# OPTIMAL STRATEGIES TO IMPROVE RAILROAD TRAIN SAFETY AND REDUCE HAZARDOUS MATERIALS TRANSPORTATION RISK 

## BY

## XIANG LIU

## DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Civil Engineering in the Graduate College of the
University of Illinois at Urbana-Champaign, 2013

Urbana, Illinois

Doctoral Committee:
Professor Christopher P. L. Barkan, Chair and Director of Research Research Assistant Professor M. Rapik Saat, Co-Director of Research
Associate Professor Yanfeng Ouyang
Associate Professor Junho Song
Todd Treichel, RSI-AAR Tank Car Safety Project


#### Abstract

Railways offer a safe and efficient means to transport hazardous materials in North America. Nonetheless, the potential risk of hazardous materials release incidents to human health, property and the environment must be appropriately managed and minimized to the extent feasible. Railroad train safety is affected by a variety of infrastructure, rolling stock and operational factors. These factors, individually and in combination, should be properly accounted for in hazardous materials transportation risk management. This dissertation presents analytical models to analyze train derailment probability and severity, evaluate hazardous materials release risk and identify optimal solutions to efficiently reduce the risk.

This research is organized into three principal parts: derailment analysis, risk analysis and decision analysis. First, statistical models are developed to estimate train derailment rates. The analysis shows the variation of train and car derailment rates under different infrastructure and traffic characteristics. Understanding derailment severity is also important. A quantile regression model is developed to analyze derailment severity distribution at various quantiles. After a hazardous materials train is derailed, the probability distribution of the number of tank cars releasing is estimated based on a series of related events using the Law of Total Probability. Based on train derailment probability and the number of tank cars releasing per derailment, the risk of a hazardous materials release incident is estimated in the second part of this dissertation. The model accounts for FRA track class, method of operation, traffic density, train length, derailment speed, number and placement of tank cars in a train, tank car safety design and population density along rail lines. In the final part, this dissertation evaluates the safety effectiveness of several risk reduction strategies, individually and in combination. An integrated


risk reduction framework is developed to determine the optimal portfolio of investments for multiple risk reduction strategies.

In summary, this dissertation presents statistical analysis, risk analysis and decision analysis models to assist decision makers to develop, evaluate, prioritize and implement potential strategies for reducing railroad hazardous materials transportation risk. The methodology and results provide insights that have been implemented or could potentially be used to develop better-informed rail safety policies and practices.

## ACKNOWLEDGEMENTS

This dissertation is dedicated to my father LIU Xiaoping, mother SU Hua, grandfather LIU Zhen and grandma MA Yunlan. Thanks for giving me the very best you ever had. More importantly, from each of you, I felt the power of honesty, integrity and responsibility since my childhood.

I am extremely grateful to my co-advisors, teachers, mentors and trustworthy friends, Dr. Christopher P.L. Barkan and Dr. Mohd Rapik Saat, for providing me with unparalleled opportunities to work on important research topics that are academically challenging and also useful for the rail industry. Your talent, vision, passion and wisdom for rail research and education motivates me to uplift my limit to achieve more.

I would also like to thank Dr. Yanfeng Ouyang, Dr. Junho Song, Mr. Todd Treichel and Dr. Xiao Qin for their collaboration, review and comment on my research work. Your professional achievement exemplifies the value of innovative thinking and restless effort to do better.

I also want to thank all my past and present colleagues in the Rail Transportation and Engineering Center (previously Railroad Engineering Program) at the University of Illinois at Urbana-Champaign. Your friendship and assistance helped completion of this dissertation. Special thanks to Dr. Rapik Saat and Dr. Athaphon Kawprasert for helping me to get started during my first two semesters at the University of Illinois.

Finally, I acknowledge the Association of American Railroads, BNSF Railway Technical Research and Development Program, NEXTRANS University Transportation Center, National University Rail (NURail) Center, USDOT Dwight David Eisenhower Transportation Fellowship Program, Monsanto and ABS Consulting for financial support during completion of my doctoral studies.

## TABLE OF CONTENTS

Chapter 1 Introduction ..... 1
1.1 Introduction ..... 1
1.2 Dissertation organization and research objectives ..... 3
1.3 Contribution summary ..... 6
1.4 Summary of each chapter ..... 7
Chapter 2 Analysis of Freight-Train Derailment Rates in the United States ..... 12
2.1 Introduction ..... 12
2.2 Data and variables ..... 14
2.3 Methodology ..... 19
2.4 Results and discussion ..... 27
2.5 Conclusion ..... 33
Chapter 3 Analysis of U.S. Freight-Train Derailment Severity Using Zero-Truncated Negative Binomial Regression and Quantile Regression ..... 39
3.1 Introduction ..... 39
3.2 Literature review ..... 40
3.3 Derailment severity ..... 41
3.4 Input data for analysis ..... 42
3.5 Zero-truncated negative binomial (ZTNB) model ..... 45
3.6 Quantile regression ..... 51
3.7 Comparison between conditional-mean and conditional-quantile models ..... 55
3.8 Conclusion ..... 56
Chapter 4 Probability Analysis of Multiple-Tank-Car Release Incidents in Railway Hazardous
Materials Transportation ..... 60
4.1 Introduction. ..... 60
4.2 Literature review ..... 61
4.3 Methodology ..... 65
4.4 Numerical example ..... 75
4.5 Conclusion ..... 84
4.6 Future research ..... 85
Chapter 5 Risk Analysis of Multiple-Tank-Car Release Accidents ..... 92
5.1 Introduction. ..... 92
5.2 Methodology ..... 94
5.3 Case study ..... 98
5.4 Flammable liquid transportation risk ..... 101
5.5 Tank car safety design improvement for risk reduction ..... 108
5.6 Discussion ..... 116
5.7 Conclusion ..... 118
Chapter 6 Integrated Risk Reduction Framework to Improve Railway Hazardous Materials Transporation Safety ..... 123
6.1 Introduction ..... 123
6.2 Multi-attribute model for hazardous materials transportation risk management ..... 124
6.3 Railroad hazardous materials transportation risk analysis model ..... 127
6.4 Cost-effectiveness of risk reduction strategies ..... 129
6.5 Pareto-optimality between risk and cost ..... 141
6.6 Optimal investment ..... 143
6.7 Conclusion ..... 145
6.8 Future research ..... 146
Chapter 7 Future Research ..... 152
7.1 Introduction. ..... 152
7.2 Future research ..... 152
Appendix A Variance in car derailment rate ..... 156
Appendix B Variable selection of zero-truncated negative binomial model ..... 158
Appendix C Truncated geometric model for derailment severity ..... 159
Appendix D Tank car derailment probability by position-in-train ..... 161
Appendix E Probability of multiple-car release incidents ..... 162
Appendix F Data for modeling broken-rail-caused car derailment rate ..... 164
Appendix G Broken-rail-caused car derailment rate by maintenance cost per ton ..... 165
Appendix H Car derailment rate reduction by increased maintenance cost ..... 166
Appendix I Derivation of optimization model ..... 167

## Chapter 1

## Introduction

### 1.1 Introduction

Railroads play an important role in safely and efficiently transporting hazardous materials (hazmat) in North America. Rail shipment of hazardous materials in the U.S. and Canada has steadily increased from approximately 1.8 million carloads in 2002 to over 2.4 million carloads in 2012 (BOE 2013). In light of growing hazardous materials traffic, managing the associated risk of a release is important to both the public and private sector (DOT 2012).

The majority of train-accident-caused hazardous materials release incidents are caused by train derailments (Barkan et al. 2003). The probability and severity of train derailments are affected by a variety of infrastructure and rolling stock characteristics. An overarching goal of the research described in this dissertation is development of a broader understanding of the most important factors affecting derailment probability, severity and consequences. Beyond this, my research attempts to understand and quantify the benefits of various train operation, infrastructure and tank car improvement strategies, both alone and in combination. The ultimate goal is development of a methodology to optimize investment in risk reduction strategies that maximizes safety improvement.

Derailment rate is a metric in train safety and risk analyses (Nayak et al. 1983; Treichel and Barkan 1993; Anderson and Barkan 2004). It has been shown to be statistically correlated with FRA track class (Nayak et al. 1983; Treichel and Barkan 1993; Anderson and Barkan
2004). These track classes include specifications for track structure, geometry, inspection frequency and method of inspection, with more stringent requirements for higher track classes. However, FRA track class does not account for all factors affecting derailment rates. Developing up-to-date freight-train derailment rates using multivariate regression techniques accounting for additional factors is the focus in chapter 2. Derailment severity, measured by number of cars derailed, was estimated as a function of derailment speed (Nayak et al. 1983). Nayak et al.'s model was improved by use of the Truncated Geometric model developed by Saccomanno and El-Hage (1989, 1991), which was further modified by Anderson (2005) and Bagheri (2009). However, previous studies have focused on estimating the mean number of cars derailed, without accounting for other important distributional statistics, such as quantiles. Chapter 3 introduces a new modeling technique called quantile regression to analyze train derailment severity distribution.

Railroad transport of hazardous materials differs from highway transport in several important ways. Notably rail transport involves trains of multiple cars, sometimes over 100 in a single train. Some or all of these may be tank cars. By contrast, highway transport generally involves a single tank trailer. Unlike highway transport, derailment of a hazardous materials train may result in releases from multiple tank cars. In chapter 4, a probabilistic approach is developed to estimate the distribution of the number of tank cars releasing per train derailment. In chapter 5, I extend the analysis in chapter 4 and develop a multiple-car release risk model. This risk model is used to evaluate the nationwide risk of rail transportation of flammable liquids.

Train safety and risk analyses are used to evaluate the cost-effectiveness of potential risk reduction strategies. Each risk reduction strategy has a direct effect on both risk and cost, and they have interactive effects as well. In chapter 6, a model is developed to optimize the integration of multiple risk reduction strategies, including broken rail prevention and tank car safety design improvement. The model is intended to provide a tool that can be adapted to optimize allocation of risk reduction resources under a wide range of scenarios.

In summary, I aim to develop analytical frameworks to inform both public and private sector in developing, evaluating, prioritizing and implementing cost-efficient policies and strategies for improving train safety and reducing railroad hazardous materials transportation risk. This dissertation develops new applications of statistical analysis, risk analysis and decision analysis in the context of railroad transportation. Finally, this dissertation has several implications regarding rail transportation safety policies and practices.

### 1.2 Dissertation organization and research objectives

This dissertation includes seven chapters (Figure 1.1).


Figure 1.1: Hierarchical organization of the analytical framework developed in the dissertation

The specific objectives of each chapter are as follows:

1. Introduction (Chapter 1)
a) Introduce the objectives and problems to address
b) Briefly discuss the essential contents of each chapter and present potential contributions to the field of railroad safety and hazardous materials transportation risk analysis
2. Derailment Rates (Chapter 2)
a) Estimate train and car derailment rates
b) Estimate the confidence intervals of freight train and car derailment rates
c) Illustrate the application of the results regarding railroad hazardous materials transportation risk management
3. Derailment Severity (Chapter 3)
a) Develop a zero-truncated negative binomial (ZTNB) model to analyze the mean of derailment severity
b) Develop a quantile regression $(\mathrm{QR})$ model to analyze the quantiles of derailment severity

## 4. Probability Analysis of Multiple Car Releases (Chapter 4)

a) Develop a probabilistic approach to estimate the distribution of the number of tank cars releasing per derailment
b) Evaluate the effect of speed reduction, tank car upgrade and changing tank car placement in trains on multiple-tank-car release incidents
5. Risk Analysis of Multiple Car Releases (Chapter 5)
a) Develop a multiple-car release risk analysis model accounting for FRA track class, method of operation, traffic density, train length, speed, number and placement of tank cars in a train, tank car safety design and population density
b) Evaluate the risk of unit train shipments of flammable liquids on U.S. Class I mainlines
6. Integrated Risk Reduction (Chapter 6)
a) Develop an optimization model to maximize risk reduction given budget
b) Develop a multi-attribute decision model to determine the optimal investment for safety improvement
7. Future Research (Chapter 7)

### 1.3 Contribution summary

The potential contributions of my dissertation are summarized below:

1. Contributions to government, policy makers and rail industry
a) Develop new methodologies to estimate train derailment probability and severity
b) Provide a risk-based approach to evaluate rail transportation safety policies
c) Provide a better-informed decisions through identification, evaluation, prioritization and implementation of risk reduction strategies
d) Support the development of Rail Corridor Risk Management System (RCRMS), a web-based software that is used by railroads to perform route-specific hazardous materials transportation risk analysis
e) Support AAR's petition to the FRA for amending PTC regulations
2. Contributions to academics and researchers
a) Develop several statistical models to advance the understanding of train derailment probability and severity
b) Develop a multiple-car release risk analysis model to evaluate the nationwide risk of rail transportation of flammable liquids
c) Develop a decision analysis model to optimally allocate railroad transportation risk reduction resources

### 1.4 Summary of each chapter

## Chapter 1: Introduction

This chapter presents the objectives of my dissertation and a description of each chapter.

## Chapter 2: Derailment Rates

Derailments are the most common type of train accident in the United States. They cause damage to infrastructure, rolling stock and lading, disrupt service, and have the potential to cause casualties, and harm the environment. Train safety and risk analysis relies on accurate assessment of derailment likelihood. Derailment rate, the number of derailments normalized by traffic exposure, is a variable to represent the likelihood of a derailment. Despite its importance, up-to-date derailment rate analysis using multiple factors has not been previously developed. In this chapter, I present an analysis of the latest derailment rates on Class I railroad mainlines based on data from the U.S. Federal Railroad Administration and the major freight railroads. The point estimator and confidence interval of train and car derailment rates are developed by FRA track class, method of operation and annual traffic density. The analysis shows that signaled track with higher FRA track class and higher traffic density is associated with a lower derailment rate. The new accident rates have important implications for safety and risk management decisions, such as the routing of hazardous materials. A numerical example shows that using the new multiple-factor-based accident rates leads to a more accurate routing decision, compared to use of the earlier single-factor-based accident rates.

## Chapter 3: Derailment Severity

Despite the low probability of a train derailment, it could potentially cause severe consequences. In this chapter, a zero-truncated negative binomial (ZTNB) regression model is developed to estimate the conditional mean of train derailment severity (measured by number of cars derailed). Recognizing that the mean is not the only statistic describing derailment severity distribution, a quantile regression $(\mathrm{QR})$ model is also developed to estimate derailment severity at different quantiles. The two regression models together provide a better understanding of train derailment severity distribution. Results of this work can be used to estimate train derailment severity under various operational conditions and by different accident causes.

## Chapter 4: Probability of Multiple-Car Releases

Hazardous materials trains are fundamentally different from highway transport because they routinely transport multiple tank cars, sometimes over 100 in a single train. Therefore, train derailments may cause releases from multiple tank cars, thereby requiring a different quantitative approach to assess the likelihood of events of various magnitudes. This has become more important for government and the industry due to the rapid growth in hazardous materials traffic, especially in unit trains, and growing regulatory and public attention to the safety and risk implications. In this chapter, a probabilistic model is developed to estimate the distribution of the number of tank cars derailing and releasing. Principal factors considered include train length, speed, position of the first car derailed, number and placement of tank cars in the train and tank car safety design. The results show that reducing train speed, enhancing tank car safety design and changing tank car placement all affect the number of tank cars releasing per derailment. This research provides insights regarding the circumstances likely to result in multiple-tank-car
release incidents and the strategies to reduce their occurrence. The model can be incorporated into a risk management framework to enable local, regional and national safety management of hazardous materials transportation by rail.

## Chapter 5: Risk Analysis of Multiple-Tank-Car Release Accidents

Over the past decade, alcohol and petroleum crude oil have dramatically increased. Much of this traffic is transported in large blocks or unit trains. The increase in this traffic has raised questions regarding transportation safety, especially after the Lac-Mégantic accident in July 2013. In this chapter, I develop a risk analysis model to evaluate the nationwide risk of rail transportation of flammable liquids and quantify the safety effectiveness of tank car safety design improvement. The probability distribution of the number of tank cars releasing is estimated based on a series of probability distributions accounting for a variety of infrastructure and rolling-stock-related factors. The consequence of a release incident is measured by the affected population, which is analyzed using a geographic information system (GIS) platform. The analysis shows that tank car safety design improvement has a more substantial effect on prevention of large, multiple-car release incidents than a small number of tank car releases. Use of more robust tank car safety designs can have a strong effect on prevention of large, catastrophic release incidents.

## Chapter 6: Integrated Risk Reduction Strategies

A number of strategies have been implemented or are being developed to reduce the risk of hazardous materials releases from train accidents. Each of these risk reduction strategies has corresponding safety benefit and implementation costs. However, the cost effectiveness of the integration of different risk reduction strategies is not well understood. This chapter presents an
optimization model that considers the combination of two risk reduction strategies, broken rail prevention and tank car safety design enhancement. The Pareto-optimality technique is used to maximize risk reduction at a given level of investment. The framework can be adapted to address a broader set of risk reduction strategies.

Chapter 7: Future Research
This chapter presents potential future research topics based on the findings of my dissertation.

## References

Anderson, R.T., Barkan, C.P.L. (2004). Railroad accident rates for use in transportation risk analysis. Transportation Research Record 1863, 88-98.

Anderson, R.T. (2005). Quantitative analysis of factors affecting railroad accident probability and severity. M.S. Thesis, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Bagheri, M. (2009). Risk-based model for effective marshalling of dangerous goods railway cars. Ph.D. Dissertation, Department of Civil Engineering, University of Waterloo, Ontario.

