Relationship Between Train Length and Accident Causes and Rates

Darwin H. Schafer II and Christopher P. L. Barkan

Train accident rates are a critical metric of railroad transportation safety and risk performance. Understanding factors affecting accident rates is also important for evaluating the effectiveness of various accident prevention measures. Accident rates have been the subject of a number of analyses, but these analyses have generally not considered the effect of train length on accident rate. Accident causes can be classified into two groups: those dependent on train length and correlated with the number of cars in the train, and those independent of train length, corresponding to the number of train miles operated. These classifications have implications for the quantitative effect of various changes in railroad operating practices on railroad safety performance. Whether an accident cause is a function of car miles or train miles affects how safety measures that might reduce the likelihood of that cause will affect overall train accident rate. Accident causes have been classified as car mile or train mile related based on expert opinion, but these classifications have not been quantitatively tested. FRA accident data were used to develop a metric to classify objectively accident causes, and 11 causes were reclassified from the previous classification. Based on the results of the study, a sensitivity analysis was conducted to evaluate how changes in train length affect individual trains' accident likelihood and systemwide accident rate. The concept of causes of car mile versus train mile accidents leads to the premise that, although longer trains are expected to experience more accidents than shorter trains, operation of longer trains results in a lower system-level accident rate.

Train accident rates are a critical measure of rail transportation safety and risk, and understanding them is necessary to evaluate the effect of accident prevention measures. Accident rates have been calculated by various organizations and railroads and aggregated statistics for all U.S. railroads are published annually by the Federal Railroad Administration (FRA) Office of Safety (*1, 2*). Rates have been used to assess various factors such as track class, geographic location, train speed, and track type (*3, 4*; T. T. Treichel and C. P. L. Barkan, Working Paper on Mainline Freight Train Accident Rates, Association of American Railroads, 1993). However, these analyses have generally not considered the effect of train length on train accident rate. Train length is thought to have an effect on accident rate because more cars in a train increase the likelihood that a car or track component in or under a train may fail. Based on this premise, it has been suggested that accident causes can be classified into two types, those that are a function of the number of train miles operated and those that are a function of car miles operated (*5, 6*). The initial classification into these two categories was developed by Arthur D. Little Inc. (ADL) based on the opinions of railroad industry experts. These classifications have implications for the quantitative effect of various changes in practice on railroad safety performance and have been used in subsequent studies of railroad safety (*7–9*). Therefore, statistical evaluation of the classifications will enhance their utility and may also clarify our understanding of them. Furthermore, this classification has implications for an accurate understanding of the relationship between train length and accident rate and consequent policy implications for railroad operating practices.

D. H. Schafer II, Room B-118, and C. P. L. Barkan, Room 1203, Newmark Civil Engineering Laboratory, MC-250, Railroad Engineering Program, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 205 North Mathews Avenue, Urbana, IL 61801. Corresponding author: D. H. Schafer II, dschafer@uiuc.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2043, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 73–82.
DOI: 10.3141/2043-09

A study was undertaken to investigate and evaluate the ADL accident cause classifications with the goal of understanding how operating practices, such as train length, affect the likelihood of a train accident. The objectives of this analysis were

- To present the methodology for calculating train accident rates based on car mile and train mile accident causes;
- To develop a metric to evaluate quantitatively the classification of accident causes as car mile or train mile related;
- To use the metric to properly classify train accident causes;
- To develop new, up-to-date train accident rates based on train length; and
- To conduct a sensitivity analysis on this model to illustrate how changes in train length may affect train accident rate.

TRAIN LENGTH-BASED ACCIDENT RATES

Train accidents include derailments, collisions, highway–rail grade crossing accidents, and other accident types. The likelihood that a train will be involved in an accident is a function of both car miles and train miles operated (*6, 10;* T. T. Treichel and C. P. I. Barkan, unpublished working paper on mainline freight train accident rates, Association of American Railroads, 1993). The number of car miles operated for a particular train is affected by train length; longer trains accumulate more car miles per train mile. However, not all accident causes are directly related to the length of the train; instead, some are related only to the operation of a train, irrespective of its length. Train accident causes that are a function of car miles and train miles can be defined as follows:

> 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 equipment failures for which accident likelihood is directly proportional to the number of components (e.g., bearing failure) and also 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 welds).

> Train mile–related causes are those for 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 exposure (e.g., grade crossing collisions). (*10*)

Car Mile Versus Train Mile Expectations

The car mile cause and train mile cause definitions lead to the premise that longer trains will experience more accidents than shorter trains. This is because longer trains are more susceptible to car mile-related accidents than shorter trains due to the additional cars in the train. Conversely, a train should experience accidents due to train mile–related causes regardless of train length. The length of a train, referred to here and throughout the paper, is the number of cars in the train and not the linear measure of a train's actual length.

This premise leads to two expectations that should be evident when examining accident data and can be used to evaluate different train accident causes. The first expectation is that the average length of a train involved in an accident should be greater for car mile-related causes compared with train mile–related causes, because longer trains will experience a greater proportion of car mile-related accidents. Conversely, train mile–related accidents are independent of train length and should not be biased toward long or short trains.

The second expectation is that the percentage of accidents for car mile-related accidents should be an asymptotically increasing function of train length, whereas the percentage of train mile–related accidents should be an asymptotically decreasing function of train length. Longer trains should experience a higher percentage of accidents from car mile-related causes because of their higher percentage of car miles per train mile operated. Conversely, shorter trains are expected to experience a greater percentage of accidents from train mile–related causes.

Accident Rate Equation

Under the premise that train accidents can be separated into two distinct groups, car mile-related causes and train mile–related causes, a new accident rate model that takes into account the two types of classifications can be developed. The new accident rate equation must include a factor for train length to account for accidents that are dependent on the number of car miles operated.

