous material shipments, has caused great loss of lives and property damage every year. Although abundant studies have been conducted to identify the relationship between different factors on crash severity, the analysis of the severity of hazard material transportation (HMT) crashes is very limited. Factors including road, vehicle, driver, and environment are not well considered in previous studies. This article analyzed the influence of various factors on HMT crash severity using Highway Safety Information System data. The random forest combined with the ordered logistic model is used for factor analysis. The results showed that annual average daily traffic, fatigues/asleep, number of lanes, speeding, adverse weather, and light are the six most important factors affecting HMT crash severity. Different from the non-HMT crashes, driver factor (e.g., driver age, gender, and drug/alcohol influence) was found to be not significantly related to crash severity. Speeding should be strictly forbidden for HMT drivers, considering the potential increased crash severity. Increasing the level of lighting can help reduce the number of severe crashes. The corresponding recommendations were provided based on the regression results.

Keywords
hazard material transportation, factor analysis, HSIS

Introduction
Transport of hazardous materials plays a critical role in the constant development of economy. The characteristics of hazardous materials are becoming more complex and the transport volume of hazard material is continuously increasing (Nam & Manering, 2000). Depending on the type of hazardous material being shipped, a crash may mean an explosion and fire, significant injuries to the motorists involved, or environmental damage if the truck rolls over and a spill occurs. More than 13% of trucks on the roads carry some type of hazardous material (Schmidt & Price, 1979). It is estimated that about 200 hazmat trucks annually are involved in fatal crashes and 5,000 in nonfatal crashes (Craft, 2004). Among those hazardous crashes, more than half happened on roads (Oggero et al., 2006), releasing hazardous material alongside the road. People’s health and the environment can be directly or indirectly affected by hazardous material crashes (Nicolet-Monnier & Gheorghe, 2013). The hazardous material release often requires a special person to deal with the event. Although the hazardous material crashes are rare events, the potential adverse consequences raise serious concerns to all stakeholders affected by the hazard material transportation (HMT; Zografos & Androutspoulos, 2008). Thus, to ensure safe transport, HMT has its unique features. The selection of route for HMT has to meet three main objectives: to minimize the total risk of crashes, to minimize the operation time, and to minimize the sensitive number of people (Ma et al., 2013). Ma et al. (2018) applied a road screening algorithm based on genetic algorithm and Levenberg–Marquardt neural network for HMT route selection. A total of 15 features of each road were involved in the developed algorithm. Fabiano et al. (2002) developed a model to estimate the risk of HMT and to select better routes. Later, Fabiano and Palazzi (2010) analyzed the relationship between HMT and road tunnels. They suggested providing adequate and well-designed ventilation in the tunnel to tackle fires related to HMT. They also suggested designing the tunnel with positive slope from a safety viewpoint. The crash risk evaluation has been the
principal criterion for determining the optimal route for HMT (List et al., 1991; Zografos & Davis, 1989). As the route selection needs to balance three objectives, there is not an “absolutely safe” route (Pradhananga et al., 2010). Crash frequency and crash severity are two concerns in understanding the relationship between crash occurrences and various risk factors (Wu & Xu, 2018b). In this article, crash severity means crash injury severity, and only the injury of the driver is considered. As the frequency of crashes related to HMT is relatively low, crash severity has been a major concern of the HMT system. Crash severity analysis is widely used for safety evaluation, traffic planning, and traffic management. It would be very useful if a comprehensive understanding of the effects of risk factors on crash severity is available (Goniewicz et al., 2016; Huang et al., 2008). Uddin and Huynh (2018) used the crash data from the Highway Safety Information System (HSIS) to investigate the influence of different factors on HMT crash severity. It was found that rural locations, light conditions, and gender were related to high probability of major injuries. Iranitalab et al. (2018) analyzed HMT crash risk at highway-rail grade crossings using the Federal Railroad Administration (FRA) crash data set with a logistic regression model. The results showed that temperature, weather condition, and vehicle types were highly related to HMT crash risk. In another study, Iranitalab and Khattak (2018) explored the impacts of different factors on the probability of HMT crashes and found that signal and communication causes were key factors for hazmat release. Wilkinson (2018) analyzed the factors related to the release of radioactive materials in severe air crash environments.