Barkan, C.P.L., Dick, C.T., Anderson, R.T. (2003). Analysis of railroad derailment factors affecting hazardous materials transportation risk. Transportation Research Record 1825, 64-74.

Bureau of Explosives (BOE) (2013). Annual report of hazardous materials transported by rail: calendar year 2012. Bureau of Explosives, Transportation Technology Center, Inc., Pueblo, CO.

Department of Transportation (DOT) (2012). Transportation for a new generation strategic plan for fiscal years 2012-2016. U.S. Department of Transportation, Washington D.C.

Nayak, P.R., Rosenfield, D.B., Hagopian, J.H. (1983). Event probabilities and impact zones for hazardous materials accidents on railroads. Report DOT/FRA/ORD-83/20, FRA, U.S.
Department of Transportation, Washington D.C.
Saccomanno, F.F., El-Hage, S.M. (1989). Minimizing derailments of railcars carrying dangerous commodities through effective marshaling strategies. Transportation Research Record 1245, 3451.

Saccomanno, F.F., El-Hage, S.M. (1991). Establishing derailment profile by position for corridor shipments of dangerous goods. Canadian Journal of Civil Engineering 18(1), 67-75.

Treichel, T.T., Barkan, C.P.L. (1993). Working paper on mainline freight train accident rates. Research and Test Department, Association of American Railroads, Washington D.C.

Treichel, T.T., Hughes, J.P., Barkan, C.P.L., Sims, R.D., Phillips, E.A., Saat, M.R. (2006). Safety performance of tank cars in accidents: probability of lading loss, report RA-05-02. RSIAAR railroad tank car safety research and test project, Association of American Railroads, Washington, D.C.

## Chapter 2

# Analysis of Freight-Train Derailment Rates in the United States 

Adapted from

Liu, X., Saat, M.R., Barkan, C.P.L. Analysis of freight-train derailment rates in the United States. Accident Analysis and Prevention (under review).

### 2.1 Introduction

Derailments are the most common type of train accident in the United States. They cause damage to infrastructure, rolling stock and lading, disrupt service, and have the potential to cause casualties and harm the environment. Understanding the most important factors affecting derailments is critical to development of effective risk reduction strategies. Train safety and risk analysis relies on accurate estimation of derailment rate, which is defined as the number of derailments normalized by some metric of traffic exposure, such as train-miles, car-miles or gross ton-miles (Nayak et al. 1983; Treichel and Barkan 1993; Anderson and Barkan 2004; Liu et al. 2011).

Highway safety researchers have conducted a number of studies quantifying the relationship between accident rates and roadway design. These studies have considered the effects of road curvature, traffic volume, grade, shoulder width, number of lanes and other factors (e.g., Miaou 1994; Maher and Summersgill 1996; Hauer 2001; Lord et al. 2005; Lord 2006; Mitra and Washington 2007). The earliest example of a comprehensive analogous study of railroad accident rates in the United States was conducted by Nayak et al. in the 1980s (Nayak et al. 1983). Using analyses of accident frequency and rail traffic volume, they found a strong statistical correlation between FRA track class and derailment rate. A subsequent unpublished study by Treichel and Barkan (1993) found a similar result. Anderson and Barkan (2004) used
new data to develop updated estimates. All of these studies found that higher FRA track classes had lower derailment rates, varying by more than an order of magnitude. This relationship was not surprising; higher FRA track classes are intended to ensure safe operation at higher operating speeds and therefore require a variety of more stringent engineering safety and maintenance standards.

These derailment rate estimates have been used by railroads, government agencies and researchers to address a variety of risk analysis and management questions (CCPS 1995; ADL 1996; STB 2003; Kawprasert and Barkan 2008, 2010). As the importance and sophistication of these questions has grown, so too has the importance of their accuracy. This led to closer scrutiny of the factors affecting the relationship between FRA track class and derailment rate. Developing a better understanding of the underlying factors affecting the relationship is important if the information is to be used to improve railroad risk management practices.

Over the three decades since Nayak et al (1983) completed their work, the U.S. mainline train derailment rate has declined ten-fold (Barkan et al. 2013). This affects the absolute rate on different track classes, and might also affect the relative rate between track classes, because the relative importance of accident causes correlated with different track classes may have shifted (Anderson and Barkan 2004). There is ongoing interest in improving rail safety and new concerns have been raised regarding the risk of rail transport of hazardous materials due to several fatal release accidents in the mid-2000s. This prompted renewed interest in a more detailed understanding of the factors affecting derailment rate (Barkan et al. 2013). In the same time frame, post-9/11 security concerns led the U.S. Department of Transportation to promulgate
new regulations that required railroads to conduct, "transportation route analysis, alternative route analysis, and route selection" for toxic inhalation hazard materials (DOT 2008). This led to new consideration of how to calculate derailment rate and whether it provided a sufficiently detailed means of assessing localized risk. Research in this chapter suggested that other factors not previously considered might be affecting it as well, notably method of operation (i.e., traffic control system) and traffic density.

Previous constraints on data systems and availability had limited the ability to consider more fine-grained questions regarding factors that might co-vary with track class and affect derailment rate. Furthermore, these analyses had used relatively simple statistical techniques that were not capable of detecting the complex relationship between derailment rates and multiple influencing factors. To address these questions, a new dataset was developed that contained information on FRA track class, method of operation and traffic density.

### 2.2 Data and variables

### 2.2.1 Derailment data

Data for the derailment rate analysis were obtained for the major freight railroads operating in the U.S. for the years 2005 to 2009 . These railroads account for approximately $69 \%$ of route miles and $88 \%$ of carloads transported on U.S. railroads (AAR 2013).

Data on the number and cause of derailments came from the U.S. Federal Railroad Administration (FRA) Rail Equipment Accident (REA) database. This database records all accidents that exceed a specified monetary damage cost to on-track equipment, signals, track,
track structures, or roadbed (FRA 2011). Each accident record includes information on approximately 50 different variables detailing the circumstances of the accident. Among these are the method of operation and the annual traffic density measured in annual gross tonnage at the accident location. However, having traffic density data only for FRA-reportable accident locations is insufficient for proper estimation of derailment rates because it does not permit understanding of the entire network under consideration. In particular, comprehensive data on the exposure of rail traffic to different combinations of infrastructure and operating conditions are needed to develop accurate estimates of accident rates. Therefore, each railroad provided additional data for their entire, mainline network. In total, there were 1,420 freight-train derailments and 17.5 trillion gross ton-miles of traffic reported for the period covered in this analysis. Derailment and traffic data were categorized into a three-dimensional matrix defined by the explanatory variables: FRA track class, method of operation and annual traffic density. A brief description of each variable is presented in the following subsections.

### 2.2.2 Explanatory variables

## FRA track class

The FRA specifies track quality standards or "track classes" required to operate freight and passenger trains at different maximum speeds (FRA 2003). There are five principal track classes commonly used by U.S. freight railroads, ranging from track class 1 with the lowest maximum speed ( 10 mph ), to class 5 with the highest ( 80 mph ). These classes include specifications for track structure, geometry, inspection frequency and method of inspection, with more stringent requirements for higher track classes. The FRA standards represent minimum requirements; in fact, railroads often maintain various sections of their infrastructure to standards that exceed the
minimum required by the FRA. This introduces additional variance in analyses of the relationship between track quality and derailment rates within the same track class (El-Sibaie and Zhang 2004).

## Method of operation

The FRA records multiple different values for method of operation; however, for the purposes of this analysis, I am interested in a somewhat higher level, specifically, whether the track has a system of automatic signaling in place or not. Approximately 60 percent of U.S. mileage and 80 percent of rail traffic operates on signaled trackage (FRA 2008). Signaled tracks use electrical track circuits to detect the presence of trains. However, an important secondary benefit of track circuits is that they also enable detection of several types of infrastructure problems including broken rails, which are the leading cause of train derailments on U.S. railroads (Dick et al. 2003; Barkan et al. 2003; Liu et al. 2012).

## Traffic density

Traffic density is the third variable included in the model. Track with a higher traffic density receives more frequent track maintenance leading to higher track quality (FRA 2003; Peng 2011). Railroad traffic density represents the total weight of all locomotives, rolling stock and lading traversing a given section of track and is commonly measured in million gross tons (MGT). The traffic density variable was assigned two values, $<20$ MGT annual traffic and $\geqslant 20$ MGT. The demarcation at 20 MGT was selected because it represents the average annual track traffic density on U.S. Class I railroad mainlines (AAR 2005-2009) so the two classifications indicate, below average traffic density and above average, respectively.

As mentioned above, FRA track class is determined by speed of operation on a given segment of track. In circumstances where there are civil speed restrictions due to curves or other features, these segments will generally be classified as lower FRA track classes. However, these lower-speed sections on higher-traffic-density routes are generally designed and maintained to the same high standards as adjacent sections of track with higher speeds and track classes, commensurate with the higher volume of traffic using them.

Each segment of track in this study could be classified by various combinations of FRA track class, method of operation and traffic density allowing us to investigate whether there was a statistical relationship with derailment rate. Thus, a $5 \times 2 \times 2$ matrix was constructed using the three categorical explanatory variables:

- track classes: $1,2,3,4,5$;
- method of operation: signaled and non-signaled;
- annual traffic density: <20 MGT and $\geqslant 20$ MGT.

Table 2.1 presents the distribution of freight-train derailment and traffic data by the predictor factors used in this study. In the next section, I describe a negative binomial regression model that was developed to analyze mainline freight-train derailment rate.

Table 2.1
Distribution of freight-train derailment and traffic data by predictor variables
(a) Freight-train derailment distribution

| Annual Traffic Density (MGT) | Method of Operation (MO) | FRA Track Class (TC) |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 1 | 2 | 3 | 4 | 5 | TC Total |
| <20 | Non-Signaled | 2.0\% | 3.5\% | 4.4\% | 3.7\% | *n/a | 13.7\% |
|  | Signaled | 1.3\% | 2.5\% | 3.3\% | 4.6\% | 0.4\% | 12.2\% |
|  | MO Total | 3.4\% | 6.1\% | 7.7\% | 8.3\% | 0.4\% | 25.8\% |
| $\geq 20$ | Non-Signaled | 0.7\% | 1.8\% | 2.0\% | 6.0\% | 0.5\% | 11.0\% |
|  | Signaled | 2.6\% | 6.7\% | 11.3\% | 31.0\% | 11.6\% | 63.2\% |
|  | MO Total | 3.3\% | 8.5\% | 13.2\% | 37.0\% | 12.1\% | 74.2\% |
| Total | Non-Signaled | 2.7\% | 5.4\% | 6.3\% | 9.7\% | 0.5\% | 24.6\% |
|  | Signaled | 3.9\% | 9.2\% | 14.6\% | 35.6\% | 12.0\% | 75.4\% |
|  | MO Total | 6.7\% | 14.6\% | 20.9\% | 45.3\% | 12.5\% | 100.0\% |

*There were no instances of non-signaled, Class 5 track with less than 20 MGT of annual traffic
(b) Freight traffic

| Annual Traffic <br> Density <br> (MGT) | Method of Operation (MO) | FRA Track Class (TC) |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 1 | 2 | 3 | 4 | 5 | TC Total |
| <20 | Non-Signaled | 0.1\% | 0.5\% | 0.9\% | 1.6\% | *n/a | 3.2\% |
|  | Signaled | 0.1\% | 0.3\% | 1.2\% | 3.3\% | 0.3\% | 5.2\% |
|  | MO Total | 0.2\% | 0.9\% | 2.1\% | 4.9\% | 0.3\% | 8.4\% |
| $\geq 20$ | Non-Signaled | 0.2\% | 0.4\% | 0.8\% | 2.1\% | 0.2\% | 3.7\% |
|  | Signaled | 0.5\% | 2.0\% | 8.2\% | 47.8\% | 29.4\% | 88.0\% |
|  | MO Total | 0.7\% | 2.4\% | 9.0\% | 49.9\% | 29.7\% | 91.6\% |
| Total | Non-Signaled | 0.3\% | 0.9\% | 1.7\% | 3.7\% | 0.2\% | 6.9\% |
|  | Signaled | 0.5\% | 2.4\% | 9.4\% | 51.1\% | 29.7\% | 93.1\% |
|  | MO Total | 0.8\% | 3.3\% | 11.1\% | 54.7\% | 30.0\% | 100.0\% |

* There were no instances of non-signaled, Class 5 track with less than 20 MGT of annual traffic


### 2.3 Methodology

### 2.3.1 Train derailment rate analysis

The negative binomial regression model has been widely used in accident rate analysis in highway transportation (e.g., Miaou 1994; Hauer 2001; Wood 2002; Lord et al. 2005; Lord 2006;

Oh et al. 2006; Mitra and Washington 2007); however, no previously published research has attempted to apply this analytical approach to multivariate estimation of railroad accident rate.

Equations (2.1) to (2.3) present the basic framework for a negative binomial model:
$Y \sim \operatorname{Poisson}(\lambda)$
$\lambda \sim \operatorname{Gamma}\left(f, \frac{f}{m}\right)$
$m=\exp \left(\sum_{p=0}^{k} b_{P} X_{P}\right) M$

Where:
Y = observed number of derailments
m = estimated number of derailments
$\mathrm{b}_{\mathrm{p}} \quad=\mathrm{p}^{\text {th }}$ parameter coefficient
$\mathrm{X}_{\mathrm{p}} \quad=\mathrm{p}^{\text {th }}$ explanatory variable
$\mathrm{M} \quad=$ traffic exposure (gross ton-miles)
f $\quad=$ inverse dispersion parameter

The confidence intervals of estimated derailment rates using the Poisson regression or negative binomial regression models were developed by Wood (2005) (Table 2.2).

Table 2.2
$95 \%$ confidence interval for accident models (Wood 2005)
Poisson
m

$$
\exp \left[X^{T} b^{*} \pm 1.96 \sqrt{\operatorname{Var}\left(h^{*}\right)}\right]
$$

y

$$
\left[0,\left\lfloor m^{*}+\sqrt{19} \sqrt{\left(m^{*}\right)^{2} \operatorname{Var}\left(h^{*}\right)+m^{*}}\right]\right]
$$

Negative binomial
m

$$
\exp \left[X^{T} b^{*} \pm 1.96 \sqrt{\operatorname{Var}\left(h^{*}\right)}\right]
$$

$\lambda \max \left\{0, m^{*}-1.96 \sqrt{\left(m^{*}\right)^{2} \operatorname{Var}\left(h^{*}\right)+\frac{\left(m^{*}\right)^{2} \operatorname{Var}\left(h^{*}\right)+\left(m^{*}\right)^{2}}{f^{*}}}\right\}$,

$$
m^{*}+1.96 \sqrt{\left(m^{*}\right)^{2} \operatorname{Var}\left(h^{*}\right)+\frac{\left(m^{*}\right)^{2} \operatorname{Var}\left(h^{*}\right)+\left(m^{*}\right)^{2}}{f^{*}}}
$$

y

$$
\left[0,\left\lfloor m^{*}+\sqrt{19} \sqrt{\left(m^{*}\right)^{2} \operatorname{Var}\left(h^{*}\right)+\frac{\left(m^{*}\right)^{2} \operatorname{Var}\left(h^{*}\right)+\left(m^{*}\right)^{2}}{f^{*}}+m^{*}}\right]\right]
$$

Note: $h^{*}=b_{0}^{*}+b_{1}^{*} X_{1}+\ldots+b_{k}^{*} X_{k}+\log (M)$ (the parameter with * represents an estimator)

### 2.3.2 Train derailment rate

$Z=\exp \left(b_{0}+b_{t r k} X_{t r k}+b_{\text {moo }} X_{\text {moo }}+b_{\text {den }} X_{d e n}\right)$
Where:
Z = estimated derailment rate per gross ton-miles
$X_{\text {trk }}=F R A$ track class (1 to 5 )
$\mathrm{X}_{\text {moo }}=$ method of operation ( 1 for signaled, 0 for non-signaled )
$\mathrm{X}_{\text {den }}=$ annual traffic density level ( 1 for $\geqslant 20 \mathrm{MGT}, 0$ for $<20 \mathrm{MGT}$ )
b = parameter coefficients

The estimated parameter coefficients were developed using the maximum likelihood method (Agresti 2007). All three variables were found to significantly affect freight-train derailment rates (Table 2.3). Although FRA track class is a categorical variable, the preliminary data analysis suggested that there was an inverse linear relationship between estimated derailment rate and FRA track class, given the other two predictor variables, therefore I treated it as a continuous variable in this analysis. A similar treatment of FRA track class was described by English et al. (2007).

