Statistical Causal Analysis of Freight-Train Derailments in the United States

Xiang Liu, Ph.D.[1]

Abstract: Freight railroads contribute to the national economy by moving over 40% of intercity ton-miles of freight. Meanwhile, train accidents can damage infrastructure and rolling stock, disrupt operations, and possibly cause casualties and harm the environment. Understanding major accident causes is the first step in developing and prioritizing effective accident prevention strategies. The literature has predominantly focused on nationwide train accident cause analysis, without accounting for possible variation in accident cause distributions by railroad and season. This research develops a log-linear statistical model that can estimate the number of freight-train derailments accounting for railroad, accident cause, season, and traffic volume. The analysis shows that broken rails and track geometry defects are the two leading freight-train derailment causes on four major U.S. freight railroads. Fall and winter appear to have a higher likelihood of a broken-rail-caused derailment than spring and summer, given the same railroad and traffic level. By contrast, track-geometry-defect-caused derailments occur more frequently in spring and summer than in fall and winter, given all else being equal. The statistical modeling techniques in this paper can be adapted to other types of train accidents or accident causes, ultimately leading to the prioritization of train safety improvement resources on various spatial and temporal scales. **DOI:** 10.1061/JTEPBS.0000014.

Author keywords: Railroad; Train derailment; Statistical analysis; Accident cause.

Introduction

The United States has an extensive freight railroad network in the world. Its 225,308 km (140,000 mi) of track transport over 40% of intercity-cargo ton-miles (FRA 2010). Ensuring the safe, efficient operation of the American rail transport system is an economic necessity. The reduction of train accident risk has long been a top priority for both the U.S. Department of Transportation (USDOT) and the railroad industry. Understanding major accident causes is the first step in developing and prioritizing effective accident prevention strategies. To this end, this research identifies major derailment causes in the United States and uses statistical techniques to quantify their derailment frequency distributions. The prior research has largely focused on empirical, nationwide train accident cause analysis. By contrast, this paper aims to advance existing knowledge in the following areas. First, major derailment causes are identified for each of the four major freight railroads in the United States, which are Burlington Northern and Santa Fe Railway (BNSF), Union Pacific Railroad (UP), CSX Transportation (CSX), and Norfolk Southern Railway (NS). Second, this paper aims to understand the seasonal effect on the distribution of track-related accident causes, particularly broken rails and track geometry defects. Third, this paper develops a log-linear statistical model to predict the number of freight train derailments by railroad, season, and traffic volume for each major derailment cause. Finally, this research identifies the existing knowledge gaps regarding train accident causal analysis and suggests future research directions.

The exposition of this paper is as follows. First, this paper reviews the relevant literature on rail safety and risk analysis to understand the state of the art in the literature and to identify and fill knowledge gaps. Second, this paper introduces the collection of accident data from the Federal Railroad Administration (FRA) and traffic exposure data from the Surface Transportation Board (STB). Third, this paper analyzes accident-cause-specific freight-train derailment rate by railroad and season on mainlines in the United States between 2000 and 2014. Finally, this paper presents its major findings and suggests directions for further research.

Literature Review

Railroad safety and risk analysis is an increasingly active research field. We review those studies that are directly related to the objectives of this paper, focusing particularly on studies based on the FRA train accident data. Many other rail-engineering studies (e.g., mechanics, lab and field test, simulation) were not included in the literature review of this paper. In general, previous train accident data analysis studies can be classified into the following categories: (1) accident cause analysis, (2) accident frequency analysis, (3) accident severity analysis, and (4) accident prevention and risk analysis. The following subsections briefly introduce each stream of research. Interested readers can refer to the cited references for technical details.

Accident Causes

The FRA identifies distinct accident causes organized into the categories of track defects, rolling stock failures, signaling failures, human errors, and other causes (FRA 2011). Each accident cause describes a specific circumstance or contributing factor that may lead to a train accident. It was found that track failures were among the common freight-train derailment causes in the United States (Barkan et al. 2003; Liu et al. 2011, 2012), most likely because heavy-haul rail operations impose cyclic high-load impact loads to

[1]Assistant Professor, Dept. of Civil and Environmental Engineering, Rutgers, State Univ. of New Jersey, CoRE 606, 96 Frelinghuysen Rd., Piscataway, NJ 08854-8018. E-mail: xiang.liu@rutgers.edu

Note. This manuscript was submitted on April 11, 2016; approved on September 19, 2016; published online on November 21, 2016. Discussion period open until April 21, 2017; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Transportation Engineering, Part A: Systems***, © ASCE, ISSN 2473-2907.

the track infrastructure. Furthermore, the extensive use of equipment may lead to some mechanical problems (e.g., overheated bearings, broken wheels), which may result in train derailments. For example, when the oil leaked or dried out, the bearings are overheated and could destroy the entire railroad car if not detected. Also, thermal stress and high impact load may damage railroad wheels and possibly cause wheel failures.

In addition to derailments, the causes for train collisions and highway–rail grade crossing incidents were also studied. Evans (2007, 2010, 2011) and Liu (2016) analyzed train collisions due to human error in Great Britain and the United States, respectively. Mok and Savage (2005) and Savage (2006) discussed improvements in rail-highway grade crossing safety resulting from engineering, law enforcement, and the educating of the public about the risks.

