Quantitative Analysis of Changes in Freight Train Derailment Causes and Rates

Brandon Z. Wang, Ph.D.[1]; Christopher P. L. Barkan, Ph.D.[2]; and M. Rapik Saat, Ph.D.[3]

Abstract: The mainline freight train derailment rate of major railroads in the United States declined 49% from 2006 to 2015. Nevertheless, derailments remain the leading cause of major railroad accidents. Identification and quantification of the types of train accidents, the trackage where they occur, and the causes having the greatest effect on train safety and risk is critical to determine the most effective strategies for further improvement. Federal Railroad Administration data were used to quantify factors contributing to the declining accident trend. Most derailment causes declined with the greatest reductions in broken rails and welds, track geometry, and other axle and journal defects. Of the few causes that increased, extreme weather was the largest. An updated statistical model of the relationship between track class, traffic density, method of operation, and derailment rate is also developed. Derailments declined uniformly with respect to all combinations of the three factors, indicating a broad general decline across the network. The new model also provides up-to-date derailment rate estimates for use in risk analysis of railroad freight and hazardous materials transportation. **DOI: 10.1061/JTEPBS.0000453.**

Introduction

Railroad train safety in the United States has improved considerably over the past decade. One measure of this is that the mainline freight train derailment rate for US Class 1 railroads declined 49% from 2006 to 2015 (FRA 2016). Despite this improvement, further reduction in railroad accidents is an ongoing objective of the rail industry and government. Train accident rates are affected by infrastructure, equipment, operating characteristics, traffic volume, and other factors. Rail operation involves a variety of potential hazards and risks and some safety measures are more effective than others so different mitigation strategies and levels may be appropriate (Evans 2013). It is in the interest of all stakeholders that risk reduction resources be allocated as efficiently as possible. As train accidents become less frequent, understanding which improvements will most effectively improve safety requires more sophisticated quantitative approaches.

Background and Objectives

Extensive research has been conducted on factors affecting highway safety (Karlaftis and Golias 2002; Vanlaar and Yannis 2006; Milton et al. 2008; Lord and Mannering 2010). However, railroad accidents have several intrinsic differences compared to highway accidents. These differences affect the pertinent questions, as well as the data, methodology, and statistics used to address them.

One key difference is accident severity. Most highway vehicles operate singly, whereas the average freight train has more than 70 cars (AAR 2018). Consequently, train derailments can vary in size from a single vehicle (railcar or locomotive) up to many dozens of vehicles involved in a single accident (in some cases as many as 80). Although very large highway accidents do occur, only 12% of highway crashes involve three or more vehicles (NHTSA 2008), whereas 64% of Class 1 railroads' mainline accidents involve three or more rail vehicles.

Another difference is that railroad accidents have many more potential failure modes. This is a result of the large size of trains and the complexity of the equipment and its interaction with infrastructure. The US DOT Federal Railroad Administration (FRA) identifies more than 400 specific cause codes for railroad accidents. Highway and railroad accident causes differ at a more general level as well. The National Highway Traffic Safety Administration (NHTSA 2008) reported that 88% of the vehicles involved in accidents had no adverse conditions, suggesting that most were the result of driver behavior. Although human factors are an important aspect of railroad train safety, their relative frequency is almost exactly reversed compared to highway accidents; infrastructure and equipment failures caused 87% of Class 1 mainline freight train derailments.

Quantitative analyses of railroad train safety and risk must account for these factors in order to properly understand how to measure the impact of different accident causes, the efficacy of possible solutions, and their potential effect on rail transportation risk. This in turn requires different analytical and statistical approaches in order to understand the frequency, severity, and causes of accidents. The FRA Office of Safety Analysis compiles detailed data on a range of railroad safety metrics and publishes aggregated summary statistics for accident rates (FRA 2016); however, a more fine-grained understanding of specific causes is needed in order to focus resources most effectively.

The substantial reduction in train accidents indicates that there have been major changes since 2006 that affected railroad

[1]Graduate Research Assistant, Rail Transportation and Engineering Center–RailTEC, Dept. of Civil and Environmental Engineering, Univ. of Illinois at Urbana–Champaign, 205 N. Mathews Ave., Urbana, IL 61801 (corresponding author). Email: zwang144@illinois.edu

[2]Professor, Rail Transportation and Engineering Center–RailTEC, Dept. of Civil and Environmental Engineering, Univ. of Illinois at Urbana–Champaign, 205 N. Mathews Ave., Urbana, IL 61801. Email: cbarkan@illinois.edu

[3]Research Assistant Professor, Director of Operations Analysis, Rail Transportation and Engineering Center–RailTEC, Dept. of Civil and Environmental Engineering, Univ. of Illinois at Urbana–Champaign, 205 N. Mathews Ave., Urbana, IL 61801; Director of Operations Analysis, Association of American Railroads, 425 3rd St. SW, Washington, DC 20024. Email: rapik@aar.org

Note. This manuscript was submitted on October 14, 2019; approved on June 24, 2020; published online on August 31, 2020. Discussion period open until January 31, 2021; 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.

train safety; however, no detailed analysis has been conducted to understand and quantify these changes. Consequently, a principal objective of the research described in this paper was to address a series of questions to inform technical and policy questions regarding which aspects of railroad train safety have improved, and what are the most important opportunities for further improvement. With this in mind, this paper addresses the following specific objectives:

- Identify the types of train accidents, the type of trackage where they occurred, and the changes over the 10-year period of this study. This will provide insight into which accident types and locations contribute the most to risk, which have experienced the greatest change, and the most important sources of remaining risk.
- Develop several graphical and quantitative methods to compare different causes of train derailments in terms of their relative risk as measured by their frequency of occurrence and severity.
- Quantify the magnitude of changes in different train derailment causes in the time period studied to provide the rail sector with insight into which actions have had the greatest effect on rail safety.
- Use improved methodology and up-to-date data to develop estimates of train derailment rates based on a previously published statistical model that correlates derailment rates with three specific railroad infrastructure and operating parameters, and quantitatively evaluate how the improvement in train safety may have affected the model estimates. Develop new train and car derailment rate estimates that can be used for railroad transportation risk analysis.

The paper is structured as follows: first the data sources are described. We then present three separate sections describing the pertinent methodology and results addressing the objectives described, and then close with a summary and conclusions section.

Data and Methodology

The analysis focused on freight train accidents on Class I (i.e., major) US railroads during the interval 2006–2015. These railroads account for the largest portion of the US rail network and operations, with 69% of total freight railroad mileage, 90% of employees, and 94% of total revenue (AAR 2018). They also handle the majority of hazardous materials shipments, further underscoring their importance to US railroad safety and risk. The time period was selected based on several considerations including the most recent years for which complete data were available when the study began, roughly equal railroad freight traffic volume over the two halves of the study period 2006–2010 and 2011–2015 that were used for comparison, and the coincidence of the first half, 2006–2010, with the time period studied in related research by Liu et al. (2017).

Overview of Methodology

This paper uses multiple applied statistics methods to analyze FRA accident data. The first section uses various statistical tests and summary statistics to understand the overall trend and changes in derailment accidents. The second section uses different data visualization techniques and applied regression methods to investigate derailment causes and attempt to quantify them. The last section uses applied statistical methods to investigate the changes in derailments in a three-factor derailment matrix.

FRA Accident Database

Accident Reporting Criteria

The FRA records data on several types of incidents. The principal database used in this study is the FRA Rail Equipment Accident or Incident Report (REAIR) 6180.54. According to FRA (2011a), "Collisions, derailments, fires, explosions, acts of God, or other events involving the operation of railroad on-track equipment (standing or moving) and causing reportable damages greater than the reporting threshold for the year in which the accident/incident occurred must be reported using Form FRA F 6180.54."

The FRA requires railroads to submit accident reports using the REAIR for all accidents or incidents that exceed a specified monetary threshold for combined damages to track, equipment, and/or structures (FRA 2011a). In order to ensure that from year to year comparable accidents involving the same real amount of damages are included in the database, the reporting threshold is periodically adjusted for inflation (FRA 2019). Over the time period covered in this study, it increased from $7,700 in 2006 to $10,500 in 2015 (FRA 2015). This database contains details on each accident, including date, location, railroad, and a number of other variables (FRA 2011b). All highway–rail grade–crossing accidents are recorded in a separate FRA highway rail accident (HRA) database using Form F 6180.57, irrespective of monetary damages. Those grade crossing accidents that exceed the monetary damage threshold criteria must also be recorded in the REAIR database using Form F 6180.54. These two databases record different but complementary information about those incidents that must be reported to both (Chadwick et al. 2012). The initial analysis (described in the "Analysis of Major Accident Types" section) included 35,389 records of Class 1 railroad accidents that were reported to the REAIR database during the 10-year study period. The analysis of mainline freight train derailments that was the principal subject of the research presented in this paper (described in the "Derailment Cause Analysis" and "Three-Factor Derailment Rate Model" sections) considered 2,860 derailment records in the REAIR database.

Types of Accident

The FRA REAIR database includes 13 types of accidents. For the purpose of the research described in this paper, they were categorized into four principal types: derailment, collision, highway–rail grade crossing accident, and other accidents. Collision accidents actually include five separate types: head-on, rear-end, side, raking, and broken-train collisions. Grade crossing accidents included in this database are those in which the ensuing damages exceeded the FRA threshold. We grouped the less frequent accident types in the other category including railroad crossings at grade, explosive-detonation, fire or violent rupture, other impacts, and an FRA category, also called other. The FRA classification of accident types is based on the initial cause but includes information on other contributing causes. For example, if a collision caused a derailment, the initial cause would be classified as a collision, but the secondary cause would be derailment. It would be classified as a collision in the research described herein.

Type of Trackage

The FRA identifies four major types of trackage where accidents may occur—mainline, yard, siding, and industrial—and we used these in our initial analysis to quantify where accidents occur and their severity. The principal information for each accident included in the data set was accident type, track type, specific accident cause, number of cars derailed, and annual rail traffic at the accident location.

J. Transp. Eng., Part A: Systems, 2020, 146(11): 04020127

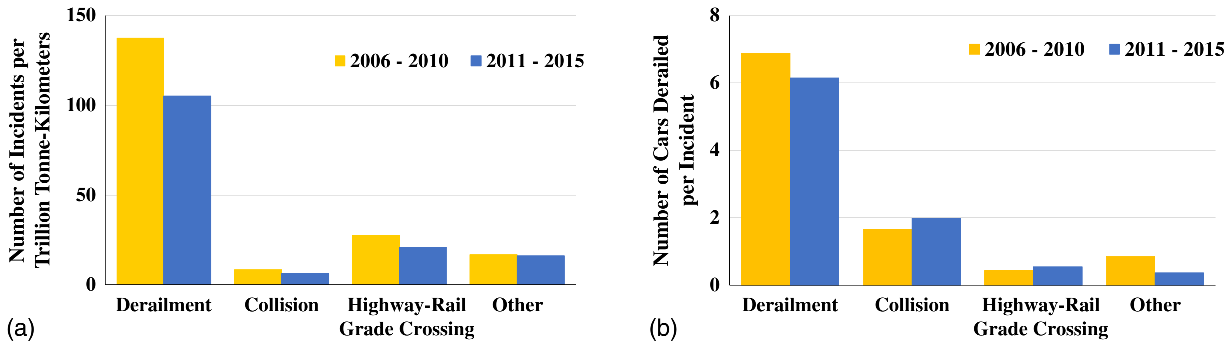

Fig. 1. (a) Incident rate for each incident type; and (b) average number of cars derailed for different incident types.