Table 2.3
Parameter coefficient estimates of freight-train derailment rates on Class I mainlines, 2005-2009

| Estimate | Standard <br> Error | Wald 95\% <br> Confidence <br> Limits |  |  |  | Wald Chi <br> Square |
| :--- | :---: | :---: | :---: | :---: | ---: | ---: |
| Parameter | Pr $>$ <br> ChiSq |  |  |  |  |  |
| $\mathrm{b}_{0}$ (Intercept) | 0.9201 | 0.1115 | 0.7016 | 1.1386 | 68.11 | $<.0001$ |
| $\mathrm{~b}_{\text {trk }}$ (Track Class) | -0.6649 | 0.0341 | -0.7318 | -0.5981 | 380.37 | $<.0001$ |
| $\mathrm{~b}_{\text {moo }}$ (Method of Operation) | -0.3377 | 0.0974 | -0.5286 | -0.1469 | 12.03 | 0.0005 |
| $\mathrm{~b}_{\text {den }}$ (Traffic Density) | -0.7524 | 0.0859 | -0.9208 | -0.5840 | 76.72 | $<.0001$ |
| Dispersion | 0.0048 | 0.0062 |  |  |  |  |

A special case of the negative binomial model is the Poisson model with a dispersion parameter of zero (Hilbe 2007). To test whether this was appropriate for our data, I calculated the Wald z-score by dividing the estimated dispersion parameter by its standard error. The calculated z-score was 0.77 ( $0.0048 / 0.0062$ ), which fails to reject the hypothesis of a zero dispersion parameter $(\mathrm{p}=0.44)$. This indicated that there is no significant difference between the Poisson model and negative binomial model in fitting the data. Thus I used the confidence intervals for
the Poisson model (Table 2.2) to estimate the $95 \%$ confidence intervals for train derailment rates (Figure 2.1 and Table 2.4). It is evident that all three variables are having a substantial effect:

1) The higher the FRA track class, the lower the train derailment rate
2) Signaled track has a lower derailment rate than non-signaled track
3) Track with higher traffic density has a lower derailment rate

The observation that a higher track class is associated with a lower train derailment rate is consistent with previous studies (Nayak et al. 1983; Treichel and Barkan 1993; Anderson and Barkan 2004). However, in addition to FRA track class, method of operation and traffic density both had a strong, significant relationship with derailment rate. As mentioned above, signaled track segments have track circuits to detect broken rails, thereby potentially reducing the likelihood of derailments due to this cause. Furthermore, given the same track class and method of operation, derailment rate is inversely correlated with traffic density level. There are several possible explanations for this. As mentioned above rail lines with higher traffic density receive more frequent track inspection and maintenance (Zarembski and Palese 2010; Sawadisavi 2010; Peng 2011) irrespective of speed (i.e., FRA track class). Busier lines may also have a greater number and variety of wayside defect detectors installed (Schlake et al. 2010) thereby reducing the incidence of certain infrastructure and equipment-caused train accidents.

Table 2.4
Estimated Class I mainline freight-train derailment rate per billion gross ton-miles, 2005-2009 (the numbers in the parenthesis represent $95 \%$ confidence intervals)

| Annual <br> Traffic Density (MGT) | Method of Operation | FRA Track Class |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Class 1 | Class 2 | Class 3 | Class 4 | Class 5 |
| <20 | Non-Signaled | $\begin{gathered} \hline 1.29 \\ (1.086,1.534) \end{gathered}$ | $\begin{gathered} 0.66 \\ (0.574,0.768) \end{gathered}$ | $\begin{gathered} 0.34 \\ (0.295,0.395) \end{gathered}$ | $\begin{gathered} 0.18 \\ (0.148,0.209) \end{gathered}$ | *n/a |
|  | Signaled | $\begin{gathered} 0.92 \\ (0.737,1.151) \end{gathered}$ | $\begin{gathered} 0.47 \\ (0.395,0.568) \end{gathered}$ | $\begin{gathered} 0.24 \\ (0.208,0.286) \end{gathered}$ | $\begin{gathered} 0.13 \\ (0.106,0.147) \end{gathered}$ | $\begin{gathered} 0.06 \\ (0.053,0.078) \end{gathered}$ |
| $\geq 20$ | Non-Signaled | $\begin{gathered} 0.61 \\ (0.495,0.747) \end{gathered}$ | $\begin{gathered} 0.31 \\ (0.260,0.376) \end{gathered}$ | $\begin{gathered} 0.16 \\ (0.134,0.194) \end{gathered}$ | $\begin{gathered} 0.08 \\ (0.067,0.102) \end{gathered}$ | $\begin{gathered} 0.04 \\ (0.033,0.055) \end{gathered}$ |
|  | Signaled | $\begin{gathered} 0.43 \\ (0.361,0.521) \end{gathered}$ | $\begin{gathered} 0.22 \\ (0.195,0.255) \end{gathered}$ | $\begin{gathered} 0.11 \\ (0.104,0.127) \end{gathered}$ | $\begin{gathered} 0.06 \\ (0.053,0.066) \end{gathered}$ | $\begin{gathered} 0.03 \\ (0.026,0.035) \end{gathered}$ |

*There were no instances of non-signaled, Class 5 track with less than 20 MGT of annual traffic


Figure 2.1: Estimated Class I mainline freight-train derailment rates by FRA track class, method of operation and annual traffic density (error bars represent 95\% confidence intervals)

### 2.3.3 Car derailment rate

### 2.3.3.1 Estimation of car derailment rate

In addition to train derailment rate, car derailment rate is also of interest. This is generally measured as the number of cars derailed per unit of traffic exposure and represents the likelihood that an individual car is involved in a derailment. Anderson and Barkan (2004) estimated average car derailment rate by multiplying train derailment rate by the average number of cars derailed per derailment:
$C^{*}=\frac{m^{*} \times D^{*}}{M}$

Where:
$C^{*} \quad=$ estimated car derailment rate per unit of traffic exposure
$m^{*} \quad=$ estimated train derailment count given traffic exposure
$D^{*} \quad=$ average number of cars derailed per train derailment
M = traffic exposure

### 2.3.3.2 Derailment severity

The average number of cars derailed in derailment is a metric of its severity (Nayak et al. 1983; Saccomanno and El-Hage 1989; Barkan et al. 2003; Anderson and Barkan 2004; Liu et al. 2013a) and can be calculated using data from the same FRA REA database used to address other questions in this chapter. I conducted an analysis of variance to determine if there was a relationship between derailment severity and the three explanatory variables being considered. I found no significant relationship with method of operation ( F statistic $=1.21$, degrees of freedom $=1, \mathrm{P}=0.27$ ) or traffic density $(\mathrm{F}$ statistic $=0.57$, degrees of freedom $=1, \mathrm{P}=0.6)$. However, I did find a significant relationship between FRA track class and derailment severity $(\mathrm{F}$ statistic $=$
4.78, degrees of freedom $=4, \mathrm{P}<0.01$ ). The higher the FRA track class, the greater the average number of cars derailed per derailment.

The lack of a relationship between severity and the first two variables is not surprising because, ceteris paribus, neither would be expected to affect the kinetic energy of a derailment and consequently its severity, whereas the third one does. Track class is directly related to maximum allowable operating speed and previous research has shown a relationship between severity and derailment speed, due in part to the greater kinetic energy (Nayak et al. 1983; Barkan et al. 2003; Anderson and Barkan 2004) and with FRA track class (Liu et al. 2011). Thus, track-class-specific number of cars derailed was estimated and used to estimate car derailment rate.

### 2.3.3.3 Variance in estimated car derailment rate

I calculated the variance in the car derailment rate estimate by accounting for the uncertainty in train derailment rate and number of cars derailed. A general approach to estimate the variance of multiple random variables was originally developed in the 1960s (Goodman 1962). Since then, it has been used in various studies of highway safety (Lord 2008; Geedipally and Lord 2010) but I am unaware of its application to estimation of car derailment rate. Train derailment frequency and severity were assumed to be independent variables, and the variance in estimated car derailment rate, denoted by $\operatorname{Var}\left(\mathrm{C}^{*}\right)$, was estimated using the following equation (Table 2.5) (derived in Appendix A):
$\operatorname{Var}\left(C^{*}\right)=\frac{\left[E\left(m^{*}\right)\right]^{2} \operatorname{Var}\left(D^{*}\right)+\left[E\left(D^{*}\right)\right]^{2} \operatorname{Var}\left(m^{*}\right)+\operatorname{Var}\left(m^{*}\right) \operatorname{Var}\left(D^{*}\right)}{M^{2}}$
where:

| $\operatorname{Var}\left(\mathrm{C}^{*}\right)$ | $=$ variance of estimated car derailment rate |
| :--- | :--- |
| $\mathrm{E}\left(\mathrm{m}^{*}\right)$ | $=$ expected value of train derailment count |
| $\operatorname{Var}\left(\mathrm{m}^{*}\right)$ | $=$ variance of estimated train derailment count |
| $\mathrm{E}\left(\mathrm{D}^{*}\right)$ | $=$ expected value of number of cars derailed per derailment |
| $\operatorname{Var}\left(\mathrm{D}^{*}\right)$ | $=$ variance of estimated number of cars derailed per derailment |
| $M$ | $=$ traffic exposure |

## Table 2.5

Estimated car derailment rate per billion car-miles, Class I freight-train mainline derailments, 2005-2009 (Italic numbers in the parentheses represent the standard errors of estimated car derailment rates)

| Annual Traffic Density (MGT) | Method of Operation | FRA Track Class |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 1 | 2 | 3 | 4 | 5 |
| <20 | Non-Signaled | $\begin{aligned} & \hline 632 \\ & \text { (66) } \end{aligned}$ | $\begin{aligned} & 446 \\ & \text { (55) } \end{aligned}$ | $\begin{aligned} & \hline 266 \\ & \text { (35) } \end{aligned}$ | $\begin{aligned} & 150 \\ & (23) \end{aligned}$ | *n/a |
|  | Signaled | $\begin{array}{r} 451 \\ \text { (57) } \\ \hline \end{array}$ | $\begin{array}{r} 318 \\ (43) \\ \hline \end{array}$ | $\begin{aligned} & 189 \\ & (25) \\ & \hline \end{aligned}$ | $\begin{aligned} & 107 \\ & (16) \\ & \hline \end{aligned}$ | $\begin{array}{r} 59 \\ (10) \\ \hline \end{array}$ |
| $\geq 20$ | Non-Signaled | $\begin{aligned} & 298 \\ & \text { (35) } \end{aligned}$ | $\begin{aligned} & 210 \\ & (29) \end{aligned}$ | $\begin{aligned} & 125 \\ & (18) \end{aligned}$ | $\begin{array}{r} 71 \\ (12) \end{array}$ | 39 $(7)$ |
|  | Signaled | $\begin{aligned} & 213 \\ & \text { (23) } \end{aligned}$ | $\begin{aligned} & 150 \\ & (18) \end{aligned}$ | $\begin{array}{r} 89 \\ \text { (11) } \end{array}$ | $\begin{aligned} & 50 \\ & (7) \end{aligned}$ | 28 (4) |

*There were no instances of non-signaled, Class 5 track with less than 20 MGT of annual traffic

### 2.4 Results and discussion

### 2.4.1 Summary of results

Multivariate statistical analyses of American train derailment data, combined with information on FRA track class, method of operation and traffic density, showed that each of these variables had a strong, significant effect on derailment rate. Previous studies had only found an effect of track class, but did not consider the other two variables. I also found that average derailment severity was unaffected by method of operation or traffic density, but was strongly related to FRA track class, consistent with several previous studies (Nayak et al.1983; Barkan et al. 2003; Anderson and Barkan 2004). Despite the larger average number of cars derailed in accidents on higher track classes, car derailment rate is still lower. This is because the reduction of train derailment rate more than offsets the increase in derailment severity.

### 2.4.2 Discussion

Accurate calculation of train accident rate has important implications for a number of railroad industry safety policy, operating practice, risk management and resource allocation decisions. It is also an important aspect of federal regulatory development, review policies and decision making. The first attempt to develop nationwide, track-class-specific accident rates was conducted by Nayak et al. (1983) in a study for the U.S. DOT Federal Railroad Administration. Railroad train safety had been deteriorating in the years prior to economic deregulation of the US rail industry in 1980 and there was interest in understanding the effect of various potential contributing factors. Meanwhile, in 1975 the FRA had implemented new train accident data recording requirements and Nayak et al. used these data, along with data from other sources, to try and understand the quantitative relationship between track class and derailment rate.

At the same time, there was increasing interest in the risk associated with rail transport of hazardous materials such as toxic-inhalation-hazard materials and flammable gases (Andrews 1980; Geffen 1980). In the absence of more specific data, these studies relied on an average railroad derailment rate. Although such an approach may enable nationwide estimates of average risk, most rail transport risk management decisions require greater precision, for example, understanding localized differences in risk due to differing track quality or development of risk profiles for a route or region. Nayak et al. recognized that both national, and geographically specific estimates of train safety and derailment risk required a more fine-grained understanding of the key factors affecting risk. Since that time, both the private and public sectors have made extensive use of the FRA database, the Nayak et al. statistics, and subsequent revisions and refinements of their analyses using track-class-specific derailment rates to conduct safety and risk assessments at both the local and national level (Glickman and Rosenfield 1984; CCPS 1995; STB 2003; Kawprasert and Barkan 2010).

Part of the track-class effects observed in previous studies was likely due to co-variance with the other two variables described in this chapter, but even when that is accounted for, FRA track class still has a strong effect. However, the results presented in this chapter indicate that track class is one of (at least) three different factors that are significantly related to derailment rate. In this section, I describe an example of how use of the new, three-factor derailment rate model has significant implications for rail safety policy and practice compared to use of the earlier single-factor, track-class-specific model.

### 2.4.3 Risk modeling of hazardous materials transportation

Railroad hazardous materials transportation risk along a route can be defined as the multiplication of tank car derailment rate, traffic exposure, conditional probability of release (CPR) of a derailed tank car and release consequence (Andrews 1980; Geffen 1980; Nayak et al. 1983; Glickman and Rosenfield 1984; CCPS 1995):

R $=\mathrm{Z} \times \mathrm{L} \times \mathrm{P} \times \mathrm{C}$
where:
R = hazardous materials transportation risk per carload
Z $\quad=$ tank car derailment rate per traffic exposure
$\mathrm{L} \quad=$ segment length (miles)
P = conditional probability of release of a derailed hazardous materials car
C = release consequence (e.g., number of persons in the affected area)

Based on the analysis described in this chapter, the tank car derailment rate will vary depending on three characteristics - FRA track class, method of operation and traffic density. Given that a tank car derails, its release probability is dependent on derailment speed and tank car safety design (Treichel et al. 2006; Kawprasert and Barkan 2010). The consequence of a release incident can be measured using various metrics, such as quantity released, number of persons affected, environmental impacts, etc. (Glickman and Rosenfield 1984; Barkan et al. 1991; Rhyne 1994; CCPS 1995; Nicolet-Monnier and Gheorghe 1996; Anand and Barkan 2006; Yoon et al. 2009; Kawprasert 2010).

### 2.4.4 Train derailment rate estimation effect on routing of hazardous materials

As discussed above, use of average values as parameters in calculation of risk may be a useful expedient in certain contexts for simple or high-level assessments. However, in other circumstances, it may lead to decisions that are not only incorrect, but may lead to the opposite of the desired policy objective. An example is routing of hazardous materials to minimize risk to the public.

Somewhat by definition, the majority of rail traffic is routed on high density traffic lines. This helps railroads achieve greater economies of scale in terms of infrastructure investment and maintenance. In general, these high-density lines travel near or through urban areas where the majority of traffic originates or terminates. Higher density lines have more intense design and maintenance standards and are more likely to have signalized traffic control systems to maximize operating efficiency. For safety reasons, railroads often have operating speed restrictions on these lines when they pass through urban areas (Kawprasert and Barkan 2010). Consequently, these segments will generally be classified as having a lower FRA track class than adjacent, higher-speed sections of the same line; however, this does not mean that most track design and maintenance standards on the lower speed sections are less than the higher-speed sections. If these low-speed segments are designed and maintained to higher standards than is evident from track class alone, it may lead to over-estimation of risk under these circumstances.

The higher estimated accident rate based only on track-class, combined with the high population density can result in estimates that these are high-risk locations in the hazardous materials transportation network. Conversely, considering routes with higher speeds but less
traffic, skirting or avoiding metropolitan areas may result in under-estimation of risk along these lines. Circumstances such as these can affect routing, emergency response, and other risk-based decisions. In short, the greater the ability to measure and differentiate the accident rates of various segments given specified railroad infrastructure and operational characteristics, the more reliable the risk-based decision making.

To illustrate this, consider the following example. Assume that there are two alternative, 20 -mile routes through a region (Figure 2.2). Route 1 is signaled, has a maximum operating speed of 25 mph (FRA track class 2) and annual traffic of 40 MGT . Route 2 is non-signaled, with a maximum speed of 40 mph (FRA track class 3 ) and an annual traffic density of 15 MGT . A hazardous material is being transported through the region in a non-insulated DOT 111A100W1 tank car. This car has an average conditional probability of release, given that it is derailed in an FRA-reportable accident, of 0.35 (Treichel et al. 2006). However, the speeddependent effect on the car's release probability if derailed is 0.24 on Route 1 , and 0.384 on Route 2 due to their different operating speeds (Kawprasert and Barkan 2010). The population densities differ along the two routes, with 800 people estimated to be affected if there is a release along Route 1, compared to 400 people on Route 2.


Figure 2.2: Hypothetical alternative routes with different track characteristics and estimated release consequences

Using the single-parameter FRA track class model, the average tank car derailment rates would be 240 and 102 per billion car-miles for track classes 2 and 3, respectively. Using the model described in Equation (2.7), risk per carload is $9.22 \mathrm{E}-04$ on the Route 1 and $3.13 \mathrm{E}-04$ on Route 2, indicating that the risk on Route 2 is about three times lower than on Route 1.

However, if we use the new, three-factor accident rate model described in this chapter with method of operation and annual traffic density accounted for, as well as track class, the estimated tank car derailment rates are 150 per billion car-miles on Route 1, and 266 per billion car-miles on Route 2. Using these accident rates, the corresponding risk per carload is 5.76E-04 on Route 1 and $8.17 \mathrm{E}-04$ on Route 2 (Table 2.6).