Although a lot of studies (Bener et al., 2003; Bochner, 1998) have investigated the influence of different factors on crash severity, none of them have highlighted the crashes related to HMT. A critical issue in crash severity analysis for HMT relates to the lack of data due to the low frequency of such events (Bener et al., 2003). The HSIS is a multistate database that contains crash, roadway inventory, and traffic volume data for a select group of states, which provides the ideal data source for HMT crash severity analysis (Pour-Rouholamin & Jalayer, 2016). This article analyzed the influence of various factors on HMT crash severity using HSIS data in four states: California (CA), Minnesota (MN), North Carolina (NC), and Ohio (OH). The corresponding recommendations were provided based on the analysis results.

**Materials and Method**

Five-year crash data of the four states were extracted from the HSIS database. The crash data, roadway data, and vehicle data were linked together using the accident report number, county route, and milepost information. A total of 2,484 HMT crash records were then identified. The HMT crash severity was divided into five ordered levels: Property Damage Only—PDO (P), Complaint of Pain (C), Other Visible Injury (B), Severe Injury (A), and Fatal (K). Among those records, 1.97% of them were fatal crashes (K), 29.79% were injury crashes (A, B, or C), and 68.27% were PDO crashes (P). There are multiple potential factors that may influence the severity of HMT crashes (Harwood et al., 1989). The potential influencing factors of crash severity can be roughly grouped into four parts according to different attributes: driver factors, vehicle factors, road factors, and environmental factors. For HMT, the vehicle types are usually limited to heavy-duty trucks; therefore, vehicle factors are not considered in this research. A total of 18 potential influencing factors were extracted from the HSIS database. To better interpret the influence of categorical factors, different codes were assigned to those categorical independent variables. After excluding those blank records among the investigated factors, 2,340 records were left for further analysis. A summary of the variables and their corresponding codes in the database are documented in Table 1.

A lot of models have been developed for the analysis of crash severity and its contributing factors. Discrete response models (e.g., multinomial logit models, nested logit models, mixed logit models, ordered logit models [OLMs], and ordered probit models) are widely applied in previous studies (Savolainen et al., 2011). In the mixed logit models, the impact of variables on severity levels can vary across observations. Chen and Chen (2011) applied a mixed logit model to identify key factors influencing crashes on rural highways. In another study, a time-varying mixed logit model was successfully developed for vehicle merging behavior analysis in work zone (Weng et al., 2018). A random parameters bivariate ordered probit model is another widely used model for crash severity analysis. This model can also address the unobserved heterogeneity. Chen et al. (2019) used the random parameters bivariate ordered probit model to analyze the factors influencing the injury severity of drivers in rear-end collisions. Some other models such as partial proportional odds (PPO), decision trees (Wu, Xu, et al., 2018), multivariate Poisson regression (Ma & Kockelman, 2006), and artificial neural network (Delen et al., 2006) have been also used for crash severity analysis by researchers. Basically, each method has its own advantages and disadvantages. Therefore, there is not a uniform standard for model selection. In this article, we used a combination of random forest (RF) and OLM named RF-OLM for factor analysis.

The RF, one of the decision tree–based methods, has attracted more interests by researchers (Das et al., 2009; Kaplan & Prato, 2012) because of its excellent performance in handling multiple independent factors (numerical and categorical factors). In this research, RF is applied to identify the ranking of the importance of suspicious factors on HMT crash severity. The principle of RF regression is the minimum mean square error (MMSE). For any random feature \( A \), the corresponding random division points can generate two
Table 1. Variable Extraction From the HSIS Database.