Traffic Exposure Data

Estimation of accident rates requires information on the exposure of trains to potential incidents. FRA provides some of these data and the Association of American Railroads (AAR) provides additional information on railroad traffic. The annual gross ton-kilometers for Class 1 railroad freight trains were used as a metric for traffic exposure for calculation of accident rates (AAR 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015).

Analysis of Major Accident Types

In order to understand the most important types of accidents and the trackage types where they occurred, we first conducted a high-level analysis of the FRA Rail Equipment Accident or Incident data. We used two quantitative metrics that are indicative of relative risk, frequency of occurrence and average number of cars derailed per incident. The latter has previously been identified as a better proxy for the physical severity of accidents than financial damages because it is independent of the wide potential variation in the value of assets that may be damaged in an accident (Barkan et al. 2003).

The decline in accident rate since the mid-2000s prompted several questions about how and why it had occurred. In order to better understand factors contributing to the trend, the 10-year study was separated into two periods: 2006–2010 and 2011–2015. There were no major differences in rail traffic exposure for the two periods, with 24.4 trillion ton-kilometers (16.7 trillion ton-miles) in the first and 25.1 trillion ton-kilometers (17.2 trillion ton-miles) in the second. The objective was to examine differences between the two periods to address two major questions:

- How had the major types and severity of accidents changed? In particular, were there uniform declines across all major accident types?
- How were different types of trackage related to the change? Had accidents declined uniformly on the FRA's four major categories of track type?

The answers to these questions were intended to identify the key changes over the past decade and which types of accidents and trackage were the largest contributors to risk.

Highway–rail grade crossing accidents were the most frequent type of accident reported to FRA; however, most were below the FRA damage threshold so they were not included in the REAIR database (Chadwick et al. 2012; Chadwick 2017). Derailments were the most common type of accident that exceeded the REAIR threshold.

There were significant differences in the distribution of accident types across different track types for the 10-year study period ($\chi^2 = 987$, degrees of freedom (df) = 9, $P < 0.01$). This is not surprising given the differing types and speeds of operation on different types of track. The same significant differences were observed when the first and second time periods were considered individually ($\chi^2 = 531$, df $= 9$, $P < 0.01$; $\chi^2 = 650$, df $= 9$, $P < 0.01$).

Derailment accidents comprised the majority of the total incidents and number of cars derailed in both time periods. From 2006 to 2010, derailments accounted for 72% of the total number of incidents on all tracks ($n = 4{,}638$) and 96% of the total number of cars derailed ($n = 24{,}011$). Similarly, from 2011 to 2015, derailments were 71% and 95% of the total incidents on all tracks and number of cars derailed, respectively. The number of incidents per trillion gross ton-kilometers for derailment, collision, highway–rail grade crossing, and others in the latter time period decreased 23%, 24%, 23%, and 3%, respectively, compared to the earlier period [Fig. 1(a)]. In terms of severity, derailments and other had 11% and 58% decreases, respectively, while collisions and highway–rail grade crossings increased 19% and 25%, respectively [Fig. 1(b)].

Of the four different accident types, derailments were the most frequent across all track types, accounting for 70% of total incidents ($n = 8{,}392$) and 96% of all cars derailed ($n = 41{,}608$) over the 10-year period. In order to account for possible changes in traffic, the number of derailments for each type of track, mainline, yard, siding, and industry, was normalized using the total gross ton-kilometers for all Class 1 railroad traffic. We found that they declined by 30%, 9%, 20%, and 5%, respectively [Fig. 2(a)]. Each of the four track types also experienced a reduction in average derailment severity, with mainline, yard, siding, and industry trackage having 9%, 11%, 18%, and 15% reductions, respectively [Fig. 2(b)]. Mainline derailments were more common than other types and more severe than on other types of track, comprising 38% of all accidents and 57% of the cars derailed across all accident types and track types. Due to their combination of higher frequency and severity, mainline derailments were the principal subject of the research described in the rest of this paper.

Derailment Cause Analysis

The previous section analyzed derailments compared to other major railroad train accident types, and the type of trackage on which they generally occur. Having established that mainline derailments pose the greatest hazard, we need to understand the major causes of these derailments in order to develop insight into the most effective prevention strategies. Such analyses should also consider derailment severity and the effect on hazardous materials transportation risk. Barkan et al. (2003) investigated the distribution of derailment causes and severity and Liu et al. (2012) investigated the relationship between train speed, FRA track class, and accident cause distribution. This section will investigate major derailment causes using several approaches.

J. Transp. Eng., Part A: Systems, 2020, 146(11): 04020127

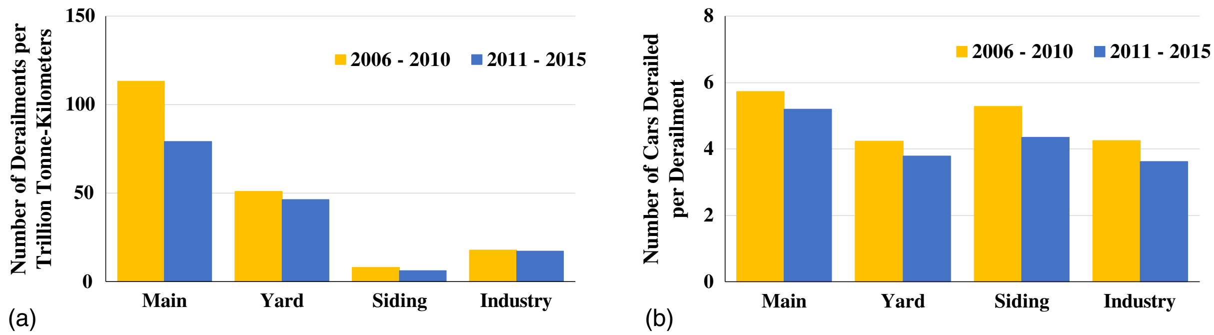

Fig. 2. (a) Derailment rate for each track type; and (b) average number of cars derailed for different track types.

Accident Cause Definitions

The FRA lists approximately 400 different accident cause codes that are separated into five broad categories denoted using single-letter codes as follows: T = track, road bed, and structure; S = signal and communication; H = train operation; E = mechanical and electrical failures; and M = miscellaneous (FRA 2011a). Each of the five categories contains subcategories indicated by three numeric digits that provide additional subcategorization and detail. An illustrative example is T202: broken rail (base). The T indicates it is a track, road bed, and structure cause, the first 2 indicates a rail failure, and the final 02 indicates the specific failure mode, in this case broken base. The suffix C or L after E cause numbers denotes a car or locomotive, respectively. This level of detail is useful for many studies; however, identification of certain trends may benefit from some degree of aggregation of related causes, so FRA also uses a system of subcategories (FRA 2011a).

In the 1990s, Arthur D. Little (ADL, Cambridge, Massachusetts) worked with the AAR developing a refinement of the FRA accident cause category grouping system (ADL 1996; Schafer and Barkan 2008) that consolidated various causes into groups while providing distinction between certain other causes (Anderson and Barkan 2004). The objective of the ADL-AAR grouping was to better link causes that could be addressed through similar or related preventative measures. For example, the FRA subcategorization combines broken rails or welds, joint bars, and rail anchors together, whereas the ADL-AAR method separates broken rails or welds, joint bar defects, and rail anchors due to the difference in the underlying factors affecting them and approaches to preventing them. Another example is that FRA combines buckled track as a subgroup to track geometry, while the ADL-AAR method separates those two causes. The ADL-AAR grouping consists of 51 cause groups. FRA updated its accident reporting methodology in 2011, including adding and removing cause codes and various subcategories as well (FRA 2011a). Most of the new additions were related to signal and human factors. All 43 new FRA causes were interpreted and assigned to the various ADL-AAR subgroups based on consultation with rail industry derailment investigators (Table 1). In addition, a new cause code, extreme weather, was added to define causes due to weather effects.

Derailment Frequency and Severity

Dick et al. (2003) and Barkan et al. (2003) introduced a graphical approach to visualizing train derailment causes to facilitate their comparison in terms of frequency (number of derailments) and severity (number of cars derailed per derailment). Each derailment cause is plotted in terms of its average frequency and average severity (Fig. 3).

Dashed lines divide the graph into four quadrants, with the horizontal line indicating average severity for all causes and the vertical line indicating average frequency for all causes. By definition, points to the right of the vertical line indicate above-average frequency and points above the horizontal line indicate above-average severity.

Derailment causes in the upper-right quadrant occur more frequently and are more severe, thereby posing the greatest risk in terms of number of cars derailed. Conversely, the lower-left quadrant indicates less frequent and less severe derailment causes, which pose the lowest risks overall. The causes in the upper-left quadrant have larger consequences, but their lower frequency makes consistent predictions less reliable. The lower-right quadrant includes less severe but higher frequency derailment causes.

Fig. 3 also includes an enhancement of the frequency–severity graph that we refer to as iso-risk contours; these represent equal levels of risk in terms of whatever risk-generating process may be of interest (Fig. 3). The units on the axes can be defined to best address a particular question. This might involve differential weighting if this helps an investigator or decision maker better interpret the results. Iso-risk contours are an inverse function of the frequency and severity associated with the process.

This concept can be specifically applied to train accidents in the form of iso-car contours representing equal levels of risk in terms of number of cars derailed. Number of cars derailed is a good measure of the impact of a derailment in terms of its physical severity, financial impact, and the potential harm to infrastructure, the environment, and nearby populations. As such, iso-cars provide a means to quantitatively compare the risk associated with different derailment causes. The distance from the origin represents the magnitude of the risk; the greater the iso-car contour, the higher the risk. For example, Derailment Cause A lies on the same iso-car contour as Derailment Cause B (Fig. 3). Cause B occurs more frequently than Cause A, but accidents due to Cause A are sufficiently more severe that the difference in frequency is overcome and the consequent risk is equal. On the other hand, accidents due to Derailment Cause C are on a lower iso-car line because it has lower severity than Cause A and lower frequency than Cause B.

This study used a similar approach to evaluate frequency and severity in which the normalized derailment frequency and average severity per derailment were plotted. Mainline derailment causes over the period 2006–2015 were compared using a frequency–severity plot such as described previously (Fig. 4).

