International Journal of Rail Transportation

ISSN: 2324-8378 (Print) 2324-8386 (Online) Journal homepage: https://www.tandfonline.com/loi/tjrt20

Modeling railroad trespassing crash frequency using a mixed-effects negative binomial model

Y. Kang, A. Iranitalab & A. Khattak

To cite this article: Y. Kang, A. Iranitalab & A. Khattak (2019) Modeling railroad trespassing crash frequency using a mixed-effects negative binomial model, International Journal of Rail Transportation, 7:3, 208-218, DOI: 10.1080/23248378.2018.1550626

To link to this article: https://doi.org/10.1080/23248378.2018.1550626

INTERNATIONAL JOURNAL OF RAIL TRANSPORTATION
2019, VOL. 7, NO. 3, 208–218
https://doi.org/10.1080/23248378.2018.1550626

Modeling railroad trespassing crash frequency using a mixed-effects negative binomial model

Y. Kang, A. Iranitalab and A. Khattak

Department of Civil Engineering and Nebraska Transportation Center, University of Nebraska-Lincoln, Lincoln, NE, USA

ABSTRACT

A better understanding of rail trespass crashes is needed as more than 400 trespassing related fatalities occur along rail tracks each year in the United States (U.S.). The objective of this research was to investigate factors associated with the occurrence of rail trespass crashes. Yearly crash frequency for counties in the U.S. with train tracks was modeled using a Mixed-effects Negative Binomial Model based on 2012–2016 datasets from the Federal Railroad Administration, the U. S. Census Bureau and National Historical Geographic Information System. Results revealed that key factors affecting rail trespassing crashes include county population density, length of rail tracks in a county, median age and male proportion of the county population, and average train traffic within a county. The findings provided useful information on improving public safety along railroad tracks.

ARTICLE HISTORY
Received 4 March 2018
Revised 16 November 2018
Accepted 17 November 2018

KEYWORDS
Rail safety; trespassing; mixed-effects negative binomial

1. Introduction

The United States (U.S.) rail network consists of about 155,000 miles of operating routes. An important element in the transportation network is the highway-rail grade crossing (HRGC), which facilitates the movement of trains and highway traffic. Investments and dedicated efforts in improving railroad safety at HRGCs has reduced crash occurrence and injury severities at these important junctions in the transportation network. However, trespassing on railroad property and facilities and the resulting crashes need attention in the interest of improving public safety along railroad tracks. According to Federal Railroad Administration (FRA), everyday there are two trespassing fatalities or injuries on railroad property on average and more than 412 people died as a result of rail trespassing in 2011 [1]. The impacts of crashes caused by rail trespassing are often high as they not only lead to injuries and fatalities but also affect railroad traffic.

Research on railroad trespassing has not received as much attention in transportation safety literature as some of the other aspects of railroad safety. For example, various statistical models have been utilized to investigate crash frequencies at HRGCs and a plethora of factors affecting crash frequencies and injuries identified [2,3]. In comparison,

CONTACT A. Khattak khattak@unl.edu Department of Civil Engineering and Nebraska Transportation Center, University of Nebraska-Lincoln, Lincoln, NE, USA

little is known about the factors on the frequency of trespassing-related crashes. Related issues addressed by some previous studies included train-pedestrian crashes, measures on preventing trespassing, etc [2]. There is lack of statistical analysis on the particular issue of trespassing-related crashes in the literature. Since 2011, FRA started including spatial information in its trespasser casualty database, which can uncover useful information for safety research [4]. The study described herein attempted to investigate contributing factors to rail trespassing crash frequency by estimation of a Mixed-effect Negative Binomial Model (MNBM). The benefit of MNBM is accounting for correlations among county-based frequency of trespassing crashes, which are also referred to as unobserved heterogeneity. As addressed by various research studies, unobserved heterogeneity remains an important issue in many traffic safety datasets [2,3]. Failure to consider these correlations within the data can lead to inaccuracies when estimating and interpreting models.