Although this example is hypothetical, it is based on typical circumstances available to railroads regarding routing of hazardous materials in metropolitan regions and demonstrates the effect of accident rate on risk analysis and the corresponding decision making. The more finegrained accident rate estimate leads to a more accurate risk modeling, thereby facilitating more reliable safety decisions.

## Table 2.6

Different risk comparison using old and new accident rate estimates

|  | Old Accident Rate (Account for Track Class Only) |  | New Accident Rate (Account for Track Class, Method of Operation and Traffic Density) |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Route 1 (Class 2, Signaled, 40 MGT Annual Traffic Density) | Route 2 (Class 3, NonSignaled, 15 MGT Annual Traffic Density) | Route 1 (Class 2, Signaled, 40 MGT Annual Traffic Density) | Route 2 (Class 3, NonSignaled, 15 MGT Annual Traffic Density) |
| Tank Car Derailment Rate per Billion Car-Miles, Z | 240 | 102 | 150 | 266 |
| Conditional Probability of Release of a Derailed Tank Car, P | 0.24 | 0.384 | 0.24 | 0.384 |
| Release Consequence, C (Number of Persons Affected per Release Incident) | 800 | 400 | 800 | 400 |
| Risk per Carload (Number of Persons Affected) | $9.22 \mathrm{E}-04$ | 3.13E-04 | $5.76 \mathrm{E}-04$ | 8.17E-04 |

### 2.5 Conclusion

This chapter develops the latest train and car derailment rates on Class I railroad mainline tracks in the United States. FRA track class has been the principal factor used to quantitatively assess, location-specific derailment rate in railroad transportation safety and risk studies for over three decades. The analysis described here accounts for two new factors (method of operation and annual traffic density) that also have a strong and significant effect. The methodology and statistics presented herein can be used for more accurate train safety and risk analyses, thereby enabling more precise estimates of local and route-specific risk, and contributing to development of more effective risk reduction strategies to improve rail safety.

## References

Agresti, A. (2007). An introduction to categorical data analysis. John Wiley \& Sons, Inc., New Jersey.

Anand, P., Barkan, C.P.L. (2006). Exposure of soil and groundwater to spills of hazardous materials transported by rail: A geographic information system analysis. Transportation Research Record 1943, 12-19.

Anderson, R.T., Barkan, C.P.L. (2004). Railroad accident rates for use in transportation risk analysis. Transportation Research Record 1863, 88-98.

Andrews, W.B. (1980). An assessment of the risk of transporting liquid chlorine by rail. Battelle Pacific Northwest Labs, Richland, WA. http://www.osti.gov/scitech/biblio/5054139.

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

Association of American Railroads (AAR) (2005-2009). Analysis of Class I railroads. Association of American Railroads, Washington D.C.

Association of American Railroads (AAR) (2013). Class I railroad statistics. https://www.aar.org/StatisticsAndPublications/Documents/AAR-Stats-2013-01-10.pdf. Association of American Railroads, Washington D.C.

Barkan, C.P.L., Glickman, T.S. Harvey, A.E. (1991). Benefit cost evaluation of using different specification tank cars to reduce the risk of transporting environmentally sensitive chemicals. Transportation Research Record 1313, 33-43.

Barkan, C.P.L., Saat, M.R., González, III, F., Treichel, T.T. (2013). Cooperative research in tank car safety design: how science and engineering are reducing the risk of rail transport of hazardous materials. TR News 286, 12-19.

Center for Chemical Process Safety (CCPS) (1995). Guidelines for chemical transportation risk analysis. Center for Chemical Process Safety of the American Institute of Chemical Engineers, New York.

Department of Transportation (DOT), 2008. Pipeline and Hazardous Materials Safety Administration, Federal Railroad Administration 49 CFR Parts 172, 174 and 209 Hazardous Materials: Enhancing Rail Transportation Safety and Security for Hazardous Materials Shipments; Railroad Safety Enforcement Procedures; Enforcement, Appeal and Hearing Procedures for Rail Routing Decisions; Final Rules. Federal Register 73, 229, 72182-72201.

Dick, C.T., Barkan, C.P.L., Chapman, E., Stehly, M.P. (2003). A multivariate statistical model for predicting the occurrence and location of broken rails. Transportation Research Record 1825, 64-74.

El-Sibaie, M., Zhang, Y.J. (2004). Objective track quality indices. Transportation Research Record 1863, 81-87.

English, G.W., Higham, G., Bagheri, M., Moynihan, T.W., Saccomanno, F.F. (2007).
Evaluation of risk associated with stationary dangerous goods railway cars. Transport Canada Report, No. TP 14690E, prepared for the Transportation Development Centre (TDC), Montreal, Canada.

Federal Railroad Administration (FRA) (2003). Track safety standards. U.S. Department of Transportation, Washington D.C.

Federal Railroad Administration (FRA) (2008). Signals and train control fact sheet. U.S. Department of Transportation, Washington D.C.

Geedipally, S.R., Lord, D. (2010). Investigating the effect of modeling single-vehicle and multivehicle crashes separately on confidence intervals of Poisson-gamma models. Accident Analysis and Prevention 42(4), 1273-1282.

Geffen, C.E. (1980). An assessment of the risk of transporting propane by truck and train. Battelle Pacific Northwest Labs, Richland, WA. http://www.osti.gov/scitech/biblio/6994240.

Glickman T.S., Rosenfield, D.B. (1984). Risks of catastrophic derailments involving the release of hazardous materials. Management Science 30(4), 503-511.

Goodman, L.A. (1962). On the exact variance of products. Journal of the American Statistical Association 55(292), 708-713.

Hauer, E. (2001). Overdispersion in modeling accidents on road sections and in empirical Bayes estimation. Accident Analysis and Prevention 33(6), 799-808.

Hilbe, J.M. (2007). Negative binomial regression. Cambridge University Press, Cambridge, United Kingdom.

Kawprasert, A., Barkan, C.P.L. (2008). Effects of route rationalization on hazardous materials transportation risk. Transportation Research Record 2043, 65-72.

Kawprasert, A., Barkan, C.P.L. (2010). Effect of train speed on risk analysis of transporting hazardous materials by rail. Transportation Research Record 2159, 59-68.

Kawprasert (2010). Quantitative analysis of options to reduce risk of hazardous materials transportation by railroad. Ph.D. Dissertation, University of Illinois at Urbana-Champaign, Urbana, IL.

Liu, X., Barkan, C.P.L., Saat, M.R. (2011). Analysis of derailments by accident cause: evaluating railroad track upgrades to reduce transportation risk. Transportation Research Record 2261, 178-185.

Liu, X., Saat, M.R., Barkan, C.P.L. (2012). Analysis of causes of major train derailment and their effect on accident rates. Transportation Research Record 2289, 154-163.

Liu, X., Saat, M.R., Qin, X., Barkan, C.P.L. (2013a). Analysis of U.S. freight-train derailment severity using zero-truncated negative binomial regression and quantile regression. Accident Analysis and Prevention 59, 87-93.

Liu, X., Saat, M.R., Barkan, C.P.L. (2013b). Safety effectiveness of integrated risk reduction strategies for rail transport of hazardous materials. Transportation Research Record 2374, 102110.

Lord, D., Washington, S.P., Ivan, J.N. (2005). Poisson, Poisson-gamma and zero inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Accident Analysis and Prevention 37(1), 35-46.

Lord, D. (2006). Modeling motor vehicle crashes using Poisson-gamma models: examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter. Accident Analysis and Prevention 38(4), 751-766.

Lord, D. (2008). Methodology for estimating the variance and confidence intervals of the estimate of the product of baseline models and AMFs. Accident Analysis and Prevention 40(3), 1013-1017.

Maher, M.J., Summersgill, I. (1996). A comprehensive methodology for the fitting of predictive accident models. Accident Analysis and Prevention 28(3), 281-296.

Miaou, S.P. (1994). The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. Accident Analysis and Prevention 26(4), 471-482.

Miaou, S.P., Lord, D. (2003). Modeling traffic crash-flow relationships for inter-sections: dispersion parameter, functional form, and Bayes versus empirical Bayes methods. Transportation Research Record 1840, 31-40.

Mitra, S., Washington, S. (2007). On the nature of over-dispersion in motor vehicle crash prediction models. Accident Analysis and Prevention 39(3), 459-568.

Nayak, P.R., Rosenfield, D.B., Hagopian, J.H. (1983). Event probabilities and impact zones for hazardous materials accidents on railroads. Report DOT/FRA/ORD-83/20, FRA, U.S. Department of Transportation, Washington D.C.

Nicolet-Monnier, M., Gheorghe, A.V. (1996). Quantitative risk assessment of hazardous materials transport systems: rail, road, pipelines and ship. Volume 5 of Topics in Safety, Risk, Reliability and Quality, Kluwer Academic Publishers, Dordrecht.

Oh, J., Washington, S.P., Nam, D. (2006). Accident prediction model for railway-highway interfaces. Accident Analysis and Prevention 38(4), 346-356.

Peng, F. (2011). Scheduling of track inspection and maintenance activities in railroad networks. Ph.D. Dissertation, University of Illinois at Urbana-Champaign, Urbana, IL.

Rhyne, W.R. (1994). Hazardous materials transportation risk analysis. Van Nostrand-Reinhold, New York.

Saat, M.R. (2009). Optimizing railroad tank car safety design to reduce hazardous materials transportation risk. Ph.D. Dissertation, University of Illinois at Urbana-Champaign, Urbana, IL.

Saccomanno, F.F., El-Hage, S.M. (1989). Minimizing derailments of railcars carrying dangerous commodities through effective marshaling strategies. Transportation Research Record 1245, 34-51.

Sawadisavi, S. (2010). Development of machine-vision technology for inspection of railroad track. M.S. Thesis, University of Illinois at Urbana-Champaign, Urbana, IL.

Schlake, B. (2010). Impact of automated condition monitoring technologies on railroad safety and efficiency. M.S. Thesis, University of Illinois at Urbana-Champaign, Urbana, IL.

Surface Transportation Board (STB) (2003). Hazardous Materials Transportation, 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, Surface Transportation Board, Section of Environmental Analysis.
http://www.stb.dot.gov/boundvolumes7.nsf/38f502d6898daf3385256811004b5a12/c8e3911c765 f5939852576da006464ee/\$FILE/Pages\%2031\%20-\%2048.pdf.

Treichel, T.T., Barkan, C.P.L. (1993). Working paper on mainline freight train accident rates. Research and Test Department, Association of American Railroads, Washington D.C.

Treichel, T.T., Hughes, J.P., Barkan, C.P.L., Sims, R.D., Phillips, E.A., Saat, M.R. (2006). Safety performance of tank cars in accidents: probability of lading loss, report RA-05-02. RSIAAR railroad tank car safety research and test project, Association of American Railroads, Washington, D.C.

Wood, G.R. (2002). Generalised linear accident models and goodness of fit testing. Accident Analysis and Prevention 34(4), 417-427.

Wood, G.R. (2005). Confidence and prediction intervals for generalised linear accident models. Accident Analysis and Prevention 37(2), 267-273.

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

Zarembski, A.M., Palese, J.W. (2010). Evaluation of the effectiveness of rail grinding on reducing rail defects on North American Class 1 Railroad. American Railway Engineering and Maintenance-of-Way Association Annual Meeting, Florida, FL.

## Chapter 3

# Analysis of U.S. Freight-Train Derailment Severity Using Zero-Truncated Negative Binomial Regression and Quantile Regression 

Adapted from
Liu, X., Saat, M.R., Qin, X., Barkan, C.P.L. (2013). Analysis of U.S. freight-train derailment severity using zero-truncated negative binomial regression and quantile regression. Accident Analysis and Prevention 59, 87-93.

### 3.1 Introduction

Railways are vital to the U.S. national economy. While society derives significant benefits from rail transportation, there are also certain safety risks that must be managed and minimized to a feasible extent. Train accidents cause damage to infrastructure and rolling stock, disrupt services, and may cause casualties and harm the environment. Derailments accounted for $72 \%$ of freighttrain accidents in the United States from 2001 to 2010 (Liu et al. 2012), causing 1,257 casualties, $\$ 1.8$ billion in damages and resulting in 344 hazardous materials cars releasing at least some of their contents. Consequently, understanding the causes and preventing train derailments is a high priority for the rail industry and government.

Train derailment likelihood has been studied by previous researchers (e.g., Nayak et al. 1983; Treichel and Barkan 1993; Dennis 2002; Anderson and Barkan 2004; Liu et al. 2011, 2012). In addition, understanding the magnitude and variability of derailment severity is also important. In this chapter, severity is measured by number of cars derailed. The generic use of "cars" refers to all vehicles (including locomotives, railcars and cabooses), unless specifically stated otherwise.

### 3.2 Literature review

Simulation and statistical analysis are the two basic approaches used to model train derailment severity. Simulation models predict the response of railroad vehicles to specific track and environmental conditions. These models are typically based on detailed nonlinear wheel-rail interaction models. For example, Yang et al. $(1972,1973)$ developed a simulation model to determine the effect of ground friction, mating coupler moment and brake retarding force on the number of cars derailed. They found that the position of the first car involved in the derailment (called point-of-derailment, or POD) and derailment speed could affect the number of cars derailed (Yang et al. 1972, 1973). In the late 1980s, Yang et al.'s model was extended by accounting for coupler failure and independent car motion (Coppens et al. 1988; Birk et al. 1990). The precision of simulation models are subject to the accuracy of models of train derailment dynamics.

In addition to simulation models, derailment severity has also been estimated using statistical methods. Saccomanno and El-Hage $(1989,1991)$ developed a truncated geometric model to estimate the mean number of cars derailed as a function of derailment speed, residual train length and accident cause. The model was later modified by Anderson (2005) and Bagheri (2009). There is interest in considering other factors that may affect derailment severity. Last but not least, all previous derailment severity models have focused on the mean number of cars derailed. Depending on factors discussed later in this chapter, other distributional statistics may also need to be understood, such as quantiles.

### 3.3 Derailment severity

The number of cars derailed per freight-train derailment on U.S. Class I railroads, from 2001 to 2010, is plotted in Figure 3.1. On average, a freight-train derailment resulted in approximately 10 cars derailed, and the median number derailed was six.


Figure 3.1: Distribution of number of cars derailed per FRA-reportable freight-train derailment, all accident causes combined, U.S. Class I freight railroad mainlines, 2001-2010

Derailment severity varies by accident cause (Saccomanno and El-Hage 1989, 1991; Barkan et al. 2003; Anderson 2005; Bagheri 2009; Bagheri et al. 2011; Liu et al. 2011, 2012). Broken rails were the most common cause of freight-train derailments on U.S. Class I mainlines (Barkan et al. 2003; Liu et al. 2012). The average broken-rail-caused freight-train derailment severity was 13 cars, compared to an average of six cars derailed when the cause was bearing failure. Although bearing failures sometimes cause severe train derailments, $55 \%$ of them caused only a single car to derail (Figure 3.2). Because broken rails are likely to pose greater risk than other causes due to
their high frequency and severity, this research focuses on this accident cause, but the methodology can be adapted to other accident causes as well.


Figure 3.2: Freight-train derailment severity distribution by major accident cause, U.S. Class I mainlines, 2001 to 2010

### 3.4 Input data for analysis

### 3.4.1 Data source

Train derailment data came from the Rail Equipment Accident (REA) database maintained by the Federal Railroad Administration (FRA) of the U.S. Department of Transportation. The FRA REA database contains all accidents that exceeded a specified monetary threshold of damage costs to on-track equipment, signals, track, track structures and roadbed. The reporting threshold is periodically adjusted for inflation, from $\$ 5,700$ in 1990 to $\$ 9,900$ in 2013 (FRA 2013). The REA database contains detailed train accident information such as total damage costs, number of
cars derailed, track type, train length, derailment speed and others. In some previous studies, monetary damage has been used to assess derailment severity. However, the financial cost of a derailment is subject to many variables, such as the cost difference between locomotives and railcars, or the difference in repairing regular track versus special trackwork. Instead, number of cars derailed was used to measure derailment severity (Barkan et al. 2003).

### 3.4.2 Explanatory variables

Several factors may affect train derailment severity, including residual train length, derailment speed, train power distribution and proportion of loaded railcars in the train. The explanation of each variable follows.

## Residual train length

Residual train length is defined as the number of railcars following the point-of-derailment (POD) (including the POD), where POD is the position of the first car derailed. Residual train length describes the maximum number of cars potentially subject to derailment (Saccomanno and El-Hage 1989, 1991). Previous studies have found that all else being equal, a greater residual train length is expected to result in more cars derailed (Saccomanno and El-Hage 1989, 1991; Anderson 2005; Bagheri 2009; Bagheri et al. 2011).

## Derailment speed

Nayak et al. (1983), Treichel and Barkan (1993), Saccomanno and El-Hage (1989, 1991), Barkan et al. (2003), Anderson (2005), Bagheri (2009), Bagheri et al. (2011) and Liu et al. (2011) all showed a positive correlation between the mean number of cars derailed and derailment speed.

## Distributed power

No previous study has analyzed whether distributed power (i.e. placement of some locomotives at the rear and/or middle of the train) affects train derailment severity. In this study, freight-trains are classified into two types: (1) non-distributed-power trains with only head-end locomotives and (2) distributed-power trains with head-end locomotives and additional locomotives in other positions. A binary variable (1 represents a distributed-power train, 0 otherwise) was created to examine whether the two types of trains have statistically different derailment severities.