Accident Frequency Analysis

Count data regression models were widely used for transportation accident frequency analysis. Evans (2007, 2010, 2011) developed Poisson regression and negative binomial regression models to estimate train accident frequency by year and traffic volume. Liu (2015, 2016) applied similar techniques to U.S. freight-train derailments and collisions. Some other studies focused on associating train accident rate (accident frequency normalized by traffic exposure) with a set of engineering factors, such as FRA track class (Anderson and Barkan 2004). The FRA track class is determined by maximum operating speed. A higher track class denotes both a higher maximum speed and correspondingly more stringent safety standards (FRA 2014). The maximum allowable speed for freight trains on Track Class 1 to Track Class 5 are 16 km/h (10 mph), 40 km/h (25 mph), 64 km/h (40 mph), 97 km/h (60 mph), and 129 km/h (80 mph), respectively (note: nonsignaled track has a speed limit 79 km/h (49 mph) for freight trains, regardless of track class).

Accident Severity Analysis

In addition to analyzing accident frequency, researchers also analyzed accident severity. Train accident severity can be measured by different metrics, such as the number of cars or locomotives derailed, property damage, casualties, or environmental impact. A number of studies focused on analyzing the number of cars derailed because this statistic is an indication of the accumulated kinetic energy of the accident. Barkan et al. (2003) found that the average number of cars derailed per train accident is associated with accident speed. A higher speed may result in more cars derailed, given all other factors being equal. In addition to speed, Saccomanno and El-Hage (1989, 1991) found that the point of derailment (the position of the first derailed car), train length, and accident cause also affect accident severity. They developed a truncated geometric model to quantify the relationship between the expected number of cars derailed and the previously mentioned factors. Liu et al. (2012) found that, on average, derailments caused by track failure typically resulted in more cars being derailed than in the case of mechanical failures. For example, broken rails resulted in an average of 13 cars derailed per freight train derailment, as compared to seven cars derailed from bearing failures (Liu et al. 2012). Considering that the distribution of the number of cars derailed is asymmetrical, Liu et al. (2013) developed a quantile regression to model the median and other quantiles of freight-train-derailment severity. In addition to analyzing the number of cars derailed, some other studies analyzed other metrics, such as casualties (Clarke and Loeb 2005; Evans 2007), evacuated population (Erkut and Verter 1995, 1998), or environmental impact (Yoon et al. 2009).

Accident Prevention and Risk Management

It is of keen interest to both the rail industry and the FRA to identify promising strategies for the prevention of train accidents. The literature has largely concentrated on individual risk reduction options, such as track quality upgrade (Liu et al. 2011), equipment condition improvement (Schlake et al. 2011), asset condition monitoring (Schlake 2010), or advanced train control technology (Martland et al. 2001). However, there is very limited prior research analyzing the interactive effect of simultaneously applied multiple accident prevention strategies. Furthermore, few studies have comprehensively quantified the implementation costs and benefits (safety and business) of accident prevention strategies, individually or in combination. Because of the low-probability, high-severity characteristics of train accidents, researchers used risk analysis methods to account for the combination of train accident likelihood and severity. Either an expected consequence model or an expected utility model is used when the decision makers are risk-averse toward catastrophic consequences (Erkut and Verter 1995, 1998; Erkut and Ingolfsson 2000). This line of research largely focuses on the transport of passengers or hazardous materials.

Knowledge Gaps and Contributions of This Paper

To the author's knowledge, no published studies explicitly analyzed how major derailment causes vary with railroad and season in the United States. Also, no statistical model exists to predict accident-cause-specific derailment frequency with a set of influencing factors. Understanding the temporal-spatial variation in accident cause distribution can provide the rail industry with information to develop cost-justified safety improvement strategies. This paper analyzes whether and how major accident causes differ by railroad and season through a log-linear statistical model. If a certain carrier is expected to have more derailments of a specific cause in a particular season, proper inspection and remedial actions may be undertaken to improve safety. This paper focuses on the two leading freight train derailment causes (broken rails, track geometry defects), but its statistical methodology can be adapted to other accident causes as well. In essence, this research falls into the exploratory-analysis category. This type of analysis focuses on understanding what the data pattern is, instead of explaining why the pattern exists. The latter requires an *explanatory analysis*, which associates the response variable with a set of potential affecting (ideally causal) factors. Explanatory analysis, which comprises the next stage of this paper, requires extensive access to railroads' infrastructural and operational data, which are proprietary and rarely available to the public.

Data Sources

Accident Count Data

The FRA maintains three major databases, each related to a different aspect of train operating safety: train accidents, employee casualties, and grade crossing collisions. A particular reportable event may require that reports be submitted to any or all of these databases, alone or in combination, depending on the circumstances. The rail equipment accident/incident report (REAIR) form (FRA F 6180.54) is used by railroads to report all accidents that exceed a monetary threshold of damages to infrastructure and rolling stock. The form accounts for damage to on-track equipment, signals, track, track structures, and roadbed (FRA 2011). FRA compiles these reports into the Rail Equipment Accident (REA) database, which records rail equipment accident data dating back to the 1970s. In addition to the REAIR, the highway–rail grade crossing accident/incident report (FRA F 6180.57) and the death, injury, or occupational illness summary (FRA F 6180.55a) are the two other principal railroad accident and incident reporting forms. Because of the overlap of the reporting criteria, a single accident may require more than one report. This study used data exclusively from the FRA REA database. The REA database contains useful information on the type of accident or incident (e.g., derailment, collision, grade crossing incident, etc.), type of track (mainline, yard, siding, industrial), cause of accident (e.g., track failure, mechanical failure, human error), consequences (number of cars derailed, track and rolling stock damage, casualties), and other information. This information has been used to support numerous railroad safety and risk analyses.

