Research Article

Statistical Analysis of Train Derailment Severity for Unit Trains Versus Manifest Trains

Weitian Li[1] (ID)**, Zheyong Bian[2]** (ID)**, and Xiang Liu[1]** (ID)

Abstract

Using freight trains is a cost-effective and safe method for transporting numerous products over long distances. However, freight train derailments have significant consequences. Derailment severity, measured by the number of cars derailed per derailment, is an important risk factor. Based on prior literature and Federal Railroad Administration data, this research analyzes and estimates derailment severity of freight railroad transportation using a statistical method: the Truncated Geometric (TG) model. The methodology accounts for three train types: manifest train, loaded unit train, and empty unit train. Train length, speed, and gross tonnage per car are the input features, and the output variable is the number of cars derailed in a derailment incident. The TG model quantifies the influence of these factors contributing to derailment severity and estimates train derailment severity with a performance of the mean square error (MSE) of 68.91 and mean absolute error (MAE) of 6.05. When data outliers such as abnormally high or low severity are excluded, the MSE drops to 36.82 and the MAE drops to 4.22. Overall, the results indicate that the train derailment severity estimation performance based on the given factors is satisfactory. With all other factors being equal, a loaded unit train is likely to derail more cars than a manifest train and an empty unit train. When data outliers are excluded, there is no significant difference between manifest trains and empty unit trains as regards derailment severity, and both are less likely to derail than a loaded unit train.

Keywords

rail, freight rail transportation, freight, rail safety

Railroads are some of society's most critical infrastructures for transporting goods across great distances. It is essential to manage and minimize railroad safety risks to avoid potentially massive loss of life and public property. Freight railroads account for about 40% of the United States' long-distance freight volume, which is more than any other mode of transportation (*1*). In recent years, there have been several incidents of train derailments on freight railways, some of which were associated with hazardous material leaks. In October 2022, twenty tanker cars of paraffin wax from a 101-car train derailed in Sandusky, Ohio (*2*). In January 2023, sixteen total cars derailed from a train and two of them leaked chemicals in Keachi, Louisiana (*3*). In February 2023, a total of thirty-eight trains derailed in East Palestine, near the Ohio–Pennsylvania border, and eleven out of twenty tank cars carrying hazardous material subsequently ignited (*4*).

Unit trains and manifest trains are the two main types of freight trains in the U.S. Unit trains are designed to carry a single commodity, such as coal or grain, and are shipped from one shared origin to one destination. Unit trains are cost-effective and efficient because of reduced operating expenses, bulk loading, improved asset utilization, and cost savings through increased scale (*5–9*). Unit trains have been used to transport an increasing proportion of hazmat over the past two decades, and unit trains accounted for approximately 21% of hazmat carloads in 2018 (*9*). Manifest trains, as opposed to unit trains, are

[1]Department of Civil and Environmental Engineering, Rutgers University-New Brunswick, Piscataway, NJ
[2]Department of Construction Management, University of Houston, Houston, TX

Corresponding Author:
Xiang Liu, xiang.liu@rutgers.edu

used for transporting a variety of goods and commodities and typically consist of different types of rail cars for multiple shippers (*10*).

According to train derailment data from the Federal Railroad Administration (FRA) rail equipment accident database, there were 2,669 unit train derailments and 5,931 manifest train derailments on mainlines from 1996 to 2021 (*11*). Zhang et al. conducted an empirical analysis of freight train derailment rates for unit trains and manifest trains using FRA mainline freight train derailment data for the years 1996 to 2018 and found that the weight and length of trains involved in derailments of unit trains are statistically larger than those of manifest trains (*9*). Additionally, the average number of cars derailed in unit trains and manifest trains was significantly different, with 11.3 and 7.6 cars derailed on average, respectively (*9*). Thus, the derailment risk and severity for manifest trains and unit trains are different.

Train derailment severity analysis is an essential part of hazmat transportation risk analysis. This report belongs to a broader study of the effects of train configuration on railroad hazmat transportation risk. Therefore, estimating the train derailment severity for two major trains and quantifying the relationship between train derailment severity and associated affecting factors are pivotal for comprehending risks associated with freight transportation, including hazardous material by rail, as well as for aiding the rail industry and government to develop, evaluate, prioritize, and implement cost-effective safety-enhancement strategies.

For a train derailment incident, the number of derailed cars is greater than or equal to one and less than or equal to the residual train length, including locomotives and railcars. In this paper, the number of derailed cars following a train derailment is used to measure the derailment severity. Previous studies focused on the probability of a train derailment and developed various methods to measure derailment severity (*12–15*). The truncated geometric (TG) model is one of the most mainstream and effective of these methods. The TG model can truncate the number of derailed cars starting from one and extending to the residual train length. Although previous research has focused on attributes such as residual train length, derailment speed, and loading factor, relatively few studies have investigated the derailment rates of unit trains and manifest trains on the mainlines of freight networks. There remains a lack of more fine-grained quantity analysis of derailment severity for different types of trains. In this research, we conduct a detailed exploratory data analysis of the most recent train accident data from the FRA and investigate and clarify the corresponding impact of train types and factors, including train length, gross tonnage, and train speed. Based on selected train derailment data from the

FRA, the TG model is utilized to estimate derailment severity for different types of freight trains. There might be a significant difference between loaded unit trains, empty unit trains, and manifest trains in train length, gross tonnage, and train speed, which might inversely affect derailment severity. Ultimately, the relationship between train derailment severity and train types and related affecting factors can be quantified for the government and railroad industry to develop more cost-effective accident-prevention strategies.

Literature Review

Statistical models have been developed using historical accident data to estimate expected severity based on associated factors such as train speed, position of the derailment, and the proportion of loaded and empty cars within a train (*12–15*). These models are effective at estimating train derailment severity as well as identifying risk factors that influence train derailment severity.

For instance, Liu et al. developed zero-truncated negative binomial regression and quantile regression models to estimate the conditional mean of freight train derailment severity based on residual train length, derailment speed, and loading factor, using data from historical derailments on U.S. Class I railroad mainlines from 2001 to 2010 (*12*). Additionally, Saccomanno and El-Hage developed a TG model to estimate the mean number of cars derailed as a function of derailment speed, residual train length, and accident cause (*16, 17*). Martey and Attoh-Okine developed a joint mixed copula-based model to estimate derailed cars and monetary damage, and conducted a combined analysis of their relationship with a set of covariates that might affect both outcomes (*13*). Martey and Attoh-Okine employed a vine copula quantile regression model, an interval estimation approach, to predict conditional mean and quantiles of derailment severity outcomes, considering derailment speed, residual length, loading factor, derailed cars, and monetary damage (*14*). Bagheri used a TG model to estimate the number of cars derailed per train derailment (*18*). Liu built a negative binomial regression model to estimate the number of cars derailed in the United States between 2000 and 2014 based on year, track type, annual traffic volume, and accident cause (*19*). Li et al. conducted a quantitative analysis of the derailment characteristics of loaded and empty unit trains to determine the frequency of derailments for both (*7*). Kang et al. developed a probabilistic risk analysis model for railroad hazmat transportation risk modeling and presented a practical case study comparing unit and manifest trains transporting hazardous materials (*20*).