A review of previous studies is presented after the introduction, followed by methodology in the third section. The data and the descriptive statistics are presented in the fourth section, which also explains the process of merging data from different sources. In the fifth section the modeling results are reported and interpreted. In the end, conclusions complete the paper.

2. Literature review

As one of the leading causes of injuries and fatalities in rail-related crashes, train-pedestrian crashes are a major concern in railway safety [2]. While train-pedestrian crashes that occurred at HRGCs are investigated in the literature, trespassing at non-crossings is significantly different with relatively few similarities in rail characteristics [3]. In this section, previous literature covering trespassing safety is discussed. In addition, background information on application of mixed models in transportation safety is presented.

2.1. *Rail trespassing studies*

Trespassing on rail property is a worldwide concern. A European Union project, RESTRAIL (REduction of Suicides and Trespasses on RAILway property) focused on the reduction of frequency and impact of suicide and trespass-related incidents [5–9]. A decision-making system was specifically developed to deal with trespassing incidents for railway authorities. By conducting pilot studies at two sites in Finland, Kallberg and Silla suggested that automatic sound warning systems could reduce trespassing occurrence by 18% and 44% when fencing is not a viable option [8]. They also found that local circumstances and safety cultures influenced people's reasons for trespassing on rail property, which included the use of shortest or fastest route and recreational purposes, etc.

The intent of most trespassers is typically looking for the shortest and fastest route or using an existing path [10], while some trespassing behaviors are suicidal. Suicidal trespassing is fundamentally different since people purposefully endanger themselves. However, Limosin et al. showed that in terms of train-person incidents, suicides and accidental crashes did not show much difference in the mechanism of injury and severity of consequences [11].

With respect to characteristics of train-person crashes, Silla and Luoma compared data from Finnish railroads with Sweden and found that most incidents occurred at the end of week [12]. It was also concluded that fatalities were more likely to happen in densely populated areas, and younger groups were more likely to be involved in train-pedestrian crashes. Patterson examined a database in New Zealand from 1994 to 2003, and statistically summarized various features which included causes of the incidents, locations, times, gender, severity, ethnicity and contributing factors [13]. The research concluded that trespassers were most likely to be killed or injured during daytime, between Wednesday and Saturday in December. Some cases possibly involved suicides or influence by alcohol based on the pre-crash behaviors, which was consistent with other studies [12–14].

The FRA trespassing casualty data provides a detailed report format regarding pre-crash behaviors, including working, walking, running, laying, sitting, etc. on or around the rail tracks. Wang et al. found that lying or sleeping were highly associated with severe injuries or fatalities [15]. Spatial modeling was also applied to estimate trespass-related crash severity factors in their study and the estimated coefficients for different variables were not stationary over space but varied across the country. Prevention of rail trespassing crashes was investigated in existing literature. Countermeasures, including surveillance, fencing, audio warning and public education, could reduce the frequency of trespassing. Though fencing can limit the public's access to railroad property and was proven to be effective in reducing rail trespassing frequency by 94.6%, it is costly and subject to vandalism [16]. Silla and Luoma used a survey to investigate people's opinions on rail trespassing and found that 68.9% of the respondents had personal experience of trespassing [17].

A number of existing studies focused on descriptive analysis of trespassing reports in addition to proposing prevention strategies, instead of in-depth statistical analyses of trespassing frequency [18,19]. As was mentioned, FRA included spatial information in the trespasser casualty dataset since 2011, and Wang et al. took advantage of the geospatial information and explored the trespass-related crash severity [15].