Traffic Exposure Data

Traffic exposure is another important variable in railroad safety and risk analysis. U.S. railroads typically use train-miles, car-miles, or gross ton-miles to represent traffic volume (Nayak et al. 1983; Anderson and Barkan 2004). Previous studies found that some derailment causes are train-mile related, whereas other derailment causes are related to the number of car-miles (Schafer and Barkan 2008). The differences between train-miles and car-miles vary with the number of cars per train. According to Schafer and Barkan (2008), car-mile-related causes are those for which the likelihood of an accident is proportional to the number of car-miles operated. These include most track component failures for which accident likelihood is proportional to the number of load cycles imposed on the track (e.g., broken rails or track geometry defects). By comparison, train-mile-related causes are those in which the accident likelihood is proportional to the number of train-miles operated. These include most human-error failures for which accident likelihood is independent of train length and depends only on train exposure.

Because most track-related accident causes (e.g., broken rails, track geometry defects) are car-mile related (Schafer and Barkan 2008), it is more proper to calculate the derailment rate per car-mile. However, monthly car-mile data are not available in FRA databases. The only information from the FRA Operational Safety Database is the monthly train mile data reported by each railroad. In addition to the FRA database, the Surface Transportation Board (STB), an economic regulator of the railroad industry, requires each Class 1 railroad to submit its operational data (including annual car miles) through the Form R-1. These data are publicly available on the STB website (STB 2014).

Because of data limitations, this paper infers monthly car-mile data by using the distribution of monthly train-mile traffic data. For example, if 10% of annual train-miles for a specific railroad occur in January (from the FRA Operational Safety Database), we assume that January also has 10% of annual car-miles. Because annual car-mile data are available through the STB database, we can estimate car-miles for that railroad in January. The underlying assumption is that there is no substantial variation in train length by month (train length affects car-miles). This assumption might be reasonable if a railroad does not substantially change train length in the fleet within different months in a year. When railroad carriers use the model in this paper, they should update the analysis based on their actual car-mile data in each month because this information is not available to the author at the time of writing this paper.

Table 1. Accident Frequency by Accident Type and Track Type, U.S. Freight Railroads, 2000–2014 (Data from Liu 2016)

Track type	Derailment	Collision	Highway–rail	Other	Total
Main	6,026	429	1,929	874	9,258
Yard	4,220	524	14	518	5,276
Siding	632	33	7	66	738
Industry	1,286	76	9	190	1,561
Total	12,164	1,062	1,959	1,648	16,833

Derailment Analysis

Scope of the Analysis

In the FRA REA database, there are four types of track: main, siding, yard, and industrial. Operational functions differ among these four track types. Each one can have different types of accident, causes, and consequences. In addition, train accidents are classified into derailment, collision, highway–rail grade crossing incident, and several other, less frequent types. Table 1 presents an analysis of train derailment frequency and severity both by track type and accident type, including data from 2000 to 2014 (Liu 2016).

Table 1 shows that derailment was the most common type of accident on each track type, accounting for 72% of train accidents across all types of track. Because of the prevalence of derailment on U.S. freight railroads, the following sections focus exclusively on the analysis of derailments. The methodology can be adapted to other types of train accident as well.

Major Derailment Causes

The FRA REA database records hundreds of accident cause codes. Each cause code describes a specific accident circumstance. The train accident cause codes are hierarchically organized and categorized into five major cause groups: track, equipment, human factors, signal, and miscellaneous (FRA 2011). Within each cause group, the FRA organizes individual cause codes into subgroups of related causes. A variation on the FRA subgroups was developed by Arthur D. Little (ADL), in which similar cause codes were combined into groups on the basis of expert opinion (ADL 1996). The ADL groupings are similar to FRA's subgroups but are more fine-grained for certain causes, allowing greater resolution in some cases. These groups were used to analyze cause-specific derailment frequency and severity. Note that the ADL accident cause grouping might not be the only grouping approach. Additionally, the same cause may fall into multiple groups. Therefore, if analysts use a different accident cause grouping scheme, the analyses should be adapted accordingly.

The prior literature has focused exclusively on nationwide statistics on accident cause distribution. For example, Barkan et al. (2003) and Liu et al. (2012) found that track and mechanical failures are common accident causes on U.S. freight railroads. In particular, broken rails were identified as the leading track-related freight train derailment cause (Barkan et al. 2003; Liu et al. 2012; Liu and Dick 2016). However, there is no prior published research that addresses railroad-wide major derailment causes.

The four largest freight railroads operating in the United States include UP, BNSF, NS, and CSX. In 2012, these four railroads accounted for 76% of network mileage, over 80% of traffic carloads, and 89% of revenue among all the railroads (AAR 2014). Table 2

Table 2. Railroad-Specific Freight Train Derailment Cause Distribution on Mainlines, 2000 to 2014

			Accident cause									
			08T	04T	10E	12E	09H	05T	03T			
Rail-road	Top 1	Top 2	Broken rails	Track geometry defects	Bearing failures	Broken wheels	Train handling	Buckled track	Wide gauge	Others	Total	Percentage by top 2 causes
BNSF	04T	08T	132	142	113	109	54	81	64	940	1635	17
CSX	08T	10E	219	46	65	23	42	33	30	548	1,006	28
NS	08T	12E	116	23	34	49	37	5	24	369	657	25
UP	08T	04T	232	138	92	109	90	58	53	1,076	1,848	20
Total	08T	04T	699	349	304	290	223	177	171	2,933	5,146	20

Note: 08T = broken rails or welds; 04T = track geometry defects (excl. wide gauge); 03T = wide gauge; 10E = bearing failures (car); 09H = train handling error (excl. brakes); 05T = buckled track; 12E = broken wheels (car).

presents the top derailment causes on each of the four selected freight railroads.