To develop the new model, all FRA train accident causes were examined (*11, 12*). A previous study by ADL classified each accident cause as either car mile- or train mile-related (*6*). The purpose of this study was to quantify the risk of hazardous material transportation by examining all accident causes. The ADL study showed that accident types should be classified as either car mile or train mile related to quantify properly the car mile- and train mile–related risk. By determining the number of accidents that have occurred due to each cause, two independent and mutually exclusive accident rates can be calculated, the car mile accident rate and the train mile accident rate. The expected number of accidents that a train will be involved in is the sum of the car mile accident rate multiplied by number of car miles and the train mile accident rate multiplied by the number of train miles. The expected number of train accidents that will occur can be calculated as follows:

$$A_{\mathrm{EXP}} = R_C M_C + R_T M_T$$

where

A_{EXP} = accidents expected,
R_C = car mile accident rate (accidents per car mile),
M_C = number of car miles,
R_T = train mile accident rate (accidents per train mile), and
M_T = number of train miles.

Under this model one would expect that longer trains will experience more train accidents. As a train's length increases, the number of train miles operated remains constant, but the number of car miles increases with each additional car. Therefore, the number of expected accidents for a single train increases due to the additional car miles (Figure 1*a*).

If one extends this model to any given number of cars that must be transported, it suggests the general result that operating longer trains should result in fewer accidents. As train length decreases, more trains are required to move the same number of cars, thereby leading to more train mile–related accidents. Under this simple scenario, accidents will be minimized by running the longest trains feasible given infrastructure and other constraints (Figure 1*b*).

It should be noted that there are limits to the validity of this result for very long train lengths (>150). This is because the expectations presented, as well as the data used in this analysis, apply to trains less than this length. In practice, it is possible that accident rates for certain train mile–related accidents may increase as train length becomes very long due to causes such as train handling, train braking, and other factors. The intention of this analysis is not to suggest that longer trains will necessarily improve safety; instead the purpose is

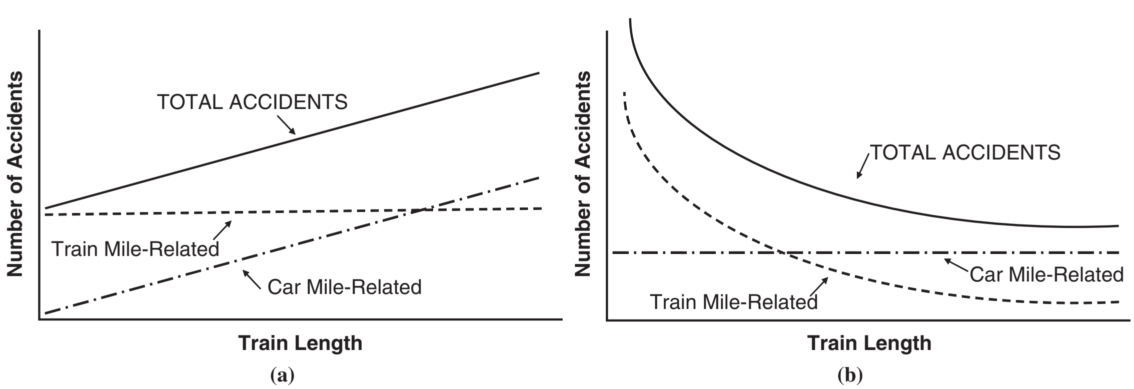

FIGURE 1 Expected accidents from car mile– and train mile–related causes as a function of train length: (*a*) for a single train and (*b*) for a fixed amount of traffic.

to develop a better quantitative understanding of how changes that affect various accident causes, such as number of trains and train length, may affect overall accident rates.

CLASSIFICATION OF ACCIDENT CAUSES

To determine accurately the car mile and train mile accident rates, proper classification of each FRA accident cause is needed. The FRA accident cause classification system is very detailed and often includes several variations of one related group of causes. This is a useful attribute of the database, but is more detailed than necessary for the purpose of this analysis. Consequently, ADL combined similar accident causes into 51 unique groups, 34 of which they classified as car mile related and 17 as train mile related (Table 1) (*6*). The FRA accident causes are separated into five main groups: mechanical, human, signal, track, and miscellaneous causes. ADL defined most track and mechanical failures as car mile related, while most human and signal errors were defined as train mile related. The various miscellaneous causes were assigned to either car mile or train mile related.

The authors used FRA accident data, "Rail Equipment Accidents" from the FRA Office of Safety, to evaluate the ADL classification of accident causes for the period 1990 to 2005 (*11*). These data included all accidents occurring on either mainline or siding tracks for all classes of railroads. Accidents on yard and industry tracks were excluded because the average train length for these types of accidents is comparatively shorter because of yard operations. Mainline and siding accidents were combined because of similar accident causes and train length. Car and train mile relationship predictions for each cause group were compared with the corresponding data from the FRA database. Train lengths were grouped into 10-car bins and the percentage of all car mile-related and train mile–related accident causes was graphed versus train length (Figure 2).

A regression analysis was conducted in which a power function, of the form $y = ax^b$, was fitted to the data to evaluate how well they conformed to an asymptotically increasing or decreasing functional form. The critical term regarding the curve form of the power function is the exponent b. If $b > 0$, the data are more representative of an asymptotically increasing function (Figure 3*a*). If $b < 0$, the data are more representative of an asymptotically decreasing function (Figure 3*b*). As b approaches zero, the power curve approaches linearity, whereas for larger absolute values of b, the power function curves more sharply. In the case of $b > 0$, the function will be convex for $b > 1$ or concave for $b < 1$. The residual error from the fitted power curves was also calculated as a function of train length (Figure 2).

The results were generally consistent with the car mile and train mile premises developed. The average length of trains involved in an accident due to car mile-related causes was 68.3 cars, whereas the average for train mile–related causes was 52.5 cars. Also, the percentage of train mile–related accidents declined asymptotically as a function of train length. However, although the R^2 values for the regression analysis were significant, it was evident that there were some discrepancies between the observed data and the predicted relationships, as indicated by the large residual error for the extreme train lengths (Figure 2).

