THE RELATIONSHIP BETWEEN TRAIN LENGTH AND ACCIDENT CAUSES AND RATES

Darwin H. Schafer II – Corresponding Author
Graduate Research Assistant, Railroad Engineering Program
Department of Civil and Environmental Engineering
University of Illinois at Urbana-Champaign
B-118 Newmark Civil Engineering Laboratory
205 N. Matthews Ave.
Urbana, IL, USA 61801
Phone: (217) 244-6063
Fax: (217) 333-1924
Email: dschafer@uiuc.edu

Christopher P.L. Barkan
Associate Professor
Director - Railroad Engineering Program
Department of Civil and Environmental Engineering
University of Illinois at Urbana-Champaign
1201 Newmark Civil Engineering Laboratory
205 N. Matthews Ave.
Urbana, IL, USA 61801
Phone: (217) 244-6338
Fax: (217) 333-1924
Email: cbarkan@uiuc.edu

Submission Date: 11/15/07

Word Count: 7,750 (5,250 words, 5 Tables, 5 Figures)

ABSTRACT

Train accident rates are a critical metric of railroad transportation safety and risk performance. Understanding the factors that affect 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 in general these have not considered the effect of train length on train accident rate. It has been suggested that train 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 that cause will affect overall train accident rate. Accident causes have been classified as car or train-mile correlated based on expert opinion but no quantitative test of these classifications has been conducted. The definition of car-mile versus train-mile causes leads to the hypothesis that longer trains should experience more accidents than shorter trains. FRA accident data were used to develop and test a quantitative metric to objectively characterize different accident causes as either car-mile or train-mile correlated. Based on the results of the study a sensitivity analysis was conducted to evaluate how changes in train length affect individual trains' accident rate and system-wide 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 on a location specific scale 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 (3-5). However, these analyses have generally not considered the effect of train length on train accident rate. It has been suggested that train length has an effect on accident rate because more cars in a train increase the likelihood that a car or track component may fail and that accident causes can be classified into two types of causes, those that are a function of the number of train-miles operated and those that are a function of car-miles operated (6, 7). 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 (4, 8). 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.

We undertook a study 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:

- Present the methodology for calculating train accident rates based on car-mile and train-mile accident causes,
- Develop a metric to quantitatively evaluate the classification of accident causes as car or train-mile related,
- Use the metric to properly classify train accident causes,
- Provide new train accident rates based on train length using current data, and
- Conduct a sensitivity analysis on our model to illustrate how changes in train length may affect train accident rate.

TRAIN LENGTH BASED ACCIDENT RATES

Train accident rates are composed of 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 (7, 9, 10). The number of car-miles operated for a particular train is affected by train length; longer trains accumulate more car-miles. However, not all accident causes are directly related to the length of the train, and instead are related only to the operation of the train. This leads to the concept that train accident causes can be separated into two groups, those dependent on train length, corresponding to the number of car-miles operated, and those independent of train length, corresponding to the number of train-miles operated. They 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 vs. Train-Mile Expectations

The car-mile cause and train-mile cause definitions lead to the hypothesis that longer trains should 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, corresponds to the number of cars in the train and not the linear measure of a train's actual length.

The hypothesis leads to two predictions that should be evident when examining accident data and can be used to evaluate different train accident causes. The first prediction is that the average length of a train involved in an accident should be greater for car-mile-related causes compared to 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 towards long or short trains.

The second prediction is that the proportion of accidents for car-mile-related accidents should be an asymptotically increasing function of train length, whereas train-mile-related accidents should be an asymptotically decreasing function. Longer trains should experience a higher percentage of accidents from car-mile-related causes due to 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 hypothesis 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. The FRA accident database contains 389 unique accident causes (11, 12). A previous study by ADL classified each accident cause as either car-mile or train-mile-related (7). 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 properly quantify 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:

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

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)
M_T = Number of train miles

Under this model we expect that longer trains will experience more train accidents. As a train's length increases, 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 1a).

