Statistical Modeling of Freight Train Derailments at Highway-Rail Level Crossings

Samantha G. Chadwick, C. Tyler Dick, Mohd Rapik Saat, Christopher P.L. Barkan

Rail Transportation and Engineering Center (RailTEC)
University of Illinois at Urbana-Champaign, Urbana, IL, USA

2013 World Congress on Railway Research
25-28 November 2013
Sydney, Australia

ABSTRACT
Implementation of highway-rail level crossing warning systems, educational programs and research on crossings have all contributed to a steady reduction in the risk of level crossings to North American highway users over the past several decades. Much less attention has been given to understanding the effect of level crossings on train safety and risk. Incidents at highway-rail level crossings can have serious consequences for the public and the railroads alike, especially if an incident results in a train derailment involving passengers or hazardous materials. The goal of past research has been to identify and understand the physical factors leading to these derailments. These physical factors point to a subset of level crossings with the highest derailment risk including train speed, highway vehicle speed, highway vehicle type, and incident type. The results of this paper can be used to provide railway decision makers with a metric for prioritizing level crossing upgrades and closures.

1.   Introduction

Considerable research has been conducted to understand the impact of highway-rail level crossings on highway users (1). Much less research has focused on understanding the risk that highway users pose to trains at highway-rail level crossings. In our previous work, we identified the factors that cause level crossing incidents to result in derailment (2). This paper discusses level-crossing-caused freight train derailments in a larger risk-analysis context. We generalize our incident-specific models to crossing-specific models by evaluating proxy variables. While this paper focuses only on freight train derailments, we have applied the same methodology to passenger train derailments and will presented the results in our future work. These results can be used to develop an additional tool for decision makers to prioritize level crossings upgrades and/or closures.

2.   Development of Derailment Model

Our results suggest that four principal factors have a significant effect on level crossing derailment probability (Table 1). Chadwick et al. (2) discussed three of them: highway vehicle speed at impact (VEHSPD), train speed at impact (TRNSPD), and highway vehicle size (LGVEH). We subsequently identified a fourth factor, incident type (TRNSTK), i.e. whether the train strikes the highway vehicle (TSV) or the vehicle strikes the train (VST).

Table 1
Description of Predictor Variables

Variable Name	Definition (*3*)	Variable Type	Range of Values
VEHSPD	Estimated highway vehicle speed	Continuous	Range: 0-105 mph (0-169 km/h) Average: 10.50 mph (16.89 km/h) Standard Deviation: 13.57 mph (21.84 km/h)
TRNSPD	Train speed	Continuous	Range: 0-80 mph (0-129 km/h) Average: 31.45 mph (50.61 km/h) Standard Deviation: 15.58 mph (25.07 km/h)
LGVEH	Was a large highway vehicle involved?	Binary	N if no; Y if yes
TRNSTK	Incident type	Binary	VST if highway user struck train; TSV if train struck highway user

Data collected in the Federal Railroad Administration (FRA) Highway Rail Grade Crossing Accident/Incident Database (4) was analyzed using SAS software and a logistic regression was performed. The data set included 213 derailments and 43,326 non-derailment events. Of those events, 110 derailments and 8,639 non-derailments occurred when a vehicle struck a train. Using two statistical models for prediction of freight derailments, one for each incident type, produced a good statistical fit. The model selection process is discussed in detail in Chadwick et al (5). The resulting models are as follows:

For VST incidents: $p_{VST} = \frac{1}{e^{-x_{VST}}+1}$

$$x_{VST} = -6.4039 + \begin{cases} 0, & LGVEH = Y \\ -1.5044, & LGVEH = N \end{cases} + 0.00101\, VEHSPD^2$$

For TSV incidents: $p_{TSV} = \frac{1}{e^{-x_{TSV}}+1}$

$$x_{TSV} = -7.1789 + \begin{cases} 0, & LGVEH = Y \\ -1.8687, & LGVEH = N \end{cases} + 0.0166\,TRNSPD$$

where VEHSPD, TRNSPD and LGVEH are as described in Table 1.