TABLE 1 ADL/AAR Accident Cause Groups and Classification of FRA Accident Causes

Group	CM/TM	Cause Description	Group	CM/TM	Cause Description
01E	CM	Air hose defect (car)	06H	TM	Radio communications error
02E	CM	Brake rigging defect (car)	07H	TM	Switching rules
03E	CM	Handbrake defects (car)	08H	TM	Mainline rules
04E	CM	UDE (car or loco)	09H	TM	Train handling (excl. brakes)
05E	CM	Other brake defect (car)	10H	TM	Train speed
06E	CM	Centerplate/carbody defects (car)	11H	TM	Use of switches
07E	CM	Coupler defects (car)	12H	TM	Misc. track and structure defects
08E	CM	Truck structure defects (car)	01M	TM	Obstructions
09E	CM	Sidebearing, suspension defects (car)	02M	TM	Grade crossing collisions
10E	CM	Bearing failure (car)	03M	CM	Lading problems
11E	CM	Other axle/journal defects (car)	04M	CM	Track–train interaction
12E	CM	Broken wheels (car)	05M	TM	Other miscellaneous
13E	CM	Other wheel defects (car)	01S	TM	Signal failures
14E	CM	TOFC/COFC defects	01T	CM	Roadbed defects
15E	CM	Loco trucks/bearings/wheels	02T	TM	Nontraffic, weather causes
16E	CM	Loco electrical and fires	03T	CM	Wide gauge
17E	CM	All other locomotive defects	04T	CM	Track geometry (excl. wide gauge)
18E	CM	All other car defects	05T	CM	Buckled track
19E	CM	Stiff truck (car)	06T	CM	Rail defects at bolted joint
20E	CM	Track/train interaction (hunting) (car)	07T	CM	Joint bar defects
21E	CM	Current collection equipment (loco)	08T	CM	Broken rails or welds
01H	TM	Brake operation (main line)	09T	CM	Other rail and joint defects
02H	TM	Handbrake operations	10T	CM	Turnout defects–switches
03H	TM	Brake operations (other)	11T	CM	Turnout defects–frogs
04H	TM	Employee physical condition	12T	CM	Misc. track and structure defects
05H	TM	Failure to obey/display signals			

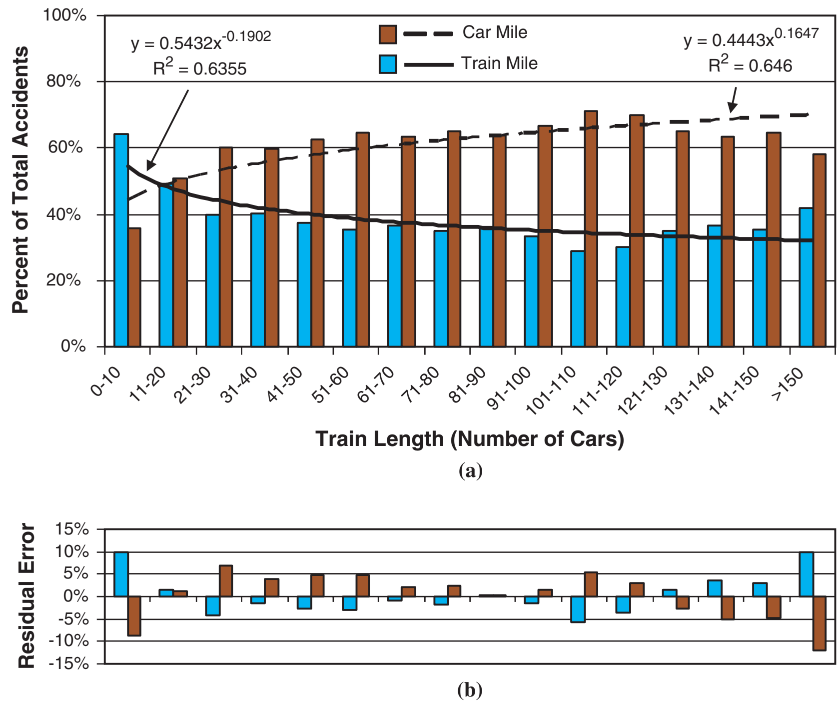

FIGURE 2 (*a*) Percentage of car mile– and train mile–related accidents versus train length using the ADL accident cause classification. (*b*) A power function with residual error is also shown.

These discrepancies suggested that the previous classification of accident causes should be reevaluated to see if they could be improved based on newer data and analysis. Therefore, a more detailed analysis of individual accident causes was conducted. The relationships between number of accidents and percentage of accidents as a function of train length were graphed for each cause group. Although there were not enough data for accurate assessment of all the accident cause groups, many of them conformed well to the expectations for train mile- or car mile-related causes, examples of which were grade crossing collisions and air hose defects, respectively (Figures 4*a* and 4*b*). However, examination of the data also suggested that some of the cause groups needed to be reclassified because the results were inconsistent with the car and train mile expectations (Figures 4*c* and 4*d*).