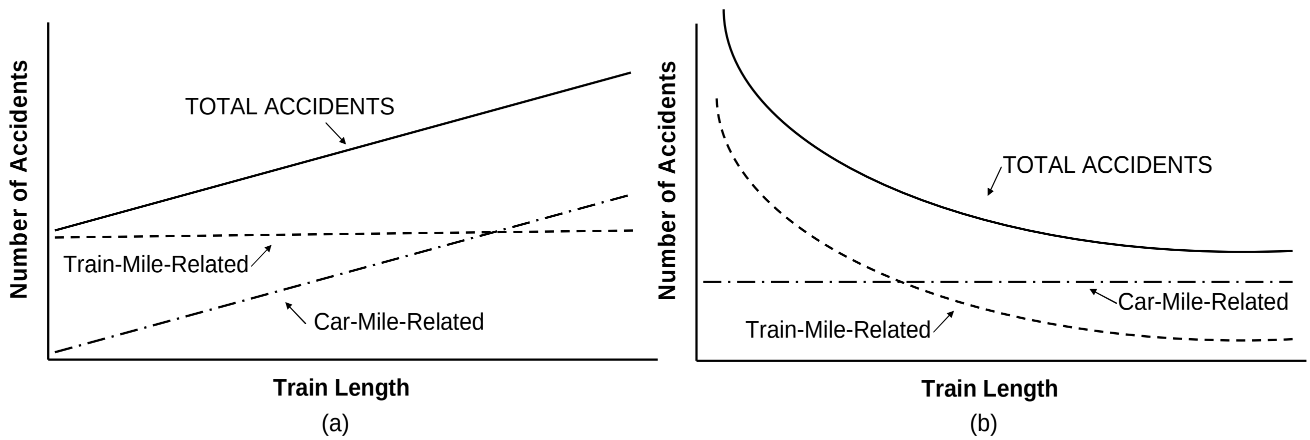

FIGURE 1 Expected accidents from car-mile and train-mile-related causes as a function of train length for a single train (a) and for a fixed amount of traffic (b).

If one extends this model system wide, 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 1b).

It should be noted that there are limits to the validity of this result for very long train lengths (>150). This is because the hypothesis presented, as well as the data used in our 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 and train braking. The intention of this analysis is not to suggest that longer trains will necessarily improve safety; instead the purpose is to develop a better quantitative understanding of how changes that affect various accident causes, such as number of trains and train length, will affect overall accident rates.

CLASSIFICATION OF ACCIDENT CAUSES

To accurately determine 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 is 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 (CM) and 17 as train-mile-related (TM) (Table 1) (7). 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.

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	Non-Traffic, 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			

We 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 due to 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).