2.2. *Mixed models in transportation safety*

To account for the latent information in space-varying explanatory variables, data are often considered in small time and space intervals [20,21]. For instance, one may collect crash data and divide the data based on spatial codes and consider crash frequencies per area. Multiple crashes may belong to each roadway segment or intersection. These observations will be correlated over space because many of the unobserved effects remain the same for roadway segments or intersections in close proximity [21]. Similarly, there can be correlation over time. This will set up a correlation of disturbances among observations and introduce bias during the model parameter estimations [22]. To deal with the temporal and spatial correlation issues associated with such data, various methods have been applied such as generalized estimating equation, random effects models, hierarchical/multi-level models, etc [22–26]. In 1984, Hausman et al. first examined random-effects and fixed-effects negative binomial models for panel data which had temporal considerations [23].

Traffic data based on 8019 census wards in England was investigated by Wang et al. using negative binomial models [24]. To address the unobserved spatial variations (e.g. topography and weather) within England, they added a series of dummy (indicator) variables representing different regions of England into the negative binomial (NB) models. Guo et al. applied a mixed-effect model in which a corridor-specific random effect was included to account for the corridor-level correlation to a signalized intersection safety study [25]. By comparing the Deviance Information Criterion performance of various models, they concluded that mixed-effect NB model improved model fit over the simple NB model.

3. Methodology

Various statistical models have been used to model crash frequency data over the years, including negative binomial regression models, Poisson regression models, multiple linear regression models, etc. The negative binomial regression is an alternative to Poisson regression model that assumes the Poisson parameter follows a gamma probability distribution to avoid over-dispersion, i.e., the Poisson distribution requirement that the mean and variance of the variable under consideration be equal [26,27]. This research utilized a variation of the negative binomial (NB) regression model, the mixed-effect negative binomial model (MNBM) for modeling the frequency of rail trespassing crashes.

On the basis of a national level database, frequency of rail trespassing crashes reported within the same county in different years are potentially more ‘similar’, compared to those reported in other counties. Mixed effects models can accommodate correlated data, by introducing a random county-specific effects term into the relationship between the explanatory variables and the response variable. The specifications are as follows:

$$\tilde{\mu}_{ij} = \mu_{ij}\delta_i \tag{1}$$

where $\tilde{\mu}_i$ denotes the expected number of crashes, δ_i is the random location-specific effect. Similar to NB, the value is guaranteed non-negative values by assuming:

$$\log \tilde{\mu}_{ij} = X_{ij}\beta + u_i \tag{2}$$

where β denotes the estimated coefficients, $\exp(u_i)$ is gamma-distributed error term with mean 1 and variance k, where k is also the overdispersion parameter in the NB model. If k is significantly different from zero, the data are over-dispersed or under-dispersed. The resultant relationship is given by:

$$Var\left(n_{ij}\right) = E\left(n_{ij}\right)\left[1 + kE\left(n_{ij}\right)\right] \tag{3}$$

The number of crashes n_{ij} has a negative binomial distribution, for a given county i and year j, with parameters $\mu_{ij}\delta_i$ where $\mu_{ij} = \exp(X_{ij}\beta)$. Hence n_{ij} has mean $\mu_{ij}\delta_i/\phi_i$ and variance $\left(\mu_{ij}\delta_i/\phi_i\right)/z$, where $z = 1/\left(1 + \delta_i/\phi_i\right)$., z is assumed to be a beta-distributed random variable with parameters (a, b), in order to illustrate the variation of location over time. Based on the derivation from Hausman et al. [23], the probability density function of the MNBM for county $i$is:

$$P(n_{i1},\ldots,n_{iJ}|X_{i1},\ldots,X_{iJ}) = \frac{\sqrt{a+b}\sqrt{a+\sum_J \mu_{ij}}\sqrt{b+\sum_J \mu_{ij}}}{\sqrt{a}\sqrt{b}\sqrt{a+b+\sum_J \mu_{ij}+\sum_J n_{ij}}} * \prod_J \frac{\sqrt{\mu_{ij}+n_{ij}}}{\sqrt{\mu_{ij}}\sqrt{n_{ij}+1}} \tag{4}$$