A possible explanation exists for the cause group "train handling." This type of accident is caused by a locomotive engineer improperly handling the train and is commonly attributed to excessive horsepower use. This had been previously defined as a train mile–related cause because it is due to human error. However, accidents caused by the use of excessive horsepower are in fact more common in long trains than short trains and therefore resemble a car mile-related cause. Conversely, the cause group "all other locomotive defects" had been classified as a car mile cause because it involves a mechanical failure. However, the number of locomotives, and therefore the

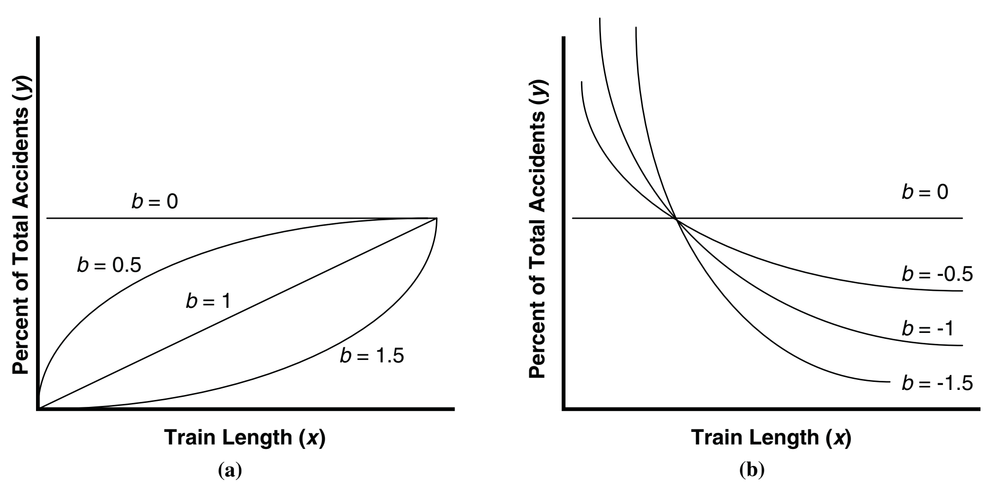

FIGURE 3 Characteristics of exponential term, b, of power function $y = ax^b$, where (*a*) represents a car mile–related cause and (*b*) represents a train mile–related cause.

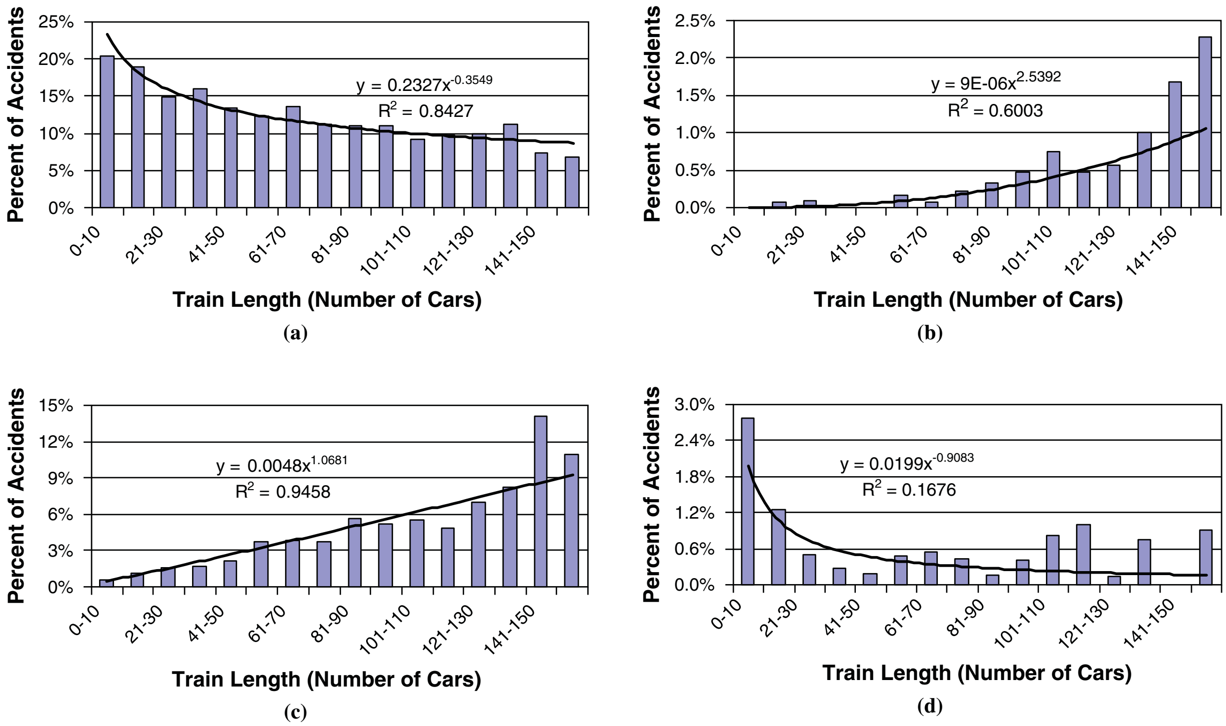

FIGURE 4 Percentage of accidents versus train length for four example cause groups: (*a* and *b*) correctly classified and (*c* and *d*) incorrectly classified.

likelihood of a locomotive defect, is not necessarily affected by an increased number of cars. Several discrepancies were also observed in other accident cause groups. Therefore, a quantitative metric was developed to classify objectively each accident cause group as train mile or car mile related.

Development of Classification Metric

The previously stated premise about car mile- and train mile–related causes was used to develop a quantitative metric to classify each of the ADL accident cause groups. Car mile accidents should be more prevalent in longer trains and should be an asymptotically increasing function of the percentage of accidents as train length increases, and the reverse should be true for train mile–related causes.

Two parameters were calculated for each accident cause to characterize the causes as either car mile or train mile related. The first parameter is the average length of trains involved in an accident for each cause group. The second parameter is derived from the power function curve and its goodness of fit to the data for the percentage of accidents for each cause group as a function of train length. The exponent in the power function was used to assess the asymptotical increase or decrease in the data (Figure 3). The greater the difference between the calculated value of b and zero, the stronger the asymptotically increasing or decreasing function, and therefore the indication of either a car mile- or a train mile–related cause. For example, Cause Group 2T, nontraffic and nonweather causes ($b = -0.8666$), showed a much stronger indication of a train mile–related cause than 1M, obstructions ($b = -0.3322$).