| Factors                    | Variables          | Values | Code | Frequency | Percentage |
|----------------------------|--------------------|--------|------|-----------|------------|
| Dependent variable         | Severity           | PDO    | PDO  | 1,590     | 67.9       |
|                            |                    | C      | C    | 334       | 14.3       |
|                            |                    | B      | B    | 282       | 12.1       |
|                            |                    | A      | A    | 87        | 3.7        |
|                            |                    | K      | K    | 47        | 2.0        |
| Driver factors             | Driver age         | 20–91 years | Numerical | 1,736 | 74.2 |
|                            |                    | Average: 45.9 | SD: 12.6 |       |      |
|                            | Driver gender      | Male   | 0    | 2,067     | 88.3       |
|                            |                    | Female | 1    | 273       | 11.7       |
|                            | Drug/Alcohol influence | Yes | 1    | 19        | 0.8        |
|                            |                    | No     | 0    | 2,321     | 99.2       |
|                            | Fatigue/Asleep     | Yes    | 1    | 19        | 0.8        |
|                            |                    | No     | 0    | 2,321     | 99.2       |
|                            | Physical impairment | Yes | 1    | 170       | 7.3        |
|                            |                    | No     | 0    | 2,170     | 92.7       |
|                            | Speeding           | Yes    | 1    | 103       | 4.4        |
|                            |                    | No     | 0    | 2,237     | 95.6       |
| Road factors               | Road surface       | Dry    | 0    | 1,736     | 74.2       |
|                            |                    | Other condition | 1 | 604       | 25.8       |
|                            | Horizontal alignment | Straight | 0 | 2,046 | 87.4 |
|                            |                    | Curve | 1 | 294 | 12.6 |
|                            | Vertical alignment | Level | 0 | 1,741 | 74.4 |
|                            |                    | Grade | 1 | 599 | 25.6 |
|                            | Divided            | Yes    | 1    | 1,820     | 77.8       |
|                            |                    | No     | 0    | 520       | 22.2       |
|                            | Intersection related | Yes | 1 | 737 | 31.5 |
|                            |                    | No     | 0 | 1,603 | 68.5 |
|                            | Number of lanes    | ≤2     | 0 | 736 | 31.5 |
|                            |                    | 3–6    | 1 | 1,367 | 58.4 |
|                            |                    | 7–12   | 2 | 237 | 10.1 |
|                            | AADT               | 0–5,000 | 0 | 448 | 19.1 |
|                            |                    | 5,001–10,000 | 1 | 269 | 11.5 |
|                            |                    | 10,001–20,000 | 2 | 382 | 16.3 |
|                            |                    | 20,001–50,000 | 3 | 597 | 25.5 |
|                            |                    | 50,001+ | 4 | 644 | 27.6 |
|                            | Rural or urban     | Urban | 1 | 1,553 | 66.4 |
|                            |                    | Rural | 0 | 787 | 33.6 |
| Environmental factors      | Light              | Yes    | 1 | 1,871 | 81.0 |
|                            |                    | No     | 0 | 469 | 19.0 |
|                            | Adverse weather    | Yes    | 1 | 1,945 | 83.1 |
|                            |                    | No     | 0 | 395 | 16.9 |

Note. HSIS = Highway Safety Information System; AADT = annual average daily traffic; PDO = Property Damage Only; C = Complaint of Pain; B = Other Visible Injury; A = Severe Injury; K = Fatal.

The RF calculates the situation when both $D_1$ and $D_2$ have MMSE. Then the corresponding feature and division point can be obtained when the sum of MMSE of $D_1$ and $D_2$ reaches the minimum value. This procedure can be presented as

$$
\min_{c_1, c_2} \left[ \sum_{x \in D_1} (y_i - c_1)^2 + \min_{c} \sum_{x \in D_2} (y_i - c_2)^2 \right]
$$

(1)

where $c_1$ and $c_2$ are the average value of the samples in $D_1$ and $D_2$, respectively. RF can reduce the overfitting issue by averaging several trees. In this research, RF is applied for variable importance examination before OLM regression. The results from RF are used in the following stepwise regression for OLM. The “mean decrease in accuracy” can be generated from RF to rank the importance of the variables. The strategy can be illustrated as
OLM should be very limited. If the feature is important, disordering the values can influence the accuracy change of the model, which is named as out-of-bag. If the feature is important, disordering the values can reduce the accuracy of the model (Wu & Xu, 2018a).

The idea of the mean decrease in accuracy is to disorder the values of each feature and identify the impact of the change on the accuracy change of the model, which is named as out-of-bag. If the feature is important, disordering the values can reduce the accuracy of the model (Wu & Xu, 2018a).