As explained, in the MNBM, random effects were added to NB model structure to address the correlations (also referred to as unobserved heterogeneity). Hence, the model can be also derived as follows:

$$\log \tilde{\mu}_{ij} = \log E_{ij}^{\rho} e^{\lambda_{ij}} \tag{5}$$

$$\lambda_{ij} = X_{ij}'\beta + b_i \tag{6}$$

where E_{ij} denotes the total train traffic volume, ρ is the exposure coefficient of train traffic to the population; b_i is a county-specific random effect that conforms to a normal distribution with mean 0 and variance τ. In this study, the estimations of coefficient vectors, b_i and evaluation results were reported [28].

Likelihood Ratio Tests (LRT) were used in this study to investigate the significance of estimated regression coefficients for each explanatory variable [29]. The LRT statistic approximates a χ^2 distribution (with degrees of freedom equal to the number of estimated parameters) and is written as

$$LRT = -2\log\left[\frac{L_s(\hat{\beta}^{(0)}|y_1,\ldots,y_n)}{L_g(\hat{\beta}^{(a)}|y_1,\ldots,y_n)}\right] \tag{7}$$

where L_s and L_g indicate the estimated probabilities of success from estimating different models, with vectors of parameters $\hat{\beta}^{(0)}$ and $\hat{\beta}^{(a)}$ respectively [30]. Akaike information criterion (AIC) and Bayesian information criterion (BIC) were used for model and variable selection. Percentage Change (PC) was adopted as a tool for model interpretation: a c-unit increase in the explanatory variable results in a $100(e^{c\beta_1}-1)\%$ change in the response variable, in which c was selected accordingly.

4. Study data

The study utilized different datasets pertaining to the 2012–2016 period that were aggregated to obtain the analysis dataset. The individual datasets included the highway rail crossing inventory dataset and railroad casualties (Form 6180.55A) obtained from FRA, the U.S. railroad information, census data and county boundaries obtained from U.S. Census Bureau and National Historical Geographic Information System, were merged together using unique county FIPS codes [31,32].

The trespassing data included crash characteristics, information of trespasser, pre-crash behavior, nature of injury and location (latitude and longitude, county, state). Among these, as a crash frequency study, location information of the crashes was used in the analysis. Average train traffic volume information was extracted from FRA rail

Table 1. Variable descriptions.

Variable	Name	Average	Minimum	Maximum	Standard Deviation
Number of trespass-related crashes in a year	Count_by_year	0.43	0	43	1.4
Number of trains per day	AveTrainVolume	3.53	0	58.91	7.09
Rail line length (miles)	RailLength	106.8	0.20	1194	90.04
Median household income ($)	HHincome	47,070	19,330	112,600	11,775.8
Population Density (people/square mile)	PopDensity	312.80	0.53	70,450	2022.29
Median age	MedianAge	40.33	23.20	65.30	4.87
Percentage of male population	MaleRatio	0.50	0.41	0.73	0.02

crossing inventory database, by averaging the train traffic of all the crossings within each county. This represents an average of the train traffic at rail crossings in a county as actual county-based train traffic data was not readily available. The length of railroad within each county was calculated using the US railroad data and county boundary data obtained from the Census Bureau. And lastly the census provided the socio-demographics of each county, including population, median household income, age, gender, etc. The data were sorted by county and year, and yielded 11,155 observations in total. Table 1 presents the descriptive statistics of the variables used in this study.

5. MNBM modeling results

This section presents the results of the MNBM. Fixed effects were considered for all the explanatory variables, while random effects were accounted for by the model intercept, *RailLength*, and *PopDensity*. Different combinations of variables with random effects were tested in the model and estimated models compared based on AIC and BIC, along with an LRT for the random effects. These three random effects were found as statistically significant based on the LRTs, and keeping all the explanatory variables in the model resulted in the lowest value in AIC and BIC (the lower values of these two criteria indicate a better model [33]). Table 2 presents the estimated coefficients and their standard errors, p-values for LRTs, and percentage change-related measures including c, point estimates and 95% confidence intervals. The table also includes standard deviations for the variables with random effects. Figure 1 presents the correlation matrix for the fixed effects. Though the correlation between intercept and *MaleRatio* was high, it does not imply a correlation between an unknown parameter being estimated. It simply resulted from positive covariate values. For easier observation, Figure 2 also illustrates the PC point estimates and the 95% confidence intervals.