Over the 10-year study period, broken rails or welds were the most frequent mainline derailment cause with the highest iso-car level, consistent with previous studies (Anderson 2005; Liu 2013). Other rail and joint defects were the most severe derailment cause, but they occurred much less frequently.

J. Transp. Eng., Part A: Systems, 2020, 146(11): 04020127

Table 1. Modified ADL-AAR cause groups

Cause group	Description	FRA cause codes
01T	Roadbed defects	T001, T099
02T	Nontraffic, weather causes	T002, T003, T004, T005
03T	Wide gauge	T110, T111, T112, T113
04T	Track geometry (excluding wide gauge)	T101, T102, T103, T104, T105, T106, T107, T108, T199
05T	Buckled track	T109
06T	Rail defects at bolted joint	T201, T211
07T	Joint bar defects	T213, T214, T215, T216
08T	Broken rails or welds	T202, T203, T204, T207, T208, T210, T212, T218, T219, T220, T221
09T	Other rail and joint defects	T299
10T	Turnout defects: switches	T307, T308, T309, T310, T311, T312, T313, T314, T315, T319
11T	Turnout defects: frogs	T304, T316, T317, T318
12T	Miscellaneous track and structure defects	T205, T206, T217, T222, T301, T302, T303, T305, T306, T399, T499, T223, T224, T404
01S	Signal failures	S001, S002, S003, S004, S005, S006, S007, S008, S009, S010, S011, S012, S013, S099, S101, S103, S014, S015, S016, S102, S104
01E	Air hose defect (car)	E00C
02E	Brake rigging defect (car)	E07C
03E	Handbrake defects (car)	E08C, E0HC
04E	UDE (car or locomotive)	E05C, E05L
05E	Other brake defect (car)	E01C, E02C, E03C, E04C, E06C, E09C
06E	Centerplate or car body defects (car)	E20C, E21C, E22C, E23C, E24C, E25C, E26C, E27C, E29C
07E	Coupler defects (car)	E30C, E31C, E32C, E33C, E34C, E35C, E36C, E37C, E39C
08E	Truck structure defects (car)	E44C, E45C
09E	Side bearing and suspension defects (car)	E40C, E41C, E42C, E43C, E47C, E48C
10E	Bearing failure (car)	E52C, E53C
11E	Other axle or journal defects (car)	E51C, E54C, E55C, E59C
12E	Broken wheels (car)	E60C, E61C, E62C, E63C, E6AC
13E	Other wheel defects (car)	E64C, E65C, E66C, E67C, E68C, E69C
14E	TOFC or COFC defects	E11C, E12C, E13C, E19C
15E	Locomotive trucks, bearings, and wheels	E07L, E40L, E41L, E42L, E43L, E44L, E45L, E46L, E47L, E48L, E4TL, E49L, E51L, E52L, E53L, E54L, E55L, E59L, E60L, E61L, E62L, E63L, E64L, E65L, E66L, E67L, E68L, E6AL, E69L, E70L, E77L, E78L, E7BL
16E	Locomotive electrical and fires	E71L, E72L, E73L, E74L, E76L, E7AL
17E	All other locomotive defects	E00L, E01L, E02L, E03L, E04L, E06L, E08L, E0HL, E09L, E20L, E21L, E22L, E23L, E24L, E25L, E26L, E27L, E29L, E30L, E31L, E32L, E33L, E34L, E35L, E36L, E37L, E39L, E79L, E99L, E10L
18E	All other car defects	E49C, E80C, E81C, E82C, E83C, E84C, E85C, E86C, E89C, E99C, E4AC
19E	Stiff truck (car)	E46C, E4BC
20E	Track–train interaction (hunting) (car)	E4TC
21E	Current collection equipment (locomotive)	E75L
01H	Brake operation (main line)	H510, H511, H512, H513, H514, H515, H516, H517, H518, H519, H520, H521, H525, H526
02H	Handbrake operations	H017, H018, H019, H020, H021, H022, H025, M504
03H	Brake operations (other)	H008, H099
04H	Employee physical condition	H101, H102, H103, H104, H199
05H	Failure to obey or display signals	H201, H202, H203, H204, H205, H206, H207, H208, H209, H215, H216, H217, H299, H218, H219, H220, H221, H222
06H	Radio communications error	H210, H211, H212, H405
07H	Switching rules	H301, H302, H303, H304, H305, H306, H307, H308, H309, H310, H311, H312, H313, H314, H315, H399, H318, H316, H317
08H	Mainline rules	H401, H402, H403, H404, H406, H499
09H	Train handling (excluding brakes)	H501, H502, H503, H504, H505, H506, H507, H508, H509, H522, H523.H524, H599
10H	Train speed	H601, H602, H603, H604, H605, H606, H699, H607
11H	Use of switches	H701, H702, H703, H704, H705, H799, H706, H707
12H	Miscellaneous human factors	H821, H822, H823, H824, H899, H991, H992, H993, H994, H995, H999, H996, H99E, H99A, H99B, H99C, H99D, H997
01M	Obstructions	M101, M402, M403, M404
02M	Grade crossing collisions	M301, M302, M303, M304, M305, M306, M307, M399, M308, M309, M310
03M	Lading problems	M201, M202, M203, M204, M205, M206, M207, M299, M409, M410, M208
04M	Track–train interaction	M405
05M	Other miscellaneous	M401, M406, M407, M408, M501, M502, M503, M505, M599, M506, M507, M411, M509, M510
06M	Extreme weather	M102, M103, M104, M105, M199

Note: Detailed descriptions of each FRA cause code can be found in the *FRA Guide for Preparing Accident/Incident Reports* (FRA 2011a). UDE = Undesired emergency brake application; TOFC = Trailer on flatcar; and COFC = Container on flatcar.

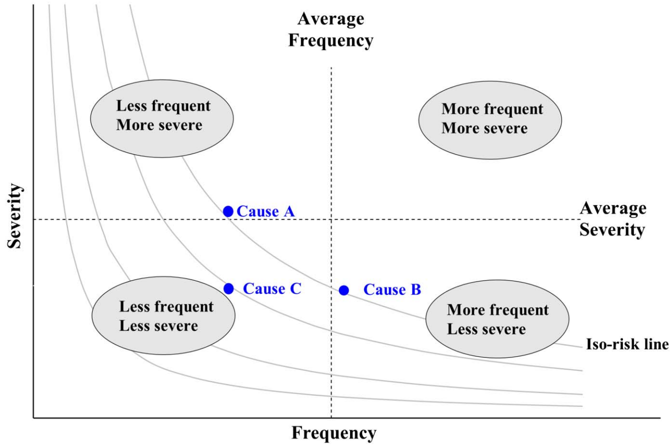

Fig. 3. Example of a train accident frequency–severity graph including example iso-risk contours.

As discussed, comparison of cause-specific derailment trends is an objective of this paper, so a frequency–severity plot was adapted for this comparison by using different symbols to plot the two time periods on the same graph (Fig. 5). Considering broken rails or welds again, there was a small increase in severity but a substantial reduction in frequency when comparing 2006–2010 to 2011–2015. Despite this decrease, broken rails or welds remained the leading cause of large derailments and were on the highest iso-car contour of any cause group. Track geometry (excluding wide gauge) also declined substantially. Other rail and joint defects were supplanted by joint bar defects as the cause with the highest severity. Derailments caused by buckled track decreased in terms of frequency, but their severity increased slightly.

Iso-car graphs such as Fig. 5 offer a means to simultaneously compare the relative changes among different causes' relative frequency and severity. Quantitative comparison of changes can be further enhanced by graphing the change in the number of derailments and number of cars derailed per trillion ton-kilometers between the two time periods [Figs. 6(a and b)]. Three candidates for traffic exposure were considered: car-miles, train-miles, and ton-miles (ton-kilometers). The latter were chosen because previous researchers found that train-miles had certain limitations compared to car-miles or ton-miles (Nayak and Palmer 1980) and only ton-mile data were available for our study.

Broken rails or welds showed the greatest reduction in both number of derailments and number of cars derailed, and track geometry (excluding wide gauge) showed the second largest reduction in both categories [Figs. 6(a and b)]. This reduction in number of derailments is consistent with recent studies (Liu 2015). Other than the two top-ranked causes, the rank order of the causes showing the greatest decline differed between the two measures, number of derailments versus number of cars derailed. This reflects differing average severities associated with different causes. For example, bearing failure was the fourth-ranked cause in terms of decline in derailment rate, but ranked seventh in terms of numbers of cars derailed. This is consistent with previous research that found that bearing failure was among the most frequent derailment causes but had considerably lower average derailment severity (Barkan et al. 2003).

Although most derailment causes declined in their rate of occurrence between the two time periods, extreme weather showed an increase using number of accidents and obstructions showed an increase in number of cars derailed. Finally, wide gauge had a lower rank in terms of decline in the rate of derailments, but the rate of cars derailing ranked higher. This suggests that although there were fewer derailments due to these causes, some of them were high-severity events.

A final question considered was whether most accident causes declined in proportion to their relative frequency, or whether some declined disproportionately, i.e., more or less than average. This was addressed by comparing the magnitude of change of each cause group to its frequency in the first time period [Fig. 7(a)]. The linear regression line represents the average change for all cause groups combined. Those above the regression line had relatively

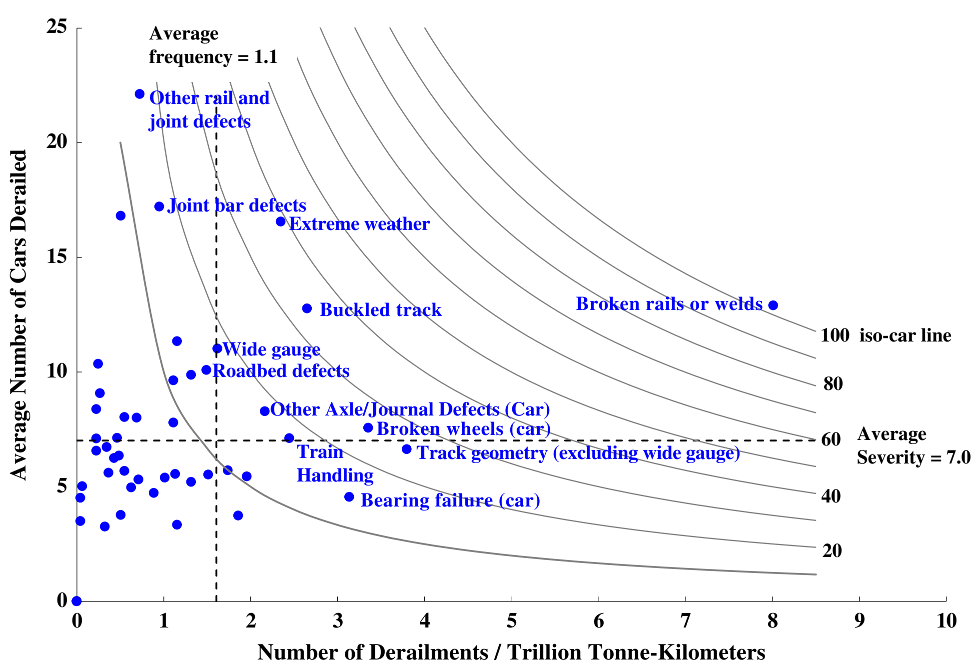

Fig. 4. Frequency–severity graph from 2006 to 2015 (causes with iso-car greater than 15 are labeled).