The difference in these two models reflects the differing physics of the two types of incidents. In VST incidents, the equation shows that derailment likelihood increases with increased vehicle speed (+ 0.00101 VEHSPD2), whereas in TSV incidents derailment likelihood increases with increased train speed (+ 0.0166 TRNSPD). For both types of incident, the likelihood of derailment increases if a large vehicle is involved (i.e. a truck).

3. Level-Crossing Derailment Risk Model

In its simplest form, risk can be defined as the probability of an event occurring multiplied by the consequence of that event. Train derailments caused by level crossing incidents are the result of a chain of events, each of which influences the overall probability term in the risk equation (Figure 1). Thus, development of a level-crossing derailment risk model requires understanding and quantifying the conditional probabilities associated with each of these events. The results presented in this paper are for freight trains, but a similar analysis can be performed for passenger trains.
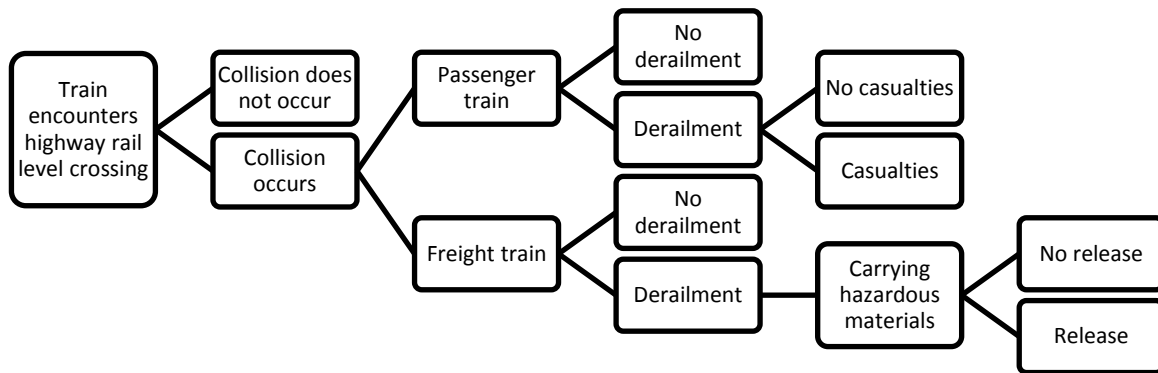

Figure 1. Simplified Risk Flowchart for Level Crossing Incidents

3.1 Prospective Risk Model Variables

The model described above is based on records from thousands of past incidents; however, development of a predictive or prospective model requires variables that can be used to estimate the probability of future events. For example, knowing that derailment likelihood increases with vehicle speed is not useful unless crossings where vehicles travel faster can be reliably identified. Therefore, we identified "proxy" variables that would be good predictors of incident circumstances to provide insight about the risk associated with different crossings. For each incident-specific variable used in the model, Table 2 indicates a crossing-specific variable (or variables) that are used as proxy variables for the different model terms.

Table 2
Definition of Model Variables

Incident-Specific Variable	Crossing-Specific Variable
Vehicle Speed	Highway Speed Limit
Train Speed	Timetable Speed

Large Vehicle Involvement	Percent Truck Traffic Annual Average Daily Traffic (AADT)
Incident Type	Average Ratio TSV:VST

The first three variables are physical or operational characteristics that can be uniquely estimated on a crossing-specific basis based on data available from government or industry. For "Incident Type", we will use a fixed ratio based on historical data to define the relative likelihood of TSV and VST incidents, instead of using a proxy variable. The development of crossing-specific statistics for this ratio is not feasible as the likelihood of an incident being TSV or VST is dependent on a number of complex factors. Our analysis indicates that over the 20-year period of this study, about 80% of incidents involved a train striking a vehicle (TSV) and 20% involved a vehicle striking a train (VST). There may be factors that influence whether a crossing is more likely to experience a TSV or VST incident, including human factors, operating characteristics, crossing geometry, sight distances and other factors; to the authors' knowledge these effects have not been quantified. A detailed review of research on human factors and drivers' behavior at level crossings was written by Yeh and Multer in 2008 (6).