With 95% confidence and keeping all the variables constant except the variable that is subject to interpretation, the model interpretation is as follows.

Length of railroad tracks was statistically significant in the model. A five-mile increase in the length of rail tracks in a county resulted in a percentage between 3.24% to 4.05% increase in the mean annual frequency of rail trespassing crashes. This variable was taken into account because increase in the length of railroad tracks essentially raises trespassing exposure while previous literature merely considered surrounding population and train frequencies [8,15].

Household income was the only variable that was not statistically significant based on LRT and percentage change 95% confidence interval (it included 0). Therefore,

Table 2. Mixed-effects negative binomial model results.

						Percentage Change		
							95% Confidence Interval	
Variables	Estimated Coefficient	Standard Error	LRT p-value		c	Point Estimate	Lower Bound	Upper Bound
Fixed Effects								
(Intercept)	7.71700	1.19400	–		–	–	–	–
RailLength	0.00716	0.00040	0.00000	***	5	3.64	3.24	4.05
HHincome	0.00001	0.00000	0.20140		5000	2.86	−0.17	5.99
PopDensity	0.00063	0.00008	0.00000	***	100	6.47	4.80	8.18
MedianAge	−0.06588	0.00767	0.00000	***	5	−28.06	−33.28	−22.45
MaleRatio	−16.86000	2.28700	0.00000	***	0.01	−15.52	−19.22	−11.64
AveTrainVolume	0.05247	0.00448	0.00000	***	1	5.39	4.47	6.32
Random Effects $\sim N(0, \sigma^2)$								
(Intercept)	$\sigma =$ 1.03710							
RailLength	$\sigma =$ 0.00149							
PopDensity	$\sigma =$ 0.00051							

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
– : Not Applicable
c: coefficient in percentage change

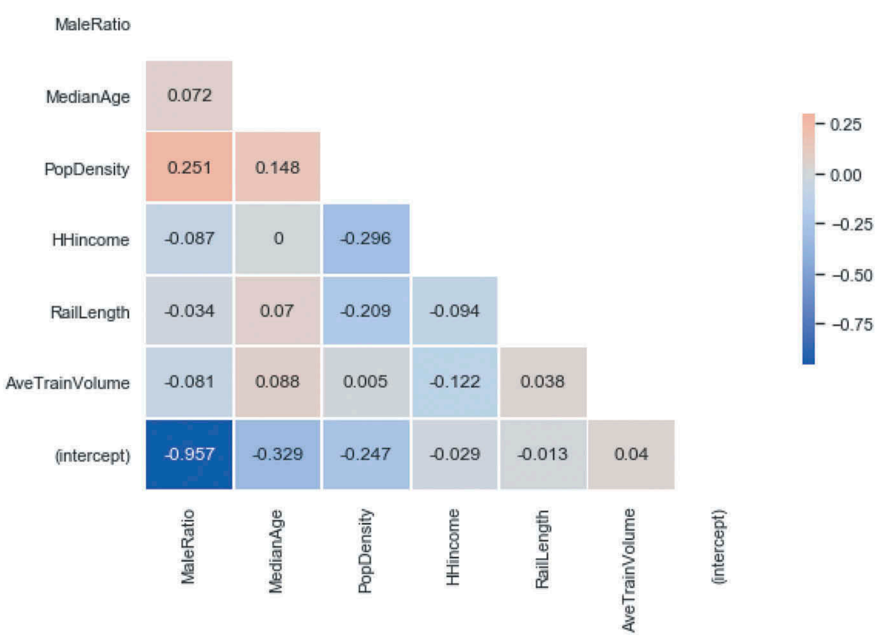

Figure 1. Correlation matrix for fixed effects.
(The annotation indicates correlation coefficient value; warm/cool colour shows positive/negative value)

sufficient evidence was not available to support a relationship between median annual household income in the county and frequency of trespassing crashes.