These studies show that statistical methods can be applied and train types can be considered in train

derailment estimation. In the past decade, some researchers have used machine learning to investigate the relationship between train derailment and associated factors (*21*, *22*). Lotfi et al. predicted the severity of train derailment utilizing decision tree, random forest, support vector machine, and other machine learning techniques (*21*). Song et al. conducted a quantitative analysis of freight train derailment severity with structured and unstructured data by combining latent dirichlet allocation (LDA) and a zero-truncated negative binomial model (*22*). Machine learning-based methods are increasingly a research focus and continue to yield valuable insights.

Knowledge Gaps and Our Goals

Previous machine learning-based studies do not show a great ability to quantify and explain the impact of each factor on train derailment. Current methods based on the TG model focus on limited attributes and only estimate the mean of freight train derailment severity. Limited previous research has estimated derailment severity pertaining to the use of different types of trains (*9*, *23*, *24*). Even fewer studies have focused on the difference between loaded and empty unit trains versus other types of trains for the same traffic demand. This research addresses these gaps by presenting a comprehensive freight train derailment factor analysis and derailment severity estimation based on the TG model, train types, and other associated factors.

This paper aims to: (1) define and identify factors that influence derailment severity, given a train derailment incident; (2) build a TG model to estimate derailment severity, given a set of influencing factors; and (3) estimate train derailment severity based on the dataset from the FRA using the TG model. The remainder of this paper is organized as follows: the next section describes the data sourced from the FRA and corresponding exploratory data analysis. Following that section is an explanation of the methodology implemented in this work, namely the TG model, to identify factors that influence derailment severity and to estimate derailment severity. Subsequently, the results are interpreted and implications are discussed. The last section concludes the estimation of train derailment severity and analysis of influencing factors, clarifies the contribution of this paper, and discusses future work.

Data Source and Exploratory Data Analysis

Data Source

The Rail Equipment Accident/Incident Form 6180.54 (REA), published by the FRA of the U.S. Department of Transportation (USDOT), records the time, cause, severity, consequence, and contributing factors of each train accident (*11*). This research uses freight train derailment data on the main tracks of Class I railroads, the largest railroads by revenue operating in the United States, from 1996 to 2021 for train derailment severity analysis, which covers all available data consistent with our derailment rate analysis.

Many previous studies have used the "number of derailed cars" to measure derailment severity (*7*, *12*, *16*, *17*). The generic term "cars" refers to all types of vehicles, such as locomotives, railcars, and cabooses, unless specifically stated otherwise. Monetary damage and the number of casualties are also used to assess derailment severity. However, prior studies have stated that monetary damage is prone to substantial variation because of factors such as the cost difference between locomotives and railcars and differences in repair costs between regular track and special track (like turnouts and crossings), and that casualties are more appropriately included in the measurement of passenger train derailments (*13*, *25*). As this research focuses on freight train derailments, this section primarily uses the number of cars derailed as a derailment severity metric in its study of unit train and manifest train derailment severity analysis.

Figure 1 shows the average train derailment severity for each year to demonstrate whether there is a statistical trend over the years from 1996 to 2021. This figure does not show an obvious increasing or decreasing trend in average derailment severity from 1996 to 2021. We use the Mann–Kendall Trend Test to verify the significance of increasing or decreasing trends in derailment severity. The Mann–Kendall Trend Test (*26*) is a non-parametric test for analyzing time series data with consistent increasing or decreasing trends. We find that train derailment severity does not have a significant increasing or decreasing trend over time. This is because the p-value (0.895) is greater than the 0.05 confidence level, and we can thus reject the null hypothesis that the data have a significant increasing or decreasing trend. Therefore, we can conclude that, statistically, derailment severity for the 25 years from 1996 to 2021 does not change significantly over time, and we can conduct an integrated analysis of derailment severity for all years from 1996 to 2021 instead of a separate derailment severity analysis for each individual year.

Using the freight train derailment data collected from 1996 to 2021, the distribution of the number of cars derailed is plotted in Figure 2, excluding data with derailment speed equal to "0 mph" or the number of derailed cars equal to "0". Figure 3 presents the same data separated by train type for the unit trains and manifest trains' distributions of the number of derailed cars, and the comparison between them. Table 1 presents descriptive statistics for train derailment severity analysis. Most derailment accidents are manifest train derailments
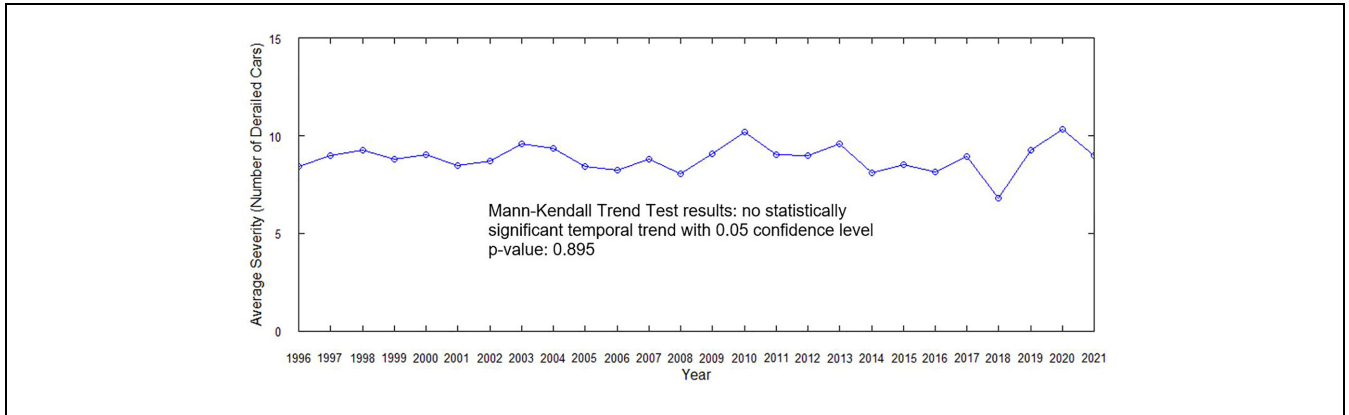
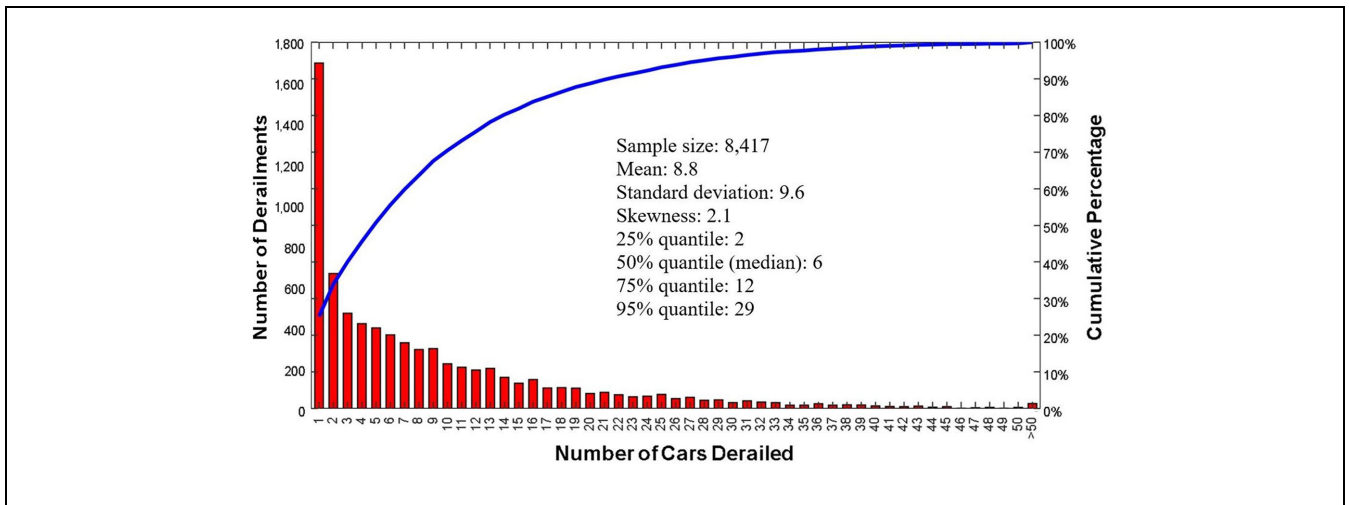**Figure 1.** Average train derailment severity each year.