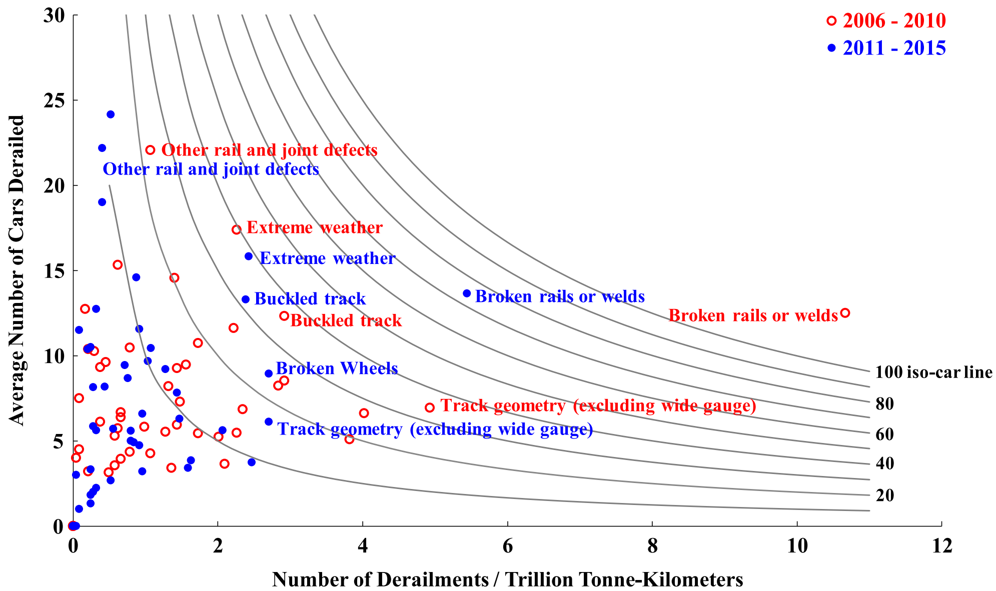

Fig. 5. Frequency–severity graph of two time periods (causes in the first time period with iso-car greater than 30 are labeled).

less change between the two periods, whereas causes below the regression line indicate disproportionately greater reduction in derailment frequency. The same approach was used to evaluate changes between the two time periods in number of cars derailed due to different cause groups [Fig. 7(b)].

Most cause groups were relatively near the average, indicating that they declined roughly in proportion to their relative frequency. However, there were a few, such as extreme weather and broken rails or welds, that were farther above or below the regression line. A way to check if cause groups significantly deviated from the average decline is to determine if there is evidence of significant variance in the residuals for the cause groups. We investigated this using the Breusch-Pagan test (Krämer and Sonnberger 1986) to determine if the data for derailment frequency and number of cars derailed were homoscedastic or heteroscedastic [Figs. 8(a and b)]. The null hypothesis for each data set was homoscedasticity. For derailment frequency and number of cars derailed, the test p-values were 0.0006 and 0.003, respectively, indicating significant variability and that the disproportionately greater or lesser changes in some of the causes were significant.

Given the significant heteroscedasticity, we wanted to know which causes had contributed most, so the standardized residuals were computed using the following equation:

$$s_i = \frac{e_i}{\sqrt{\widehat{\mathrm{VAR}}(e_i)}} \tag{1}$$

where s_i = standardized residual of point i; e_i = raw residual of point i; and $\widehat{\mathrm{VAR}}$ = variance of point i.

Observations with a standardized residual $>|2|$ are considered outliers (Peck and Devore 2008), indicating they contributed the most to the heteroscedasticity. The cause groups whose derailment frequency and number of cars derailed had absolute standardized residuals greater than 1 are given in Tables 2 and 3.

In terms of derailment frequency, three cause groups—extreme weather, broken rails or welds, and buckled track—had absolute standardized residuals greater than 2, indicating that their change significantly differed from the average for all cause groups combined. Broken rails or welds declined significantly more than average, whereas buckled track declined significantly less than average and extreme weather increased between the two time periods.

The cause groups that deviated from average the most in terms of number of cars derailed per accident were not all the same as those for derailment frequency. Four cause groups significantly differed from the average and contributed most to the heteroscedasticity in terms of number of cars derailed: extreme weather, buckled track, and broken wheels. All three causes declined less than average.

Three-Factor Derailment Rate Model

The previous sections examined the rate and severity of accidents for various track types. Such macrolevel analyses provide insight on overall accident trends and are useful for focusing attention on reducing the incidence of the most important causes. However, other uses of derailment rate data include estimation of the risk associated with different portions of a network or a particular route. FRA data alone do not permit estimation of the location-specific rates needed for this because simply knowing the number of derailments at a particular location or on a route does not account for possible differences in traffic levels. A location might have more derailments but a lower rate of occurrence due to a high volume of traffic.

Liu et al. (2017) extended previous analyses quantifying the relationship between FRA track class and derailment rate (Nayak et al. 1983; T. T. Treichel and C. P. L. Barkan, "Mainline freight train accident rates," working paper, Research and Test Department: Association of American Railroads, Washington, DC; Anderson and Barkan 2005) by introducing two new variables, method of operation (MOO) and annual traffic density, in addition to FRA track class. Using rail industry traffic data combined with FRA derailment data for the period 2005–2009, they found that all three variables were significantly correlated with derailment rate. The inverse relationship between FRA track class and derailment rate was still evident, but they also found that within each track class, signaled track had a significantly lower derailment rate than nonsignaled track, and trackage with above-average traffic density

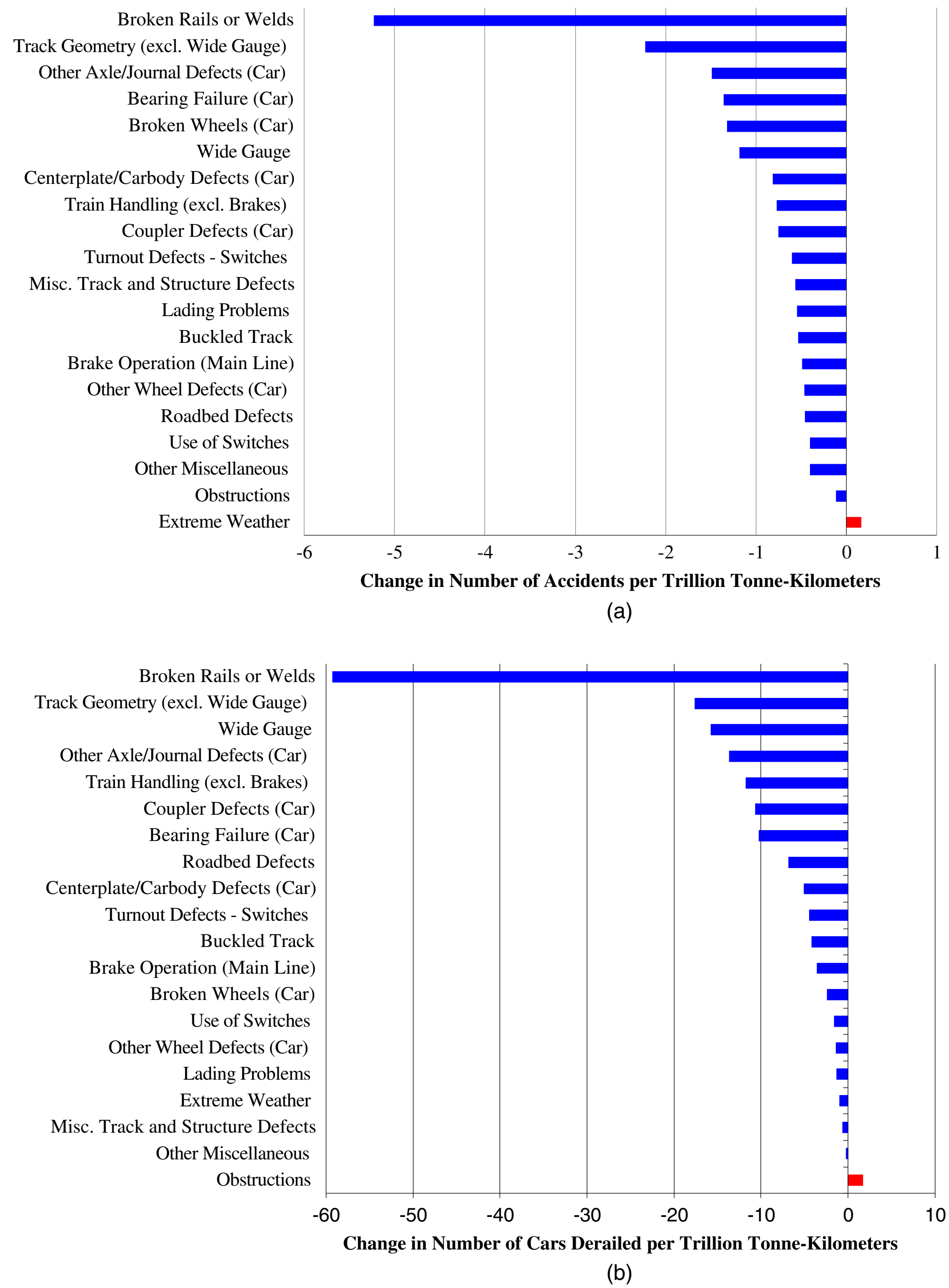

Fig. 6. (a) Change in number of derailments by cause group from 2006–2010 to 2011–2015; and (b) change in number of cars derailed by cause group from 2006–2010 to 2011–2015.

had a significantly lower derailment rate than lines with below-average traffic (Liu et al. 2017). The effect of the two new factors probably contributed to previous results that had shown the relationship with FRA track class because they covary in the manner expected, i.e., higher FRA track class covaries with traffic density and the use of signals for traffic control. The significance of Liu et al.'s results is that their data and analytical method enabled them to identify and quantify the separate effect of these new variables. Given the substantial reduction in derailment rates discussed, another objective of this paper was to investigate how they might have affected the Liu et al. three-factor model results.