Regarding population density, an increase of 100 people per square-miles led to an increase of 4.80% to 8.18% in the mean trespassing crash frequency. This finding is consistent with previous studies [12,17], and indicates that trespassing behaviors concentrate in areas with higher population density.

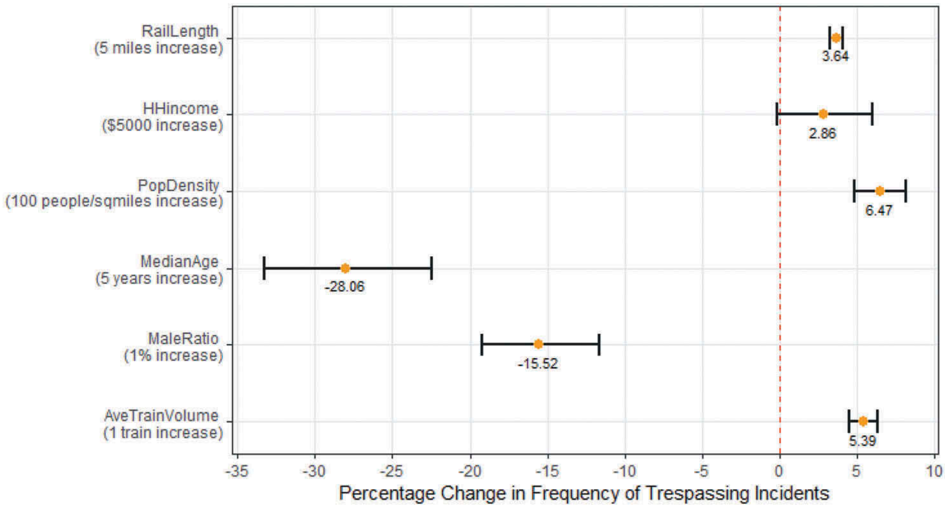

Figure 2. Percentage change and 95% confidence intervals.

The median age within a county was also statistically significant in the model. A five-year increase in the age of trespasser decreased the mean trespassing crash frequency per county by 22.45% to 33.28%. The finding is that younger age groups are more prone to rail trespassing crash involvement. This is similar to findings from other studies; Patterson suggested that around 60 percent of the trespassers killed were younger than 40 years [13]. And a survey conducted by Silla and Luoma indicated people aged 10–29 years were involved in 51.4% of trespass incidents [17].

The model results showed that a 1% increase in the proportion of males in the county population decreased rail trespassing crash frequency by 11.64% to 19.22%. However, previous research reported that males had less safe attitudes and self-stated behavior with regards to trespassing (63% of the trespassers were men) [10,17]. This could be due to the small size of observations considered in previous studies. Also, FRA trespasser casualty dataset do not reveal the gender of the trespassers, and proportion of males in each county was used in the model.

The effects of train traffic were captured by averaging the number of daily trains passing the highway-rail grade crossings within counties as an approximation, since nationwide train traffic volume for all U.S. counties was not available. An increase of one train per day in the average county train traffic led to an increase in the mean of trespassing crash frequency from 4.47% to 6.32%. The result appears valid intuitively as greater train traffic in a county may result in more trespass-related crashes.