3.2   Methodology

The U.S. Department of Transportation maintains two databases pertinent to this study. The Highway-Rail Grade Crossing Inventory History File (GCI) provides information about the characteristics of crossings (7) and the Federal Railroad Administration Highway Rail Grade Crossing Accident/Incident database (HRA) provides information about incidents (4). We developed a dataset based on a merge of these two databases for the 20-year period from 1991 to 2010. Of the 44,000 complete records in the HRA database, 2,221 had complete, accurate records in the level crossing inventory (GCI) database. Therefore, the analysis used only these records.

In order to evaluate the relationship between the variables identified as important in our model, we conducted a series of quantitative comparisons between the values recorded for incidents and the value of the proxy variable for specific crossings. For each record with a valid vehicle speed (VS) and highway speed limit (HSL), the difference between the two was calculated. This value, called "percent deviation from highway speed limit" (PDHSL), is defined as follows:

$$PDHSL = \frac{VS - HSL}{HSL}$$

Figure 2 shows the distribution of PDHSL values. A value of -100% means the vehicle was stopped on the crossing, a value of 0% means the vehicle was traveling at the HSL, and a value of 100% means the vehicle was traveling at twice the HSL. Less than 0.2% of all incidents occurred with the vehicle traveling more than twice the posted speed limit, so the chart was cut off at 100%.
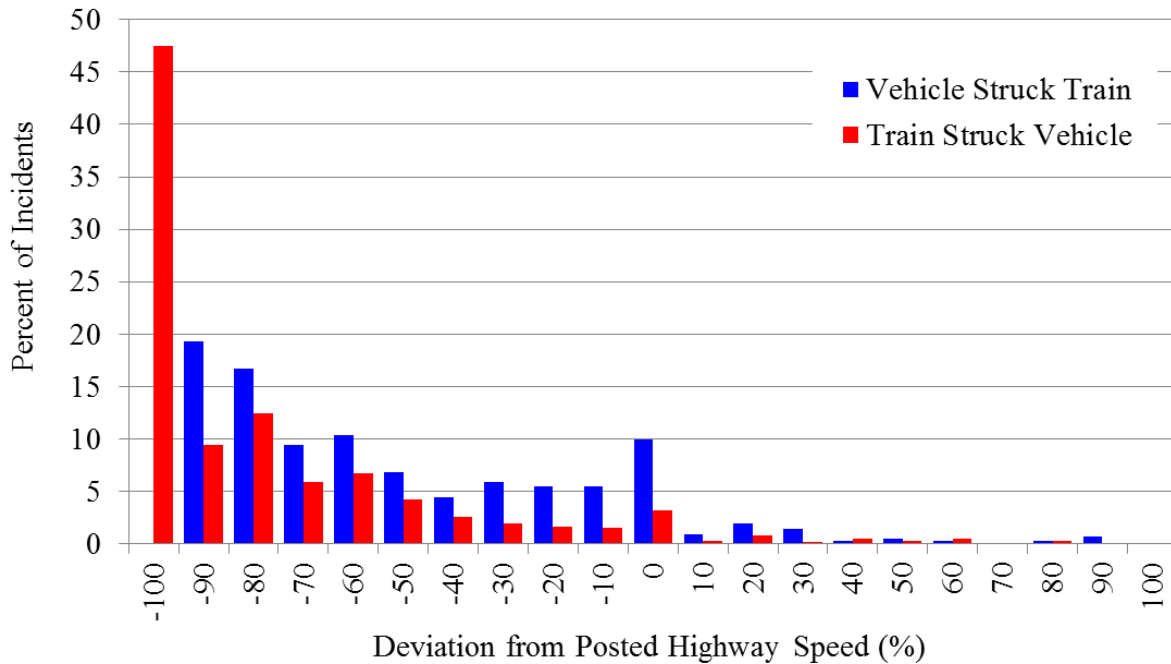
Figure 2. Distribution of Percent Deviation from Posted Highway Speed (PDHSL)

In cases where the train struck the vehicle, about 47% of incidents involved a vehicle that was stopped on a crossing. If vehicles were traveling at the speed limit, they were more likely to strike the train, than to be struck by it. Based on this analysis, the likelihood that a vehicle at a crossing will be traveling at a certain percent of the HSL when involved in an incident was calculated. This speed distribution can be used as part of the input data to calculate probability of derailment.