With the exception of BNSF Railway, broken rails were the top derailment causes for each railroad. In BNSF, track geometry defects were slightly more frequent than broken rails (142 versus 132). Track geometry defects in this study include cross level, profile, improper superelevation, and other defects (FRA 2011). On Eastern Railroads (CSX and NS), bearing failures (10E) and broken wheels (12E) were the second most common causes. For all four railroads combined, the top five derailment causes were broken rails or welds (08T), track geometry defects (04T), bearing failures (10E), broken wheels (12E), and train handling error (09H).

Overall, the top two causes (broken rails and track geometry defects) accounted for 20% of mainline freight-train derailments on four railroads combined. It indicates that prevention of these major causes may lead to a substantial reduction in derailment risk. Over the past decades, the U.S. freight rail industry has continually invested in technologies and operational changes to prevent track and equipment failures. For example, the industry employs advanced nondestructive detection technologies (e.g., ultrasonic, eddy current) to detect rail defects (Garcia and Zhang 2006) and develop risk-based approaches to prioritize track inspection frequencies (Zarembski and Joseph 2005). In terms of track geometry defects, the Federal Railroad Administration and railroads developed automated or autonomous systems to collect track geometry data (Carr et al. 2009). With respect to bearing failures and broken wheels, the industry uses wayside or train-borne sensors for condition monitoring of specific mechanical components and develop risk-based alert thresholds to guide rolling stock inspection and maintenance (Schlake 2010). In future research, it would be interesting to explore the development and use of more advanced technologies to continue to prevent these major accident causes related to infrastructure and rolling stock on heavy-haul freight railroads.

Statistical Modeling of Derailment Frequency

Model Development

After identifying major derailment causes on freight railroads, of interest is to develop a parametric model that can estimate the number of train derailments for each major cause, given affecting factors such as railroad, season, and traffic exposure. These predictor variables were selected to understand whether there is a spatial-temporal variation of derailment occurrence by accident cause given different traffic volumes. Depending on questions of interest and data availability, future research can include additional factors. The derailment frequency data by railroad, cause, and season came from the FRA REA database. Seasonal car-mile data were estimated based on the STB database. The season *spring* includes March to May, *summer* includes June to August, *fall* includes September to November, and *winter* includes December to February.

Table 3 is a contingency table displaying the distribution of derailment frequency by railroad and season.

The multicategory table contains four railroads (BNSF, CSX, NS, UP) and four seasonal levels (spring, summer, fall, winter); their complex interactive relationship is difficult to discern visually. Therefore, a statistical methodology called *log-linear modeling* is used to quantify and assess the effects of season, railroad, and traffic volume on derailment frequency. Log-linear modeling is a widely adopted approach to model Poisson-based accident count data (Abdel-Aty and Abdelwahab 2000). Eq. (1) presents a general form of the log-linear model for estimating freight train derailment frequency

$$\log(\mu) = \beta_0 + \beta_1 Season + \beta_2 Railroad + \beta_3 Railroad \times Season + \log(Traffic) \tag{1}$$

where μ = derailment frequency for a specific accident cause; *Season* = categorical variable (Spring, Summer, Fall, Winter); *Railroad* = categorical variable (BNSF, CSX, NS, UP), $Railroad \times Season$ = interaction term between railroad and season variables; *Traffic* = traffic volume (million car-miles); and β_0, β_1, β_2, and β_3 = parameter coefficients.

In the model, the variable *Season* represents the seasonal effect. Let one season (e.g., winter) be the baseline case where the parameter coefficient β_1 represents the difference between other seasons and the baseline season (i.e., winter). Similarly, *Railroad* represents railroad differences in derailment frequency, given all else being equal. The interaction term $Season \times Railroad$ describes whether the distribution of derailment frequency by season would also vary by railroad. If its coefficient β_3 is zero, there is no interactive effect between these two variables (i.e., railroad and season). Traffic volume is an exposure variable (Evans 2007), describing the fact that given each specific scenario of season and railroad, derailment frequency increases with traffic volume.

Table 4 shows the statistical significance test for parameter coefficients, for the top two derailment causes (broken rails and track geometry defects). The P-value indicates whether a variable is significant. In general, when the P-value is larger than 0.05, the variable is not statistically significant. As shown in Table 4, the interaction term between railroad and season is not statistically significant, meaning that the conditional effect of season or railroad is not mutually dependent.