Track Class, Method of Operation, and Traffic Density

In Liu et al.'s (2017) analysis, FRA track class was a five-level categorical variable ranging from 1 to 5. These five track classes are the principal ones used by the freight railroads with maximum allowable speeds ranging from 16.1 km/h (10 mi/h) for Track Class 1, up to 128.8 km/h (80 mi/h) for Track Class 5 (FRA 2014). FRA specifies minimum requirements for track structure, geometry, and maintenance for each track class. The higher the allowable speed, the more stringent the corresponding standards. However, these standards are minimums; Class 1 railroad internal maintenance

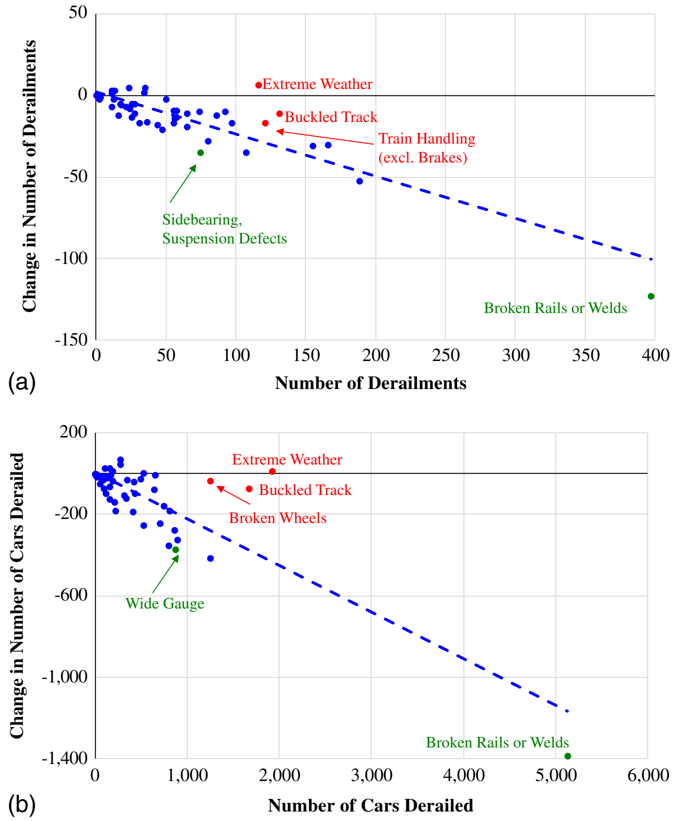

Fig. 7. (a) Change in number of derailments from 2011 to 2015 compared to number of derailments from 2006 to 2015; and (b) change in number of cars derailed from 2011 to 2015 compared to number of cars derailed from 2006 to 2015.

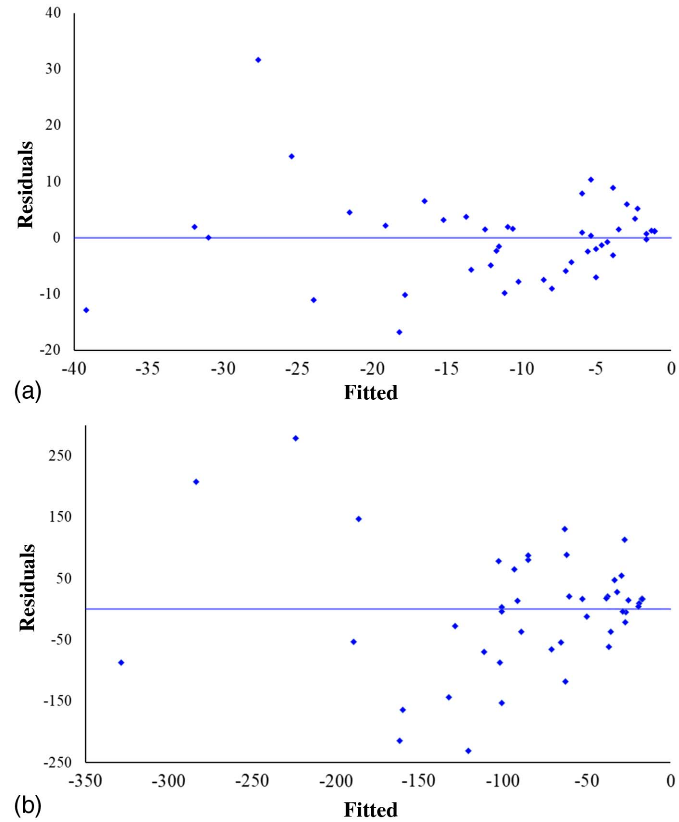

Fig. 8. (a) Fitted versus residual plot for derailment frequency, 2006–2010 to 2011–2015; and (b) fitted versus residual plot for number of cars derailed, 2006–2010 to 2011–2015.

standards commonly exceed the FRA standards for a given track class. They also may include other criteria and maintenance-related activities beyond the FRA requirements.

Method of operation was treated as a binary categorical variable, signaled or nonsignaled, indicating whether a section of track has electric track circuits and wayside signals or not [49 C.F.R. 236 (2011)]. Over the time period covered in this study, FRA changed the way it recorded MOO. Prior to May 31, 2011, the REAIR had 12 categories for method of operation. Within each category, one could determine if the MOO included signals or not. After that date, FRA collapsed the 12 categories into just two, either signaled or nonsignaled (FRA 2011a). Although the loss of granularity in FRA's data recording system for this variable is regrettable, it did not affect this research.

Annual traffic density was also treated as a binary categorical variable with two levels, low and high, with low indicating less than 20 million gross tons (MGT) and high indicating greater than or equal to 20 MGT (18.1 million gross metric tons). Gross tonnage is a measure of the total weight of the locomotives, rolling stock, and lading traveling over a given section of track. The 20 MGT threshold represents the mean value for Class 1 freight railroad mainline trackage (AAR 2017a, b). Due to higher traffic volumes, railroads may invest greater resources in track maintenance, even if the allowable speed and consequent track class is low.

Three-Factor Analysis and Data

The three variables discussed in the preceding section and the respective values for each were used to create a $5 \times 2 \times 2$ matrix with a total of 20 unique cells identical in form to Liu et al.'s (2017).

Table 2. Standardized residuals for derailment frequency data

Cause	Standardized residual	Absolute standardized residual
Extreme weather	3.64	3.64
Broken rails or welds	−2.41	2.41
Buckled track	2.23	2.23
Side bearing and suspension defects (car)	−1.89	1.89
Train handling (excluding brakes)	1.31	1.31
Other brake defect (car)	1.29	1.29
Other wheel defects (car)	1.26	1.26
All other car defects	−1.17	1.17
Broken wheels (car)	1.17	1.17
Joint bar defects	−1.16	1.16

Table 3. Standardized residuals for number of cars derailed data

Cause	Standardized residual	Absolute standardized residual
Extreme weather	3.83	3.83
Buckled track	2.61	2.61
Broken wheels (car)	2.10	2.10
Broken rails or welds	−1.83	1.83
Wide gauge	−1.56	1.56
Other rail and joint defects	−1.52	1.52
Coupler defects (car)	−1.20	1.20
Miscellaneous human factors	−1.19	1.19
Miscellaneous track and structure defects	1.19	1.19
Track geometry (excluding wide gauge)	−1.16	1.16
Other axle and journal defects (car)	−1.09	1.09

Table 4. Three-factor matrix with the number of mainline derailments in each cell during the time period 2006–2010

Number of derailments, 2006–2010							
		FRA track class					
Traffic density (MGT)	MOO	1	2	3	4	5	Total
<20	Nonsignaled	49	91	73	55	0[a]	268
	Signaled	17	31	49	52	11	160
≥20	Nonsignaled	8	22	30	77	0[a]	137
	Signaled	31	94	130	387	146	788
Total		105	238	282	571	157	1,353

[a]FRA regulations [49 C.F.R. 236 (2011)] do not permit freight train operation greater than 79 kph (49 mi/h) on nonsignaled track, so in general such track would not be higher than FRA Track Class 4, which has a maximum speed of 97 kph (60 mi/h) for freight trains; consequently, derailments on such trackage are expected to be rare.

Table 5. Three-factor matrix with the number of mainline derailments in each cell during the time period 2011–2015

Number of derailments, 2011–2015							
		FRA track class					
Traffic density (MGT)	MOO	1	2	3	4	5	Total
<20	Nonsignaled	28	48	43	43	0[a]	162
	Signaled	17	31	44	62	10	164
≥20	Nonsignaled	7	10	8	27	0[a]	52
	Signaled	25	61	97	312	102	597
Total		77	150	192	444	112	975

[a]FRA regulations [49 C.F.R. 236 (2011)] do not permit freight train operation greater than 79 kph (49 mi/h) on nonsignaled track, so in general such track would not be higher than FRA Track Class 4, which has a maximum speed of 97 kph (60 mi/h) for freight trains; consequently, derailments on such trackage are expected to be rare.

Table 6. Three-factor matrix with the number of mainline derailments in each cell for the difference during the two time periods: 2006–2010 and 2011–2015

Difference in number of derailments between the two time periods, 2006–2010 and 2011–2015							
		FRA track class					
Traffic density (MGT)	MOO	1	2	3	4	5	Total
<20	Nonsignaled	−21	−43	−30	−12	0[a]	−106
	Signaled	0	0	−5	10	−1	4
≥20	Nonsignaled	−1	−12	−22	−50	0[a]	−85
	Signaled	−6	−33	−33	−75	−44	−191
Total		−28	−88	−90	−127	−45	−378

[a]FRA regulations [49 C.F.R. 236 (2011)] do not permit freight train operation greater than 79 kph (49 mi/h) on nonsignaled track, so in general such track would not be higher than FRA Track Class 4, which has a maximum speed of 97 kph (60 mi/h) for freight trains; consequently, derailments on such trackage are expected to be rare.

Table 7. Distribution of mainline traffic data 2005–2009

		FRA track class					
Traffic density (MGT)	MOO	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)	Total (%)
<20	Nonsignaled	0.1	0.5	0.9	1.6	0[a]	3.2
	Signaled	0.1	0.3	1.2	3.3	0.3	5.2
≥20	Nonsignaled	0.2	0.4	0.8	2.1	0.2[a]	3.7
	Signaled	0.5	2.0	8.2	47.8	29.4	88.0
Total		0.8	3.3	11.1	54.7	30.0	100.0

Source: Data from Liu et al. (2017).
[a]FRA regulations [49 C.F.R. 236 (2011)] do not permit freight train operation greater than 79 kph (49 mi/h) on nonsignaled track, so in general such track would not be higher than FRA Track Class 4, which has a maximum speed of 97 kph (60 mi/h) for freight trains; consequently, derailments on such trackage are expected to be rare.