In addition to characterizing the shape of the curves for each accident cause group, it was also important to quantify how well they fit the data. In some cases there were insufficient data to fit a curve and in others the data showed no trend. In order to assess the goodness of fit, the coefficient of determination, R^2, for each data set was calculated. R^2 values range from 0 to 1 and quantify the goodness of fit with higher values indicating a better fit. Therefore, the accident causes with a high R^2 value are weighted more strongly in the metric than those with a low R^2 value. In summary, the accident metric, termed AM_i, incorporates three characteristics associated with each accident cause group, i: the average length of trains, l; the "shape" of the curve as a function of train length as indicated by the exponential term, b; and the goodness of fit of the data to the curve, as indicated by the R^2 value. This accident metric is expressed as follows:

$$AM_i = \frac{l_i}{L} + \left(b_i R_i^2\right)$$

where

AM_i = accident cause metric for cause group i;
l_i = average train length for cause group i;
L = overall average length of trains involved in accidents in data set = 61.79;
b_i = value of exponential term in power curve equation, $y = ax^b$, for cause group i; and
R_i^2 = coefficient of determination for a power curve fit to the data for cause group i.

If the average length of trains in accidents due to cause i (l_i) is greater than L, AM_i is increased and vice versa. The greater the difference between l_i and L, the more AM_i is affected. The second term of the metric is the power function exponent, b. If $b_i > 0$ for cause I, it increases AM_i, and vice versa. Similarly, the greater the difference between b_i and 0, the greater the effect is on AM_i. Finally, b is multiplied by R^2 to account for how well the function fits the data. If R^2 is close to 1, the second term will influence the metric more strongly.

If the function is a poor fit (low R^2), b will have little effect on AM_i. Therefore, for R^2 values close to 1, AM_i will be calculated based on both average train length and b, whereas for low R^2 values, AM_i will be calculated primarily based on average train length.

The metric AM_i was used to classify and rank the cause groups (Table 2). Not all cause groups included enough data to classify them properly as either car mile or train mile related and these were excluded from the analysis. In particular, Cause Group 21E, current collection equipment, was excluded because only short passenger trains (<10 cars) were involved in this cause group, with none of the accidents resulting in a derailment. The cause groups in Table 2 are ordered from most car mile related at the top to most train mile related at the bottom. Cause groups with rankings in the middle are not represented strongly by either car mile or train mile classifications.

Reclassification of Accident Causes

AM_i is used to classify accident causes as either more consistent with characteristics of car mile-related accidents or train mile–related accidents. If $AM_i > 1$, the cause group is classified as a car mile accident; conversely, if $AM_i < 1$, the cause group is classified as a train mile–related accident (Table 2). If the classification based on the metric is different from the previous ADL classification, this is indicated by a "YES" in the column heading "Change." Using the metric, the authors reclassified 11 cause groups. Cause Groups 1H, 9H, and 1S were changed from train mile to car mile causes. Groups 16E, 17E, 18E, 19E, 1T, 3T, 4T, and 12T were changed from car mile to train mile causes. Cause Groups 3E, 4E, 14E, 21E, 4H, and 11T were not evaluated using the metric because of the small number of accidents for each group. The highest-ranked car mile-related accident cause is 1E—air hose defect—with a score of 3.277; whereas the highest ranked train mile–related accident cause is 02H—handbrake operations—with a score of –0.0275.

As previously discussed, there were instances in which the accuracy of the initial classification based on the characteristics of the car mile and train mile premise could be improved. Using the calculated values for AM_i, the authors reexamined the overall train mile- and car mile-related causes for comparison with the ADL classification. After reclassifying the data, the values were now more clearly representative of car mile- and train mile–related causes (Figure 5). The average train lengths for car mile-related causes increased from 68.3 to 79.0 cars, while the average train length of train mile–related causes decreased from 52.5 to 48.4 cars. Also, b increased to 0.6175 and $R^2 = 0.9147$ for car mile-related causes, whereas b decreased to –0.4063 and $R^2 = 0.9201$ for train mile–related causes. Overall, the new classification is more consistent with expectations from the stated car mile and train mile premise.

CALCULATION OF ACCIDENT RATES

As stated earlier, train accident rates can be determined by summing the car mile- and train mile–related rates. The two rates can be calculated using known accident data, the number of car miles and train miles operated, and the new classification of accident causes. Data on car miles and train miles operated are available from the Association of American Railroads (AAR) (*13*). Car miles and train miles are defined as the movement of a car or train the distance of 1 mi and is based on the distance run between terminals or stations. Accident information was downloaded from the FRA Office of Safety for the time period 1990 to 2005 (*11*). Data for all accident types for Class I railroads operating on mainline and siding tracks were used to ensure consistency with the AAR definition of car miles and train miles for this portion of the analysis. The developed classification metric was used to classify each accident cause.

The car mile- and train mile–related accident rates from 1990 to 2005 were calculated by dividing the number of accidents by the number of miles operated (Table 3). In 2005, the accident rate for car mile-related causes was 1.05×10^{-8} or about 0.011 accidents per million car miles, and the train mile–related accident rate was 8.62×10^{-7} or about 0.86 accidents per million train miles. The expected number of train accidents, based on 2005 data, can be calculated as follows:

$$A_{\text{EXP}} = 1.05 \times 10^{-8} M_C + 8.62 \times 10^{-7} M_T$$

where

A_{EXP} = accidents expected,
M_C = number of car miles, and
M_T = number of train miles.

It is clear, based on this equation, that if the number of cars per train is increased, the consequent increase in car miles operated leads to an increase in the accident rate for each train so affected. Similarly, an increase in the number of trains operated on a system will increase the number of train miles operated, and thus increase the number of train mile-caused accidents. To understand the effect of train length on accident likelihood, the accident rate equation can be expanded to include the term for train length:

$$\begin{aligned} A_{\text{EXP}} &= 1.05 \times 10^{-8} ndT_L + 8.62 \times 10^{-7} nd \\ &= nd\left(1.05 \times 10^{-8} T_L + 8.62 \times 10^{-7}\right) \end{aligned}$$

where

A_{EXP} = accidents expected,
n = number of trains operated,
d = number of miles operated, and
T_L = average cars per train (train length).