Table 3. Derailment Frequency by Cause, Season, and Railroad on Mainlines, 2000 to 2014

	Accident cause									
RR/season	08T	04T	10E	12E	09H	05T	03T	Others	Total	Billion car-miles
BNSF	132	142	113	109	54	81	64	940	1,635	160
Spring	22	41	35	27	17	12	23	240	417	40
Summer	21	46	20	11	16	63	10	262	449	40
Fall	55	34	22	21	10	6	11	211	370	41
Winter	34	21	36	50	11	0	20	227	399	39
CSX	219	46	65	23	42	33	30	548	1,006	80
Spring	37	9	17	7	17	9	9	151	256	21
Summer	33	21	13	4	7	21	7	125	231	20
Fall	72	8	12	3	11	3	0	125	234	20
Winter	77	8	23	9	7	0	14	147	285	19
NS	116	23	34	49	37	5	24	369	657	67
Spring	19	7	6	22	9	3	8	101	175	17
Summer	14	8	7	3	11	2	5	108	158	17
Fall	38	4	6	7	8	0	1	65	129	17
Winter	45	4	15	17	9	0	10	95	195	16
UP	232	138	92	109	90	58	53	1,076	1,848	197
Spring	41	41	26	27	25	8	16	284	468	50
Summer	41	51	17	21	25	46	9	270	480	50
Fall	77	27	18	28	22	4	9	257	442	50
Winter	73	19	31	33	18	0	19	265	458	47
Total	699	349	304	290	223	177	171	2,933	5,146	504

Note: 08T = broken rails or welds; 04T = track geometry defects (excl. wide gauge); 03T = wide gauge; 10E = bearing failures (car); 09H = train handling error (excl. brakes); 05T = buckled track; 12E = broken wheels (car).

Table 4. Statistical Significance of Predictor Variables

Variables	Degree of freedom	Chi-square	Pr > ChiSq
Broken rails			
Season	3	88.00	<0.0001
Railroad	3	113.89	<0.0001
Railroad × Season	9	7.58	0.5766
Track geometry defects			
Season	3	19.18	0.0003
Railroad	3	23.20	<0.0001
Railroad × Season	9	4.77	0.8540

After removing the insignificant variable ($Railroad \times Season$), a simplified log-linear model is as follows:

$$\log(\mu) = \beta_0 + \beta_1 Season + \beta_2 Railroad + \log(Traffic) \quad (2)$$

Based on the assumption that the number of derailments follows a Poisson distribution (Liu 2015), parameter coefficients (β_0, β_1, β_2) can be estimated by maximizing the likelihood function L:

$$L = \prod_{i=1}^{N} \frac{\mu_i^{y_i} \exp(-\mu_i)}{y_i!} \quad (3)$$

where μ_i = estimated number of derailments in the ith category from the log-linear model; y_i = empirical number of derailments in the ith category; and N = number of categories.

The parameter coefficients were estimated using a statistical tool called statistical analysis system (SAS). Table 5 shows parameter coefficients and their standard errors. The P-value for each parameter coefficient indicates its statistical significance in relation to the reference level (for the season variable, the reference level is winter; for the railroad variable, the reference level is Union Pacific Railroad). When the P-value is larger than 5%, a variable is generally deemed to be nonsignificant. For example, for broken-rail-caused derailments (Table 5), the P-value for the parameter coefficient of season *Fall* is 0.9905, indicating that there is no statistical difference between fall and winter in terms of broken-rail-caused derailment frequency, given all other factors (railroad, traffic volume) being the same. Similarly, Table 5 shows that there is no statistical difference between CSX Transportation and Union Pacific Railroad in terms of track-geometry-defect-caused derailment likelihood, given season and traffic volume. Based on parameter coefficient estimates in Table 5, the predicted number of freight-train derailments by cause is as follows:

- Broken rail

$$\log(\mu_B) = 0.48 - 0.71 I_{\text{Spring}} - 0.79 I_{\text{Summer}} + 0.001 I_{\text{Fall}} - 0.36 I_{\text{BNSF}} + 0.84 I_{\text{CSX}} + 0.39 I_{\text{NS}} + \log(Traffic) \quad (4)$$

- Track geometry defects

$$\log(\mu_T) = -0.83 + 0.58 I_{\text{Spring}} + 0.84 I_{\text{Summer}} + 0.28 I_{\text{Fall}} + 0.23 I_{\text{BNSF}} - 0.20 I_{\text{CSX}} - 0.71 I_{\text{NS}} + \log(Traffic) \quad (5)$$

where μ_B = estimated derailment frequency due to broken rails; μ_T = estimated derailment frequency due to track geometry defects; $I_{\text{Spring}} = 1$ if the season is spring, 0 otherwise, similar notations for other seasonal indicators; $I_{\text{BNSF}} = 1$ if the railroad is BNSF, 0 otherwise, similar notations for other railroad indicators; and *Traffic* = billion car-miles.

For example, from 2000 to 2014, when there were a total of 40 billion car-miles in spring on BNSF Railway ($Traffic = 40$), broken-rail-caused derailment frequency can be estimated using Eq. (4):

Table 5. Parameter Coefficient Estimates in Log-Linear Modeling

Parameter	Estimate	Standard error	Wald 95% confidence limits		Pr > ChiSq
Broken rails					
Intercept	0.4760	0.0852	0.3091	0.6429	<0.0001
Season					
Fall	0.0011	0.0922	−0.1796	0.1818	0.9905
Spring	−0.713	0.113	−0.9345	−0.4915	<0.0001
Summer	−0.7904	0.1164	−1.0185	−0.5623	<0.0001
Winter (reference)	0	0	0	0	—
Railroad					
BNSF	−0.3631	0.109	−0.5767	−0.1494	0.0009
CSX	0.8395	0.0942	0.6548	1.0241	<0.0001
NS	0.3879	0.1137	0.165	0.6108	0.0006
UP (reference)	0	0	0	0	—
Track geometry defects					
Intercept	−0.8327	0.1537	−1.134	−0.5314	<0.0001
Season					
Fall	0.2802	0.1815	−0.0754	0.6359	0.1225
Spring	0.5832	0.1716	0.247	0.9195	0.0007
Summer	0.8371	0.1648	0.514	1.1601	<0.0001
Winter (reference)	0	0	0	0	—
Railroad					
BNSF	0.2344	0.1195	0.0001	0.4686	0.0499
CSX	−0.2023	0.1703	−0.536	0.1314	0.2347
NS	−0.7141	0.2252	−1.1556	−0.2727	0.0015
UP (reference)	0	0	0	0	—

$$\log(\mu_B) = 0.48 - 0.71(I_{\mathrm{Spring}} = 1) - 0.36(I_{\mathrm{BNSF}} = 1) + \log(40) = 3.099$$

where log represents the natural logarithmic operator.