Similar to the analysis of highway vehicle speed, for each record with a valid train speed (TS) and timetable speed (TTS), we calculated the difference between the actual speed of the train and the timetable speed. This value, called "percent deviation from timetable speed" (PDTTS) is defined as follows:

$$PDTTS = \frac{TS - TTS}{TTS}$$

The distribution of PDTTS was compared for the two incident types, VST and TSV (Figure 3). A value of -100% means the train was stopped on the crossing, a value of 0% means the train was traveling at the TTS, and a value of 100% means the train was traveling at twice the TTS. Although a small number of trains (less than 1%) were reported to be traveling more than twice the timetable speed, these records are believed to be reporting errors. Most of these records reported a timetable speed of 0 or 1 mph, which is unlikely to reflect the actual timetable speed.
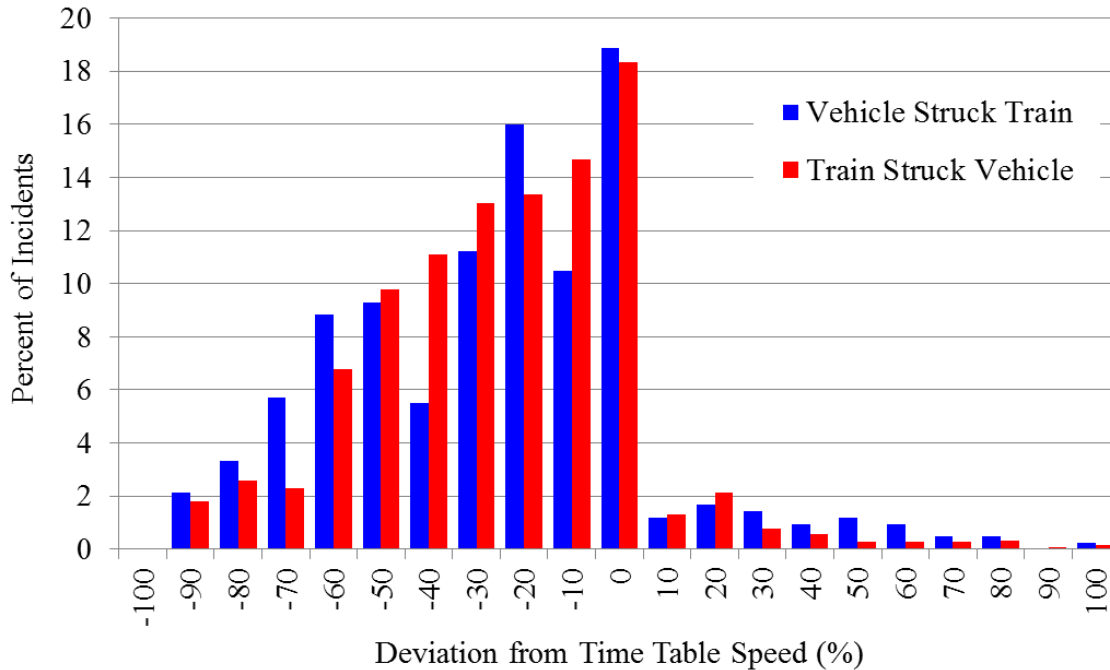
Figure 3. Distribution of Percent Deviation from Timetable Speed

Almost 20% of all trains were traveling at TTS when incidents occurred, and there were no records of trains stopped on the crossing (meaning the PDTTS would be -100%). About 8% of incidents involved a train traveling above the timetable speed. While some of these events can be confirmed as over-speed incidents by examining the narrative field, most are believed to be reporting errors. This could be determined on a case-by-case basis if actual historical timetable data were obtained from the railroad. Based on this analysis, the likelihood that a train at a crossing will be traveling at a certain percent of the TTS when involved in an incident was calculated and this speed distribution was used to calculate the probability of derailment.