This equation is useful for understanding how changes in operating procedures, such as train length or number of trains operated, will affect the expected number of train accidents.

ACCIDENT RATE SENSITIVITY ANALYSIS

Two simple sensitivity analyses were conducted to illustrate the effect of changes in train length on train accident rate. In the first, an operational choice of train length given a fixed number of car movements was examined. The analysis parameters are intended to represent a typical high-density, long-distance, Class I railroad mainline with 25,000 car movements per week and a distance of 2,000 mi with train length and number of trains as the variables. The estimated number of accidents based on 2005 data is 1.05×10^{-8} accidents per car mile plus 8.62×10^{-7} accidents per train mile as calculated by the previous reclassification of accident causes. Train length varied from 10 cars to 150 cars per train (Table 4).

As train length increases, the likelihood that a train will be involved in an accident increases due to the increase in car miles per train; however, because of the reduction in train miles, the net effect is a reduction in the total number of accidents. So all other things being equal,

TABLE 2 Classification, Score, and Rank of Accident Cause Groups Using Metric

		Trendline			Distribution		Metric		
Cause	Description	*a*	*b*	R^2	Cases	Avg. Length	Score	Rank	Change
Car Mile Causes									
01E	Air hose defect (car)	0.000	2.539	0.600	50	108.30	3.2770	1	—
12E	Broken wheels (car)	0.001	1.631	0.942	372	96.90	3.1054	2	—
10E	Bearing failure (car)	0.002	1.409	0.893	780	89.24	2.7025	3	—
11E	Other axle/journal defects (car)	0.001	1.218	0.863	156	95.81	2.6022	4	—
09H	Train handling (excl. brakes)	0.005	1.068	0.946	647	89.34	2.4561	5	Yes
01H	Brake operation (main line)	0.002	1.047	0.822	209	90.43	2.3238	6	Yes
07E	Coupler defects (car)	0.002	0.998	0.859	274	89.39	2.3043	7	—
13E	Other wheel defects (car)	0.003	0.924	0.886	324	88.38	2.2486	8	—
06E	Centerplate/carbody defects (car)	0.003	0.838	0.896	281	85.99	2.1423	9	—
05T	Buckled track	0.006	0.697	0.726	438	78.95	1.7842	10	—
08E	Truck structure defects (car)	0.000	0.834	0.059	61	94.66	1.5807	11	—
09T	Other rail and joint defects	0.003	0.498	0.667	153	75.65	1.5562	12	—
04M	Track–train interaction	0.008	0.616	0.536	483	74.36	1.5337	13	—
05E	Other brake defect (car)	0.002	0.517	0.320	109	77.73	1.4233	14	—
08T	Broken rails or welds	0.046	0.391	0.369	1,798	71.66	1.3040	15	—
02E	Brake rigging defect (car)	0.001	0.384	0.014	73	79.15	1.2863	16	—
20E	Track–train interaction (hunting) (car)	0.002	0.369	0.233	80	73.79	1.2799	17	—
07T	Joint bar defects	0.004	–0.180	0.004	115	78.44	1.2688	18	—
09E	Sidebearing, suspension defects (car)	0.006	0.355	0.149	267	71.65	1.2125	19	—
06T	Rail defects at bolted joint	0.004	–0.018	0.000	110	72.82	1.1785	20	—
01S	Signal failures	0.000	0.724	0.053	64	69.27	1.1592	21	Yes
10T	Turnout defects–switches	0.026	0.034	0.009	528	65.37	1.0583	22	—
03M	Lading problems	0.020	0.131	0.082	469	64.60	1.0563	23	—
15E	Loco trucks/bearings/wheels	0.009	–0.415	0.038	127	64.59	1.0294	24	—
Train Mile Causes									
10H	Train speed	0.002	0.113	0.014	64	61.67	0.9996	21	—
19E	Stiff truck (car)	0.021	–0.601	0.067	212	62.58	0.9728	20	Yes
04T	Track geometry (excl. wide gauge)	0.040	–0.796	0.113	1,064	63.69	0.9405	19	Yes
03H	Brake operations (other)	0.005	–0.122	0.060	80	58.05	0.9321	18	—
01T	Roadbed defects	0.040	–0.796	0.113	274	55.18	0.8028	17	Yes
05H	Failure to obey/display signals	0.040	–1.134	0.138	213	56.79	0.7621	16	—
11H	Use of switches	0.098	–0.901	0.124	561	53.41	0.7526	15	—
02T	Nontraffic, weather causes	0.027	–0.867	0.159	155	53.28	0.7242	14	—
05M	Other miscellaneous	0.061	–0.255	0.294	814	48.16	0.7045	13	—
18E	All other car defects	0.017	–0.353	0.223	254	45.41	0.6562	12	Yes
12H	Misc. track and structure defects	0.018	–0.308	0.347	248	45.14	0.6237	11	—
03T	Wide gauge	0.101	–0.480	0.407	933	49.68	0.6090	10	Yes
06H	Radio communications error	0.015	–1.196	0.214	67	52.39	0.5915	9	—
16E	Locomotive electrical and fires	0.018	–0.799	0.139	161	43.12	0.5867	8	Yes
01M	Obstructions	0.057	–0.332	0.626	686	46.41	0.5430	7	—
02M	Grade crossing collisions	0.233	–0.355	0.843	2,546	50.27	0.5145	6	—
17E	All other locomotive defects	0.020	–0.908	0.168	169	38.56	0.4718	5	Yes
07H	Switching rules	0.053	–0.601	0.678	411	44.72	0.3165	4	—
08H	Mainline rules	0.026	–0.473	0.475	349	31.64	0.2873	3	—
12T	Misc. track and structure defects	0.148	–1.379	0.303	569	30.30	0.0730	2	Yes
02H	Handbrake operations	0.144	–1.475	0.349	442	30.13	–0.0275	1	—

(continued on next page)

TABLE 2 *(continued)* **Classification, Score, and Rank of Accident Cause Groups Using Metric**

		Trendline			Distribution		Metric		
Cause	Description	*a*	*b*	R^2	Cases	Avg. Length	Score	Rank	Change
Not Evaluated Using Metric									
04H	Employee physical condition				27	59.56			
11T	Turnout defects–frogs				25	76.00			
03E	Handbrake defects (car)				25	32.80			
04E	UDE (car or locomotive)				39	103.72			
14E	TOFC/COFC defects				19	54.26			
21E	Current collection equipment (locomotive)				86	7.62			

train accidents will be minimized when train length is maximized or the number of trains operated is minimized.