Figure 2. Number of railcars derailed per freight train derailment (all train types) from 1996 to 2021.

(69%), and the corresponding average number of derailed cars is 7.7, whereas the average numbers of derailed cars for all train types and for unit trains specifically are 8.8 and 11.3, respectively. These facts indicate that manifest trains may primarily affect the estimation of train derailment severity.

Exploratory Data Analysis

Before developing statistical models to estimate train derailment severity, we conduct data screening to exclude data with inappropriate attributes which may degrade the model's performance. The following screening conditions are applied; 8,160 train derailments (94.88% of all 8,600 derailments) are selected for analysis.

- Speed ⩾ 1: In the dataset, 145 derailment incidents' speeds are 0 miles per hour. We assume that a train will not derail when train speed is

extremely slow. Thus, we set the lowest derailment speed as "1 mile per hour" and these 145 derailment incidents (1.68% of all 8,600 derailments) are excluded from subsequent analysis.

- Number of derailed cars ⩾ 1: In the original dataset, some numbers of derailed cars and locomotives were zero. However, when the number of derailed cars is zero, we deem that the train did not derail, and thus the incidents with zero derailed cars are excluded from our analysis. This screening criterion removes sixty-seven derailment incidents (0.78% of all 8,600 derailments) from the dataset.

- Number of cars behind the point of derailment (POD) ⩾ Number of derailed cars: In this paper, POD refers to the position of the first derailed vehicle (FDV). It is impossible for the number of cars behind the POD to be less than the number of derailed cars, and thus we only consider data
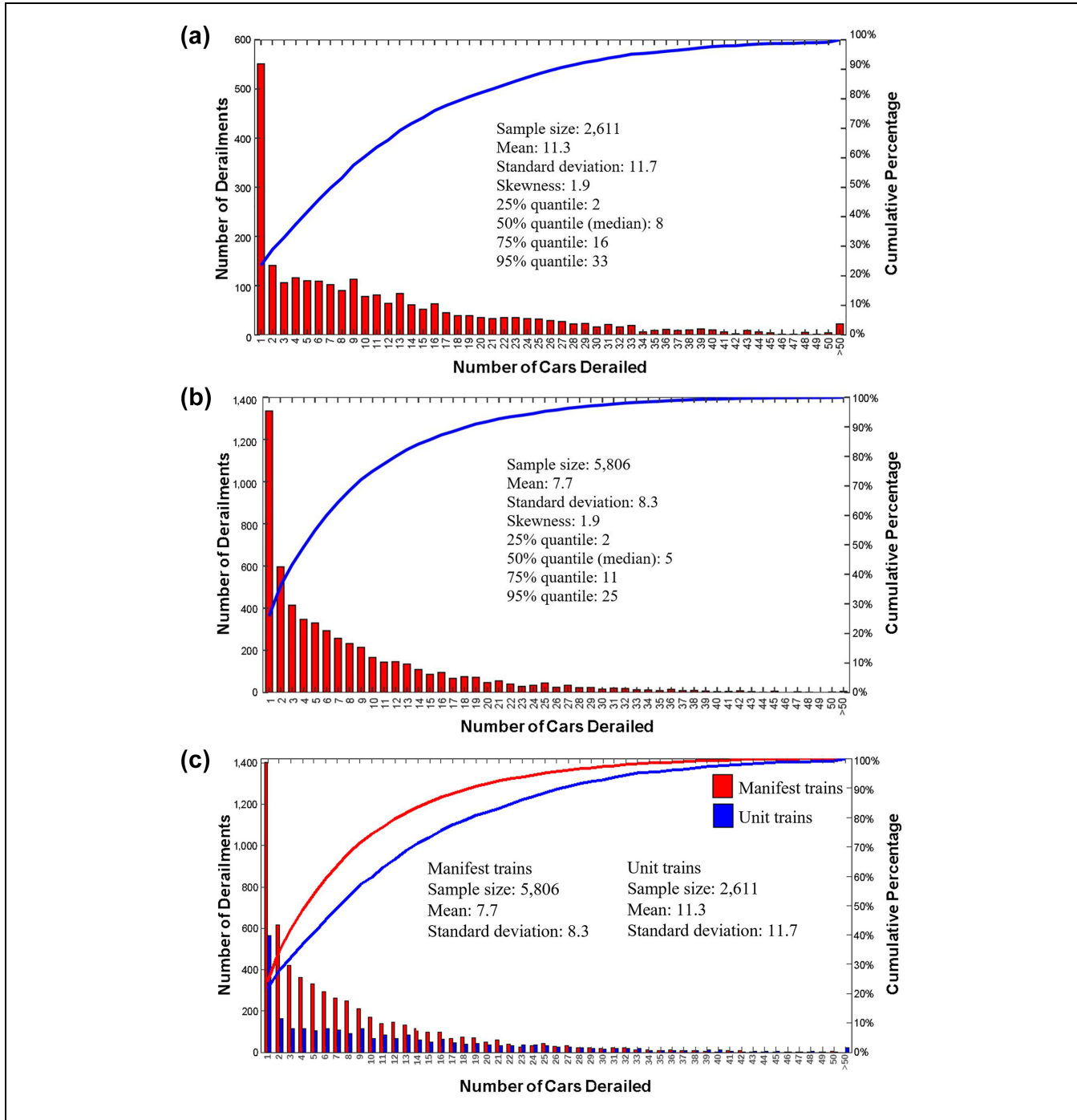
Figure 3. Number of railcars derailed per freight train derailment in: (*a*) unit trains, (*b*) manifest trains, and (*c*) comparison.