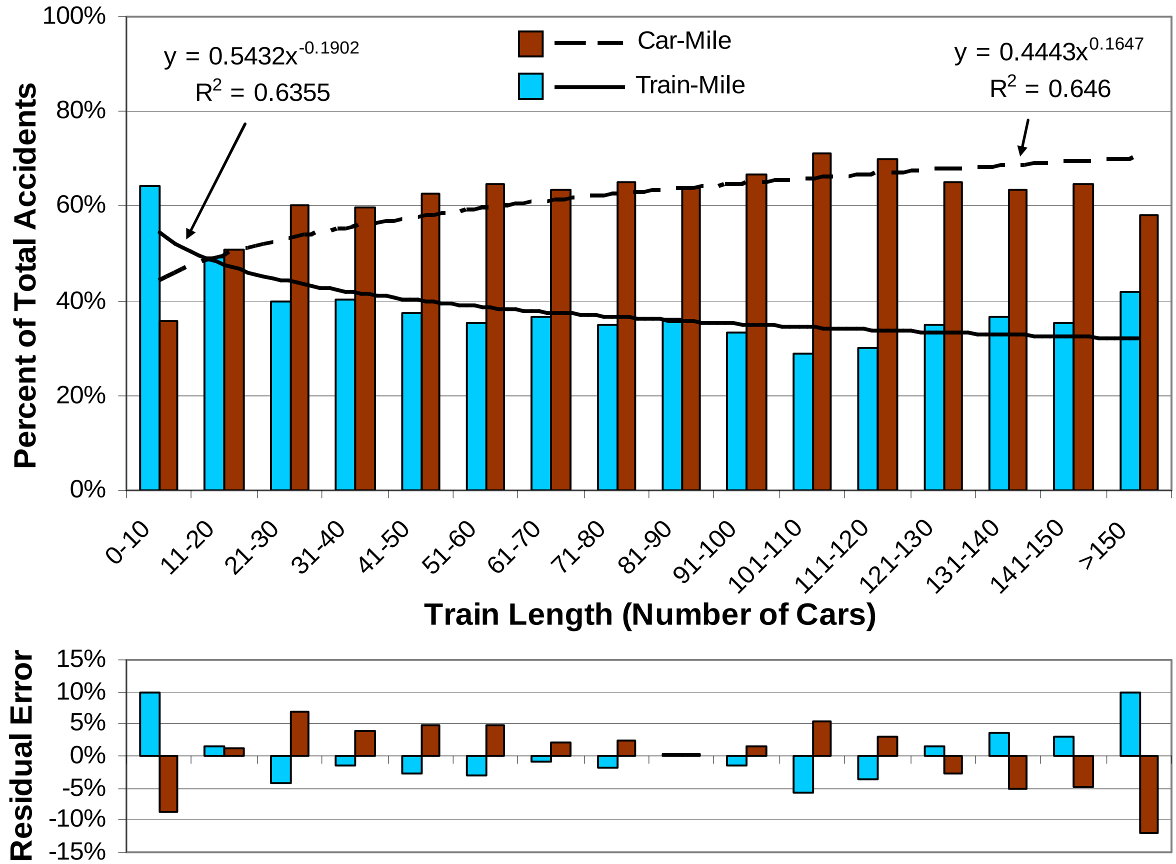

FIGURE 2 Percentage of car and train-mile-related accidents versus train length using the ADL accident cause classification. A power function with residual error is also shown.

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 3a). If $b < 0$, the data are more representative of an asymptotically decreasing function (Figure 3b). As b approaches zero the power curve becomes less curved and more representative of a horizontal, flat line; 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 for the various train lengths (Figure 2). The residual error was greatest for long train lengths and trains of less than 10 cars.

The results are generally consistent with the car and train-mile predictions. 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. 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 shown in the residual error graph. The error is particularly evident for trains longer than 110 cars.

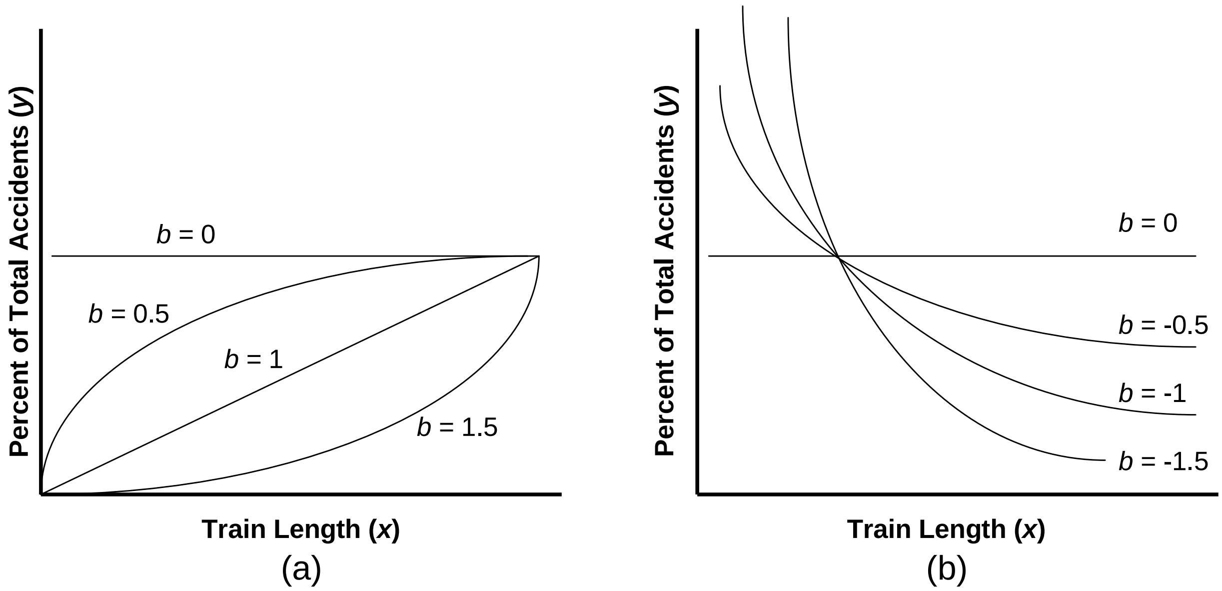

FIGURE 3 Characteristics of exponential term, *b*, of power function $y = ax^b$, where (a) represents a car-mile-related cause and (b) represents a train-mile-related cause.