The second study examines how an increase in traffic levels may affect train accident rates. The analysis parameters are similar to those from the previous study of a 2,000-mi Class I railroad freight mainline with the same weekly traffic level of 25,000 car movements. The railroad is currently operating trains with an average length of 100 cars. The car movements are expected to increase by 10% to a new total of 27,500 movements. The operational choice in this study is either to continue operating the same number but longer trains or to maintain the current train length and operate more trains. The traffic increase will lead to an increase in overall accidents; however, this effect can be minimized by increasing the length of trains instead of increasing the number of trains operated (Table 5). Again, this study suggests for this type of scenario that a railroad may be able to reduce the overall number of accidents by running fewer, longer trains as opposed to a higher number of shorter trains.

CONCLUSIONS

Accident rates are affected by both car mile- and train mile–related accident causes. A consequence of this is that the length of trains affects the likelihood of accidents. Previous research combined the FRA-defined accident causes into 51 unique cause groups based on expert opinion. The authors developed a new quantitative metric to classify the causes as either car mile- or train mile-related. Use of the new metric led to the reclassification of 11 of the cause groups and

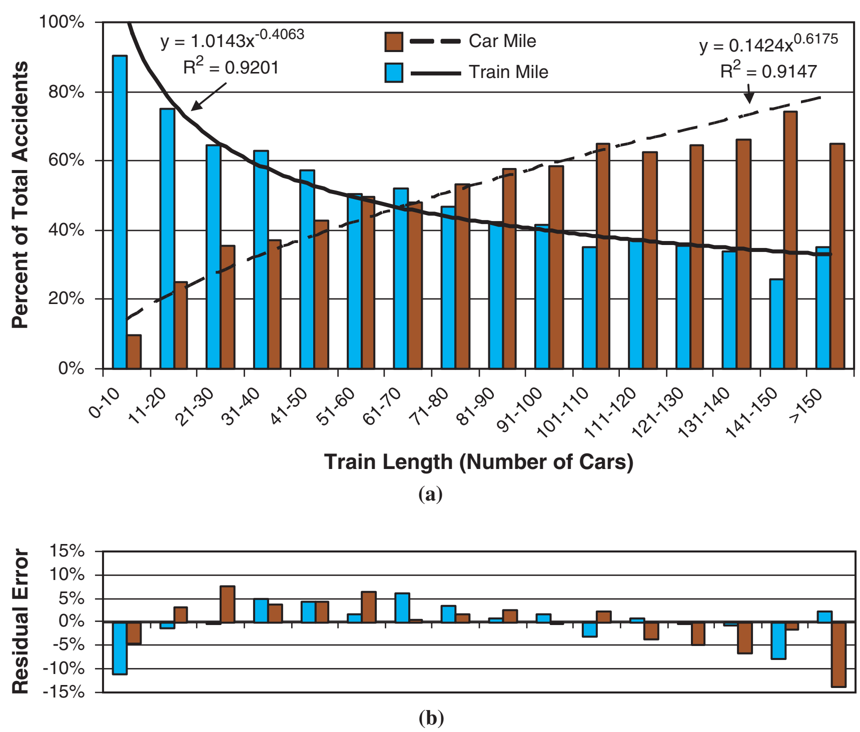

FIGURE 5 (*a*) Percentage of car mile– and train mile–related accidents versus train length using the new accident cause classification. (*b*) A power function with residual error is also shown.

TABLE 3 Car and Train Mainline Accident Rates Using Reclassification of Accident Causes, Class I Freight Railroads, 1990–2005

Year	Car Mile–Caused Accidents	Car Miles Operated (millions)	Car Mile Accident Rate (per million car miles)	Train Mile–Caused Accidents	Train Miles Operated (millions)	Train Mile Accident Rate (per million train miles)
1990	510	26,159	0.0195	486	380	1.280
1991	479	25,628	0.0187	465	375	1.240
1992	360	26,128	0.0138	414	390	1.061
1993	370	26,883	0.0138	432	405	1.065
1994	315	28,485	0.0111	418	441	0.948
1995	362	30,383	0.0119	457	458	0.997
1996	379	31,715	0.0120	402	469	0.858
1997	343	31,660	0.0108	418	475	0.880
1998	378	32,657	0.0116	422	475	0.889
1999	367	33,851	0.0108	362	490	0.738
2000	420	34,590	0.0121	433	504	0.859
2001	400	34,243	0.0117	468	500	0.937
2002	374	34,680	0.0108	380	500	0.761
2003	392	35,555	0.0110	431	516	0.835
2004	424	37,071	0.0114	453	535	0.847
2005	395	37,712	0.0105	472	548	0.862
1990–2005	6,268	507,400	**0.0124**	6,913	7,460	**0.927**

was found to be more representative of car mile and train mile expectations. Therefore, using the new classification and recent accident data, updated mainline car mile- and train mile–related accident rates were calculated for Class I freight railroads. These rates, as evaluated in a sensitivity analysis, showed that the decision to dispatch the same number of shipments in fewer, longer trains versus more, shorter trains may affect the overall accident likelihood.