OLM is a regression model for ordinal dependent variables. We assume the model takes the form of Equation (3):

$$y^* = X^T \beta + \varepsilon$$  \hspace{1cm} (3)

where $y^*$ is the exact but unobserved dependent variable, $x$ is the vector of independent variables, $\varepsilon$ is the error term, and $\beta$ is the vector of regression coefficients in the estimation. Assume $y$ represents the categories of $y^*$; then, OLM uses the observations on $y$, which are a form of censored data on $y^*$, to fit the parameter vector $\beta$. It should be noted that the OLM is essentially a linear classifier. The outlier can greatly impact the effect of the OLM. But in this article as the data were already pre-processed, the influence of data variance on the OLM should be very limited.

**Results**

The results of importance ranking of 16 independent variables with RF are illustrated in Figure 1. The top 5 prioritized factors influencing crash severity were annual average daily traffic (AADT), number of lanes, road surface, adverse weather, and speeding. Removing AADT can decrease the accuracy by 26.0%, indicating AADT has the highest influence on HMT crash severity (similar to non-HMT crashes). Driver age and drug/alcohol influence have limited impact on crash severity as removing driver age and drug/alcohol influence can only decrease the accuracy by 2.5% and 2.6%, respectively. As drug/alcohol usage is usually strictly prohibited for drivers of HMT by the law and regulations of HMT companies, the percentage of drug/alcohol influence is pretty low in the records (0.8%). Therefore, drug/alcohol influence showed less importance in Figure 1.

Previous studies (Goniewicz et al., 2016; Iranitalab et al., 2018) have found that road surface condition and weather have a strong correlation with each other. “Road surface” was excluded from the potential contributing factors to eliminate the collinearity among the independent variables. As drug/alcohol influence and driver age have limited influence on the mean decrease in accuracy, those two variables were also removed before OLM regression. The package “polr” in R was used to implement the model. As there are 13 independent variables involved in the initial model, it is necessary to examine the significance of their influence on HMT crash severity. The analysis of variance (ANOVA; Van Houten & Malenfant, 2001) is used for null hypothesis examination in this article. Table 2 shows the results of ANOVA of the initial model.

The results in Table 2 show that a lot of independent variables have a $p$ value higher than .1, indicating the null hypothesis could not be rejected. Only AADT and fatigues/asleep showed a $p$ value at .05 significant level. Apparently, some unnecessary predictors exist in the current OLM model. Unnecessary predictors can add noise to the estimation of other quantities that we are interested in and can waste degrees of freedom. The unnecessary predictors can also cause a collinearity issue. A backward elimination method was applied for variable selection (Michael Shenoda, 1998). The $p$ value and Akaike information criterion (AIC) are used for predictor removal. AIC is an estimator of the relative quality of statistical models for a given set of data, relative to each of the other models. A good model should have a lower AIC value and a lower $p$ value for each independent variable.

Our proposed backward elimination is briefly introduced as follows: (a) start with all predictors in the model; (b) remove the predictor with the highest $p$ value; (c) refit the model, calculate AIC, and go to step (b); and (d) stop when the minimum AIC is found. The generated AIC of different models is summarized in Table 3.

It is shown that Model 8 has the lowest AIC. When removing light from Model 8, the new generated Model 9 has a higher AIC. This means that light should be included in the regression model. The final selected Model 8 contains six independent variables: fatigues/asleep, speeding, number of lanes, AADT, light, and weather. The regression result is illustrated in Table 4.