These discrepancies suggest that the previous classifications of accident causes by ADL should be evaluated as they may have changed due to the inclusion of this new data and analysis. Therefore, a more detailed analysis of individual accident causes was conducted. The relationships between number of accidents versus train length and percentage of accidents as a function of train length were graphed for each cause group. Although, not all of the accident cause groups contained enough data to allow an accurate evaluation; many of the cause groups conformed well to the predictions for train-mile or car-mile-related causes, examples of which were grade crossing collisions and air hose defects, respectively (Figures 4a and 4b).

However, examination of the data also suggested that some of the cause groups need to be reclassified because the results were inconsistent with the car and train-mile predictions (Figures 4c and 4d). A possible explanation exists for the cause group “train handling”, which is caused by a locomotive engineer improperly handling the train, commonly attributed to excessive horsepower use. ADL defined this 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” was classified by ADL as a car-mile cause because it is a mechanical failure. However, the number of locomotives, and therefore the likelihood of a locomotive defect, is not significantly affected by an increase in cars. Several discrepancies were also observed in other accident cause groups. Therefore a quantitative metric was developed to objectively classify each accident cause group as train-mile or car-mile-related.

Development of Classification Metric

We used the two expectations about car and train-mile related causes 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.

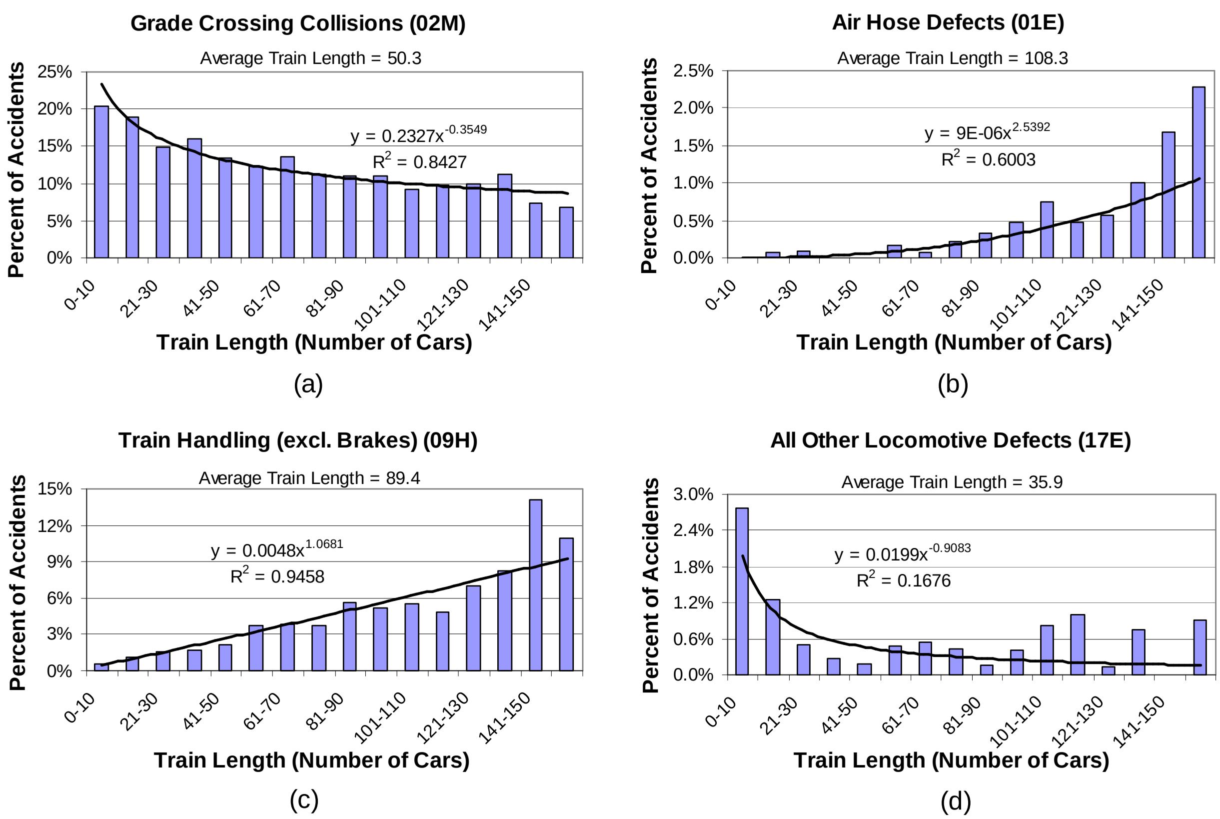

FIGURE 4 Percentage of accidents versus train length for four example cause groups; correctly classified (a) and (b), and incorrectly classified (c) and (d).