6. Discussions and conclusions

This study examined contributing factors affecting rail trespassing crash frequencies in the U.S. using a Mixed-effect Negative Binomial Model. Based on the 11,155 observations from 2012–2016 nationwide rail trespass crashes, the modeling results may be used for spatial prioritization of countermeasures and policy-making regarding rail trespassing crashes. Counties or areas with relatively longer rail tracks, higher population density, younger population, higher female proportion in population and with higher average train traffic may be considered a priority in rail trespassing safety improvements. These areas may be the first candidates for fencing, monitoring and using detection systems, signage, lighting systems, education and enforcement, and emergency response preparation. It can be also noted that, pre-crash action was considered a key factor pertaining to trespass-related crash injury severities [15]. However, this factor cannot be considered in the modeling because this study is concerned about frequency. Implementation of the policies, in terms of enforcement, punishment and patrol assignment may be also considered in these areas. Female youth may be considered for education regarding avoidance of rail trespassing on a priority basis.

According to previous studies, suicides constitute a significant number of rail trespassing crashes. In the rail trespassing dataset used in this study, determination of suicide/non-suicide crashes was not included even though the narratives of some instances implied suicidal incidents. Therefore, these crashes were not differentiated in the estimated models. This might affect the estimated model parameters and is a limitation of this study. Also, other potentially significant factors on frequency of rail trespassing crashes were missing in the dataset and thus not used in this study. These included rail-related factors such as train speed, length of trains, available safety devices, proportion of different types of trains (i.e. freight/passenger trains), and demographic factors such as mental conditions of the population, consumption of alcohol and drugs and type of land-use in the vicinity of crash locations. This resulted in the limited number of explanatory variables available for use in this study. Another limitation is that while counties were used as the unit of analysis, areas within a county may not be necessarily similar in terms of rail trespassing crash frequency. This may enter a level of inaccuracy in modeling results.

The usage of MNBM addressed the issue of correlation among observations, which could lead to biased results and conclusions given the use of conventional methods. Even though the included random parameters may not have conclusive interpretations, the proposed model is an improvement in capturing the nature of the incidents. Future studies may attempt to address the aforementioned limitations of this research by obtaining more comprehensive datasets that include variables that were missing in this study and at more detailed geographic levels (e.g., city-level or census tract-level). Other data analysis methods may be explored to fully investigate rail trespassing crashes. Suicide/non-suicide crashes may be studied separately and a comparison of the results may lead to uncovering of useful information. Severity of rail trespassing crashes is another topic that may be investigated for further improvement of public safety.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