Other than the six predictors, there are four intercepts (4 = number of levels in severity − 1) in the results. The order of the variables (importance from high to low) is AADT, fatigues/asleep, number of lanes, speeding, adverse weather, and light. The order is somewhat different from the results of RF in Figure 1. This also verified that the unnecessary predictors in the initial model can binary the regression results.
The estimates in the output are given in units of ordered logits or ordered log odds. For example, for a one-unit increase in fatigued/asleep (going from No to Yes), there is a 0.6919 increase in severity on the log odds scale, given all of the other variables in the model are held constant. The coefficients from the model are somewhat difficult to interpret because they are scaled in terms of logs. Another way to interpret OLM is to convert the coefficients into odds ratios. Table 5 illustrates the odds ratios of the independent variables. As for fatigued/asleep, for a one-unit increase in fatigued/asleep, that is, going from 0 (No) to 1 (Yes), the odds of increased severity is 2.00, given that all of the other variables in the model are held constant. This means that fatigued/asleep can increase the severity of HMT crashes. The odds ratio of speeding is 1.31, indicating that speeding can increase HMT crash severity. This is common sense. The odds ratio of the number of lanes is less than 1.00. These results showed that the presence of more lanes on the road can decrease HMT crash severity as there are conflicts between HMT vehicles and other vehicles. When AADT is less than 50,000, the odds ratio varies from 1.10 to 1.42, indicating crashes tend to be more severe with increased AADT. When AADT is more than 50,000, the odds ratio drops to 0.94. This may be caused by the speed reduction in heavy AADT (Rangel et al., 2013). The calculated odds ratio for light indicated that darkness can increase the severity of HMT crashes. The adverse weather has a 0.86 odds ratio. The result shows that adverse weather is not necessary to increase the severity of HMT crashes. It should be noted that HMT drivers are required to have a commercial driver license (CDL). Drivers are usually well trained to drive under adverse weather. At this point, HMT drivers’ behavior is different from non-HMT drivers’ behavior.

Table 2. ANOVA of Initial Model.

| Response: Severity | LR $\chi^2$ | df | Pr($>\chi^2$) | Significance |
|--------------------|-------------|----|--------------|-------------|
| Driver gender      | 0.5371      | 1  | 0.46365      |             |
| Fatigues/Asleep    | 5.0991      | 1  | 0.02394      | **          |
| Physical impairment| 0.0561      | 1  | 0.81269      |             |
| Speeding           | 2.5286      | 1  | 0.1118       |             |
| Horizontal alignment| 0.3844     | 1  | 0.53527      |             |
| Vertical alignment | 0.0514      | 1  | 0.82069      |             |
| Divided            | 0.0069      | 1  | 0.93364      |             |
| Number of lanes    | 4.3591      | 2  | 0.11309      |             |
| AADT               | 11.98       | 4  | 0.0175       | **          |
| Light              | 3.1552      | 1  | 0.07568      | *           |
| Intersection related| 0.5886     | 1  | 0.44298      |             |
| Rural/Urban        | 1.2032      | 1  | 0.27268      |             |
| Adverse weather    | 3.429       | 1  | 0.06406      |             |

Note. ANOVA = analysis of variance; LR = likelihood ratio; AADT = annual average daily traffic.

*p < .1. **p < .05.

Table 3. Summary of AIC.

| Model ID | Predictor removal | AIC      |
|----------|-------------------|----------|
| 1        | None              | 4,687.59 |
| 2        | Divided           | 4,666.766|
| 3        | Divided; physical impairment | 4,664.824|
| 4        | Divided; physical impairment; vertical alignment | 4,662.872|
| 5        | Divided; physical impairment; vertical alignment; horizontal alignment | 4,661.323|
| 6        | Divided; physical impairment; vertical alignment; horizontal alignment; driver gender | 4,659.888|
| 7        | Divided; physical impairment; vertical alignment; horizontal alignment; driver gender; intersection related | 4,658.428|
| 8        | Divided; physical impairment; vertical alignment; horizontal alignment; driver gender; intersection related; rural/urban | 4,657.57|
| 9        | Divided; physical impairment; vertical alignment; horizontal alignment; driver gender; intersection related; rural/urban; light | 4,658.08|

Note. AIC = Akaike information criterion.
Conclusion

The factors influencing HMT crash severity levels were investigated in this article. A total of 2,340 crashes were used for analysis. RF was used to initially examine the relationship between 16 suspected factors and HMT crash severity. And it was found that AADT had the highest influence on HMT crash severity, whereas driver age and drug/alcohol influence had limited impact among the investigated factors. The OLM was applied for regression. A backward elimination process was proposed to filter the important factors. According to the regression results, AADT, fatigues/asleep, number of lanes, speeding, adverse weather, and light are the six most important factors influencing HMT crash severity. Different from the non-HMT crashes, driver factor (e.g., driver age, gender, and drug/alcohol influence) was found to be not significantly related to crash severity. AADT is another important factor influencing the severity. The impact of AADT on HMT crash severity does not follow a linear trend. In general, the increasing AADT (<50,000) generates more severe crashes. However, when the AADT is heavy (>50,000), the severity of crashes decreases. The results in this method being used for HMT route selection.