Table 1. Descriptive Statistical Results

	Sample size (number of train derailments)	Mean (average number of derailed cars)	Median (the median number of derailed cars)	Variance (the variance of numbers of derailed cars)	Standard deviation of numbers of derailed cars
All train types	8,417	8.8	6	92.3	9.6
Unit trains	2,611	11.3	8	135.9	11.7
Manifest trains	5,806	7.7	5	68.8	8.3

with the number of cars behind the POD greater than or equal to the number of derailed cars. In addition, we assume there is only one POD per accident. The POD information is from the FRA safety database and its accuracy could potentially affect the results. In the dataset, there are 179 derailment incidents (2.08% of all 8,600 derailments) whose numbers of cars behind the POD are less than the numbers of derailed cars, and these data are excluded from our analysis.

- Other data with apparent errors (seventy-nine data records [0.92% of all 8,600 derailments], e.g., tons per car equal to 0, two data records for one derailment incident) are excluded.

Out of the 8,160 train derailments, 2,120 data points refer to loaded unit trains, 414 data points refer to empty unit trains, and 5,626 data points refer to manifest trains.

The response variable is the total number of railcars derailed, including loaded railcars, empty railcars, and locomotives. Several engineering and operational factors may affect train derailment severity. Based on the literature, estimator variables are identified for statistical analysis, and include train derailment speed, residual train length, gross tonnage per car, and train type.

Train Speed. Train derailment speed is the speed of train operation when the accident occurs. This factor's effect on derailment severity is most widely studied in the literature (*12*, *17*, *27*, *28*). It has been found that, given all other factors are equal, derailment speed is positively associated with the number of cars derailed. This finding is reasonable because speed is an indicator of an accident's kinetic energy.

Residual Train Length. Residual train length is defined as the number of railcars behind the POD, which is the maximum number of cars potentially subject to derailment. Figure 4 illustrates the POD and residual train length. Saccomanno and El-Hage (*17*), Anderson (*27*), and Liu et al. (*12*) found that a greater residual train length is associated with more derailed cars if a derailment occurs.

Gross Tonnage per Car. Liu et al. verified the hypothesis that a train carrying a larger proportion of loaded cars is expected to derail more cars (*12*). Therefore, we consider the gross tonnage per car as a factor that influences derailment severity. Higher gross tonnage per car in the train may also indicate greater kinetic energy in the derailment, thereby causing more cars to derail, given all else being equal.

Figure 4. An illustrative example of point of derailment and residual train length.

Train Type. Train type is the new factor considered in this research. There is no data field in the FRA REA database to directly identify train type (i.e., unit or manifest train). In this research, we classify train types based on previous studies (*7*, *9*). We identify unit trains and manifest trains using the railroad code, train symbol ID, causing car reporting mark and number, number of empty cars, number of loaded cars, number of locomotives, and narratives recorded in the REA database. According to several variables extracted from the accident database, including the number of empty cars, the number of loaded cars, the number of locomotives, the length of the train considering the total number of cars and locomotives, and the percentage of loaded or empty railcars in the train, a train is classified as loaded if 95% or more of its cars are loaded, or as empty if 95% or more of its cars are empty (Figure 5) (*9*). These percentages were calculated by dividing the number of loaded or empty cars by the total number of cars in the train. The percentage of "95%" of loaded or empty cars is used as the threshold to measure whether a train is a unit train or a manifest train (instead of the percentage of "100%") because of the buffer cars required by federal regulation (*29*) whose loading condition is independent of the loading condition of the rest of the train. This same threshold was used in the previous study (*7*) to determine the loading conditions of freight trains. Buffer cars can either be empty or loaded with an inert material. In this research, a train that is not more than 95% loaded or not more than 95% empty is directly defined as manifest train. To determine whether a train is a unit train or a manifest train, further steps are required to identify train type when it is more than 95% loaded or empty, while the remaining trains with partially loaded consists are generally manifest trains. For example, train numbers with prefixes of "C," "G," or "U" stand for loaded unit coal trains, loaded unit grain trains, and loaded unit trains carrying materials other than coal or grain, respectively,
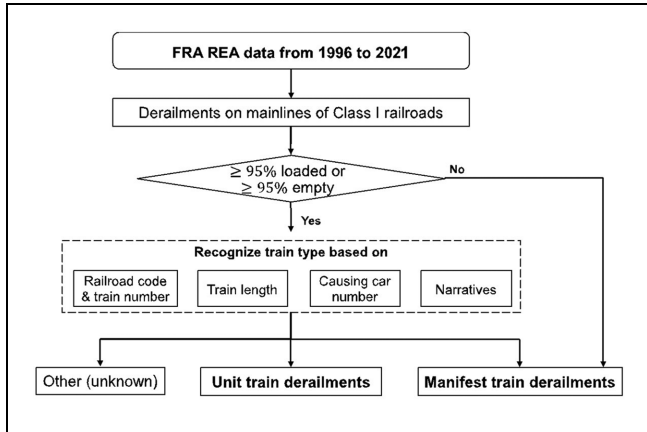
Figure 5. Methodology for classifying type of derailed trains (9).

whereas "M" represents manifest trains, following BNSF Railway's symbol guide (30).

We consider three types of trains: loaded unit trains, empty unit trains, and manifest trains. There may be significant differences in train length, gross tonnage, and train speed for loaded unit trains, empty unit trains, and manifest trains, which may influence derailment severity, as shown in Table 2. The average total numbers of cars per train for both the loaded unit train and empty unit train are greater than the manifest train, with a small difference between the loaded unit train and empty unit train. The average tonnage per car of a loaded unit train is remarkably larger than empty unit trains and manifest trains, which may increase the corresponding train derailment severity. We use the one-way analysis of variance (ANOVA) to analyze the differences between the

means of two or more groups. The one-way ANOVA (31) compares the variance within groups with the variance between groups and, based on the results, it determines whether there is a statistically significant difference between the means of the groups being compared. As shown in Table 3, one-way ANOVAs are performed to evaluate the existence of statistical differences in these attributes among train types. We find that these attributes among train types considered in the one-way ANOVA are significant, with all p-values being less than 0.05. The results indicate that there was a statistically significant difference in these attributes between all three train types. These attributes of all three train types with different distributions indicate that these factors have significant influences on derailment severity. Moreover, train type represents some unobserved, hidden factors besides tonnage, train length, speed, and so forth. Therefore, train type is also treated as a variable. Table 4 summarizes the input variables considered for the severity estimation model.

Methodology

Truncated Geometric Model

Given that a train derailment incident occurs, the number of cars derailed must be greater than or equal to "1" and less than or equal to the residual train length. As verified by Bagheri (18), the number of cars derailed follows a TG distribution and the TG model can truncate the number of derailed cars beyond the range from "1" to the residual train length. For a train derailment event with train length L and POD j (the residual train length

Table 2. Attributes of Loaded Unit Trains, Empty Unit Trains, and Manifest Trains

Train types	Average residual train length (number of cars behind point of derailment)	Average tonnage per car	Average total number of cars per train	Average derailment speed (mph)
Loaded unit trains	58	131	109	25.3
Empty unit trains	72	32	111	26.7
Manifest trains	50	85	86	24.5

Table 3. One-way ANOVA Result of Train Types and Attributes

Factor	DF	F-value	p-value
Average residual train length (number of cars behind point of derailment)	2	89.97	$< 2 \times 10^{-16}$
Average tonnage per car	2	2060	$< 2 \times 10^{-16}$
Average total number of cars per train	2	392	$< 2 \times 10^{-16}$
Average derailment speed (mph)	2	4.961	0.00703

Note: ANOVA = analysis of variance; DF = degrees of freedom.