[1] Federal Railroad Administration (FRA). Rail trespassing fact sheet. Washington, DC: U.S. Department of Transportation; 2013.
[2] Zhao S, Iranitalab A, Khattak A. A clustering approach to injury severity in pedestrian-train crashes at highway-rail grade crossings. J Transp Saf Secur. 2018;5:1–8.
[3] Kang Y, Khattak A. A cluster based approach to analyze crash injury severity at highway-rail grade crossings. Transportation research record 2608. Washington, DC. Transp Res Board, National Res Counc. 2017;2608(1):58–69.
[4] Federal Railroad Administration (FRA). Railroad casualty user guide. Washington, DC: U.S. Department of Transportation; 2012.
[5] Havârneanu GM, Bonneau MH, Colliard J. Lessons learned from the collaborative European project RESTRAIL: rEduction of suicides and trespasses on RAILway property. Eur Transp Res Rev. 2016;8(2):16.
[6] Havârneanu GM, Burkhardt JM, Paran F. A systematic review of the literature on safety measures to prevent railway suicides and trespassing accidents. Accid Anal Prev. 2015;81:30–50.
[7] Havârneanu GM, Burkhardt JM, Silla A. Optimizing suicide and trespass prevention on railways: a problem-solving model from the RESTRAIL project. Int J Inj Contr Saf Promot. 2017;24(4):469–486.
[8] Kallberg VP, Silla A. Prevention of railway trespassing by automatic sound warning—A pilot study. Traffic Inj Prev. 2017;18(3):330–335.
[9] Burkhardt JM, Rådbo H, Silla A, et al. A model of suicide and trespassing processes to support the analysis and decision related to preventing railway suicides and trespassing accidents at railways. Transp Res Arena 2014.2014;1–9.
[10] Silla A, Luoma J. Trespassing on finnish railways: identification of problem sites and characteristics of trespassing behaviour. Eur Transp Res Rev. 2009;1(1):47–53.
[11] Limosin F, Loze JY, Cothereau C, et al. A prospective study of the psychological effects of "person under train" incidents on drivers. J Psychiatr Res. 2006;40(8):755–761.
[12] Silla A, Luoma J. Main characteristics of train-pedestrian fatalities on finnish railroads. Accid Anal Prev. 2012;45:61–66.
[13] Patterson T. Analysis of trespasser accidents. Land Transp Saf Authority, NZ. 2004;1–15.
[14] Savage I. Trespassing on the railroad. Res Transp Econ. 2007;20:199–224.
[15] Wang X, Liu J, Khattak AJ, et al. Non-crossing rail-trespassing crashes in the past decade: a spatial approach to analyzing injury severity. Saf Sci. 2016;82:44–55.
[16] Silla A, Luoma J. Effect of three countermeasures against the illegal crossing of railway tracks. Accid Anal Prev. 2011;43(3):1089–1094.
[17] Silla A, Luoma J. Opinions on railway trespassing of people living close to a railway line. Saf Sci. 2012;50(1):62–67.
[18] Skládaná P, Skládaný P, Tučka P, et al. Trespassing railway property - Typology of risk localities. Transp Res Procedia. 2016;14:2091–2100.
[19] Rådbo H, Andersson R. Patterns of suicide and other trespassing fatalities on state-owned railways in greater Stockholm; Implications for prevention. Int J Environ Res Public Health. 2012;9(3):772–780.
[20] Mannering FL, Bhat CR. Analytic methods in accident research: methodological frontier and future directions. Anal Methods Accid Res. 2014;1:1–22.
[21] Lord D, Mannering F. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transp Res A Policy Pract. 2010;44(5):291–305.
[22] Washington SP, Karlaftis MG, Mannering FL. Statistical and econometric methods for transportation data analysis. second ed. Boca Raton, FL: Chapman and Hall/CRC; 2010.
[23] Hausman J, Hall BH, Griliches Z. Econometric models for count data with an application to the patents-R&D relationship. Econometrica. 1984;52(4):909–938.
[24] Wang C, Quddus M, Ison S. The effects of area-wide road speed and curvature on traffic casualties in England. J Transp Geogr. 2009;17(5):385–395.

[25] Guo F, Wang X, Abdel-Aty MA. Modeling signalized intersection safety with corridor-level spatial correlations. Accid Anal Prev. 2010;42(1):84–92.
[26] Lord D, Mahlawat M. Examining application of aggregated and disaggregated Poisson-Gamma models subjected to low sample mean bias. Transp Res Rec. 2009;1(2136):1–10.
[27] Shankar V, Albin R, Milton J, et al. Evaluating median crossover likelihoods with clustered accident counts: an empirical inquiry using the random effects negative binomial model. Transp Res Rec. 1635;1998:44–48.
[28] Bates D, Maechler M, Bolker B, et al. Fitting linear mixed-effects models using lme4. J Stat Softw. 2015;67(1):1–48.
[29] John A, Weisberg S, Adler D, et al. Package car. Vienna: R foundation for statistical computing; 2012.
[30] Dickey DA, Fuller WA. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica. 1981;49(4):1057–1072.
[31] Minnesota Population Center. National historical geographic information system: version 11.0 [Database]. Minneapolis: University of Minnesota; 2016.
[32] Federal Railroad Administration (FRA). FRA guide for preparing accident/incident reports. Washington, DC: U.S. Department of Transportation; 2011.
[33] Burnham KP, Anderson DR. Multimodel inference: understanding AIC and BIC in model selection. Sociol Methods Res. 2004;33(2):261–304.