We used this matrix as a framework to develop data for the number of derailments and the volume of traffic. Mainline freight train derailment count data were developed in a manner similar to Liu et al. The REAIR database contains all the information needed to categorize each derailment into the appropriate cell in the matrix. The derailment data are the numerator in the derailment rate estimation (Tables 4 and 5). The number of derailments in most cells in the matrix declined from the first time period to the second (Table 6). All of the exceptions were on signaled trackage with less than 20 MGT annual traffic as follows: FRA Classes 1 and 2 tracks showed no change, and the number of derailments increased on Class 4 track. However, these cells with no change or an increase in derailments accounted for only 3.6% of the total traffic compared to the cells accounting for the remaining 96.4% of the traffic, all of which experienced a reduction in the number of derailments.

Overall, the total traffic for the two time periods was similar (24.4 trillion gross ton-kilometers for 2006–2010 and 25.1 trillion ton-kilometers for 2011–2015), so the general decline in mainline derailments across most cells in the three-factor matrix is consistent with the results described in the previous sections of this paper.

Development of the requisite rate estimates required denominator data. Specifically, how was the Class 1 railroad mainline freight traffic volume distributed over the cells in the matrix for a comparable period of time as the derailment data. In Liu et al.'s (2017) study, the railroads were able to provide traffic data for the period 2005–2009 that were categorized in a manner that allowed Liu et al. to classify them using the matrix parameters. Compilation of data in this manner is not a routine process for railroads so comparable data were not available for the more recent time period considered in this study. Consequently, we used an alternative approach developed by Anderson and Barkan (2004) to estimate the traffic distribution (Table 7). Liu et al.'s data were combined with overall traffic volume data from the Association of American Railroads (AAR 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015). The 2005–2009 distribution of traffic was extrapolated using the overall traffic data to develop estimated traffic distributions for the missing years (Table 8).

Comparison of Derailment Distributions

As discussed, the total number of derailments in the second time period was lower than the first. The question was whether the decline in derailments was uniformly distributed throughout the matrix or had some combinations of conditions experienced larger or smaller reductions than expected.

To address this question, the two distributions were statistically compared. We used log-linear models to test the independence between the explanatory variables (Agresti 2002). This method allows analysis of discrete categorical variables for count or rate data. A traditional log-linear model enables comparison of the exact count

Table 8. Estimated mainline traffic distribution 2011–2015 in billions of ton-kilometers

Traffic density (MGT)	MOO	FRA track class 1	2	3	4	5	Total
<20	Nonsignaled	29 (20)	127 (87)	215 (147)	377 (258)	0	748 (512)
	Signaled	12 (8)	81 (56)	279 (191)	761 (522)	70 (48)	1204 (824)
≥20	Nonsignaled	43 (29)	86 (59)	192 (131)	481 (329)	53 (36)	854 (585)
	Signaled	112 (77)	475 (326)	1,911 (1,309)	11,141 (7,831)	6,862 (4,700)	20,502 (14,043)
Total		196 (134)	770 (527)	2,597 (1,779)	12,760 (8,740)	6,985 (4,784)	23,308 (15,964)

Note: Values in parentheses are the same data presented in ton-miles.

Table 9. Estimated mainline derailment rate per billion ton-kilometers for the time period 2006–2010

Traffic density (MGT)	MOO	Estimated derailment rates, 2006–2010: FRA track class 1	2	3	4	5	Total
<20	Nonsignaled	1.549 (2.264)	0.662 (0.968)	0.313 (0.458)	0.135 (0.197)	0[a]	0.331 (0.484)
	Signaled	1.318 (1.927)	0.352 (0.515)	0.162 (0.237)	0.063 (0.092)	0.145(0.212)	0.123 (0.179)
≥20	Nonsignaled	0.172 (0.252)	0.236 (0.345)	0.145 (0.211)	0.148 (0.216)	0[a]	0.148 (0.216)
	Signaled	0.256 (0.374)	0.183 (0.267)	0.063 (0.092)	0.032 (0.047)	0.020(0.029)	0.036 (0.052)
Total		0.495 (0.724)	0.286 (0.417)	0.100 (0.147)	0.041 (0.060)	0.021(0.030)	0.054 (0.078)

Note: Values in parentheses are rates normalized by ton-miles.
[a]FRA regulations [49 C.F.R. 236 (2011)] do not permit freight train operation greater than 79 kph (49 mi/h) on nonsignaled track, so in general such track would not be higher than FRA Track Class 4, which has a maximum speed of 97 kph (60 mi/h) for freight trains; consequently, derailments on such trackage are expected to be rare.

number in each cell when total counts are equal (Fienberg 1978). This study compared two time periods—$\mu_{ijk}^{2006\text{–}2010}$ and $\mu_{ijk}^{2011\text{–}2015}$—so in order to compare corresponding proportions, a minor data transformation of the count data for 2006–2010 was conducted to match the total count for 2011–2015, as shown in the following, where μ represents the specific frequency for traffic density i, method of operation j, and FRA track class k:

$$\mu_{ijk}^{\text{time 1 new}} = \mu_{ijk}^{\text{time 1}} \times \frac{\text{total count}^{\text{time 2}}}{\text{total count}^{\text{time 1}}} \tag{2}$$

To understand the effect of time, it was set as a fourth categorical variable in the contingency table, resulting in a four-way contingency table.

Log-linear model fitting generally employs two approaches: forward or backward (Agresti 2002). The forward approach starts with a simple model and adds variables until the best-fit model is achieved. The backward approach starts with a fully saturated model including all interaction terms between variables and removes insignificant interactions based on goodness-of-fit tests. For this study, the objective was to examine the relationship between the two time variables and no additional variables were required. Therefore, a forward approach using a complete independence model should capture the effect of time. The log-linear model used was as follows:

$$\log(\mu_{ijkl}) = \beta_{\text{traffic}}\text{traffic}_i + \beta_{\text{MOO}}\text{MOO}_j + \beta_{\text{track}}\text{track}_k + \beta_{\text{time}}\text{time}_l + \beta_0 \tag{3}$$

where $\log(\mu_{ijkl})$ = log of the expected cell frequency; β = overall effect, or the grand mean of the logarithms of the expected counts; traffic_i = traffic density; MOO_j = method of operation; track_k = FRA track class; time_l = time period; and β_0 = intercept.

The model consists of traffic density, method of operation, FRA track class, time period, and the intercept. The intercept is the overall mean of the log of the expected frequencies. The null hypothesis is that there is no difference in the cell distributions between the two time periods. SAS version 9.4 was used to perform the log-linear analysis. A p-value of 0.9825 for the time variable indicates that the null hypothesis cannot be rejected, i.e., that the distributions for the two time periods do not significantly differ. This implies that the 10-year period can be used for future work because the distributions are consistent and no individual cell contributed more than others to the overall reduction.

Traffic Exposure and Derailment Rate Estimation

Traffic exposure data provided by the railroads were used as denominator values for derailment rate calculation between 2005 and 2009 (Liu 2013). As mentioned in the "Data and Methodology" section, traffic exposure data for 2010–2015 were unavailable so estimates were extrapolated for those years with the assumption that the traffic exposure distribution did not change substantially between the two time periods.

Comparison of the estimated rates between the earlier and the later time periods showed a reduction in derailment rates for most cells in the matrix (Tables 9–11). Track Classes 1–5 all showed a reduction in derailment rate, of 21%, 32%, 26%, 18%, and 23%, respectively. Nonsignaled and signaled trackage showed 44% and 14% reductions, respectively. Finally, low traffic density and high traffic density trackage had 18% and 25% respective reductions in derailment rate. Although there was an overall reduction in the estimated derailment rates across the matrix, there were a few exceptions. The rates for three cells for signaled trackage and annual traffic less than 20 MGT increased. We do not know what to attribute this to; however, these cells accounted for just 3.6% of the total traffic so they did not have a major impact on the general trend.

J. Transp. Eng., Part A: Systems, 2020, 146(11): 04020127

Table 10. Estimated mainline derailment rate per billion ton-kilometers for the time period 2011–2015

Traffic density (MGT)	MOO	Estimated derailment rates, 2011–2015 FRA track class 1	2	3	4	5	Total
<20	Nonsignaled	0.958 (1.399)	0.378 (0.552)	0.200 (0.292)	0.114 (0.167)	0[a]	0.217 (0.316)
	Signaled	1.427 (2.083)	0.381 (0.557)	0.158 (0.230)	0.081 (0.119)	0.143(0.208)	0.136 (0.199)
≥20	Nonsignaled	0.163 (0.238)	0.116 (0.169)	0.042 (0.061)	0.056 (0.082)	0[a]	0.061 (0.089)
	Signaled	0.223 (0.326)	0.128 (0.187)	0.051 (0.074)	0.028 (0.040)	0.015(0.022)	0.029 (0.042)
Total		0.393 (0.575)	0.194 (0.284)	0.074 (0.108)	0.034 (0.050)	0.016(0.023)	0.041 (0.060)

Note: Values in parentheses are rates normalized by ton-miles.
[a]FRA regulations [49 C.F.R. 236 (2011)] do not permit freight train operation greater than 79 kph (49 mi/h) on nonsignaled track, so in general such track would not be higher than FRA Track Class 4, which has a maximum speed of 97 kph (60 mi/h) for freight trains; consequently, derailments on such trackage are expected to be rare.

Table 11. Estimated mainline derailment rate per billion ton-kilometers for the difference between two time periods, 2006–2010 and 2011–2015

Traffic density (MGT)	MOO	Difference in derailment rate between the two time periods, 2006–2010 and 2011–2015 FRA track class 1	2	3	4	5	Total
<20	Nonsignaled	−0.591 (−0.864)	−0.285 (−0.416)	−0.113 (−0.166)	−0.021 (−0.031)	0[a]	−0.115 (−0.167)
	Signaled	0.136 (0.198)	0.027 (0.039)	−0.005 (−0.007)	0.018 (0.027)	−0.002 (−0.004)	0.013 (0.019)
≥20	Nonsignaled	−0.007 (−0.010)	−0.120 (−0.175)	−0.103 (−0.150)	−0.092 (−0.134)	0[a]	−0.087 (−0.127)
	Signaled	−0.034 (−0.049)	−0.055 (−0.080)	−0.012 (−0.018)	−0.005 (−0.007)	−0.005 (−0.007)	−0.007 (−0.010)
Total		−0.102 (−0.149)	−0.091 (−0.133)	−0.026 (−0.039)	−0.007 (−0.011)	−0.005 (−0.007)	−0.012 (−0.018)

Note: Values in parentheses are rates normalized by ton-miles.
[a]FRA regulations [49 C.F.R. 236 (2011)] do not permit freight train operation greater than 79 kph (49 mi/h) on nonsignaled track, so in general such track would not be higher than FRA Track Class 4, which has a maximum speed of 97 kph (60 mi/h) for freight trains; consequently, derailments on such trackage are expected to be rare.