Discussion

As the trip of HMT is usually long, fatigues/asleep is a critical issue and can greatly increase the severity of HMT crashes. To reduce fatigues/asleep, the HMT companies or HMT policy-makers may need to force the drivers to rest after they reach a pre-defined working time. Speeding should be strictly forbidden for HMT drivers, considering the potentially increased crash severity. The number of lanes should be considered when planning the HMT routes. A higher proportion of more lanes on the road is helpful to reduce crash severity. Other geometric factors, such as the presence of intersections, divided/undivided roads, and horizontal/vertical alignment,

| Table 4. Regression Result of the Finally Selected OLM. |
|-------------------------------------------------------|
| polr(formula = severity ~ fatigues/asleep + speeding + number of lanes + AADT + light + weather) |
| **Coefficients** | **Estimates** | **SE** | **t value** |
| Fatigues/Asleep (Yes) | 0.6919 | 0.30679 | 2.2553 |
| Speeding (Yes) | 0.2663 | 0.14403 | 1.8489 |
| Number of lanes (3–6) | −0.1094 | 0.17639 | −0.6203 |
| Number of lanes (7–12) | −0.207 | 0.10219 | −2.026 |
| AADT (5,001–10,000) | 0.3444 | 0.17957 | −1.9181 |
| AADT (10,001–20,000) | 0.3482 | 0.11038 | 3.1546 |
| AADT (20,001–50,000) | 0.111 | 0.11033 | 1.0058 |
| AADT (50,000+) | −0.0639 | 0.11057 | −0.5779 |
| Light (Yes) | −0.1248 | 0.07825 | −1.5949 |
| Adverse weather (Yes) | −0.1529 | 0.08552 | −1.7874 |
| **Intercepts:** | **Estimates** | **SE** | **t value** |
| PDO|C | 0.1835 | 0.2425 | 0.7568 |
| C|B | 0.9731 | 0.2432 | 4.0012 |
| B|A | 2.2556 | 0.2514 | 8.9714 |
| A|P | 3.3458 | 0.2774 | 12.0593 |

Residual deviance: 4,629.57  
AIC: 4,657.57

Note. OLM = ordered logit model; PDO = Property Damage Only (P); AIC = Akaike information criterion; AADT = annual average daily traffic;  
C = Complaint of Pain; B = Other Visible Injury; A = Severe Injury.

| Table 5. Odds Ratios of the Independent Variables. |
|--------------------------------------------------|
| Variables | Odds ratio |
| Fatigues/Asleep (Yes) | 2.00 |
| Speeding (Yes) | 1.31 |
| Number of lanes (3–6) | 0.90 |
| Number of lanes (7–12) | 0.81 |
| AADT (5,001–10,000) | 1.10 |
| AADT (10,001–20,000) | 1.42 |
| AADT (20,001–50,000) | 1.12 |
| AADT (50,000+) | 0.94 |
| Light (Yes) | 0.88 |
| Adverse weather (Yes) | 0.86 |

AADT = annual average daily traffic.
have limited influence on crash severity. It should be noted that although crash severity may not increase when AADT becomes high, the high AADT can definitely increase the crash frequency based on the *Highway Safety Manual* (Goniewicz et al., 2016). Increasing the level of lighting can help reduce the number of severe crashes. This is in good agreement with the finding by Pour-Rouholamin and Jalayer (Pradhananga et al., 2010). As for weather, the odds ratio is less than 1, indicating the adverse weather may not increase the number of severe crashes. This is somewhat different from general crashes.