Two parameters were calculated for each accident cause to characterize them 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, non-traffic/weather 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. Higher values indicate that the curve fits the data better, whereas low values of R^2 indicate a curve that does not. Therefore the lines with a high R^2 are weighted more strongly in the metric than those with a low R^2 value. In summary, the accident metric, which we term AM_i, needs to incorporate three characteristics: average length of trains involved in an accident due to a particular accident cause group, the "shape" of the curve as a function of train length as indicated by the exponent, b, and the goodness of fit of the data to the curve, as indicated by the R^2. The metric is as follows:

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

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 dataset = 61.79
b_i = Value of exponential term in power curve equation, $y=ax^b$, for cause group *i*
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 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 equally based on average train length and b; whereas for low R^2 values AM_i will be calculated primarily based on average train length.

AM_i was used to classify and rank the cause groups (Table 2). Not all cause groups included enough data to properly classify them 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 we 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-miles causes. Cause groups 3E, 4E, 14E, 4H, and 11T were not evaluated using the metric due to the small number of accidents for each group. Also, cause group 21E, "current collection equipment", was not evaluated because these accidents involved only very short trains that did not typically result in a derailment.

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.

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

CAR-MILE-CAUSES		*Trendline $y=ax^b$*			*Distribution*		*Metric*		
Cause	**Description**	***a***	***b***	***R^2***	**Cases**	**Avg. Length**	**Score**	**Rank**	**Change**
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	1798	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		*Trendline $y=ax^b$*			*Distribution*		*Metric*		
Cause	**Description**	***a***	***b***	***R^2***	**Cases**	**Avg. Length**	**Score**	**Rank**	**Change**
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	1064	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	Non-Traffic, 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	2546	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	--
NOT EVALUATED USING METRIC		*Trendline $y=ax^b$*			*Distribution*		*Metric*		
Cause	**Description**	***a***	***b***	***R^2***	**Cases**	**Avg. Length**	**Score**	**Rank**	**Change**
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 Loco)				39	103.72			
14E	TOFC/COFC Defects				19	54.26			
21E	Current Collection Equipment (Loco)				86	7.62			

Using the calculated values for AM_i we reexamined the overall train-mile and car-mile-related causes for comparison to the ADL classification. Figure 2 indicated that the initial classification was not entirely accurate based on the car and train-mile expectation. After reclassifying the data, the values are 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 the car-mile and train-mile accident predictions.