Therefore, the estimated number of broken-rail-caused derailments at this season on this railroad during the study period is

$$\mu_B = \exp(3.099) = 22.17$$

where the empirical derailment count is 22, from Table 3.

Model Validation

For a log-linear regression model, the goodness of fit can be evaluated using a statistical criterion called *Deviance* (Agresti 2007). The Deviance approximately follows a chi-squared distribution. The Deviance values for the broken-rail-related model [Eq. (4)] and the track-geometry-defect-related model [Eq. (5)] are 7.58 and 4.77, respectively. The corresponding P-values are 0.58 and 0.85, respectively. When the P-value is larger than 0.05, the overall fit of the model is adequate.

In addition to the overall goodness of fit, the empirical-versus-predicted value is conducted by railroad and season (Table 6).

A chi-squared test is used to evaluate whether the model prediction matches the empirical value

$$\chi^2 = \sum_{i=1}^{N} \frac{(O_i - E_i)^2}{E_i} \tag{6}$$

where O_i = observed number of derailments in year i; E_i = estimated number of derailments in year i; and N = number of categories.

The P-value of the chi-squared test for broken-rail-caused derailment analysis is 0.2490, and for track-geometry-defect-caused derailment analysis is 0.2732, both of which are greater than 0.05, suggesting that both models are statistically adequate.

Table 6. Empirical versus Predicted Derailment Frequency by Railroad and Season

		Broken rails		Track geometry defects	
Railroad	Season	Empirical	Predicted	Empirical	Predicted
BNSF	Spring	22	22	41	40
BNSF	Summer	21	20	46	51
BNSF	Fall	55	46	34	30
BNSF	Winter	34	43	21	21
CSX	Spring	37	38	9	13
CSX	Summer	33	34	21	17
CSX	Fall	72	75	8	10
CSX	Winter	77	72	8	7
NS	Spring	19	20	7	7
NS	Summer	14	18	8	8
NS	Fall	38	40	4	5
NS	Winter	45	38	4	3
UP	Spring	41	39	41	39
UP	Summer	41	36	51	50
UP	Fall	77	81	27	29
UP	Winter	73	76	19	21

Model Application Using Conditional Odds Ratio Analysis

One useful output of the model is the conditional odds ratio (COR), which provides a comparison of the effect of different levels of a variable on derailment frequency (Table 7). For example, for broken-rail-caused derailment, the conditional odds ratio for spring versus winter is 0.49. It indicates that, given all else being equal (including traffic volume), the derailment frequency in spring is 49% of that in winter, with a 95% confidence interval between 0.39 and 0.61. Similarly, summer has a lower derailment frequency due to broken rails than winter does, given all else being equal. This discrepancy probably exists because the thermal contraction forces in rails under lower temperature will likely pull apart internal rail

Table 7. Comparison of Derailment Distribution by Cause, Season, and Railroad

Conditional odds ratio	Broken rails	Track geometry defects
$\mu_{\text{spring}}/\mu_{\text{winter}}$	0.49 (0.39, 0.61)	1.79 (1.28, 2.51)
$\mu_{\text{summer}}/\mu_{\text{winter}}$	0.45 (0.36, 0.57)	2.31(1.67, 3.19)
$\mu_{\text{fall}}/\mu_{\text{winter}}$	1.00 (0.84, 1.20)	1.32 (0.93, 1.89)
$\mu_{\text{BNSF}}/\mu_{\text{UP}}$	0.70 (0.56, 0.86)	1.26 (1.00, 1.60)
$\mu_{\text{CSX}}/\mu_{\text{UP}}$	2.32 (1.92, 2.78)	0.82 (0.59, 1.14)
$\mu_{\text{NS}}/\mu_{\text{UP}}$	1.47 (1.18, 1.84)	0.49 (0.31, 0.76)

Note: The values in the parenthesis represent the 95% confidence interval of the conditional odds ratio.

defects, causing more broken rails. From the COR analysis, there is no statistical difference between fall and winter in the rate of broken-rail-caused derailment. An example manual calculation (comparing broken-rail-caused derailment frequency in spring versus in winter, given railroad and traffic volume being equal) of the COR for broken rails is as follows, using Eq. (4):

$$\log(\mu_{\text{Spring}}) = 0.48 - 0.71 I_{\text{Spring}} - 0.36 I_{\text{BNSF}} + 0.84 I_{\text{CSX}} + 0.39 I_{\text{NS}} + \log(Traffic)$$

$$\log(\mu_{\text{Winter}}) = 0.48 - 0.36 I_{\text{BNSF}} + 0.84 I_{\text{CSX}} + 0.39 I_{\text{NS}} + \log(Traffic)$$

$$\log(\mu_{\text{Spring}}) - \log(\mu_{\text{Winter}}) = -0.71$$

$$\frac{\mu_{\text{Spring}}}{\mu_{\text{Winter}}} = \exp(-0.71) = 0.49$$

Regarding railroads, taking Union Pacific Railroad (UP) as a reference, broken-rail-caused derailment rate on BNSF Railway is 70% of that of Union Pacific Railroad. The two Eastern railroads (CSX and NS) have a higher derailment likelihood due to broken rails than the two Western railroads, given the same season and traffic volume.