## Proportion of loaded cars

Another new factor considered in this study is the proportion of loaded cars in the train. It is defined as the ratio of number of loaded cars, normalized by total number of cars (both empty and loaded) in the train. The null hypothesis is that a train carrying a larger proportion of loaded cars may derail more cars. A larger proportion of loaded cars in the train may indicate greater kinetic energy in the derailment, thereby causing more cars to derail.

Table 3.1 presents some descriptive statistics of the explanatory variables. The Spearman correlation coefficients were calculated (Table 3.2). There is a significant correlation between train power distribution and the proportion of loaded cars in the train ( $\mathrm{P}<0.05$ ). It indicates that a derailed train having a higher proportion of loaded cars is more likely to be equipped with distributed power.

## Table 3.1

Descriptive statistics of explanatory variables, broken-rail-caused freight-train derailments, U.S. Class I mainlines, 2001 to 2010

| Variables | Mean | Standard Deviation | Minimum | Maximum | Type |
| :--- | ---: | :---: | ---: | ---: | ---: |
| residual train length | 51.1 | 35.2 | 1 | 152 | Count |
| derailment speed (mph) | 28.8 | 14.1 | 4 | 70 | Continuous |
| proportion of loaded cars in the train | 0.8 | 0.3 | 0 | 1 | Continuous |
| train power distribution | 0.1 | 0.3 | 0 | 1 | Binary |

## Table 3.2

Spearman correlation coefficients between explanatory variables, broken-rail-caused freight-train derailments, U.S. Class I mainlines, 2001 to 2010

|  | derailment <br> speed | proportion of <br> loading cars | train power <br> distribution |
| :--- | :---: | :---: | :---: |
| residual train length | -0.04 | -0.03 | 0.11 |
| derailment speed |  | -0.09 | 0.05 |
| proportion of loading cars |  |  | $0.29^{\star}$ |

### 3.5 Zero-truncated negative binomial (ZTNB) model

### 3.5.1 Model development

The number of cars derailed represents non-negative count data, whose mean value can be estimated using regression techniques. Poisson regression and negative binomial (NB) regression are among the commonly used count data regression methods in accident analyses (e.g., Maccullagh and Nelder 1989; Miaou 1994; Hauer 2001; Wood 2002, 2005; Lord et al. 2005; Lord and Mannering 2010). The Poisson model is suitable for data whose mean is equal to its variance, whereas the NB model assumes that the Poisson mean follows a gamma distribution. The NB model has been used for analyzing over-dispersed data (the variance is greater than the mean) by previous researchers (e.g., Maccullagh and Nelder 1989; Long 1997; Hauer 2001; Lord et al. 2005; Hilbe 2007).

Both the Poisson and NB distributions include zeros, so they cannot be directly used to analyze data excluding zero counts, such as the data in this study. The minimum number of cars derailed is 1 . The Poisson or NB probability functions and their respective log-likelihood functions need to be modified to account for the exclusion of zeros. Gurmu (1991) and Grogger and Carson (1991) discussed methodologies to analyze zero-truncated count data. Compared to the traditional count data models (Poisson or negative binomial), the zero-truncated models calculate the probability of a positive response variable of count data using Bayes's Theorem (Gurmu 1991; Grogger and Carson 1991; Long 1997; Hilbe 2007). The following equations show the probability mass function (Equation 3.1), mean (Equation 3.2), variance (Equation 3.3), likelihood function (Equation 3.4) and response surface (Equation 3.5) of a zero-truncated negative binomial (ZTNB) model. The comparison of a ZTNB and NB model shows that the ZTNB model accounts for exclusion of zeros. A detailed discussion of ZTNB model can be found in Gurmu (1991), and Grogger and Carson (1991).
$\operatorname{Pr}\left(y_{i} \mid y_{i}>0\right)=\frac{\frac{\Gamma\left(y_{i}+\alpha^{-1}\right)}{y_{i}!\Gamma\left(\alpha^{-1}\right)}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_{i}}\right)^{\alpha^{-1}}\left(\frac{\mu_{i}}{\alpha^{-1}+\mu_{i}}\right)^{y_{i}}}{1-\left(1+\alpha \mu_{i}\right)^{-\alpha^{-1}}}$
$E\left(y_{i} \mid y_{i}>0\right)=\frac{\mu_{i}}{\operatorname{Pr}\left(y_{i}>0\right)}=\frac{\mu_{i}}{1-\left(1+\alpha \mu_{i}\right)^{-\alpha^{-1}}}$
$\operatorname{Var}\left(y_{i} \mid y_{i}>0\right)=\frac{E\left(y_{i} \mid y_{i}>0\right)}{\operatorname{Pr}\left(y_{i}>0\right)^{\alpha}}\left[1-\operatorname{Pr}\left(y_{i}=0\right)^{1+\alpha} E\left(y_{i} \mid y_{i}>0\right)\right]$
$L=\prod_{i=1}^{N} \operatorname{Pr}\left(y_{i} \mid y_{i}>0\right)=\prod_{i=1}^{N} \frac{\frac{\Gamma\left(y_{i}+\alpha^{-1}\right)}{y_{i}!\Gamma\left(\alpha^{-1}\right)}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_{i}}\right)^{\alpha^{-1}}\left(\frac{\mu_{i}}{\alpha^{-1}+\mu_{i}}\right)^{y_{i}}}{1-\left(1+\alpha \mu_{i}\right)^{-\alpha^{-1}}}$
$\log \left(\mu_{i}\right)=\beta_{0}+\beta_{1} X_{1 i}+\ldots+\beta_{k} X_{k i}$
Where:
$P_{r}\left(y_{i} \mid y_{i}>0\right) \quad=$ probability mass function of zero-truncated negative binomial distribution $\mathrm{E}\left(\mathrm{y}_{\mathrm{i}} \mid \mathrm{y}_{\mathrm{i}}>0\right) \quad=$ expectation of zero-truncated negative binomial distribution $\operatorname{Var}\left(y_{i} \mid y_{i}>0\right) \quad=$ variance of zero-truncated negative binomial distribution $\alpha \quad=$ over-dispersion parameter in the zero-truncated negative binomial distribution $\mathrm{L} \quad=$ likelihood function
$\mu_{\mathrm{i}} \quad=$ estimated derailment severity for the $\mathrm{i}^{\text {th }}$ observation
$y_{i} \quad=$ observed derailment severity for the $\mathrm{i}^{\text {th }}$ observation
$\beta_{\mathrm{k}} \quad=$ parameter coefficient of the $\mathrm{k}^{\text {th }}$ predictor variable $(\mathrm{k}=0$ for intercept $)$
$\mathrm{X}_{\mathrm{ki}} \quad=$ value of the $\mathrm{k}^{\text {th }}$ predictor variable for the $\mathrm{i}^{\text {th }}$ observation

A ZTNB model was developed based on 458 broken-rail-caused freight-train derailments on Class I mainlines from 2001 to 2010. The model accounts for main effect, higher-order components and their interaction terms of explanatory variables. A logarithmic transformation of residual train length and derailment speed provided a better fit to derailment severity data in Canada (Saccomanno and El-Hage 1989, 1991) and the same was observed in the U.S. data used in this study as well. Table 3.3 shows the "final" model using a hierarchical variable selection technique based on deviance (Agresti 2007) (the detailed variable selection process is presented in Appendix B).

Table 3.3
ZTNB modeling results for broken-rail-caused freight-train derailments, U.S. Class I mainlines, 2001 to 2010

| Variable | Coefficient | Standard Error | P-value |
| :--- | :---: | :---: | :---: |
| Intercept | 1.38 | 0.71 | 0.05 |
| RL | -0.03 | 0.23 | 0.91 |
| DS | -0.26 | 0.17 | 0.13 |
| LO | -0.57 | 0.34 | 0.09 |
| $(\text { RL })^{2}$ | -0.06 | 0.02 | 0.01 |
| RL $\times$ DS | 0.24 | 0.04 | 0.00 |
| RL $\times$ LO | 0.21 | 0.08 | 0.01 |

Notes:
RL = logarithmic residual train length
DS = logarithmic derailment speed
LO = loading factor (proportion of loaded cars in a train)

The train derailment severity model is:
$Z=\exp \left[1.38-0.03 R L-0.26 D S-0.57 L O-0.06(R L)^{2}+0.24 R L \times D S+0.21 R L \times L O\right]$

Where, $Z=$ estimated number of cars derailed per broken-rail-caused freight-train derailment on Class I mainlines. Using Equation (3.6), the first-order partial-derivative of derailment severity with respect to each predictor variable is:

$$
\begin{align*}
& \frac{\partial Z}{\partial R L}=Z(-0.03-2 \times 0.06 \times R L+0.24 D S+0.21 L O)  \tag{3.7}\\
& \frac{\partial Z}{\partial D S}=Z(-0.26+0.24 R L)  \tag{3.8}\\
& \frac{\partial Z}{\partial L O}=Z(-0.57+0.21 R L) \tag{3.9}
\end{align*}
$$

There are several observations based on Equations (3.7) to (3.9):

1) The sensitivity of derailment severity to residual train length is also dependent on derailment speed and loading factor. If $R L<2 \mathrm{DS}+1.75 \mathrm{LO}-0.25$, then $\frac{\partial Z}{\partial R L}>0$ ( RL is the logarithmic residual train length). This means that derailment severity increases with residual train length when it is below a threshold value that is determined by derailment speed and loading factor. The greater the speed and/or loading factor, the more likely that derailment severity has a positive correlation with residual train length, all else being equal. Whereas, given certain values of low speed and/or low loading factor, it is possible that derailment severity decreases when residual train length increases. However, the latter scenario is rare in historical accident data (4 out of 458 , or $0.08 \%$ ).
2) The sensitivity of derailment severity with respect to derailment speed is dependent on residual train length. If $R L>1.1$, then $\frac{\partial Z}{\partial D S}>0$. It means that a greater derailment speed is associated with more cars derailed. This applies to $99 \%$ of observations in historical data (453 out of 458).
3) The relationship between derailment severity and loading factor is dependent on residual train length as well. When $\mathrm{RL}>2.7, \frac{\partial Z}{\partial L O}>0$. It means that derailment severity increases if there is a larger proportion of loaded cars in the train. This applies to the majority of observations ( 394 out of 458 , or $86 \%$ ). Note that the remaining $14 \%$ of accidents had either short train length or the POD is in the rear end of the train. Note that loading factor herein does not provide information regarding the actual position of each loaded car,
which may affect train derailment dynamics; however, this information was not available for this study.

To summarize, there is a significant interaction between residual train length, derailment speed and loading factor. This interaction may be related to the energy accumulation in train accidents. Because of the interaction between explanatory variables, caution should be taken when interpreting the effect of an individual factor. Future research is needed to better understand the train derailment process based on additional possible affecting factors.

### 3.5.2 Limitations of ZTNB model

The ZTNB model is suitable for analyzing the data excluding zeros; however, like other traditional count data models, it has several limitations:

1) It analyzes the mean response variable; however, the mean may not fully represent the data distribution. For example, the mean derailment severity is 7 for bearing-failurecaused derailments (Figure 3.2b), while the median is 1 . Using the conditional mean may over-estimate derailment severity for a large proportion of observations. It raises a statistical question, how to analyze other distributional statistics (such as quantiles) in the regression?
2) The ZTNB or NB model is based on the homogeneous Bernoulli assumption (Long 1997; Hilbe 2007). Whether this assumption is valid in reality may be subject to uncertainty. The ZTNB model requires that each car has approximately equal derailment probability following the POD. While this assumption is difficult to verify in the field, some previous
models adopted this assumption and reported reasonable goodness-of-fit (Saccomanno and El-Hage 1989, 1991; Anderson 2005; Bagheri 2009; Bagheri et al. 2011). It is of interest to understand whether certain restrictive assumptions regarding the data distribution can be relaxed, thus motivating the introduction of quantile regression, which is the focus of the remaining sections.

### 3.6 Quantile regression

In addition to the mean, additional distributional statistics, specifically quantiles, were analyzed in this section. Quantiles are taken from the cumulative distribution function (CDF) of a random variable. Let p be a number between 0 and 1 , and the p quantile of the distribution of a random variable Y is denoted by $\mathrm{Q}(\mathrm{p})$.
$p=F(Q(p))=\int_{-\infty}^{Q(p)} f(y) d y$
$Q(p)=F^{-1}(p)=\inf \{y: F(y) \geq p\}, 0 \leqslant \mathrm{p} \leqslant 1$

Where:
$\mathrm{F}^{-1}(\mathrm{p}) \quad=$ inverse function of the cumulative distribution function (CDF)
Inf $\quad=$ the smallest $y$ that satisfies $F(y) \geqslant p$

Compared to the ZTNB model that analyzes the conditional mean, quantile regression $(\mathrm{QR})$ estimates the conditional quantiles (such as the median). QR was originally developed by Koenker and Bassett (1978), and it has been applied in various research fields (Taylor 1999; Arias et al. 2001; Machado and Mata 2001; Nielson and Rosholm 2001). The few applications of QR in transportation safety research include Hewson (2008), Qin et al. (2010), Qin and Reyes
(2011), and Qin (2012); however, I am unaware of any previous applications to rail safety research. Compared to the ZTNB model, quantile regression provides additional understanding of derailment severity. I used the following linear programming proposed by Koenker and Bassett (1978) to estimate the parameter coefficients in quantile regression:
$\min \left[\sum_{y_{i}<Q_{Y}(p)} p\left|y_{i}-Q_{Y}(p \mid X)\right|+\sum_{y_{i}<Q_{Y}(p)}(1-p)\left|y_{i}-Q_{Y}(p \mid X)\right|\right]$
Where:
$\log \left(Q_{Y}(p \mid X)\right)=\beta_{0}+\beta_{1} X_{1}+\ldots+\beta_{k} X_{k}$

To illustrate the concept of quantile regression, consider the following example in which I analyze the conditional quantile of derailment severity (number of cars derailed), denoted by $\mathrm{Qr}_{\mathrm{y}}(\mathrm{p} \mid \mathrm{X})$, based on explanatory variables. It is assumed that the logarithmic transformation of $\mathrm{Q}_{\mathrm{Y}}(\mathrm{p} \mid \mathrm{X})$ is a linear combination of predictor variables, that is $\log \left(Q_{Y}(p \mid X)\right)=\beta_{0}+\beta_{1} X_{1}+\ldots+\beta_{k} X_{k}$, where X is the main effect, higher-order components or interaction terms of predictor variables. In order to determine the "optimal" set of parameter coefficients, an objective function is defined in the following way (Koenker and Bassett 1978). First, all observations (train derailments) are categorized into two groups based on derailment severity relative to a quantile value, which is determined by specific values of explanatory variables and corresponding parameter coefficients. In the first group, all train derailments have derailment severity greater than the specified quantile value. The other group has derailment severity below the quantile. Within each group, we can calculate the total absolute deviation between the observation-specific derailment severity and specified quantile value. The objective
function is defined as the weighted average of total deviation within each group. The "optimal" set of parameter coefficients are estimated by optimizing the objective function. This optimization can be performed by various methods such as simplex algorithm, interior point method or smooth algorithm (Chen 2005). QR exhibits several advantages over traditional count data regression models according to Qin et al. (2010), Qin and Reyes (2011), and Qin (2012):

- First, QR does not require a specific data distribution, such as a normal or Poisson distribution. Therefore, it is suitable for analyzing strongly skewed data.
- Second, QR is more robust against outliers. Compared to the mean, certain quantiles (e.g., median) are less sensitive to skewed data, outliers and multi-modality.
- QR estimates multiple rates of change on different parts of the response distribution, and provides a complete view regarding the effects of predictors.

QR is implemented in several statistical software packages, such as QREG in Stata and QUANTREG in SAS. However, none of these standard statistical procedures were designed for discrete variables. In terms of count data, the objective function in the optimization is not differentiable, making it difficult to model quantiles directly as a continuous function of predictor variables (Machado and Silva 2005). Smoothing approaches were used to apply QR to count data analysis (Machado and Silva 2005). In this study, I used the jittering method developed by Machado and Silva (2005). Machado and Silva's jittering algorithm was implemented in a procedure called QCOUNT in Stata by Miranda (2007). Using the QCOUNT in Stata, I developed a quantile regression model based on the freight-train derailment severity data. Table 3.4 presents the QR estimates for selected quantiles and the associated standard errors. For comparison, I used the same parameters in developing the ZTNB model as shown in

Table 3.3. The analysis shows that the relationship between derailment severity and affecting factors may differ at different quantiles. There are two principal observations:

- First, the signs of coefficients for significant parameters are the same at the studied quantiles and the mean. So change of a factor may consistently affect train derailments for all severities.
- The parameter coefficient in modeling the mean derailment severity using ZTNB model is within the range of coefficients using quantile regression to analyze $20^{\text {th }}$ and $60^{\text {th }}$ percentiles. This is not surprising because the mean is also a special quantile value.