Table 4. Variables in Derailment Severity Estimation Models

Variables	Data type
Model input	
Gross tonnage per car	Continuous variable
Train derailment speed (miles per hour)	Continuous variable
Number of cars behind point of derailment	Integer
Train types including loaded unit trains, empty unit trains, and manifest trains	Category
Model output	
Number of cars derailed in one derailment	Integer

$Lr = L-j+1$), the probability of derailing x cars can be calculated by:

$$P(X = x | POD \text{ at } j) = \begin{cases} \dfrac{p(1-p)^{x-1}}{1-(1-p)^{Lr}} & \text{if } x = 1, 2, ..., Lr \\ 0 & \text{otherwise} \end{cases}$$
$$(1)$$

where

Lr: residual train length

x: number of cars derailed

p: the probability of success at each trial, which is a constant probability. In other words, if a derailment has occurred, "x" cars derailing before the first non-derailing car is a geometric distribution and the probability of a car derailing is equal to $(1-p)$. This probability p is assumed to be related to the factors/covariates through the logit link function:

$$p = \frac{e^Z}{1 + e^Z} \quad (2)$$

where Z is a linear function (Equation 3) of influencing factors, including derailment speed, residual train length, gross tonnage per car, and train type.

$$Z = \beta_0 + \beta_1(speed) + \beta_2(cars\ residual)$$
$$+ \beta_3(tons\ per\ car) + \beta_4(empty\ unit\ train) \quad (3)$$
$$+ \beta_5(loaded\ unit\ train)$$

speed: Train derailment speed (miles per hour)

cars residual: Number of cars behind POD

tons per car: Average gross tonnage per car

empty unit train: If the train is an empty unit train, *empty unit train* = 1, otherwise *empty unit train* = 0

loaded unit train: If the train is a loaded unit train, *loaded unit train* = 1, otherwise *loaded unit train* = 0

The model uses the manifest train as a reference, which means that the manifest train type variable is not included in the model. A detailed introduction of the TG model can be found in Bagheri ([18]).

We use the VGAM package in R to fit the TG model. The VGAM package ([32]) in R aims to fit Vector Generalized Linear and Additive Models (VGLMs and VGAMs), as well as reduced-rank VGLMs (RR-VGLMs) and quadratic RR-VGLMs (QRR-VGLMs). It is a general program for maximum likelihood estimation. In this research, the VGAM package is able to fit the TG model using the function "truncgeometric()".

Model Results

Table 5 presents the results of the model. We find that except for the variable "empty unit train," all other variables considered in the TG model are significant, with all p-values being less than 0.05. Thereafter, we get the model result of the Z function in Equation 4.

$$Z = -0.908 - 0.0313(speed) - 0.00239(cars\ residual)$$
$$- 0.00240(tons\ per\ car) - 0.340(loaded\ unit\ train)$$
$$(4)$$

In Equation 4, if the coefficient of a variable is negative, then increasing this variable would tend to increase derailment severity. In contrast, if the coefficient of a variable is positive, then increasing the variable would tend to decrease derailment severity. The reason can be interpreted as follows. Based on Equation 2, the probability of a car not to derail (p) is a monotone increasing function of "Z," and thus the probability of a car to derail ($1-p$) is a monotone decreasing function of "Z." When a variable is increased, given that its coefficient in

Table 5. Truncated Geometric Model Results for All Train Types with Outliers

| Variables	Definition	Coefficient	p-value
Intercept	The constant	−0.908	$< 2 \times 10^{-16}$
speed	Train derailment speed (miles per hour)	−0.0313	$< 2 \times 10^{-16}$
cars residual	Number of cars behind point of derailment	−0.00239	3.65×10^{-8}
tons per car	Average gross tonnage per car	−0.00240	1.23×10^{-6}
empty unit train	If the train is an empty unit train, *empty unit train* = 1, otherwise *empty unit train* = 0	0.0783	0.243
loaded unit train	If the train is a loaded unit train, *loaded unit train* = 1, otherwise *loaded unit train* = 0	−0.340	5.19×10^{-16}
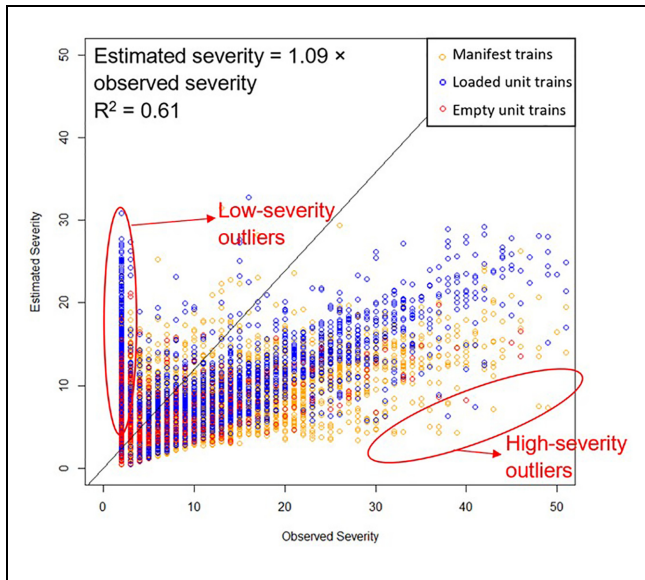
Figure 6. Estimated severity versus observed severity with outliers.

Z function is negative, Z decreases and then the probability of a car to derail increases, and thus derailment severity tends to increase, and vice versa. Based on this discussion, we offer the following interpretation of the obtained model.

- The coefficient of derailment speed is -0.0313. This indicates that higher derailment speed is associated with increased derailment severity.
- The coefficient of residual train length is -0.00239. The negative value indicates that longer residual train length is associated with increased derailment severity.
- Similarly, greater gross tonnage per car is associated with increased derailment severity.
- The coefficient of "loaded unit train type" is -0.340. This means that a loaded unit train tends to have more cars derailed compared with a manifest train or an empty unit train, given all else being equal.

We use the mean number of cars derailed as the estimated derailment severity, as formulated by Equation 5.

$$D = E[X] = \sum_{x=1}^{Lr} P(x)x = \frac{1}{p\left[1 - (1-p)^{Lr}\right]} \qquad (5)$$

Figure 6 plots the estimated severity versus the observed severity. Because of low-severity outliers (i.e., those incidents where estimated severity is high but observed severity is low) and high-severity outliers (i.e., those incidents where estimated severity is low but observed severity is high), the

R square is not sufficiently large (0.61), which indicates that the model does not estimate the severity perfectly.