TABLE 4 Sensitivity Analysis of the Effect of Train Length on Accident Rate

Average Train Length (T_L)	Number of Trains (n)	Probability of Accident for Each Individual Train	Total Expected Number of Accidents
10	2,500	0.00193	4.84
20	1,250	0.00214	2.68
30	833	0.00235	1.96
40	625	0.00256	1.60
50	500	0.00277	1.39
60	417	0.00298	1.24
70	357	0.00319	1.14
80	313	0.00340	1.06
90	278	0.00361	1.00
100	250	0.00382	0.96
110	227	0.00403	0.92
120	208	0.00424	0.88
130	192	0.00445	0.86
140	179	0.00466	0.83
150	167	0.00487	0.81

NOTE: 25,000 carloads shipped; 2,000 mi; 150-car maximum train length

FUTURE WORK

The analysis completed in this paper is based on a binary classification of accident causes as either train mile or car mile related. However, many causes may not be purely train mile or car mile related, but instead depend on a combination of both. Future work may be possible to define a function for each cause group based on both car miles and train miles. Additional information, such as the distribution of trains operated by train length, would be useful in defining the linear or nonlinear accident cause functions.

Future work may also be possible to evaluate and further refine the accident cause classification metric. For example, it may be possible to transform the current summation metric into a product metric using the same evaluation factors. The product metric may strengthen the analysis because it would multiply the classification terms and their effects on the metric instead of a simple summation. Further research into the classification metric might also reveal a better threshold for accident cause classification. An adjustment to the classification metric to include the average length of trains operated instead of average length of trains involved in accidents may remove potential bias.

TABLE 5 Sensitivity Analysis of the Effect of Traffic Increase on Accident Rate

Number of Trains (n)	Average Train Length (TL)	Probability of Accident for Each Individual Train	Total Expected Number of Accidents
250	100	0.00382	0.96
250	110	0.00403	1.01
275	100	0.00382	1.05

NOTE: 27,500 carloads shipped; 2,000 mi

Finally, it may also be possible to determine an optimal train length to minimize the number of cars derailed. Longer trains may be involved in fewer total accidents, but longer trains derail or damage more cars on average than shorter trains (*14*).

ACKNOWLEDGMENTS

The authors appreciate the helpful comments and suggestions from Daniel Peltier and Yung-Cheng Lai, graduate students at the University of Illinois, and Mike Wnek, an undergraduate at the University of Illinois, for his assistance with the analysis. Additionally, the authors thank Mark Stehly, Peter French, Todd Treichel, and Robert Anderson for their useful review and comments. Robert Anderson conducted an initial exploration of the concepts investigated here in connection with his graduate research. The first author was supported by a research grant from the BNSF Railway and a CN Research Fellowship at the University of Illinois.

REFERENCES

1. Federal Railroad Administration. *Railroad Safety Statistics: 2005 Annual Report.* Prepared for U.S. Department of Transportation, 2006. safetydata.fra.dot.gov/OfficeofSafety/Forms/Default.asp. Accessed July 31, 2007.
2. FRA Office of Safety. *Train Accident Trends.* safetydata.fra.dot.gov/OfficeofSafety/Query/Default.asp?page=inctally1.asp. Accessed July 31, 2007.
3. TranSys Research, Ltd. *Evaluation of Risk Associated with Stationary Dangerous Goods Railway Cars.* Prepared for Transport Canada, Ottawa, 2006. www.tc.gc.ca/tdc/publication/pdf/14600/14690e.pdf. Accessed July 31, 2007.
4. Anderson, R. T., and C. P. L. Barkan. Railroad Accident Rates for Use in Transportation Risk Analysis. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1863,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 88–98.
5. Center for Chemical Process Safety. *Guidelines for Chemical Transportation Risk Analysis.* American Institute of Chemical Engineers, New York, 1995.
6. Risk Assessment for the Transportation of Hazardous Materials by Rail. *Supplementary Report: Railroad Accident Rate and Risk Reduction Option Effectiveness Analysis and Data,* 2nd rev. Arthur D. Little, Inc., Cambridge, Mass., 1996.
7. Surface Transportation Board. Draft Environmental Impact Statement: Finance Docket No. 34079. San Jacinto Rail Limited and The Burlington Northern and Santa Fe Railway Company Construction and Operation of a Rail Line from the Bayport Loop in Harris County, Texas. Decision ID No. 33200, 2002.
8. Kawprasert, A., and C. P. L. Barkan. Effects of Route Rationalization on Hazardous Materials Transport Risk. In *Transportation Research Record: Journal of the Transportation Research Board,* No. 2043, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 65–72.
9. Zhao, J., A. H. C. Chan, and M. P. N. Burrow. Probabilistic Model for Predicting Rail Breaks and Controlling Risk of Derailment. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1995,* Transportation Research Board of the National Academies, Washington, D.C., 2007, pp. 76–83.
10. Anderson, R. T., and C. P. L. Barkan. Derailment Probability Analysis and Modeling of Mainline Freight Trains. In *Proc., 8th International Heavy Haul Conference,* Rio de Janeiro, June 2005, International Heavy Haul Assn., Virginia Beach, Va., pp. 491–497.
11. FRA Office of Safety. Download Data on Demand. safetydata.fra.dot.gov/officeofsafety/Downloads/Default.asp. Accessed May 10, 2007.
12. FRA Office of Safety. Train Accident Cause Codes. safetydata.fra.dot.gov/OfficeofSafety/Downloads/Default.asp. Accessed May 10, 2007.
13. *Railroad Facts,* 2000–2006 eds. Association of American Railroads, Washington, D.C.
14. Barkan, C. P. L., C. T. Dick, and R. T. Anderson. Railroad Derailment Factors Affecting Hazardous Materials Transportation Risk. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1825,* Transportation Research Board of the National Academies, Washington, D.C., 2003, pp. 64–74.

The Transportation of Hazardous Materials Committee sponsored publication of this paper.