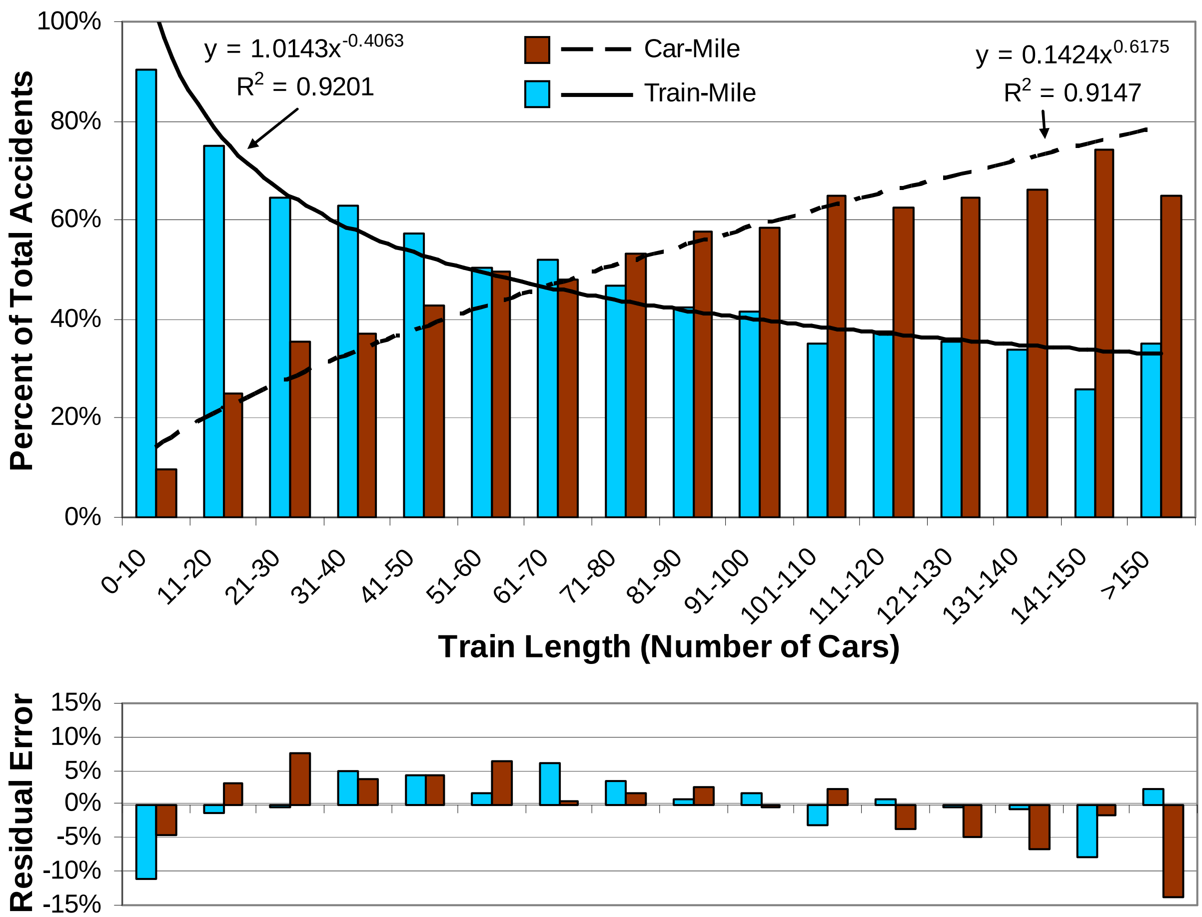

FIGURE 5 Percentage of car and train-mile-related accidents versus train length using the new accident cause classification. A power function with residual error is also shown.

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 and train-miles operated, and the new classification of accident causes. Data on car-miles and train-miles operated are available from the AAR (13). Car and train-miles are defined as the movement of a car or train the distance of one mile and is based on the distance run between terminals or stations. Accident information was downloaded and filtered for our criteria from the FRA Office of Safety for the time period 1990-2005 (11). FRA 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 and train-miles for this portion of the analysis. The developed classification metric was used to classify each accident cause.

The car 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 $1.05x10^{-8}$ or about .011 accidents per million car-miles and the train-mile-related accident rate was $8.62x10^{-7}$ or about 0.86 accidents per million train-miles. The expected number of train accidents, based on 2005 data, can be calculated as follows:

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

where:

A_{EXP} = Accidents expected
M_C = Number of car miles
M_T = Number of train miles

TABLE 3 Car and Train Mainline Accident Rates using the 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

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:

$$A_{EXP} = 1.05\text{x}10^{-8}\, n\, d\, T_L + 8.62\text{x}10^{-7}\, n\, d = n\, d\, (1.05\text{x}10^{-8}\, T_L + 8.62\text{x}10^{-7})$$

where:

A_{EXP} = Accidents expected
n = Number of trains operated
d = Number of miles operated
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

We conducted two simple analyses of the sensitivity analysis to illustrate the effect of changes in train length on train accident rate. In the first we examine an operational choice of train length given a fixed number of shipments. The analysis parameters are intended to represent a typical high density, long distance, Class I railroad mainline with 25,000 shipments per week and a distance of 2,000 miles with train length and number of trains as the variables. The estimated number of accidents based on 2005 data is $1.05\text{x}10^{-8}$ accidents per car-mile plus $8.62\text{x}10^{-7}$ accidents per train mile. We varied train length from 10 cars to 150 cars per train (Table 4).