One of the concerns identified by this result is that there are different types of behaviors in the derailment data and the variation cannot be represented well by this type of model. Thus, we analyze train derailment data through a physical model (*33*) that can model typical derailments. For example, the derailment data include a significant number (approximately 20%) of single-car derailments (severity = 1) with a wide range of derailment speeds. These often result from mechanisms like a broken wheel or axle, and are situations where a single truck might derail but the problem is identified and the train is brought to rest by the operator, initiating no larger derailment severity. All cars in the train remain connected and upright, primarily aligned with the track. Liu et al. analyzed the causes of major train derailment and revealed that equipment is the main cause of train derailment at derailment speeds greater than 25 mph (*34*).

In other extreme derailments, the train is moving very slowly and an initiating event, such as extreme weather, causes a long string of connected cars to fall over, remaining aligned with the track. According to our findings based on accident narratives in the FRA REA database, we can assume that if there are empty tank cars with double shelf couplers and one of these cars rolls off its tracks in a derailment, the strong rotational coupling between the cars could result in a long line of trailing cars rolling over. These events can have very high severities even at low speeds or no speed.

Consequently, we conducted additional analysis of the data where we excluded the high-severity and low-severity outliers. Nayak et al. proposed a positive non-linear relationship between derailment speed and the mean of derailed cars, and expressed the mean number of derailed cars as a function of the square root of the derailment speed (*15*). Martey and Attoh-Okine employed a vine copula quantile regression model to analyze the relationship between derailment speed and derailment, which showed the marginal effects of derailment speed in their model with different alpha values (*14*). Therefore, we developed a method to exclude these outliers and then used the TG model again to estimate derailment severity. Considering that we covered most of the derailment data and improved the TG model's performance, we determined the threshold for excluding outliers by inferring from previous research (*14, 15, 33, 34*) and several trials. Derailment data outside the range from $0.1 \times$ derailment speed to $1.6 \times$ speed are treated as outliers and are excluded from analysis (Figure 7). The derailments outside these bounds are assumed to have a "unique behavior" compared with typical derailments. Using this outlier elimination method, 5,849
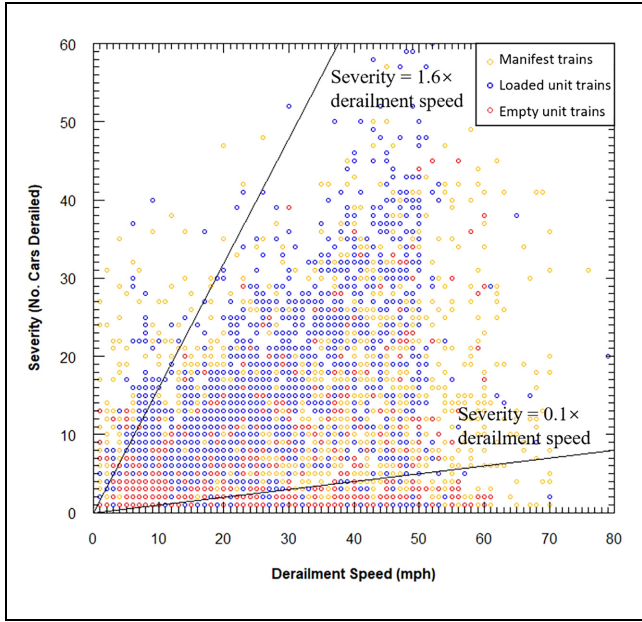
Figure 7. Method to exclude outliers.

derailment data are selected from the total 8,160 train derailments; 2,311 outliers are excluded, among which 2,004 derailment data are low-severity outliers and 307 derailment data are high-severity outliers. Among the 5,849 train derailments, there are 1,581 data records from loaded unit trains, 299 data records from empty unit trains, and 3,969 data records from manifest trains. The result of the TG model is presented in Table 6 and Equation 6.

$$Z = -0.587 - 0.0709(speed) - 0.00171(cars\ residual) \\ - 0.00252(tons\ per\ car) - 0.332(loaded\ unit\ train) \tag{6}$$

Mean squared error (MSE, Equation 7) and mean absolute error (MAE, Equation 8) are selected to measure the performance of TG models in estimating train derailment severity. For all data with outliers, the MSE and MAE of the model are 68.91 and 6.05, respectively.

For the data without outliers, the MSE and MAE of the model are 36.82 and 4.22, respectively. The MAE indicates that the average gap between the estimated number of derailed cars and the observed number of derailed cars is "6.05" and "4.22" for the data with and without outliers, respectively. Figure 8 plots the estimated severity versus the observed severity without outliers.

$$MSE = mean\left[(estimated\ severity\ -\ observed\ severity)^2\right] \tag{7}$$

$$MAE = mean\left[|estimated\ severity\ -\ observed\ severity|\right] \tag{8}$$

The TG regression model is also used to separately fit the derailment severities for the loaded unit train, empty unit train, and manifest train. Equations 9 and 10 calculate the TG model results for loaded unit train derailment severity based on the data including and excluding outliers, respectively. The effect of average gross tonnage per car on the derailment severity for the data both with and without outliers is insignificant.

$$Z = -1.492 - 0.0423(speed) - 0.00119(cars\ residual) \tag{9}$$

$$Z = -1.216 - 0.0785(speed) - 0.00504(cars\ residual) \tag{10}$$

Equations 11 and 12 calculate the TG model results for empty unit train derailment severity based on the data including and excluding outliers, respectively. In the model based on the data with outliers, the effects of average gross tonnage per car on the derailment severity are insignificant. In the model based on the data without outliers, the effects of average gross tonnage per car and residual train length on the derailment severity are insignificant.

$$Z = -0.597 - 0.0330(speed) - 0.00573(cars\ residual) \tag{11}$$

Table 6. Truncated Geometry Model Results for all Train Types without Outliers

| Variables	Definition	Coefficient	p-value
Intercept	The constant	−0.587	4.90×10^{-16}
speed	Train derailment speed (miles per hour)	−0.0709	$< 2\times10^{-16}$
cars residual	Number of cars behind POD (point of derailment)	−0.00171	0.0025
tons per car	Average gross tonnage per car	−0.00252	5.02×10^{-5}
empty unit train	If the train is an empty unit train, empty unit train = 1, otherwise empty unit train = 0	0.123	0.151
loaded unit train	If the train is a loaded unit train, loaded unit train = 1, otherwise loaded unit train = 0	−0.332	7.84×10^{-10}
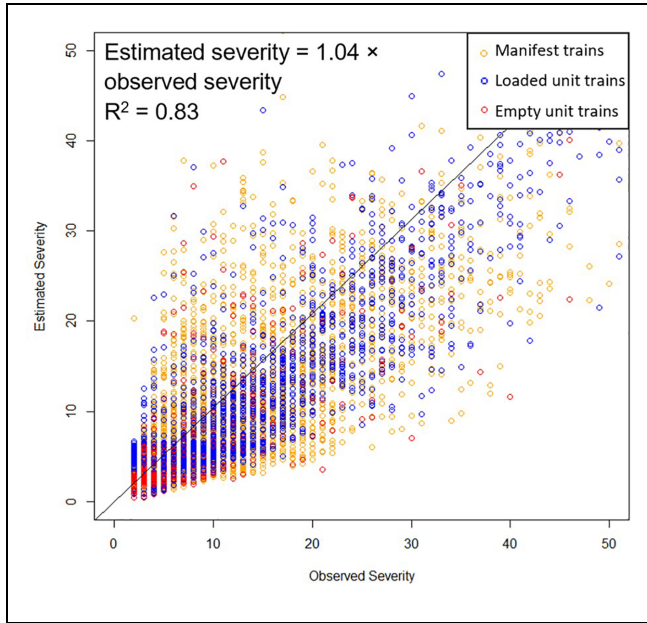
Figure 8. Estimated severity versus observed severity without outliers.