In this research, we applied the combination of RF and OLM for the analysis of HMT crash severity. The other methods may also achieve similar or even higher accuracy compared with RF-OLM. The performance of other methods requires further investigation. This article did not consider the influence of releasing hazardous materials into the nearby population. For future research, the population along the route should also be considered for HMT route planning. The HSIS did not contain different accident scenarios (e.g., toxic release, fire, explosion) and potential environmental damage; more data need to be collected to analyze the pollution related to HMT crashes. It should be noted that the same factor may have different effects on crash severity in different observations; adding random coefficients to the OLM can be a possible way to solve this limitation in future studies. Comparison of the results of this article with the findings from different countries will be conducted in the next step.

**Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Natural Science Foundation of Jiangsu Province (BK2020040885).

**ORCID iD**

Jianqing Wu https://orcid.org/0000-0003-2844-6082

**References**

Bener, A., Abu-Zidan, F. M., Bensiali, A. K., Al-Mulla, A. A., & Jadaan, K. S. (2003). Strategy to improve road safety in developing countries. *Saudi Medical Journal*, 24(6), 603–608.

Bochner, B. S. (1998). Automated enforcement reduces crashes. *ITE Journal*, 68(6), 12.

Chen, F., & Chen, S. (2011). Injury severities of truck drivers in single- and multi-vehicle accidents on rural highway. *Accident Analysis & Prevention*, 43(5), 1677–1688.

Chen, F., Song, M., & Ma, X. (2019). Investigation on the injury severity of drivers in rear-end collisions between cars using a random parameters bivariate ordered probit model. *International Journal of Environmental Research and Public Health*, 16(14), 2632.

Craft, R. (2004). Crashes involving trucks carrying hazardous materials. Federal Motor Carrier Safety Administration, U.S. Department of Transportation.

Das, A., Abdel-Aty, M., & Pande, A. (2009). Using conditional inference forests to identify the factors affecting crash severity on arterial corridors. *Journal of Safety Research*, 40(4), 317–327.

Delen, D., Sharda, R., & Bessonov, M. (2006). Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accident Analysis & Prevention*, 38(3), 434–444.

Fabiano, B., Currò, F., Palazzi, E., & Pastorino, R. (2002). A framework for risk assessment and decision-making strategies in dangerous good transportation. *Journal of Hazardous Materials*, 93, 1–15.

Fabiano, B., & Palazzi, E. (2010). HazMat transportation by heavy vehicles and road tunnels: A simplified modelling procedure to risk assessment and mitigation applied to an Italian case study. *International Journal of Heavy Vehicle Systems*, 17(3–4), 216–236.

Goniewicz, K., Goniewicz, M., Pawłowski, W., & Fiedor, P. (2016). Road accident rates: Strategies and programmes for improving road traffic safety. *European Journal of Trauma and Emergency Surgery*, 42(4), 433–438.

Harwood, D. W., Russell, E. R., & Viner, J. G. (1989). Characteristics of accidents and incidents in highway transportation of hazardous materials. *Transportation Research Record*, 1245, 23–33.

Huang, H., Chin, H. C., & Haque, M. M. (2008). Severity of driver injury and vehicle damage in traffic crashes at intersections: A Bayesian hierarchical analysis. *Accident Analysis & Prevention*, 40(1), 45–54.

Iranitalab, A., Kang, Y., & Khattak, A. (2018). Modeling the probability of hazardous materials release in crashes at highway–rail grade crossings. *Transportation Research Record*, 2672(10), 28–37.

Iranitalab, A., & Khattak, A. (2018). Train-level and car-level modeling of hazardous materials release in railroad incidents. *Transportation Research Record*, 2672(9), 249–260.

Kaplan, S., & Prato, C. G. (2012). The application of the random regret minimization model to drivers’ choice of crash avoidance maneuvers. *Transportation Research Part F: Traffic Psychology and Behaviour*, 15(6), 699–709.

List, G. F., Mirchandani, P. B., Turnquist, M. A., & Zografos, K. G. (1991). Modeling and analysis for hazardous materials transportation: Risk analysis, routing/scheduling and facility location. *Transportation Science*, 25(2), 100–114.

Ma, C., Hao, W., Pan, F., & Xiang, W. (2018). Road screening and distribution route multi-objective robust optimization for hazardous materials based on neural network and genetic algorithm. *PLOS ONE*, 13(6), Article e0198931. https://doi.org/10.1371/journal.pone.0198931

Ma, C., Li, Y., He, R., Wu, F., Qi, B., & Ye, Q. (2013). Route optimisation models and algorithms for hazardous materials transportation under different environments. *International Journal of Bio-Inspired Computation*, 5(4), 252–265.