Table 12. Estimated mainline car derailment rate per billion car-kilometers for the time period 2011–2015

Traffic density (MGT)	MOO	FRA track class 1	2	3	4	5	Total
<20	Nonsignaled	504 (811)	280 (450)	172 (276)	117 (188)	0	175 (282)
	Signaled	492 (791)	288 (464)	91 (146)	54 (87)	79 (127)	84 (136)
≥20	Nonsignaled	58 (94)	73 (117)	58 (94)	61 (99)	0	58 (93)
	Signaled	111 (178)	81 (130)	33 (53)	25 (40)	14 (23)	24 (38)
Total		182 (292)	135 (217)	53 (85)	31 (49)	15 (24)	33 (53)

Note: Values in parentheses are rates normalized by car-miles.

In addition to evaluating how rates changed between the two time periods, we also wanted to know how the relative relationships among the three factors compared to Liu et al.'s (2017) results for the earlier time period. In terms of the marginal totals, the qualitative relations were the same as Liu et al.'s: the higher the FRA track class the lower the derailment rate, signaled trackage had a lower rate than nonsignaled, and higher traffic density trackage had a lower derailment rate than lower density. However, there were some pairwise cell differences compared to Liu et al.'s results (Table 11). All of these related to signaled versus nonsignaled trackage on lower FRA track classes where there were several cases where signaled trackage had higher estimated derailment rates than nonsignaled. These cells accounted for a minority of the traffic with 11% of the total.

We also calculated another metric, the number of railcars derailed per unit of traffic exposure (Table 12). This was calculated by multiplying car derailment rate per ton-mile with the average number of gross ton-miles per car-mile conversion factor (AAR 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015). This is an improvement over the previous method because the exposure factor for each cell is derived using the known traffic data in ton-miles, whereas the previous method used an average number of cars derailed per derailment to convert the derailment rates to car derailment rates.

Conclusions

This paper provides general insights regarding the reduction in train accidents over the time period studied, as well as specific insight regarding which causes are most likely to benefit from further derailment reduction investment. It also provides new techniques to

J. Transp. Eng., Part A: Systems, 2020, 146(11): 04020127

identify and quantify the types of train accidents, the trackage where they occur, and the causes having the greatest effect on train safety and risk. These results are critical to rail-sector development of the most effective strategies to further improve railroad train safety. The main conclusions are summarized as follows.

Data Analysis Indicated a Declining Trend in Accidents

FRA-reportable accident frequency and rates have declined for the four principal FRA-defined accident types: derailments, collisions, grade crossings, and others, with the first three showing the greatest decline. In terms of track type, mainline derailments showed the largest decline, but remained the most frequent type of reportable train accident and the most severe in terms of cars derailed.

New Data Visualization Techniques Identified Main Factors Impacting Mainline Derailments

The concept of iso-risk contours was introduced and adapted to derailment analysis by use of iso-car contours providing a new approach to systematically and quantitatively compare the risk associated with widely differing accident causes. Railroad safety and risk managers may have different options available to reduce risk. The comparative insight offered here regarding how different strategies will affect the probability and/or consequences of derailments enables better informed decisions about investment in different risk reduction strategies. Use of iso-car contours also enabled comparison of the two time periods to better understand changes in risk due to previous actions taken that affected the frequency and severity of different causes.

Applied Regression Methods Developed to Investigate Derailment Causes

An applied regression approach found that the reduction in derailment rate was due to most derailment causes declining, although a few causes increased, most notably extreme-weather-related causes. Railroads have invested in improvements in infrastructure and rolling stock as well as new technologies that allow many defects to be detected and corrected before they fail (Lagnebäck 2007; Schlake et al. 2010). Consequently, the decline in derailments due to these causes is not surprising, whereas extreme weather causes are largely outside railroads' control. A regression analysis showed that most causes declined in proportion to their overall frequency in the first half of the study period, but there were some exceptions. Broken rails or welds and railcar side bearing and suspension defects showed a disproportionately greater reduction compared to their frequency, while buckled track and train handling showed disproportionately less reduction relative to their frequency. Finally, despite the substantial reduction in broken rails or welds between the two time periods, they remained the most frequent and severe cause of mainline derailments; consequently, understanding how to further reduce these derailments should remain a high priority for researchers, industry, and the government.

Statistical modeling was used to estimate the likelihood of broken-rail derailments given various track and operating conditions (Shyr and Ben-Akiva 1996; Dick et al. 2003). Their occurrence can be more effectively reduced by determining the optimal inspection frequency for rail defects (Liu et al. 2014). Adopting a risk-based approach to rail defect detection has the potential to improve both the efficiency and effectiveness of rail-flaw detection with the consequent potential to further reduce broken-rail-caused derailments (Liu and Dick 2016). Other than extreme weather and broken wheels, all of the cause groups above the iso-car 20 contour in the more recent time period (2011–2015) were track related. Track upgrades may reduce accident rates but result in higher capital and operating costs (Liu et al. 2011); however, better scheduling of such activities can improve their efficacy and cost effectiveness (Lovett et al. 2015). Research on the various stresses incurred by the track structure and its various components suggests the means to improve its strength, durability, and reliability, thereby reducing derailments due to track-related failures (Van Dyk et al. 2016; Zhu et al. 2017; Canga Ruiz et al. 2019; Dersch et al. 2019). Finally, although derailments due to broken wheels have declined, they remain a prominent cause and are the subject of extensive research on wheel–rail interface and dynamics, contact stresses, fatigue and fracture, wheel profile maintenance, materials, and wear (Katoa et al. 2019; Li et al. 2019; Klomp et al. 2020; Shi et al. 2020).

Improvement in Three-Factor Derailment Matrix and New Derailment Rate Estimates

In addition to analyzing how derailment causes changed during the study period, we also wanted to understand how the decline related to three attributes that Liu et al. (2017) found to be significantly correlated with derailment rate. Using an improved methodology and more recent data, we developed an updated three-factor statistical model to estimate derailment rate. We found that the decline in derailments did not significantly differ among the several attributes Liu et al. identified during the earlier study period. This indicates that the reduction in derailments was fairly uniform, irrespective of FRA track class, traffic density, or method of operation, suggesting a proportional reduction on all portions of the Class 1 railroad network. Comparing the 2011–2015 derailment rate estimates to the 2005–2009 study, 17 of the 20 cells in the matrix (accounting for 96.4% of the overall mainline traffic) showed a reduction in derailment rate. The only exceptions were low traffic density, signaled FRA Track Classes 1, 2, and 4, which accounted for just 3.4% of the total traffic.

Overall, we found that the relative relationships for the three factors were the same as what Liu et al. (2017) observed: higher FRA track classes had lower derailment rates than lower ones, signaled trackage derailment rates were lower than nonsignaled, and higher-density trackage derailment rates were lower than low-density trackage. However, at a more detailed level we did observe several differences compared to Liu et al. In several instances for lower FRA track classes, the estimated derailment rate for signaled trackage was higher than for nonsignaled. Combined, these cells accounted for about 11% of all the traffic analyzed in our study. We do not have an explanation except to suggest that perhaps railroads used a risk-based approach and prioritized their derailment prevention efforts on the remaining 89% of the network that includes the most densely used, highest-speed trackage, which had the lowest derailments rates overall.

These new estimates for train and car derailment rate can be used for more accurate studies of railroad train safety including hazardous materials transportation risk analysis. Finally, for higher track classes, signaled track had lower derailment rates than nonsignaled track. Our analysis indicated this was not due to broken rails or welds prevention on signaled track. A more thorough study should be conducted to investigate the effect of derailment causes on the difference in derailment rates.

Future Work

This paper presents a comprehensive statistical analysis of freight train derailment causes and rates using historical data for the 10-year period of 2006–2015. Inferential statistical methods provide valuable information on trends and causes of derailments; nevertheless, relying on these results alone would mean focusing on issues based on lagged information. As data quantity and quality increase and safety standards become more stringent, emphasis should shift from historical analysis to greater forecasting ability. The development of predictive analytics for derailment modeling could provide industry and regulators with advance insight regarding developing trends in derailment rates and causes. The FRA and railroads both collect extensive operating and accident data that could be used with statistical forecasting techniques. Two potential models are autoregressive integrated moving average (ARIMA) and long short-term memory (LSTM). Another approach that has the potential to improve railroad risk management would be to integrate predictive models and derailment causal analysis with the objective of developing probabilistic, location-specific estimates of derailment occurrence. Advanced analytic techniques using comprehensive data on train operations, derailments, rolling stock, track condition, and maintenance activities have the potential to provide such capability and research should be conducted to explore this.

Data Availability Statement

Some or all data, models, or code generated or used during the study are available in a repository online in accordance with funder data retention policies at https://safetydata.fra.dot.gov/OfficeofSafety/publicsite/on_the_fly_download.aspx. Some or all data, models, or code used during the study were provided by a third party. Direct requests for these materials may be made to the provider as indicated in the Acknowledgments. Analysis of Class I railroads, published by the Association of American Railroads, is available for purchase at https://my.aar.org/Pages/allproducts.aspx. Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request (raw, preprocessed data for the analysis; SAS code for the log-linear analysis).

Acknowledgments

Support for this research was provided by the Association of American Railroads, BNSF Railway, and the National University Rail Center, a US DOT Office of the Assistant Secretary for Research and Technology (OST-R) Tier 1 University Transportation Center. This paper is solely the work of the authors and does not necessarily reflect the opinions of the sponsors. The authors thank Samantha Chadwick and Chen-Yu Lin for providing insights and discussion of this study, and the reviewers for their helpful comments that helped improve the manuscript.

References

AAR (Association of American Railroads). 2006. *Analysis of class I railroads*. Washington, DC: AAR.

AAR (Association of American Railroads). 2007. *Analysis of class I railroads*. Washington, DC: AAR.

AAR (Association of American Railroads). 2008. *Analysis of class I railroads*. Washington, DC: AAR.

AAR (Association of American Railroads). 2009. *Analysis of class I railroads*. Washington, DC: AAR.

AAR (Association of American Railroads). 2010. *Analysis of class I railroads*. Washington, DC: AAR.

AAR (Association of American Railroads). 2011. *Analysis of class I railroads*. Washington, DC: AAR.

AAR (Association of American Railroads). 2012. *Analysis of class I railroads*. Washington, DC: AAR.

AAR (Association of American Railroads). 2013. *Analysis of class I railroads*. Washington, DC: AAR.

AAR (Association of American Railroads). 2014. *Analysis of class I railroads*. Washington, DC: AAR.

AAR (Association of American Railroads). 2015. *Analysis of class I railroads*. Washington, DC: AAR.

AAR (Association of American Railroads). 2017a. *Class I railroad statistics*. Washington, DC: AAR.

AAR (Association of American Railroads). 2017b. "Railroad facts." Accessed June 1, 2018. https://my.aar.org/Pages/allproducts.aspx.

AAR (Association of American Railroads). 2018. "Railroad facts." Accessed December 21, 2018. https://my.aar.org/Pages/Product-Details.aspx?ProductCode=RFB2018Web.