$$Z = -0.416 - 0.0614(speed) \tag{12}$$

Equations 13 and 14 calculate the model results for derailment severity of manifest trains based on the data with and without outliers, respectively. All considered variables and the intercept are significant for both models based on the data with and without outliers.

$$Z = -0.858 - 0.0289(speed) - 0.00335(cars\ residual)$$
$$- 0.00291(tons\ per\ car) \tag{13}$$

$$Z = -0.498 - 0.0699(speed) - 0.00320(tons\ per\ car) \tag{14}$$

Table 7 summarizes the performance of TG models at estimating train derailment severity for loaded unit trains, empty unit trains, and manifest trains. The MSE (Equation 7) and MAE (Equation 8) are used as criteria to measure performance. We found that the TG model has a better performance when outliers are excluded. Overall, the performance of the TG model is acceptable: the models with outliers all have MAEs less than 7.5, and the models without outliers all have MAEs less than 5. Based on these results excluding data outliers, we propose the following interpretations of the obtained model.

- The coefficient of derailment speed is -0.0709. This reveals that higher derailment speed is associated with increased derailment severity.
- The coefficient of residual train length is -0.00171, which is also negative, indicating that longer residual train length is associated with increased derailment severity.
- The coefficient of gross tonnage per car is -0.00252. Correspondingly, greater gross tonnage per car is associated with increased derailment severity.
- Similar to the situation considering the data outliers, the coefficient of "loaded unit train type" is also negative, with a value of -0.332. Given all else being equal, this means that a loaded unit train tends to have more cars derailed compared with a manifest train and an empty unit train.

Table 7. Summary of Model Performance of Truncated Geometry Model

Data	Model results	Mean square error	Mean absolute error
All train types			
With outliers	$Z = -0.908 - 0.0313(speed) - 0.00239(cars\ residual)$ $-0.00240(tons\ per\ car) - 0.340(loaded\ unit\ train)$	68.91	6.05
Without outliers	$Z = -0.587 - 0.0709(speed) - 0.00171(cars\ residual)$ $-0.00252(tons\ per\ car) - 0.332(loaded\ unit\ train)$	36.82	4.22
Loaded unit trains			
With outliers	$Z = -1.492 - 0.0423(speed) - 0.00119(cars\ residual)$	96.91	7.04
Without outliers	$Z = -1.216 - 0.0785(speed) - 0.00504(cars\ residual)$	39.46	4.34
Empty unit trains			
With outliers	$Z = -0.597 - 0.0330(speed) - 0.00573(cars\ residual)$	83.76	6.13
Without outliers	$Z = -0.416 - 0.0614(speed)$	45.71	4.80
Manifest trains			
With outliers	$Z = -0.858 - 0.0289(speed) - 0.00335(cars\ residual)$ $- 0.00291(tons\ per\ car)$	56.80	5.60
Without outliers	$Z = -0.498 - 0.0699(speed) - 0.00320(tons\ per\ car)$	34.84	4.11

Concluding Remarks

This research identifies factors that influence derailment severity, given a train derailment incident, and develops a TG model to estimate train derailment severity based on these factors. First, we conduct a comprehensive exploratory data analysis of train derailment severity and select appropriate data from the FRA for statistical analysis. Second, we systematically summarize factors that can significantly influence train derailment severity, with the train types being loaded unit train, empty unit train, and manifest train. Third, we utilize these factors to estimate train derailment severity and quantify the influence of these factors, which provides further insights into train derailment severity.

Based on the severity estimation results obtained by the TG model, factors that can significantly influence train derailment severity are identified; these include train speed, residual train length, gross tonnage per car, and train type. The results indicate that increased derailment severity is associated with higher derailment speed, longer residual train length, and greater gross tonnage per car. Given all other factors are the same, a loaded unit train tends to have more cars derailed compared with a manifest train and an empty unit train.

We conducted additional analysis excluding low and high-severity outliers, and estimated derailment severity using the TG model. The results show that the TG model achieves better estimation performance when outliers are excluded. Given all other factors being equal, a loaded unit train tends to have more cars derailed compared with a manifest train and an empty unit train, whereas there is no significant difference in derailment severity between a manifest train and an empty unit train. For the empty unit train and loaded unit train, the effect of average gross tonnage per car on the derailment severity is insignificant for the data both with and without outliers, whereas the average gross tonnage is significant for the derailment severity of the manifest train.

However, statistical models may have difficulty predicting extreme derailment events that fall outside historical experience. Physical models are able to capture the conditions of a specific train derailment scenario by mathematically describing the physical dynamics of the derailment. In the future, we will leverage physical modeling, which could help in understanding the derailment mechanism determined by train speed, tonnage, train length, and so forth, and compare it with the statistical model to estimate train derailment severity more accurately.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Zheyong Bian and Xiang Liu; data collection: Zheyong Bian and Weitian Li; analysis and interpretation of results: Zheyong Bian and Weitian Li; draft manuscript preparation: Weitian Li. All authors reviewed the results and approved the final version of the manuscript.

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.

ORCID iDs

Weitian Li (iD) https://orcid.org/0000-0003-1835-8230
Zheyong Bian (iD) https://orcid.org/0000-0001-9094-3333
Xiang Liu (iD) https://orcid.org/0000-0002-4348-7432

References

1. Association of American Railroads. Freight Rail Facts & Figures. https://www.aar.org/facts-figures. Accessed February 16, 2023.
2. Sandusky Register. Railroad Must Pay. https://sandusky register.com/news/439916/railroad-must-pay/. Accessed February 16, 2023.
3. Greer, E. Keachi Train Derailment Leaks 10,000 Gallons of Acid, Causing Evacuation. *KTALnews.com*. https://www.ktalnews.com/news/state-news/louisiana/train-derailment-causes-evacuation-road-closures-in-de-soto-parish. Accessed February 16, 2023.
4. McDaniel, J. What's Known About the Toxic Plume from the Ohio Train Derailment. *The Washington Post*. https://www.washingtonpost.com/climate-environment/2023/02/15/ohio-train-derailment-toxic-chemicals. Accessed February 16, 2023.
5. Starr, J. T. *The Evolution of the Unit Train: 1960-1969*. Committee on Geographical Studies, Research Papers. University of Chicago, 1976.
6. Kenkel, P. L., S. R. Henneberry, and H. N. Agustini. *An Economic Analysis of Unit-Train Facility Investment*. No.1364-2016-108008. Southern Agricultural Economics Association, Tulsa, OK, 2004.