## Table 3.4

Estimated coefficients in quantile regression versus ZTNB model, broken-rail-caused Class I mainline freight-train derailments, 2001-2010

| Variables | 0.2 |  | 0.4 |  | 0.6 |  | Mean |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Coefficient | P-value | Coefficient | P-value | Coefficient | P -value | Coefficient | P-value |
| Intercept | 1.90 | 0.01 | 1.47 | 0.00 | 0.97 | 0.06 | 1.38 | 0.05 |
| RL | -0.43 | 0.11 | -0.26 | 0.04 | 0.03 | 0.83 | -0.03 | 0.91 |
| DS | -0.25 | 0.16 | -0.25 | 0.01 | -0.21 | 0.06 | -0.26 | 0.13 |
| LO | -0.95 | 0.01 | -0.50 | 0.10 | -0.22 | 0.45 | -0.57 | 0.09 |
| $(\mathrm{RL})^{2}$ | -0.02 | 0.43 | -0.04 | 0.00 | -0.07 | 0.00 | -0.06 | 0.01 |
| RL×DS | 0.24 | 0.00 | 0.26 | 0.00 | 0.24 | 0.00 | 0.24 | 0.00 |
| $\underline{R L \times L O}$ | 0.32 | 0.00 | 0.21 | 0.01 | 0.15 | 0.05 | 0.21 | 0.01 |

Notes:
RL = logarithmic residual train length
DS = logarithmic derailment speed
LO = loading factor

### 3.7 Comparison between conditional-mean and conditional-quantile models

In the last three decades, several statistical models have been developed to analyze count data in transportation safety research, and most of them have focused on the conditional mean. However, the mean may be insufficient to describe highly skewed or multi-modal data distributions (Qin et al. 2010; Qin and Reyes 2011; Qin 2012). For instance, estimating the conditional mean may over-estimate train derailment severity for a large proportion of derailments if the data are skewed towards low-severity derailments. As an alternative approach, quantile regression provides estimates at different quantiles, and it provides additional views in analyzing derailment severity distribution. Another advantage of quantile regression is its robustness against outliers in the dataset. A few outliers may significantly affect the mean and variance, but may not affect certain quantiles. Consequently, quantile regression may be useful in analyzing datasets with input errors.

An accurate estimation of derailment severity is important for evaluating the safety effectiveness of accident prevention strategies. Quantile regression can be used to better understand the effects of explanatory variables on the response variable. Specifically, if an explanatory variable has similar parameter estimates at different quantiles, an accident prevention strategy affecting this variable may be effective for all derailments regardless of their severities. However, if the parameter estimates vary widely at different quantiles, it indicates that the accident prevention strategy may affect low-severity derailments differently than highseverity derailments. To conclude, zero-truncated negative binomial regression analyzes the conditional mean, while quantile regression analyzes the quantiles of a response variable.

Depending on the questions to address, the two approaches can supplement one another to provide rail transportation safety information.

Note that quantile regression has its own limitations. For instance, development of quantile regression does not rely on the maximum likelihood method, thereby certain model selection criteria (e.g. AIC or BIC) may not directly be used for identifying the "best" model. The methodologies for evaluating the goodness-of-fit of a quantile regression model need additional research. In addition, quantile regression originates from continuous variable analysis and has recently been adapted to count data (Machado and Silva 2005). More work is needed to advance the theory and practicability of quantile regression in order to supplement the existing count data models.

### 3.8 Conclusion

Zero-truncated negative binomial and quantile regression models are developed to analyze the distribution of freight-train derailment severity and the effect of influencing factors. A zerotruncated negative binomial model is used to estimate the mean severity, excluding zeros. A quantile regression model is used to estimate a specific quantile value and it is suitable for analyzing the entire data distribution accounting for possible data skewness, outliers or when certain assumptions are not met. The two models together provide a better understanding of freight derailment severity distribution. These methods can be adapted to various other related research problems.

## References

Agresti, A. (2007). An introduction to categorical data analysis. John Wiley \& Sons, Inc., New Jersey.

Anderson, R.T., Barkan, C.P.L. (2004). Railroad accident rates for use in transportation risk analysis. Transportation Research Record 1863, 88-98.

Anderson, R.T. (2005). Quantitative analysis of factors affecting railroad accident probability and severity. M.S. Thesis, University of Illinois at Urbana-Champaign, Urbana, IL.

Arias, O., Hallock, K.F., Sosa-Escudero, W. (2001). Individual heterogeneity in the returns to schooling: instrumental variables quantile regression using twin data. Empirical Economics 26, 7-40.

Association of American Railroads (AAR) (2013). Class I railroad statistics. https://www.aar.org/StatisticsAndPublications/Documents/AAR-Stats-2013-01-10.pdf. Association of American Railroads, Washington D.C.

Bagheri, M. (2009). Risk-based model for effective marshalling of dangerous goods railway cars. Ph.D. Dissertation, Department of Civil and Environmental Engineering, University of Waterloo, Ontario.

Bagheri, M., Saccomanno, F.F., Chenouri, S., Fu, L.P. (2011). Reducing the threat of in-transit derailments involving dangerous goods through effective placement along the train consist. Accident Analysis and Prevention 43, 613-620.

Barkan, C.P.L., Dick, C.T, Anderson, R.T. (2003). Analysis of railroad derailment factors affecting hazardous materials transportation risk. Transportation Research Record 1825, 64-74.

Birk, A.M., Anderson, R.J. Coopens, A.J. (1990). A computer simulation of a derailment accident, parts I \& II. Journal of Hazardous Materials 25, 121-147 \& 149-165.

Chen, C. (2005). An introduction to quantile regression and the QUANTREG procedure. SAS Institute Inc., Cary, NC.

Coppens, A.J., Wong, J.D.E., Bibby, A., Birk, A.M, Anderson, R.J. (1988). Development of a derailment accident computer simulation model. Transport Canada Report No. TP 9254E, Prepared for the Transportation Development Centre and Transport of Dangerous Goods, Ottawa, Canada.

Dennis, S.M. (2002). Changes in railroad track accident rates. Transportation Quarterly 56(4), 161-174.

Federal Railroad Administration (FRA) (2013). Railroad equipment accident/incident reporting threshold 1957-2013. FRA, U.S. Department of Transportation, Washington D.C.

Grogger, J.T., Carson, R.T. (1991). Models for truncated counts. Journal of Applied Econometrics 6(3), 225-238.

Gurmu, S. (1991). Tests for detecting over-dispersion in the positive Poisson regression model. Journal of Business and Economic Statistics 9, 215-222.

Hauer, E. (2001). Overdispersion in modeling accidents on road sections and in empirical Bayes estimation. Accident Analysis and Prevention 33(6), 799-808.

Hewson, P. (2008). Quantile regression provides a fuller analysis of speed data. Accident Analysis and Prevention 40(2), 502-510.

Hilbe, J.M. (2007). Negative binomial regression. Cambridge University Press, Cambridge.
Koenker, R., Bassett, G. (1978). Regression quantiles. Econometrica 46, 33-50.
Liu, X., Barkan, C.P.L., Saat, M.R. (2011). Analysis of derailments by accident cause: evaluating railroad track upgrades to reduce transportation risk. Transportation Research Record 2261, 178-185.

Liu, X., Saat, M.R., Barkan, C.P.L. (2012). Analysis of causes of major train derailment and their effect on accident rates. Transportation Research Record 2289, 154-163.

Long, J.S. (1997). Regression models for categorical and limited dependent variables. SAGE Publications, Inc, Thousand Oaks, California.

Lord, D., Washington, S.P., Ivan, J.N. (2005). Poisson-gamma and zero inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Accident Analysis and Prevention 37(1), 35-46.

Lord, D., Mannering, F. (2010). The statistical analysis of crash-frequency data: a review and assessment of methodologies alternatives. Transportation Research Part A 44(5), 291-305.

Maccullagh, P., Nelder, J.A. (1989). Generalized linear models (second edition). Chapman \& Hall, London.

Machado, J.A.F., Mata, J. (2001). Earning functions in Portugal 1982-1994: evidence from quantile regressions. Empirical Economics 26, 115-134.

Machado, J.A.F., Santos Silva, J.M.C. (2005). Quantiles for counts. Journal of American Statistical Association 100(472), 1226-1237.

Miaou, S.P. (1994). The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. Accident Analysis and Prevention 26(4), 471-482.

Miranda, A. (2007). QCOUNT: Stata program to fit quantile regression models for count data http://ideas.repec.org/c/boc/bocode/s456714.html.

Nayak, P.R., Rosenfield, D.B., Hagopian, J.H. (1983). Event probabilities and impact zones for hazardous materials accidents on railroads. Report DOT/FRA/ORD-83/20, FRA, U.S. Department of Transportation, Washington D.C.

Nielson, H.S., Rosholm, M. (2001). The public-private sector wage gap in Zambia in the 1990s: a quantile regression approach. Empirical Economics 26, 169-182.

Qin, X., Ng, M., Reyes, P. (2010). Identifying crash-prone locations with quantile regression. Accident Analysis and Prevention 42(6), 1531-1537.

Qin, X., Reyes, P. (2011). A conditional quantile analysis for crash count data. Journal of Transportation Engineering 137(9), 601-607.

Qin, X. (2012). Quantile effects of causal factors on crash distributions. Transportation Research Record 2279, 40-46.

Saccomanno, F.F., El-Hage, S.M. (1989). Minimizing derailments of railcars carrying dangerous commodities through effective marshaling strategies. Transportation Research Record 1245, 3451.

Saccomanno, F.F., El-Hage, S.M. (1991). Establishing derailment profile by position for corridor shipments of dangerous goods. Canadian Journal of Civil Engineering 18(1), 67-75.

Taylor, J. (1999). A quantile regression approach to estimating the distribution of multiperiod returns. Journal of Derivatives 24, 64-78.

Treichel, T.T., Barkan, C.P.L. (1993). Working paper on mainline freight train accident rates. Research and Test Department, Association of American Railroads, Washington D.C.

Wood, G.R. (2002). Generalised linear accident models and goodness of fit testing. Accident Analysis and Prevention 34(4), 417-427.

Wood, G.R. (2005). Confidence and prediction intervals for generalised linear accident models. Accident Analysis and Prevention 37(2), 267-273.

Yang, T.H., Manos, W.P., Johnstone, B. (1972). A study continuation of derailment behavior final report (phase 08 report on computer derailment study). RPI/AAR Report RA-08-1-12 (R135), Railroad Tank Car Safety Research and Test Project, Association of American Railroads, Washington, D.C.

Yang, T.H., Manos, W.P., Johnstone, B. (1973). Dynamic analysis of train derailments. 72-WA/RT-6, Presented at the 1972 Winter Annual Meeting of ASME, Rail Transportation Proceedings, The American Society of Mechanical Engineers, New York.

## Chapter 4

# Probability Analysis of Multiple-Tank-Car Release Incidents in Railway Hazardous Materials Transportation 

Adapted from

Liu, X., Saat, M.R., Barkan, C.P.L. Probability analysis of multiple-tank-car release incidents in railway hazardous materials transportation. Journal of Hazardous Materials (under review).

### 4.1 Introduction

There were more than two million carloads of hazardous materials transported by rail in North America in 2012 (BOE 2013). Although hazardous materials traffic accounts for only $6 \%$ of U.S. rail traffic and more than $99 \%$ of shipments safely reached destinations, it is responsible for a major share of liability and insurance risk (AAR 2013a). There has been growing regulatory and public attention to hazardous materials transportation safety in response to rapid growth in this traffic and several recent train-accident-caused release incidents.

Railroad transport of hazardous materials differs from highway transport in several important ways. Notably rail transport involves trains of multiple cars, sometimes over 100 in a single train. Some or all of these may be tank cars. By contrast, highway transport generally involves a single tank trailer. Unlike highway transport, derailment of a hazardous materials train may result in releases from multiple tank cars. In the event of a large, multiple-car release incident, there is the potential for considerable impact on human health, property, and the environment. Furthermore, such releases may be much more challenging for emergency response. Several recent multiple-tank-car release incidents, such as the derailments in Schellebelle, Belgium in May 2013 and in Lac-Mégantic, Canada in July 2013, underscore the importance of preventing multiple-car release incidents.

The number of tank cars releasing per derailment is affected by a number of factors such as train length, speed, accident cause, point of derailment (the position of the first car derailed), positions of tank cars in the train, and tank car safety design (Nayak et al. 1983; Glickman and Rosenfield 1984; Saccomanno and El-Hage 1989, 1991; Glickman et al. 2007; Bagheri 2009; Bagheri et al. 2011, 2013). Previous analyses have focused on one or more of these factors under specific circumstances. For example, Bagheri et al. (2013) estimated the probability of multiplecar releases assuming that all tank cars have equal release probability and are grouped together in the train. Their assumptions are suitable under certain circumstances (e.g., unit-train shipments); however, circumstances in which tank cars have different safety designs and may be distributed anywhere throughout the train require a more sophisticated approach.

In this chapter, I develop a generalized model to estimate multiple-car release probability using the Law of Total Probability (Ross 2007). The model is applicable to any type of train configuration, tank car safety design and distribution of tank car positions in the train. It extends and generalizes previous models. Using this model, I analyze the safety effectiveness of various strategies to reduce the occurrence of multiple-tank-car release incidents.

### 4.2 Literature review

Hazardous materials transportation risk assessment relies on the estimation of the probability and consequences of a release incident (HSE 1991; Erkut and Verter 1995, 1998; Erkut and Ingolfsson 2000, 2004; CCPS 2008). Each element in the event chain of a hazardous materials release incident (Figure 4.1) has been analyzed in previous studies. In this section, I review these studies, identify their applications and limitations.


Figure 4.1: Sequence of events leading to a hazardous materials release incident

### 4.2.1 Freight-train derailment rate

Most major railroad hazardous materials release incidents occur as a result of train derailments. Attempts to relate derailment rate to infrastructure characteristics date back to the early 1980s. Nayak et al. (1983) found that higher FRA track classes had lower train derailment rates. A subsequent study by Treichel and Barkan (1993) found a similar result and Anderson and Barkan (2004) used new data to develop updated estimates. All of these studies found that higher FRA track classes had lower derailment rates. Higher track classes are intended for higher operating speeds, thus require a variety of more stringent maintenance and engineering safety standards (Liu et al. 2011). In addition to FRA track class, Liu et al. (2014) analyzed two additional factors - method of operation and traffic density and found that all three factors are strongly correlated with train derailment rates. In Canada, Saccomanno et al. (1989) analyzed train derailment rates by traffic volume, track type (single track versus multiple tracks), train speed and region and also found that train derailment rate varies with track characteristics.

### 4.2.2 Number of cars derailed

The total number of cars derailed is a metric of derailment severity (Nayak et al. 1983; Barkan et al. 2003; Anderson 2005; Bagheri et al. 2011; Liu et al. 2011, 2012, 2013a). It is affected by accident cause (Saccomanno and El-Hage 1989, 1991; Barkan et al. 2003; Anderson 2005; Liu et al. 2011, 2012, 2013a), train speed (Nayak et al. 1983; Saccomanno and El-Hage 1989, 1991; Barkan et al. 2003; Anderson 2005; Liu et al. 2013a), train length (Nayak et al. 1983; Saccomanno and El-Hage 1989, 1991; Anderson 2005; Liu et al. 2013a) and point of derailment (the position of the first car derailed) (Saccomanno and El-Hage 1989, 1991; Anderson 2005; Bagheri et al. 2011; Liu et al. 2013a).

### 4.2.3 Number of tank cars derailed

The number of tank cars derailed is related to the total number of cars (both tank and non-tank cars) derailed, and the number and placement of the tank cars in the train. Glickman et al. (2007) assumed that the number of tank cars derailed follows a hyper-geometric distribution when tank cars were randomly placed in the train. Bagheri et al. (2013) estimated the total number of tank cars derailed given their positions. Although the effects of tank car safety design and tank car position in the train have been considered independently in prior work, to my knowledge, no published research has simultaneously accounted for both factors. So one objective of this research is to develop a generalized approach to estimate the total number of tank cars derailed and their probability of releasing that accounts for type of tank car and its position in a train.

### 4.2.4 Number of tank cars releasing contents

Not all derailed tank cars release their contents. The Railway Supply Institute (RSI) Association of American Railroads (AAR) Railroad Tank Car Safety Research and Test Project has developed a tank car accident database (TCAD) containing information on tank cars involved in accidents in the U.S. since 1970. Using this database, Treichel et al. (2006) developed logistic regression models to estimate the conditional probability of release for nearly all common designs, as well as new designs that incorporate existing design features. Previous analyses have used an average release probability (Glickman et al. 2007; Bagheri et al. 2011); however, Barkan et al. (2003) and Treichel et al. (2006) found a strong effect of speed on both derailment severity and release probability of hazardous materials cars derailed. Kawprasert and Barkan (2010) extended Treichel et al.'s analysis by accounting for the effect of derailment speed in estimating release probability.

### 4.2.5 Release consequence

The consequence of a hazardous material release can be expressed using several metrics, such as human impact (e.g., the number of people potentially affected by a release, i.e. evacuees, injuries and/or fatalities), monetary units for costs due to property damage, environmental damage, and litigation or other forms of financial impact (Barkan et al. 1991; CCPS 2008; Saat 2009; Kawprasert 2010). The hazard area of a release incident is affected by a variety of factors including chemical properties, quantity released, rate of release, meteorological conditions and local terrain. Toxic vapors, pool fires, explosions, soil and groundwater contamination and ecological damage are typical hazard types that a hazardous materials release incident may pose (Birk et al. 1990; Raj and Turner 1992; Yoon et al. 2009; Kawprasert 2010; Appel et al. 2011;

Marruffo et al. 2012). Previous studies have analyzed the generation, propagation and impact of these hazard types using pool fire modeling (Raj and Turner 1992) and Gaussian plume model (GPM) for airborne chemicals (Hanna et al. 1993). Based on different hazard types, simulation tools have been developed to estimate the impact of a hazardous materials release (e.g., Birk et al. 1990; Appel et al. 2011). Based in part on simulation analysis and historical data, the U.S Department of Transportation Emergency Response Guidebook recommends first responders' initial isolation and protective action distances for specific chemicals and scenarios of release (PHMSA 2012). In addition to human impact, the consequences of a hazardous materials release could also include environmental clean-up cost (Barkan et al. 1991), train delay cost and others (Saat 2009). Consequence analysis can be conducted using a Geographic Information System platform integrated with other databases such as census and track infrastructure data (Zhang et al. 2000; Verter and Kara 2001; Verma and Verter 2007).