The results indicate that a truck's involvement in a level crossing incident increases the likelihood of a derailment. Therefore, it is important to be able to identify crossings where incidents are more likely to involve trucks. It stands to reason that crossings with greater truck traffic are more likely to have train-truck incidents. This was tested using percent truck traffic (PTT) to quantify the number of trucks at a crossing. Figure 4 shows the probability distribution of PTT for cases with and without a large vehicle involved. A t-test comparing the two distributions showed that there is a small but statistically significant difference (p-value = 0.0113, α=0.025) in the likelihood of large vehicle involvement based on PTT at a crossing. Specifically, the likelihood that an incident will involve a truck is greater for crossings with higher PTT. To quantify this relationship for use in the derailment likelihood model, the ratio of truck incidents to car incidents was calculated for each value of PTT, and a linear regression was performed (Figure 4). If the ratio was zero or undefined, that point was omitted from analysis.
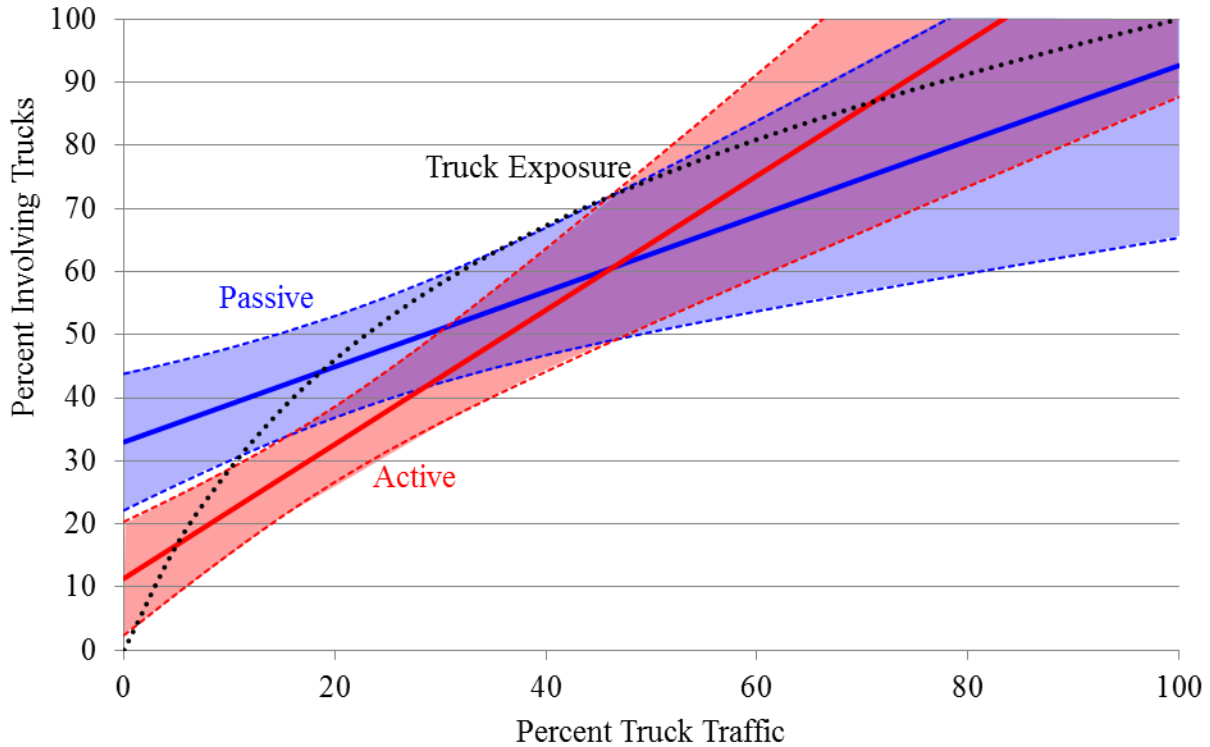
Figure 4. Distribution of Percent Truck Traffic for Incidents Involving Cars and Trucks

Next, the data were divided according to the warning device type in use at the level crossing to see if different trends would be observed for different warning device types. The following relationships were found:

Passive: $PIT_P = 33.0 + 0.597\ PTT$

Active: $PIT_A = 11.4 + 1.06\ PTT$

where *PIT* is the predicted percent of incidents involving trucks at a crossing.