With respect to track geometry defects, the seasonal effect is opposite to that of broken rails. Spring and summer have more track-geometry-defect-caused derailments than in fall and winter (given railroad and traffic volume), probably due to thermal expansion stress in rails. In terms of track geometry defects, NS appears to have a lower derailment frequency than the other three railroads.

According to the COR analysis, it appears that broken-rail-caused derailment is generally more likely to occur in fall and winter, holding all else being equal, whereas track-geometry-defect-caused derailment occurs more frequently in spring and summer. Given the same season and traffic volume, the two Eastern railroads (CSX and NS) have more derailments due to broken rails, but fewer derailments due to track geometry defects, when compared with the two Western railroads (BNSF, UP). This might partly be due to the temperature difference in each season between the western and eastern region, and/or due to different operational and infrastructure characteristics on different railroads. Because of data limitations, explaining accident rate heterogeneity by railroad is beyond the scope of this paper, but can be one direction for future research.

Conclusions

This research develops a statistical log-linear model to analyze the causal distribution of freight train derailment frequency by railroad and season. The analysis finds that broken rails and track geometry defects were the two leading freight train derailment causes on U.S. mainlines from 2000 to 2014, accounting for approximately 20% of train derailments. The odds of having a broken-rail-caused derailment in fall and winter appear to double that of spring and summer, given the same railroad and traffic level. By contrast, track-geometry-defect-caused derailments occur more frequently in spring (80% more likely) and summer (130% more likely) than in fall and winter, given the same railroad and traffic level. The two Eastern railroads (CSX Transportation and Norfolk Southern Railway) appear to have higher odds of broken-rail-caused derailment than the two Western railroads (BNSF Railway and Union Pacific Railroad), given the same season and traffic volume. In terms of track geometry defects, the two Western railroads appear to have equal or higher odds of a derailment than do the two Eastern railroads, given all else being equal. The statistical modeling techniques herein can be adapted to other types of train accidents or accident causes. Ultimately, future research can be developed to further understand why derailment likelihood varies, what the potential contributing factors are, and how these factors affect derailment risk.

Future Research

The current research focuses on exploring the statistical causal distribution of freight train derailments by railroad and season. Because of data limitations, the author was not able to perform a detailed explanatory analysis to articulate why different railroads or seasons have different causal distributions. This will be one direction for future research. Also, this research focuses on broken rails and track geometry defects, which are the two top derailment causes on freight railroads. The log-linear modeling approach can be adapted to other major derailment causes as well (e.g., wide gauge, broken wheels). In addition to derailment, collision and grade crossing are two other common types of train accident. For collisions and grade-crossing incidents, human error is the major cause. Future research can be directed to understand the distribution of human-related causes for other accidents under different circumstances. Additionally, this research analyzes cause-specific derailment frequency. The severity of an accident (e.g., casualties, number of cars derailed, damage cost) is not analyzed in this paper but can be addressed in future research. The statistical models in this paper were based on data from 2000 to 2014. Future data can be used to validate and possibly refine the current model to reflect the change in railroad safety. Additional data sources (e.g., track geometry inspection data) (Xu et al. 2016) might be used to predict train accident risk in future research. Ultimately, future research can incorporate train safety statistical analysis into an integrated accident prevention and risk management framework that optimizes the portfolio of railroad safety improvement options.

Acknowledgments

The author was partially funded by the Department of Civil and Environmental Engineering (CEE), the School of Engineering (SOE), and the Center for Advanced Infrastructure and Transportation (CAIT), all at Rutgers University. Ms. Ling Ma assisted with the development of log-linear modeling in this paper. Mr. Masoud Imaniboossejin assisted with formatting the paper. However, the author is responsible for all views and analyses in this paper.

References

AAR (Association of American Railroads). (2014). "Analysis of Class I railroads." ⟨https://www.aar.org/StatisticsAndPublications/Documents/AAR-Stats.pdf⟩ (Jun. 16, 2016).

Abdel-Aty, M. A., and Abdelwahab, H. T. (2000). "Exploring the relationship between alcohol and the driver characteristics in motor vehicle accidents." *Accid. Anal. Prev.*, 32(4), 473–482.

ADL (Arthur D. Little). (1996). "Risk assessment for the transportation of hazardous materials by rail." Boston.

Agresti, A. (2007). *An introduction to categorical data analysis*, Wiley, New York.

Anderson, R. T., and Barkan, C. P. L. (2004). "Railroad accident rates for use in transportation risk analysis." *Transp. Res. Rec.*, 1863, 88–98.

Barkan, C. P. L., Dick, C. T., and Anderson, R. T. (2003). "Analysis of railroad derailment factors affecting hazardous materials transportation risk." *Transport. Res. Rec.*, 1825, 64–74.

Carr, G. A., Tajaddini, A., and Boris, Nejikovsky. (2009). "Autonomous track inspection systems—Today and tomorrow." *2009 American Railway Engineering and Maintenance Way Association Annual Conf. and Exposition*, Lanham, MD.

Clarke, W. A., and Loeb, P. D. (2005). "The determinants of train fatalities: Keeping the model on track." *Transp. Res. Part E.*, 41(2), 145–158.