The surrounding natural and built environment at an accident location may also affect tank car derailment and release probabilities and consequences; however, the effect of these factors have not been investigated and therefore were not explicitly incorporated into our risk model.

### 4.3 Methodology

Equation (4.1) is used to estimate the probability distribution of the number of tank cars releasing after a train derailment occurs.

$$
P\left(X_{R}\right)=\sum_{X_{D}=0}^{\min \left(L_{T}, X\right)}\{\underbrace{P\left(X_{R} \mid X_{D}\right)}_{\begin{array}{c}
\text { Tank Cars }  \tag{4.1}\\
\text { Releasing }
\end{array}}\{\underbrace{\sum_{X=1}^{L-K+1} P\left(X_{D} \mid X\right)}_{\begin{array}{c}
\text { Tank Cars } \\
\text { Derailed }
\end{array}}[\underbrace{\sum_{K=1}^{L} P(X \mid K) P O D(K)}_{\begin{array}{c}
\text { All Types of } \\
\text { Cars Derailed }
\end{array}}]\}\}
$$

Where:
K = point of derailment
L = train length (total number of locomotives and railcars)
$\mathrm{X}_{\mathrm{D}} \quad=$ number of tank cars derailed
$\mathrm{X}_{\mathrm{R}} \quad=$ number of tank cars releasing
$\mathrm{L}_{\mathrm{T}} \quad=$ number of tank cars in the train

In order to estimate the probability distribution of the total number of tank cars releasing, the following distributions need to be estimated sequentially. In the following sections, each of them will be explained in more detail.

1) point-of-derailment (POD), the position of the first car derailed in the train
2) number of cars derailed (include both tank cars and other types of railcars) given a POD
3) number of tank cars derailed given the total number of cars derailed
4) number of tank cars releasing given the total number of tank cars derailed

### 4.3.1 Point of derailment, $\operatorname{POD}(K)$

Point-of-derailment is the position of the first car derailed. The first vehicle (generally the lead locomotive) in the train is frequently the first to derail (Anderson 2005). Previous studies found that, ceteris paribus, the nearer the POD is to the front of a train, the more cars derail (Saccomanno and El-Hage 1989, 1991; Anderson 2005; Bagheri 2009; Bagheri et al. 2011). Data used in this paper were from the U.S. Federal Railroad Administration (FRA) Rail Equipment

Accident (REA) database from 2002 to 2011. The FRA REA database contains all accidents that exceeded a specified monetary threshold of damage costs to on-track equipment, signals, track, track structures or roadbed. The reporting threshold is periodically adjusted for inflation, from \$5,700 in 1990 to \$9,900 in 2013 (FRA 2013). This research focuses on mainline, freight-train derailments on Class I railroads, a group of the largest U.S. railroads accounting for $69 \%$ of route miles and $88 \%$ of carloads transported on U.S. railroads (AAR 2013b). The analysis showed that approximately $25 \%$ of train derailments had the POD (the first car derailed) in the first ten positions of the train (Figure 4.2).


Figure 4.2: Distribution of point-of-derailment, Class I mainline freight-train derailments, all accident causes combined, 2002-2011 (only a portion of POD distribution is displayed)

To account for different train lengths, the normalized POD (NPOD) was calculated by dividing POD by train length (Saccomanno and El-Hage 1989, 1991). Several probability distributions (Beta, Normal, Logistic, Weibull, Uniform, Gamma) were selected to fit the NPOD
data. The goodness-of-fit of a fitted NPOD distribution was evaluated using the KolmogorovSmirnov (K-S) test (Corder and Foreman 2009). The "best-fit" for the NPOD distribution (all accident causes combined) is a Beta distribution. This is consistent with prior research based on different study periods (Saccomanno and El-Hage 1989, 1991; Anderson 2005). The beta distribution is a continuous probability distribution defined on the interval $[0,1]$ and parameterized by two positive shape parameters. In the Kolmogorov-Smirnov (K-S) test, the Pvalue evaluates the goodness of fit of empirical data compared to a theoretical distribution. If the P-value is larger than 0.05 , the distribution chosen to fit the data may be acceptable. Table 4.1 presents the "best" fitted NPOD distribution for several major derailment causes on U.S. railroads based on the FRA REA database.

Table 4.1
Parameter estimates for fitting NPOD distribution for major derailment causes of freight-trains, Class I mainlines, 2002 to 2011

| Cause <br> Group | Description | Fitted Distribution | P-value | Sample Size |
| :--- | :--- | :---: | :---: | :---: |
| $08 T$ | Broken Rails or Welds | Beta $(0.5519,0.8576)$ | 0.15 | 543 |
| $04 T$ | Track Geometry (excl. Wide Gauge) | Beta $(0.8622,0.7065)$ | 0.23 | 281 |
| 12E | Broken Wheels (Car) | Beta $(1.0497,0.9372)$ | 0.97 | 216 |
| 10E | Bearing Failure (Car) | Uniform (0, 1) | 0.71 | 204 |
| 09 H | Train Handling (excl. Brakes) | Beta $(0.7348,1.5488)$ | 0.70 | 184 |
| All Causes |  | Beta $(0.6793,0.8999)$ | 0.48 | 3,812 |

Given a train length L , the probability that the POD is at the $\mathrm{k}^{\text {th }}$ position, $\operatorname{POD}(\mathrm{k})$, can be estimated using the following equation:

$$
\begin{equation*}
P O D(k)=F\left(\frac{k}{L}\right)-F\left(\frac{k-1}{L}\right) \tag{4.2}
\end{equation*}
$$

Where:
$\operatorname{POD}(\mathrm{k})=\mathrm{POD}$ probability at the $\mathrm{k}^{\text {th }}$ position of a train

F()$\quad=$ cumulative density distribution of the fitted distribution
L = train length

Broken rails and bearing failures are the leading track-related and equipment-related train derailment causes on U.S. freight railroads (Barkan et al. 2003; Liu et al. 2012). Figure 4.3 shows that a broken-rail-caused train derailment is more likely to have the POD in the front of a train, whereas the POD is approximately uniformly distributed for a bearing-failure-caused train derailment. These POD distributions that differ by accident cause suggest that train accident analyses should account for specific causes and will be discussed in the next section.


Figure 4.3: Comparison of fitted NPOD distributions by major accident causes, Class I mainline freight-train derailments, 2002-2011

### 4.3.2 Total number of cars derailed, $\mathbf{P}(\mathbf{X} \mid \mathrm{K})$

The number of cars derailed in a train derailment has been used as a metric for train derailment severity (Saccomanno and El-Hage 1989, 1991; Barkan et al. 2003; Liu et al. 2012). Figure 4.4 shows the distribution of the number of cars derailed on U.S. Class I mainlines from 2002 to 2011. Approximately $24 \%$ of derailments resulted in a single car derailed. $50 \%$ resulted in five-or-fewer cars derailed, and the average number of cars derailed was approximately nine.


Figure 4.4: Distribution of number of cars derailed per train derailment, Class I railroad freight-train derailments on mainlines, all accident causes combined, 2002 to 2011

Derailment severity varies by accident cause (Nayak et al. 1983; Saccomanno and ElHage 1989, 1991; Barkan et al. 2003; Anderson 2005; Liu et al. 2012). Figure 4.5 compares the distribution of derailment severity due to broken rails and bearing failures, respectively. The average broken-rail-caused freight-train derailment severity was 13 cars, compared to an average
of six cars derailed when the cause was bearing failure. Although bearing failures sometimes cause severe train derailments, $55 \%$ of them resulted in only a single car derailed. The difference in accident-cause-specific derailment severity may be due to different POD distributions and possibly different derailment dynamics. Broken rails caused more hazardous materials car releases than other causes. According to the FRA REA database, 438 hazardous materials cars released on U.S. railroads from 2002 to 2011, of which 106 (24\%) were caused by broken rails. So broken rails are used as an example in this chapter to illustrate the methodology. The methodology can be adapted to other accident causes as well.


Figure 4.5 (cont.)

(b) Bearing failures

Figure 4.5: Number of cars derailed in freight-train derailments due to (a) broken rails or welds, and (b) bearing failures, Class I mainlines, 2002-2011

The statistical model for estimating train derailment severity was first developed by Saccomanno and El-Hage $(1989,1991)$ and subsequently modified by Anderson $(2005)$ and Bagheri et al. (2011). The probability of derailing $X$ cars can be estimated using the following equation (model derivation is presented in Appendix C):

$$
\begin{equation*}
\mathrm{P}(X \mid K)=\frac{\frac{\exp (z)}{1+\exp (z)}\left[\frac{1}{1+\exp (z)}\right]^{X-1}}{1-\left[\frac{1}{1+\exp (z)}\right]^{L-K+1}} \tag{4.3}
\end{equation*}
$$

Where, $Z=a+b \times \ln (S)+c \times \ln \left(L_{r}\right)+d \times L_{d}$. The notations for $S, L_{r}$ and $L_{d}$ are presented in Appendix C. Based on the POD distribution $(\operatorname{POD}(\mathrm{K}))$ and the number of tank cars derailed given the $\operatorname{POD}(\mathrm{P}(\mathrm{X} \mid \mathrm{K}))$, the distribution of the total number of cars can be estimated.

Next, I analyze the number of tank cars derailed given the total number of all types of cars derailed.

### 4.3.3 Tank car derailment, $\mathbf{P}\left(\mathbf{X D}_{\mathrm{D}} \mid \mathbf{X}\right)$

Depending on the placement and total number of tank cars in a train, a derailment may result in a number of tank cars derailing. Glickman et al. (2007) assumed that tank cars were uniformly distributed throughout the train and have the same release probability. They used a hyper-geometric distribution to calculate the probability distribution of tank car derailment. I extended Glickman et al.'s analysis by accounting for the derailment of different types of tank cars (different types of tank cars may have different release probabilities (Treichel et al. 2006; Saat 2009)) using a multivariate hyper-geometric distribution. I adopt the same assumption as Glickman et al. (2007) that tank cars are uniformly distributed throughout the train. The multivariate hyper-geometric distribution of number of tank cars derailed has the following expression:
$\mathrm{P}\left(X_{D 1}, \ldots, X_{D m} \mid X\right)=\frac{\binom{T_{1}}{X_{D 1}} \ldots\binom{T_{m}}{X_{D m}}\binom{L-\sum_{i=1}^{m} T_{D i}}{X-\sum_{i=1}^{m} X_{D i}}}{\binom{L}{X}}$

Where:
$X_{D m}=$ number of the $m^{\text {th }}$ type of tank cars derailed
m = types of tank cars in the train
X = total number of cars (both tank and non-tank cars) derailed
$\mathrm{T}_{\mathrm{m}} \quad=$ number of the $\mathrm{m}^{\text {th }}$ type of tank cars in the train
L = train length

### 4.3.4 Tank car release, $\mathbf{P}\left(\mathbf{X}_{\mathrm{R}} \mid \mathbf{X}_{\mathrm{D}}\right)$

The conditional probability of release (CPR) of a derailed tank car depends on its design characteristics and derailment speed (Treichel et al. 2006; Barkan 2008; Saat 2009; Kawprasert and Barkan 2010; Saat and Barkan 2011). Based on Kawprasert and Barkan (2010)'s results, Liu et al. (2013b) presented a linear regression model to estimate speed-dependent CPR for tank cars most commonly used to transport hazardous materials. Table 4.2 presents the minimum tank car safety design required by the USDOT (more tank car specification information can be found in Saat (2009)). In the linear model, the slope parameter A represents the change in CPR in response to a 1 mph change in derailment speed. For example, for a non-insulated 111A100W1 tank car, the speed-dependent CPR function is 0.0096 S , where S is derailment speed in mph. At 25 mph , the CPR for this tank car is $0.0096 \times 25=0.24$. Different tank car safety designs could have different release probabilities. For example, at the same derailment speed, the CPR of a derailed 111A100W1 tank car is higher than that of a 112J340W tank car by a factor of five. This difference in tank-car-design-specific release probability must be taken into account in estimating the number of tank cars releasing.

## Table 4.2

Estimated speed-dependent CPR, grouped by CPR function

|  | Minimum Tank Car Design |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Stenciled <br> Specification | Head Thickness <br> (inch) | Shell Thickness <br> (inch) | Jacket | Bottom Outlet | Test Pressure <br> (psig) | Speed-Dependent <br> CPR (A×S) |
| 111A100W1 | 0.4375 | 0.4375 | No | Optional | 100 | 0.0096 S |
| 111A100W3 | 0.4375 | 0.4375 | Yes | Optional | 100 | 0.0096 S |
| 111A100W2 | 0.4375 | 0.4375 | No | No | 100 | 0.0096 S |
| 211A100W1 | 0.4375 | 0.4375 | No | Optional | 100 | 0.0096 S |
| 111A100W5 | 0.4375 | 0.4375 | No | No | 100 | 0.0096 S |
| 112J340W | 0.6875 | 0.6875 | Yes | No | 340 | 0.0018 S |
| 105J300W | 0.6875 | 0.6875 | Yes | No | 300 | 0.0018 S |
| 112J400W | 0.6875 | 0.6875 | Yes | No | 400 | $0.0018 S$ |
| 105J400W | 0.6875 | 0.6875 | Yes | No | 400 | $0.0018 S$ |
| 105J500W | 0.6875 | 0.6875 | Yes | No | 500 | $0.0012 S$ |

" S " represents derailment speed in mph

Tank car release probability is a Bernoulli variable. The sum of Bernoulli variables with different probabilities follows a Poisson binomial distribution (Wang 1993):

$$
\begin{equation*}
P\left(X_{R}=k \mid X_{D 1}, \ldots, X_{D m}\right)=\sum_{A \in F_{k}}\left(\prod_{i \in A} P_{i} \prod_{j \in A^{c}}\left(1-P_{j}\right)\right) \tag{4.5}
\end{equation*}
$$

Where:
$X_{R} \quad=$ number of tank cars releasing per train derailment
$X_{D m}=$ number of the $m^{\text {th }}$ type of tank cars derailed
$\mathrm{F}_{\mathrm{k}} \quad=$ the set of all subsets of k integers selected from $\left\{1,2,3, \ldots, \mathrm{X}_{\mathrm{D} 1}+\ldots+\mathrm{X}_{\mathrm{Dm}}\right\}$
A $\quad=$ subset of $F_{k}$
$A^{c} \quad=$ the complement of set $A$ (i.e., $A^{c}=\left\{1,2,3, \ldots, X_{D 1}+\ldots+X_{D m}\right\} \backslash A$ )
$P_{i} \quad=$ the conditional probability release of a specific type of tank car

The probability mass function (PMF) of a Poisson binomial distribution can be derived using discrete Poisson approximation (Le Cam 1960), recursive formulae (Barlow and Heidtmann 1984), normal approximation (Volkova 1996) or Fourier transformation (Fernández and Williams 2010). A summary of the Poisson binomial distribution can be found in Hong (2013).

### 4.4 Numerical example

In this section, I present a numerical example to illustrate the application of the methodology for assessing the probability distribution of the number of tank cars releasing per train derailment. Figure 4.6 shows the analytical procedure.


Figure 4.6: Procedure for assessing the distribution of the number of tank cars releasing per train derailment

In this example, I assumed that the train has two head-end locomotives and eighty loaded cars, among which five are 111A100W1 tank cars and five are 112J340W tank cars. I also assumed that the train was derailed due to a broken rail, at a speed of 40 mph . Figure 4.7 shows the estimated distribution of the number of tank cars derailing and releasing based on these assumptions.


Figure 4.7: Estimated probability distribution of the number of tank cars derailing and releasing per broken-rail-caused train derailment for an 82-car train with 10 hazardous materials cars

In this example, the probability of no tank car releasing is 0.64 , the probability of releasing one tank car is 0.25 and the probability of more than one car releasing is 0.11 .

In the following sub-sections, I evaluate the effectiveness of several strategies, individually and in combination, to reduce the occurrence of multiple-tank-car release incidents. For comparison, the train operational information in the numerical example section above is used as the baseline scenario.

### 4.4.1 Train speed reduction

Train speed has a two-fold effect on the number of tank cars releasing. First, on average, lower speed derailments result in fewer cars derailed (Nayak et al. 1983; Saccomanno and El-Hage 1989, 1991; Barkan et al. 2003; Anderson 2005). Second, as already discussed, lower derailment speed results in a lower release probability of a derailed tank car (Barkan et al. 2003; Treichel et al. 2006; Kawprasert and Barkan 2010). Figure 4.8 compares the distribution of tank car releases by derailment speed. The train above derailing at 40 mph has a probability of releasing two-ormore tank cars of 0.12 and a mean number of tank cars releasing of 0.51 . When the speed of derailment is 25 mph , the probability of a multiple-tank-car release incident reduces to 0.04 (a $67 \%$ reduction), and the mean number of tank cars releasing reduces to 0.25 (a $51 \%$ reduction). Note that although a higher speed track segment may have a larger number of tank cars releasing in a derailment (and also have greater release quantities), the derailment rate may be lower due to higher track quality required for higher speed operation (Nayak et al. 1983; Treichel and Barkan 1993; Anderson and Barkan 2004; Kawprasert and Barkan 2010). Addressing all of the tradeoffs regarding the safety effects of train speed reduction is beyond the scope of this chapter;
however, it is important to consider them in the larger risk management framework in future applications of the methodology.