7. Li, W., G. S. Roscoe, Z. Zhang, M. R. Saat, and C. P. L. Barkan. Quantitative Analysis of the Derailment Characteristics of Loaded and Empty Unit Trains. *Transportation Research Record: Journal of the Transportation Research Board*, 2018. 2672: 156–165.

8. Dick, C. T., J. Zhao, X. Liu, and S. W. Kirkpatrick. Quantifying Recent Trends in Class 1 Freight Railroad Train Length and Weight by Train Type. *Transportation Research Record: Journal of the Transportation Research Board*, 2021. 2675: 890–903.

9. Zhang, Z., C.-Y. Lin, X. Liu, Z. Bian, C. T. Dick, J. Zhao, and S. Kirkpatrick. An Empirical Analysis of Freight Train Derailment Rates for Unit Trains and Manifest Trains. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Vol. 236, No. 10, 2022, pp. 1168–1178. https://doi.org/10.1177/09544097221080615.

10. Union Pacific. How Manifest Service Works. https://www.up.com/customers/within-reach/manifest/index.htm. Accessed February 16, 2023.

11. Federal Railroad Administration. 6180.54—Rail Equipment Accident/Incident. https://safetydata.fra.dot.gov/OfficeofSafety/publicsite/on_the_fly_download.aspx?itemno = 3.03. Accessed February 16, 2023.

12. Liu, X., M. R. Saat, X. Qin, and C. P. L. Barkan. Analysis of U.S. Freight-Train Derailment Severity Using Zero-Truncated Negative Binomial Regression and Quantile Regression. *Accident Analysis and Prevention*, Vol. 59, 2013, pp. 87–93. https://doi.org/10.1016/j.aap.2013.04.039.

13. Martey, E. N., and N. Attoh-Okine. Bivariate Severity Analysis of Train Derailments Using Copula-Based Regression Models. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, Vol. 4, No. 4, 2018, p. 04018034. https://doi.org/10.1061/ajrua6.0000982.

14. Martey, E. N., and N. Attoh-Okine. Analysis of Train Derailment Severity Using Vine Copula Quantile Regression Modeling. *Transportation Research Part C: Emerging Technologies*, Vol. 105, 2019, pp. 485–503. https://doi.org/10.1016/j.trc.2019.06.015.

15. Nayak, P. R., D. B. Rosenfield, and J. H. Hagopian. *Event Probabilities and Impact Zones For Hazardous Materials Accidents on Railroads*. Report DOT/FRA/ORD-83/20. FRA, U.S. Department of Transportation, Washington, D.C., 1983.

16. Saccomanno, F. F., and S. El-Hage. Minimizing Derailments of Railcars Carrying Dangerous Commodities Through Effective Marshaling Strategies. *Transportation Research Record: Journal of the Transportation Research Board*, 1989. 1245: 39–41.

17. Saccomanno, F. F., and S. M. El-Hage. Establishing Derailment Profiles by Position for Corridor Shipments of Dangerous Goods. *Canadian Journal of Civil Engineering*, Vol. 18, No. 1, 1991, pp. 67–75. https://doi.org/10.1139/l91-009.

18. Bagheri, M. *Risk-Based Model for Effective Marshalling of Dangerous Goods Railway Cars*. Doctoral dissertation. University of Waterloo, Canada, 2010.

19. Liu, X. Analysis of Collision Risk for Freight Trains in the United States. *Transportation Research Record: Journal of the Transportation Research Board*, 2016. 2546: 121–128.

20. Kang, D., J. Zhao, C. T. Dick, X. Liu, Z. Bian, S. W. Kirkpatrick, and C.-Y. Lin. Probabilistic Risk Analysis of Unit Trains Versus Manifest Trains for Transporting Hazardous Materials. *Accident Analysis and Prevention*, Vol. 181, 2023, p. 106950. https://doi.org/10.1016/j.aap.2022.106950.

21. Lotfi, A., M. Bagheri, and A. Ahmadi. Using Machine Learning Methods for Modeling Freight Train Derailment Severity. *Transportation Research Record: Journal of the Transportation Research Board*, 2023. 2677: 961–973.

22. Song, B., Z. Zhang, Y. Qin, X. Liu, and H. Hu. Quantitative Analysis of Freight Train Derailment Severity with Structured and Unstructured Data. *Reliability Engineering and System Safety*, Vol. 224, 2022, p. 108563. https://doi.org/10.1016/j.ress.2022.108563.

23. Zhao, J., and C. T. Dick. Quantitative Derailment Rate Comparison of Unit Trains at Transload Terminals and Manifest Trains at Railroad Switching and Hump Classification Yards. *Transportation Research Record: Journal of the Transportation Research Board*, 2023. 2677: 311–325.

24. Lin, C.-Y., X. Liu, J. Kim, and C. P. L. Barkan. General Model of the Effect of Hazardous Materials Car Placement in Trains and Their Probability of Derailment. *Transportation Research Record: Journal of the Transportation Research Board*, 2022. 2676: 28–37.

25. Barkan, C. P. L., C. T. Dick, and R. Anderson. Railroad Derailment Factors Affecting Hazardous Materials Transportation Risk. *Transportation Research Record: Journal of the Transportation Research Board*, 2003. 1825: 64–74.

26. Kendall, M. G., and J. D. Gibbons. *Rank Correlation Methods*. Oxford University Press, New York, 1990.

27. Anderson, R. T. *Quantitative Analysis of Factors Affecting Railroad Accident Probability and Severity*. MS thesis. University of Illinois at Urbana-Champaign, 2005.

28. Bagheri, M., F. Saccomanno, S. Chenouri, and L. Fu. Reducing the Threat of In-Transit Derailments Involving Dangerous Goods Through Effective Placement Along the Train Consist. *Accident Analysis and Prevention*, Vol. 43, No. 3, 2011, pp. 613–620. https://doi.org/10.1016/j.aap.2010.09.008.

29. Federal Railroad Administration. Safe Placement of Train Cars. FRA, 2005. https://railroads.dot.gov/elibrary/safe-placement-train-cars-report. Accessed February 16, 2023.

30. QStation. BNSF Symbol Guide. http://www.qstation.org/bnsf/bnsfsymbols.html. Accessed February 16, 2023.

31. Girden, E. R. *ANOVA: Repeated Measures*. No. 84. Sage, Newbury Park, CA, 1992.

32. Yee, T. W. *VGAM: Vector Generalized Linear and Additive Models*. R package version 1.1–1. 2020.

33. Kirkpatrick, S. W., C.-Y. Lin, L. Iannacone, P. Gharzouzi, T. Treichel, C. P. L. Barkan, and P. Gardoni. Derailment Analysis for Prediction of Damage and Probability of Release for Novel Railroad Tank Car Designs. *Transportation Research Record: Journal of the Transportation Research Board*, 2023. 2677: 812–828.

34. Liu, X., M. R. Saat, and C. P. L. Barkan. Analysis of Causes of Major Train Derailment and Their Effect on Accident Rates. *Transportation Research Record: Journal of the Transportation Research Board*, 2012. 2289: 154–163.