ADL (Arthur D. Little, Inc). 1996. *Risk assessment for the transportation of hazardous materials by rail, supplementary report: Railroad accident rate and risk reduction option effectiveness analysis and data*. Cambridge, MA: ADL.

Agresti, A. 2002. *Categorical data analysis*. Hoboken, NJ: Wiley.

Anderson, R. T. 2005. "Quantitative analysis of factors affecting railroad accident probability and severity." M.S. thesis, Dept. of Civil and Environmental Engineering, Univ. of Illinois at Urbana–Champaign.

Anderson, R. T., and C. P. L. Barkan. 2004. "Railroad accident rates for use in transportation risk analysis." *Transp. Res. Rec.* 1863 (1): 88–98. https://doi.org/10.3141/1863-12.

Anderson, R. T., and C. P. L. Barkan. 2005. "Derailment probability analyses and modeling of mainline freight trains." In *Proc., of 8th Int. Heavy Haul Conf.* Virginia Beach, VA: International Heavy Haul Association.

Barkan, C. P. L., C. T. Dick, and R. T. Anderson. 2003. "Analysis of railroad derailment factors affecting hazardous materials transportation risk." *Transp. Res. Rec.* 1825 (1): 64–74. https://doi.org/10.3141/1825-09.

Canga Ruiz, A. E., Y. Qian, J. R. Edwards, and M. S. Dersch. 2019. "Probabilistic framework for the assessment of concrete sleeper flexural design." *J. Rail Rapid Transit.* 234 (7): 691–701. https://doi.org/10.1177/0954409719854569.

Chadwick, S. G. 2017. "Quantitative analysis of train derailment probability at highway-rail grade crossings." Ph.D. dissertation, Dept. of Civil and Environmental Engineering, Univ. of Illinois at Urbana–Champaign.

Chadwick, S. G., M. R. Saat, and C. P. L. Barkan. 2012. "Analysis of factors affecting train derailments at highway–rail grade crossings." *In Proc., 91st Annual Meeting*. Washington, DC: Transportation Research Board.

Dersch, M. S., T. Roadcap, J. R. Edwards, Y. Qian, J. Y. Kim, and M. Trizotto. 2019. "Investigation into the effect of lateral and longitudinal loads on railroad spike stress magnitude and location using finite element analysis." *Eng. Fail. Anal.* 104 (Oct): 388–398. https://doi.org/10.1016/j.engfailanal.2019.06.009.

Dick, C. T., C. P. L. Barkan, E. R. Chapman, and M. P. Stehly. 2003. "Multivariate statistical model for predicting occurrence and location of broken rails." *Transp. Res. Rec.* 1825 (1): 48–55. https://doi.org/10.3141/1825-07.

Evans, A. W. 2013. "The economics of railway safety." *Res. Transp. Econ.* 43 (1): 137–147. https://doi.org/10.1016/j.retrec.2012.12.003.

Fienberg, S. E. 1978. *The analysis of cross-classified categorical data*. Cambridge, MA: MIT Press.

FRA (Federal Railroad Administration). 2011a. "FRA guide for preparing accident and incident reports." Accessed June 1, 2015. https://railroads.dot.gov/guides/fra-guide-preparing-accident-and-incident-reports.

FRA (Federal Railroad Administration). 2011b. "Rail equipment accident/incident form 6180.54 accident downloads on demand data file structure." Accessed June 1, 2015. https://safetydata.fra.dot.gov/OfficeofSafety/publicsite/DownloadFStructure.aspx.

FRA (Federal Railroad Administration). 2014. *Track safety standards; improving rail integrity: Final rule, 2014*. Washington, DC: FRA.

FRA (Federal Railroad Administration). 2015. *Railroad equipment accident/incident reporting threshold 1957–2015*. Washington, DC: FRA.

FRA (Federal Railroad Administration). 2016. "FRA Office of Safety Analysis accident query." Accessed May 1, 2016. https://safetydata.fra.dot.gov/OfficeofSafety/publicsite/query/AccidentByRegionStateCounty.aspx.

FRA (Federal Railroad Administration). 2019. *Proposed rules 49 CFR part 225, revision of method for calculating monetary threshold for reporting rail equipment accidents/incidents*. Washington, DC: FRA.

Karlaftis, M. G., and I. Golias. 2002. "Effects of road geometry and traffic volumes on rural roadway accident rates." *Accid. Anal. Prev.* 34 (3): 357–365. https://doi.org/10.1016/S0001-4575(01)00033-1.

Katoa, T., T. Fujimura, Y. Yamamoto, S. Dedmon, S. Hiramatsu, H. Kato, and J. Pilch. 2019. "Effect of wheel size and tread braking on subsurface crack initiation in heavy haul car wheels." *Procedia Struct. Integrity* 19 (Jan): 238–248. https://doi.org/10.1016/j.prostr.2019.12.026.

Klomp, M., F. Bruzelius, J. Nielsen, and A. Hillemyr. 2020. *Advances in dynamics of vehicles on roads and tracks*. Switzerland: Springer.

Krämer, W., and H. Sonnberger. 1986. *The linear regression model under test*. Heidelberg, Germany: Physica-Verlag.

Lagnebäck, R. 2007. "Evaluation of wayside condition monitoring technologies for condition-based maintenance of railway vehicles." Licentiate degree, Dept. of Civil, Mining and Environmental Engineering, Div. of Operation and Maintenance Engineering, Luleå Univ. of Technology.

Li, Z., M. Berg, D. Fletcher, P. Meehan, and A. Núñez. 2019. "Rolling contact fatigue: Damage function development from two-disc test data." In *Proc., 11th Int. Conf. on Contact Mechanics and Wear of Rail/Wheel Systems*, 430–431. Delft, Netherlands: TU-Delft/Delft Cluster.

Liu, X. 2013. "Optimal strategies to improve railroad train safety and reduce hazardous materials transportation risk." Ph.D. dissertation, Dept. of Civil and Environmental Engineering, Univ. of Illinois at Urbana–Champaign.

Liu, X. 2015. "Statistical temporal analysis of freight−train derailment rates in the United States: 2000 to 2012." *Transp. Res. Rec.* 2476 (1): 119–125. https://doi.org/10.3141/2476-16.

Liu, X., C. P. L. Barkan, and M. R. Saat. 2011. "Analysis of derailments by accident cause: Evaluating railroad track upgrades to reduce transportation risk." *Transp. Res. Rec.* 2261 (1): 178–185. https://doi.org/10.3141/2261-21.

Liu, X., and C. T. Dick. 2016. "Risk-based optimization of rail defect inspection frequency for petroleum crude oil transportation." *Transp. Res. Rec.* 2545 (1): 27–35. https://doi.org/10.3141/2545-04.

Liu, X., A. Lovett, T. Dick, M. R. Saat, and C. P. L. Barkan. 2014. "Optimization of ultrasonic rail-defect inspection for improving railway transportation safety and efficiency." *J. Transp. Eng.* 140 (10): 04014048. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000697.

Liu, X., M. R. Saat, and C. P. L. Barkan. 2012. "Analysis of causes of major train derailment and their effect on accident rates." *Transp. Res. Rec.* 2289 (1): 154–163. https://doi.org/10.3141/2289-20.

Liu, X., M. R. Saat, and C. P. L. Barkan. 2017. "Freight-train derailment rates for railroad safety and risk analysis." *Accid. Anal. Prev.* 98 (Jan): 1–9. https://doi.org/10.1016/j.aap.2016.09.012.

Lord, D., and F. Mannering. 2010. "The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives." *Transp. Res. Part A: Policy Pract.* 44 (5): 291–305. https://doi.org/10.1016/j.tra.2010.02.001.

Lovett, A. H., C. T. Dick, C. Ruppert, and C. P. L. Barkan. 2015. "Cost and delay of railroad timber and concrete crosstie maintenance and replacement." *Transp. Res. Rec.* 2476 (1): 37–44. https://doi.org/10.3141/2476-06.

Milton, J. C., V. N. Shankar, and F. L. Mannering. 2008. "Highway accident severities and the mixed logit model: An exploratory empirical analysis." *Accid. Anal. Prev.* 40 (1): 260–266. https://doi.org/10.1016/j.aap.2007.06.006.

Nayak, P. R., and D. W. Palmer. 1980. *Issues and dimensions of freight car size: A compendium*. Rep. No. FRA/ORD-79/56. Washington, DC: Federal Railroad Administration.

Nayak, P. R., D. B. Rosenfield, and J. H. Hagopian. 1983. *Event probabilities and impact zones for hazardous materials accidents on railroads*. Rep. No. DOT/FRA/ORD-83/20. Washington, DC: Federal Railroad Administration.

NHTSA (National Highway Traffic Safety Administration). 2008. *National motor vehicle crash causation survey: Report to Congress*. Washington, DC: NHTSA.

Peck, R., and J. Devore. 2008. *Statistics: The exploration and analysis of data*. 7th ed. Belmont, CA: Cengage Learning.

Schafer, D. H., and C. P. L. Barkan. 2008. "Relationship between train length and accident causes and rates." *Transp. Res. Rec.* 2043 (1): 73–82. https://doi.org/10.3141/2043-09.

Schlake, B. W., S. Todorovic, J. R. Edwards, J. M. Hart, N. Ahuja, and C. P. L. Barkan. 2010. "Machine vision condition monitoring of heavy-axle load railcar structural underframe components." *J. Rail Rapid Transit.* 224 (5): 499–511. https://doi.org/10.1243/09544097JRRT376.

Shi, H., J. Yang, and J. Si. 2020. "Centralized maintenance time prediction algorithm for freight train wheels based on remaining useful life prediction." *Math. Prob. Eng.* 2020 (Mar): 1–12. https://doi.org/10.1155/2020/9256312.

Shyr, F.-Y., and M. Ben-Akiva. 1996. "Modeling rail fatigue behavior with multiple hazards." *J. Infrastruct. Syst.* 2 (2): 73–82. https://doi.org/10.1061/(ASCE)1076-0342(1996)2:2(73).

Van Dyk, B. J., J. R. Edwards, M. S. Dersch, C. J. Ruppert, and C. P. L. Barkan. 2016. "Evaluation of dynamic and impact wheel load factors and their application in design processes." *J. Rail Rapid Transit.* 231 (1): 33–43. https://doi.org/10.1177/0954409715619454.

Vanlaar, W., and G. Yannis. 2006. "Perception of road accident causes." *Accid. Anal. Prev.* 38 (1): 155–161. https://doi.org/10.1016/j.aap.2005.08.007.

Zhu, K., J. R. Edwards, Y. Qian, and B. Andrawes. 2017. "Finite element analysis of rail-end bolt hole and fillet stress on bolted rail joints." *Transp. Res. Rec.* 2607 (1): 33–42. https://doi.org/10.3141/2607-06.

J. Transp. Eng., Part A: Systems, 2020, 146(11): 04020127