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

Average Train Length (T_L)	Number of Trains (n)	Probability of an 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

25,000 Carloads Shipped; 2,000 Miles; 150 Car Maximum Train Length

$$\text{A}_{\text{EXP}} = n\, d\, (1.05\text{x}10^{-8}\, T_L + 8.62\text{x}10^{-7})$$

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, 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 will affect train accident rates. The analysis parameters are similar to those from the previous study of a 2,000 mile Class I railroad freight mainline with the same weekly traffic level of 25,000 shipments. The railroad is currently operating trains with an average length of 100 cars. The shipments are expected to increase by 10% to a new total of 27,500 shipments. The operational choice in this study is either to continue operating the same number, but longer trains, or 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 can reduce the overall number of accidents by running fewer, longer trains as opposed to a high number of shorter trains.

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

Number of Trains (n)	Average Train Length (T_L)	Probability of an 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

27,500 Carloads Shipped; 2,000 Miles

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 accident rate. The decision to dispatch the same number of shipments in fewer longer trains versus more, shorter trains will affect the overall accident rate. Furthermore, since some accident causes are correlated with car-miles and others with train-miles, accurate classification of the causes is important to correctly determine the effect of changes on accident rates. The FRA accident causes were combined into 51 unique cause groups, and classified as either car-mile or train-mile related by ADL in 1996. A metric was developed to quantitatively evaluate the 51 cause groups based on accident data. Use of the metric led to a reclassification of 11 cause groups. The new classification was found to be more representative of car and train-mile expectations. Mainline car-mile and train-mile-related accident rates were calculated for Class I freight railroads. These rates were used in a sensitivity analysis to illustrate the effect of changes in train length on overall accident rate.

Future Work

The previous analysis is based on classifications of causes that are either train-mile or car-mile-related. However, many causes may not be purely train or car-mile-related, but instead may depend on a combination of both. Additionally, some causes may depend on both car and train-miles but may be strongly dominated by one or the other. Future work may be possible to define a function for each cause group based on both car-miles and train-miles. Each cause function would weight how strongly the cause is affected by the number of car-miles and the number of train-miles. The developed functions of each cause could then be added together to calculate the effect on overall accident rate.

Future work is also possible to examine the affect longer trains may have on different accident types. For example, the operation of longer train lengths may have an effect on the number of grade crossing accidents. Longer trains may lead to fewer incidents of grade crossings collisions due to fewer trains; however, drivers may be more inclined to attempt to pass in front of an oncoming train due to the increased train length and vehicle wait time.

It may also be possible to determine an optimal train length to minimize cars derailed. Longer trains may be involved in fewer total accidents, but longer trains may derail or damage more total cars than shorter trains. This is based on the idea that longer trains have more kinetic energy and therefore can derail more cars when involved in an accident.

Finally, future work could be completed on comparing the accident model presented in the paper and other accident models. Train accident rates have been developed based on various parameters (3-5). The different train accident rates can be evaluated based on current accident data to test the accuracy of each particular model. It may also be possible to study the combination of different parameters from various accident models to develop a hybrid train accident model.

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. 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.

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, 2006. www.tc.gc.ca/tdc/publication/pdf/14600/14690e.pdf. Accessed July 31, 2007.
4. Anderson, R.T. & 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*, TRB, National Research Council, Washington, D.C., 2004, pp. 88-98.
5. Treichel, T.T. & C.P.L. Barkan. Working Paper on Mainline Freight Train Accident Rates. Unpublished Report to the Association of American Railroads, 1993.
6. Center for Chemical Process Safety (CCPS). *Guidelines for Chemical Transportation Risk Analysis*. American Institute of Chemical Engineers, New York, 1995.

7. Arthur D. Little, Inc. Risk Assessment for the Transportation of Hazardous Materials by Rail. Supplementary Report: Railroad Accident Rate and Risk Reduction Option Effectiveness Analysis and Data (2^{nd} Revision), 1996.

8. Surface Transportation Board (STB). 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.
9. Anderson, R.T. & C.P.L. Barkan. Train Accidents as Functions of Train-Miles and Car-Miles. Working Paper, University of Illinois at Urbana-Champaign, 2005.
10. Anderson , R.T. and C.P.L. Barkan 2005. Derailment probability analysis and modeling of mainline freight trains. In *Proceedings of the 8th International Heavy Haul Conference*, Rio de Janeiro, June 2005, 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. Association of American Railroads (AAR). *Railroad Facts. 2000 - 2006 Editions.* Association of American Railroads, Washington, D.C.