Erkut, E., and Ingolfsson, A. (2000). "Catastrophe avoidance models for hazardous materials route planning." *Transp. Sci.*, 34(2), 165–179.

Erkut, E., and Verter, V. (1995). "A framework for hazardous materials transport risk assessment." *Risk Anal.*, 15(5), 589–601.

Erkut, E., and Verter, V. (1998). "Modeling of transport risk for hazardous materials." *Oper. Res.*, 46(5), 625–642.

Evans, A. W. (2007). "Rail safety and rail privatization in Great Britain." *Accid. Anal. Prev.*, 39(3), 510–523.

Evans, A. W. (2010). "Rail safety and rail privatization in Japan." *Accid. Anal. Prev.*, 42(4), 1296–1301.

Evans, A. W. (2011). "Fatal train accidents on Europe's railways: 1980–2009." *Accid. Anal. Prev.*, 43(1), 391–401.

FRA (Federal Railroad Administration). (2010). "National rail plan progress report." Washington, DC, 14.

FRA (Federal Railroad Administration). (2011). "FRA guide for preparing accident/incident reports." Washington, DC, 211.

FRA (Federal Railroad Administration). (2014). "Track safety standards–Classes 1 through 5." Chapter 1, *Track and rail and infrastructure integrity compliance manual*, Vol. 2, Washington, DC.

FRA (Federal Railroad Administration). (2015). "Accident data as reported by railroads." ⟨http://safetydata.fra.dot.gov/OfficeofSafety/publicsite/on_the_fly_download.aspx⟩ (Jul. 1, 2016).

Garcia, G., and Zhang, J. (2006). "Application of ultrasonic phased arrays for rail flaw inspection." *DOT/FRA/ORD-06/17*, Federal Railroad Administration, Washington, DC.

Liu, X. (2015). "Statistical temporal analysis of freight-train derailment rates in the United States: 2000 to 2012." *Transp. Res. Rec.*, 2476, 119–125.

Liu, X. (2016). "Analysis of collision risk for freight trains in the United States." *Transp. Res. Rec.*, 2546, 121–128.

Liu, X., Barkan, C. P. L., and Saat, M. R. (2011). "Analysis of derailments by accident cause: Evaluating railroad track upgrades to reduce transportation risk." *Transp. Res. Rec.*, 2261, 178–185.

Liu, X., and Dick, C. T. (2016). "Risk-based optimization of rail defect inspection frequency for petroleum crude oil transportation." *Transp. Res. Rec.*, 2545, 27–35.

Liu, X., Saat, M. R., and Barkan, C. P. L. (2012). "Analysis of causes of major train derailment and their effect on accident rates." *Transp. Res. Rec.*, 2289, 154–163.

Liu, X., Saat, M. R., and Barkan, C. P. L. (2013). "Analysis of U.S. freight-train derailment severity using zero-truncated negative binomial regression and quantile regression." *Accid. Anal. Prev.*, 59, 87–93.

Martland, C. D., Zhu, Y., Lahrech, Y., and Sussman, J. M. (2001). "Risk and train control: A framework for analysis." *Transp. Res. Rec.*, 1742, 25–33.

Mok, S., and Savage, I. (2005). "Why has safety improved at rail-highway grade crossings?" *Risk Anal.*, 25(4), 867–881.

Nayak, P. R., Rosenfield, D. B., and Hagopian, J. H. (1983). "Event probabilities and impact zones for hazardous materials accidents on railroads." *DOT/FRA/ORD-83/20*, *Federal Railroad Administration*, Washington, DC.

Saccomanno, F. F., and El-Hage, S. M. (1989). "Minimizing derailments of railcars carrying dangerous commodities through effective marshaling strategies." *Transp. Res. Rec.*, 1245, 34–51.

Saccomanno, F. F., and El-Hage, S. M. (1991). "Establishing derailment profile by position for corridor shipments of dangerous goods." *Can. J. Civ. Eng.*, 18(1), 67–75.

Savage, I. (2006). "Does public education improve rail-highway crossing safety?" *Accid. Anal. Prev.*, 38(2), 310–316.

Schafer, D. H., and Barkan, C. P. L. (2008). "Relationship between train length and accident causes and rates." *Transp. Res. Rec.*, 2043, 73–82.

Schlake, B. W. (2010). "Impact of automated condition monitoring technologies on railroad safety and efficiency." M.S. thesis, Univ. of Illinois, Urbana, IL.

Schlake, B. W., Barkan, C. P. L., and Edwards, J. R. (2011). "Train delay and economic impact of in-service failures of railroad rolling stock." *Transp. Res. Rec.*, 2261, 124–133.

STB (Surface Transportation Board). (2014). "Class I railroad R-1 report." Washington, DC.

Xu, P., Sun, Q., Liu, R., and Souleyrette, R. R. (2016). "Optimal match method for milepoint postprocessing of track condition data from subway track geometry cars." *J. Transp. Eng.*, 10.1061/(ASCE)TE.1943-5436.0000859, 04016028.

Yoon, H., Werth, C. J., Barkan, C. P. L., Schaeffer, D. J., and Anand, P. (2009). "An environmental screening model to assess the consequences to soil and groundwater from railroad-tank-car spills of light non-aqueous phase liquids." *J. Hazard. Mater.*, 165(1–3), 332–344.

Zarembski, A., and Joseph, W. (2005). "Characterization of broken rail risk for freight and passenger railway operations." *Proc., American Railway Engineering and Maintenance of Way Association (AREMA) Annual Conf.*, Chicago.