For low PTT, incidents at passive crossings are more likely to involve trucks than those at active crossings. However, for roads where trucks constitute more than 50% of traffic, incidents at active crossings are more likely to involve trucks. It should be noted that the confidence bands are distinct at truck traffic levels below 20%, but begin to overlap significantly at higher levels, indicating that the linear trend does not fit as well at the high end of the range. This is likely due to the fact that most crossings in the U.S. have PTT below 20 percent. Therefore, there are many more data points at the low end than at the high end.

Another consideration is that the average truck in the U.S. is about five times longer than the average car. This means that a truck occupies a crossing for about five times as long as a car, given a constant speed. This suggests that trucks have higher exposure than cars on a per vehicle basis. A metric representing the total length of trucks using a crossing as a percent of total length of all vehicles at that crossing is expressed as:

$Truck\ Exposure = 20(PTT/100)$

$$+ \, 80 \left(\frac{AADT \times PTT \times 5}{(AADT \times PTT \times 5) + \left(AADT \times (100 - PTT)\right)} \right)$$

where AADT is the average annual daily traffic on the intersecting highway. This value increases exponentially, not linearly, which may partially explain why the regression lines lie above the diagonal. The overall positive trend shows that likelihood of truck involvement does increase as PTT increases. These regression equations are used as a proxy for the large vehicle involvement variable in the derailment model.

## 4.	Conclusion

This paper presents a derailment probability model for freight trains at level crossings based on U.S. railroad incident data from 1991-2010. The factors shown to have a statistically significant effect on derailment likelihood – train speed, highway vehicle speed, highway vehicle type, and incident type – were correlated with proxy variables such as timetable speed, posted highway speed limit, and percent truck traffic.

The usual practice of railroads and government agencies has been to allocate resources for improvements in level crossing safety based on the risk of an incident. The model described here provides a methodology that enables an additional consideration, the risk of a derailment caused by a crossing incident. This may be an important factor for both railroads and governments to consider in decisions regarding allocation of resources to reduce risk at level crossings.

References

1.	Mok, S. C. and I. Savage 2005. Why Has Safety Improved at Rail-Highway Level Crossings? In *Risk Analysis*, 25 (4): 867-881.

2.	Chadwick, S.G., M.R. Saat and C.P.L. Barkan 2012. Analysis of Factors Affecting Train Derailments at Highway-Rail Level Crossings. In *Proceedings of Transportation Research Board 91st Annual Meeting*, Washington, DC.

3.	Federal Railroad Administration (FRA) 2011. *FRA Guide for Preparing Accident/Incident Reports*, FRA, U.S. Department of Transportation, Washington, DC.

4.	Federal Railroad Administration (FRA) 2011. Download Accident Data On Demand. http://safetydata.fra.dot.gov/OfficeofSafety/publicsite/on_the_fly_download.aspx. Accessed March 2011.

5.	Chadwick, S.G., M.R. Saat, C.T. Dick and C.P.L. Barkan 2013. Decreasing Derailment Occurrence and Severity at Highway-Rail Grade Crossings. In *Proceedings of the American Railway Engineering and Maintenance-of-Way Annual Conference*, Indianapolis, IN.

6.	Yeh, M. and Multer 2008, J. Driver Behavior at Highway-Railroad Level Crossings: A Literature Review from 1990-2006. Federal Railroad Administration Office of Research and Development, Washington, DC.

7.	Federal Railroad Administration (FRA) 2012. Crossing Inventory by State. http://safetydata.fra.dot.gov/OfficeofSafety/publicsite/Query/invtab.aspx. Accessed January 2012.