Ma, J., & Kockelman, K. M. (2006). Bayesian multivariate Poisson regression for models of injury count, by severity. *Transportation Research Record*, 1950(1), 24–34.
Michael Shenoda, P. E. (1998). Heuristic and metaheuristic approaches for optimizing traffic signal. *ITE Journal*, 58(5). https://www.huichawaii.org/assets/shenoda%2C-michael.pdf

Nam, D., & Mannering, F. (2000). An exploratory hazard-based analysis of highway incident duration. *Transportation Research Part A: Policy and Practice*, 34(2), 85–102.

Nicolet-Monnier, M., & Gheorghe, A. V. (2013). *Quantitative risk assessment of hazardous materials transport systems: Rail, road, pipelines and ship* (Vol. 5). Springer.

Oggero, A., Darbra, R. M., Munoz, M., Planas, E., & Casal, J. (2006). A survey of accidents occurring during the transport of hazardous substances by road and rail. *Journal of Hazardous Materials*, 133(1–3), 1–7.

Pour-Rouholamin, M., & Jalayer, M. (2016). Analyzing the severity of motorcycle crashes in North Carolina using highway safety information systems data. *ITE Journal*, 86(10), 45–49.

Pradhananga, R., Taniguchi, E., & Yamada, T. (2010). Ant colony system based routing and scheduling for hazardous material transportation. *Procedia – Social and Behavioral Sciences*, 2(3), 6097–6108.

Rangel, T., Vassallo, J. M., & Herraiz, I. (2013). The influence of economic incentives linked to road safety indicators on accidents: The case of toll concessions in Spain. *Accident Analysis & Prevention*, 59, 529–536.

Savolainen, P. T., Mannering, F. L., Lord, D., & Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. *Accident Analysis & Prevention*, 43(5), 1666–1676.

Schmidt, J., & Price, D. (1979). The flow of hazardous materials on Virginia highways. *Journal of Safety Research*, 11, 109–114.

Uddin, M., & Huynh, N. (2018). Factors influencing injury severity of crashes involving HAZMAT trucks. *International Journal of Transportation Science and Technology*, 7(1), 1–9.

Van Houten, R., & Malenfant, J. L. (2001). ITS animated LED signals alert drivers to pedestrian threats. *ITE Journal*, 71(7), 42–46.

Weng, J., Du, G., Li, D., & Yu, Y. (2018). Time-varying mixed logit model for vehicle merging behavior in work zone merging areas. *Accident Analysis & Prevention*, 117, 328–339.

Wilkinson, H. L. (2018). A study of severe air crash environments with particular reference to the carriage of radioactive materials by air. In G. B. Guy (Ed.), *Reliability on the move* (pp. 109–125). Chapman & Hall/CRC.

Wu, J., & Xu, H. (2018a). Driver behavior analysis on rural 2-lane, 2-way highways using SHRP 2 NDS data. *Traffic Injury Prevention*, 19(8), 838–843. https://doi.org/10.1080/15389588.2018.1524142

Wu, J., & Xu, H. (2018b). The influence of road familiarity on distracted driving activities and driving operation using naturalistic driving study data. *Transportation Research Part F: Traffic Psychology and Behaviour*, 52, 75–84.

Wu, J., Xu, H., Zheng, Y., Liu, W., Sun, Y., Yue, R., & Song, X. (2018). Driver behavior fault analysis on ramp-related crashes/near-crashes using SHRP 2 naturalistic driving study data. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 2134–2139). Institute of Electrical and Electronics Engineers. https://ieeexplore.ieee.org/document/8569985

Zografos, K. G., & Androutsopoulos, K. N. (2008). A decision support system for integrated hazardous materials routing and emergency response decisions. *Transportation Research Part C: Emerging Technologies*, 16(6), 684–703.

Zografos, K. G., & Davis, C. F. (1989). Multi-objective programming approach for routing hazardous materials. *Journal of Transportation Engineering*, 115(6), 661–673.