Figure 4.8: Estimated probability distribution of the number of tank car releasing per broken-rail-caused derailment by derailment speed for an 82-car train with 10 hazardous materials cars

### 4.4.2 Tank car safety design improvement

Tank car safety design improvement has been recognized as a risk reduction strategy (Barkan et al. 1991; Barkan 2008; Saat 2009; Saat and Barkan 2011). A 112J340W tank car is thicker and has a jacket, head shields, top fittings protection, and no bottom fittings, thus its estimated release probability is $81 \%$ lower than that of a non-insulated 111A100W1 tank car at the same derailment speed (Table 4.2). Figure 4.9 compares the number of tank cars releasing if all five 111 A 100 W 1 are replaced by 112 J 340 W cars. The probability of a multiple-tank-car release
declines from 0.12 to 0.017 (an $86 \%$ reduction). The average number of cars releasing declines from 0.51 to 0.16 , a $69 \%$ reduction.


Figure 4.9 Probability distribution of the number of tank car releasing per broken-rail-caused derailment by tank car specification for an 82-car train with 10 hazardous materials cars

Note: The baseline scenario is that a train contains five 111A100W1 tank cars and five 112 J 340 W tank cars. The upgraded scenario means that the same train carries ten 112J340W tank cars

Note that using more robust tank cars may reduce the tank cars' lading capacity. Consequently, this may increase the number of shipments required to meet the same traffic demand, thereby potentially increasing train derailment occurrences (Barkan 2008). However, the enhanced safety performance of the heavier, more damage-resistant tank cars, more than offsets the increased exposure (Barkan et al. 1991).

### 4.4.3 Change of tank car position

Previous studies have suggested that tank car derailment probability is affected by its position in the train. Thus, changing tank car placement was identified as a potential risk reduction strategy (Bagheri 2009; Bagheri et al. 2011). In the baseline scenario, I assumed that tank cars were randomly distributed throughout the train. To understand the effect of tank car placement, I analyzed the following alternative tank car placement scenarios:

- worst-case: tank cars are in the positions that are the most prone to derailment
- best-case: tank cars are in the positions that are the least prone to derailment

In order to estimate the number of tank cars released in the worst (or best) tank car placement scenario, I developed a derailment profile, which is the conditional probability of derailment by position-in-train after a train derailment occurs (Saccomanno and El-Hage 1989, 1991) (Figure 4.10). For an eighty-two-car train derailing at 40 mph due to a broken rail, the car in the $28^{\text {th }}$ position is the most prone to derail with a probability of 0.293 . By comparison, the last car in the train has the lowest derailment probability, 0.043 .


Figure 4.10: Derailment profile for a broken-rail-caused train derailment for an 82-car train with 10 hazardous materials cars

For each tank car in the train, its derailment probability is a Bernoulli variable. As stated in Section 4.3.4, the sum of Bernoulli variables with different probabilities follows a Poisson binomial distribution (Appendix D). Figure 4.11 compares the distributions of the number of tank cars releasing by different tank car placement scenarios. The "random" represents the baseline scenario in which tank cars are randomly distributed in the train and modeled by a multivariate hyper-geometric distribution. In the "best-case" scenario, the probability of a multiple-tank-car release (two-or-more cars releasing) is only 0.01 , compared to 0.14 in the "worst-case" scenario. Overall, the mean number of tank cars releasing per derailment ranges from 0.19 to 0.66 , depending on tank car placement.


Figure 4.11: Probability distribution of the number of tank cars releasing per broken-rail-caused train derailment by tank car placement for an 82-car train with 10 hazardous materials cars

The actual placement of tank cars is subject to regulatory, operational and engineering constraints (FRA 2005). A comprehensive cost-benefit analysis of "optimal" tank car marshaling strategy should account for these effects on transportation safety and efficiency.

### 4.4.4 Integrated strategies for reducing the number of tank cars releasing per derailment

In addition to analyzing the marginal effect of each strategy, it is also important to evaluate their combined safety benefit in terms of reducing the number of tank cars releasing per derailment.

Table 4.3 shows that when all three strategies are applied, the probability of a multiple-tank-car release reduces from 0.115 (baseline scenario) to 0.001 . The mean number of tank cars releasing reduces from 0.51 to 0.036 .

Table 4.3
Safety effectiveness of integrated strategies to reduce the number of tank cars releasing per train derailment

| Scenario | Strategies |  |  | Probability of Number of Tank Cars Releasing |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Speed Reduction | Tank Car Upgrade | Tank Car Placement | No Release | Single Car Release | Multiple Car Releases | Mean Number of Cars Releasing |
| 1 |  |  |  | 0.640 | 0.245 | 0.115 | 0.510 |
| 2 | Y |  |  | 0.792 | 0.171 | 0.037 | 0.250 |
| 3 |  | Y |  | 0.851 | 0.131 | 0.018 | 0.160 |
| 4 |  |  | Y | 0.827 | 0.159 | 0.014 | 0.190 |
| 5 | Y | Y |  | 0.926 | 0.069 | 0.005 | 0.079 |
| 6 | Y |  | Y | 0.908 | 0.088 | 0.004 | 0.096 |
| 7 |  | Y | Y | 0.930 | 0.068 | 0.002 | 0.073 |
| 8 | Y | Y | Y | 0.964 | 0.035 | 0.001 | 0.036 |

Notes:

1) " $Y$ " means that a strategy is implemented
2) "Speed reduction" means that train speed reduces from 40 mph to 25 mph
3) "Tank car upgrade" means that all five 111A100W1 tank cars are replaced by 112 J 340 W cars
4) "Tank car placement" means that all tank cars are in the positions that are the least likely to derail

Note that these risk reduction strategies, individually or in combination, have associated safety benefits and costs. When combined with cost information, the methodology presented here can be further developed and incorporated into a larger risk management framework to optimize the integration of multiple risk management options in the most cost-efficient manner (Liu et al. 2013c).

### 4.5 Conclusion

I develop a generalized model to evaluate the probability distribution of the number of tank cars releasing in a train derailment accounting for specified operational characteristics. The analysis shows that reducing train speed, improving tank car safety design and changing tank car placement all have the potential to reduce the number of tank cars releasing per derailment. In addition, there are interactive effects among different risk reduction strategies. This analysis accounts for a variety of factors affecting the probability of a multiple-tank-car release incident.

The model can potentially aid development of better-informed, more effective risk management policies and practices to improve hazardous materials transportation safety.

### 4.6 Future research

This study presents the first step in a systematic approach to railway hazardous materials transportation risk management. Several areas require further study:

1) Tank cars releases from multiple train derailments. The next step of this work is to incorporate the rate or total number of train derailments (Glickman and Rosenfield 1984) and evaluate total number of tank cars released from multiple train derailments.
2) Consequence modeling. Consequence modeling of hazardous materials release incidents is beyond the scope of this research, but it is an important element of a more comprehensive approach to risk assessment. Further research is needed to better evaluate the consequences of hazardous materials release incidents, thereby facilitating a more fine-grained risk analysis.
3) Other accident causes. Further analysis is needed to account for other types of accident causes, such as mechanical failure and human factors (Liu et al. 2013d). The model developed herein can be adapted to evaluate the impact of current and emerging technologies and possible safety regulations.
4) Evaluate cost-efficiency of risk reduction strategies. The next step is to evaluate the cost-effectiveness of these strategies individually and in combination. Ultimately, a decision support system will be developed to address which strategies to invest in and the optimal way to allocate resources under various conditions and circumstances.

## References

Anderson, R.T., Barkan, C.P.L. (2004). Railroad accident rates for use in transportation risk analysis. Transportation Research Record 1863, 88-98.

Anderson, R.T. (2005). Quantitative analysis of factors affecting railroad accident probability and severity. M.S. thesis, University of Illinois at Urbana-Champaign, Urbana, IL.

Appel, W., Rober, C., Davis, N., Zubrow, A., Howard, S.C. (2011). Overview of the atmospheric model evaluation tool (AMET) v1.1 for evaluating meteorological and air quality models. Environmental Modelling \& Software 26(4), 434-443.

Association of American Railroads (AAR) (2013a). Railroads and chemical. https://www.aar.org/keyissues/Documents/Background-Papers/Railroads-and Chemicals.pdf.

Association of American Railroads (AAR) (2013b). Class I railroad statistics. https://www.aar.org/StatisticsAndPublications/Documents/AAR-Stats-2013-01-10.pdf.

Bagheri, M. (2009). Risk-based model for effective marshaling of dangerous goods railway cars. Ph.D. Dissertation, Department of Civil and Environmental Engineering, University of Waterloo, Ontario.

Bagheri, M., Saccomanno, F.F., Chenouri, S., Fu, L.P. (2011). Reducing the threat of in transit derailments involving dangerous goods through effective placement along the train consist. Accident Analysis and Prevention 43(3), 613-620.

Bagheri, M., Verma, M., Verter, V. (2013). Transport mode selection for toxic gases: rail or road? Risk Analysis (in press) DOI: 10.1111/risa.12063.

Barkan, C.P.L., Glickman, T.S., Harvey, A.E. (1991). Benefit cost evaluation of using different specification tank cars to reduce the risk of transporting environmentally sensitive chemicals. Transportation Research Record 1313, 33-43.

Barkan, C.P.L., Dick, C.T., Anderson, R.T. (2003). Analysis of railroad derailment factors affecting hazardous materials transportation risk. Transportation Research Record 1825, 64-74.

Barkan, C.P.L. (2008). Improving the design of higher-capacity railway tank cars for hazardous materials transport: optimizing the trade-off between weight and safety. Journal of Hazardous Materials, 122-134.

Barlow, R.E., Heidtmann, K.D. (1984). Computing k-out-of $n$ system reliability. IEEE Transaction Reliability 33, 322-323.

Birk, A.M., Anderson, R.J., Coppens, A.J. (1990). A computer simulation of a derailment accident: part I - model basis. Journal of Hazardous Materials 25, 121-147.

Bureau of Explosives (BOE) (2013). Annual report of hazardous materials transported by rail: calendar year 2012. Association of American Railroads, Washington D.C.

Center for Chemical Process Safety (CCPS) (2008). Guidelines for chemical transportation safety, security, and risk management (second edition). American Institute of Chemical Engineers, Hoboken, NJ.

Corder, G.W., Foreman, D.I. (2009). Nonparametric statistics for non-statisticians: a step-by-step approach. Wiley, Hoboken, New Jersey.

Erkut, E., Verter, V. (1995). A framework for hazardous materials transport risk assessment. Risk Analysis 15(5), 589-601.

Erkut, E., Verter, V. (1998). Modeling of transport risk for hazardous materials. Operations Research 46(5), 625-642.

Erkut, E., Ingolfsson, A. (2000). Catastrophe avoidance models for hazardous materials route planning. Transportation Science 34(2), 165-179.

Erkut, E., Ingolfsson, A. (2004). Transportation risk models for hazardous materials: revisited. Operations Research Letters 33(1), 81-89.

Federal Railroad Administration (FRA) (2005). Safe placement of train cars: a report to the senate committee on commerce, science and transportation and the house committee on transportation and infrastructure. U.S. Department of Transportation, Washington D.C.

Federal Railroad Administration (FRA) (2013). Railroad equipment accident/incident reporting threshold 1957-2013. FRA, U.S. Department of Transportation, Washington D.C.

Fernández, M., Williams, S. (2010). Closed-form expression for the Poisson-binomial probability density function. IEEE Transactions on Aerospace and Electronic Systems 46, 803817.

Glickman, T.S., Rosenfield, D.B. (1984). Risks of catastrophic derailments involving the release of hazardous materials. Management Science 30(4), 503-511.

Glickman, T.S., Erkut, E., Zschocke, M.S. (2007). The cost and risk impacts of rerouting railroad shipments of hazardous materials. Accident Analysis and Prevention 39(5), 1015-1025.

Hanna, S.R., Chang, J.C., Strimaitis, D.G. (1993). Hazardous gas model evaluation with field observations. Atmospheric Environment 27A(15), 2265-2285.

Health and Safety Executive (HSE) (1991). Major hazard aspects of the transport of dangerous substance. Health and Safety Executive, Great Britain.

Hong, Y. (2013). On computing the distribution function for the Poisson binomial distribution. Computational Statistics and Data Analysis 59, 41-51.

Kawprasert, A., Barkan, C.P.L. (2010). Effect of train speed on risk analysis of transporting hazardous materials by rail. Transportation Research Record 2159, 59-68.

Kawprasert, A. (2010). Quantitative analysis of options to reduce risk of hazardous materials transportation by railroad. Ph.D. Dissertation, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Le Cam, L. (1960). An approximation theorem for the Poisson binomial distribution. Pacific Journal of Mathematics 10, 1181-1197.

Liu, X., Barkan, C.P.L., Saat, M.R. (2011). Analysis of derailments by accident cause: evaluating railroad track upgrades to reduce transportation risk. Transportation Research Record 2261, 178-185.

Liu, X., Saat, M.R., Barkan, C.P.L. (2012). Analysis of causes of major train derailment and their effect on accident rates. Transportation Research Record 2289, 154-163.

Liu, X., Saat, M.R., Barkan, C.P.L. (2013a). Analysis of U.S. freight-train derailment severity using zero-truncated negative binomial regression and quantile regression. Accident Analysis and Prevention 59, 87-93.

Liu, X., Saat, M.R., Barkan, C.P.L. (2013b). Safety effectiveness of integrated risk reduction strategies for rail transport of hazardous materials. Transportation Research Record 2374: 102110.

Liu, X., Saat, M.R., Barkan, C.P.L. (2013c). Integrated risk reduction framework to improve railway hazardous materials transportation safety. Journal of Hazardous Materials 260, 131-140.

Liu, X., Saat, M.R., Barkan, C.P.L. (2013d). Safety effectiveness of integrated accident prevention strategies. Proceedings of the Joint Rail Conference. Knoxville, TN.

Liu, X., Saat, M.R., Barkan, C.P.L. (2014). Analysis of freight-train derailment rates. Accident Analysis and Prevention (under review).

Marruffo, A., Yoon, H., Schaeffer, D.J., Barkan, C.P.L., Saat, M.R., Werth, C.J. (2012). NAPL source zone depletion model and its application to railroad-tank-car spills. Ground Water 50(4), 627-632.

Nayak, P.R., Rosenfield, D.B., Hagopian, J.H. (1983). Event probabilities and impact zones for hazardous materials accidents on railroads. Report DOT/FRA/ORD-83/20, FRA, U.S. Department of Transportation, Washington D.C.

Pipeline and Hazardous Materials Safety Administration (PHMSA) (2012).
Emergency response guidebook.
http://www.phmsa.dot.gov/staticfiles/PHMSA/DownloadableFiles/Files/Hazmat/ERG2012.pdf.
Raj, P.K., Turner, C.K. (1992). Hazardous material transportation in tank cars: analysis of risks. Federal Railroad Administration, DOT/FRA/ORD-92/34.

Ross, S. (2007). Introduction to probability models, Academic Press, Waltham, Massachusetts.
Saat, M.R. (2009). Optimizing railroad tank car safety design to reduce hazardous materials transportation risk. Ph.D. Dissertation, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Saat, M.R., Barkan, C.P.L. (2011). Generalized railway tank car safety design optimization for hazardous materials transport: addressing the trade-off between transportation efficiency and safety. Journal of Hazardous Materials 189, 62-68.

Saccomanno, F.F., Shortreed, J.H., Aerde, M.V., Higgs, J. (1989). Comparison of risk measures for the transport of dangerous commodities by truck and rail. Transportation Research Record 1245, 1-13.

Saccomanno, F.F., El-Hage, S.M. (1989). Minimizing derailments of railcars carrying dangerous commodities through effective marshaling strategies. Transportation Research Record 1245, 3451.

Saccomanno, F.F., El-Hage, S.M. (1991). Establishing derailment profile by position for corridor shipments of dangerous goods. Canadian Journal of Civil Engineering 18(1), 67-75.

Treichel, T.T., Barkan, C.P.L. (1993). Working paper on mainline freight train accident rates. Research and test department, Association of American Railroads, Washington, D.C.

Treichel, T.T., Hughes, J.P., Barkan, C.P.L., Sims, R.D., Phillips, E.A., Saat, M.R. (2006). Safety performance of tank cars in accidents: probability of lading loss, report RA-05-02. RSIAAR railroad tank car safety research and test project, Association of American Railroads, Washington, D.C.

Verma, M., Verter, V. (2007). Railroad transportation of dangerous goods: population exposure to airborne toxins. Computers \& Operations Research 34(5), 1287-1303.

Verter, V., Kara, B.Y. (2001). A GIS-based framework for hazardous materials transport risk assessment. Risk Analysis 21(6), 1109-1120.

Volkova, A.Y. (1996). A refinement of the central limit theorem for sum of independent random indicators. Theory of Probability and Its Applications 40, 791-794.

Wang, Y.H. (1993). On the number of successes in independent trials. Statistica Sinica 3, 295312.

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

Zhang, J., Hodgson, J., Erkut, E. (2000). Using GIS to assess the risks of hazardous materials transport in networks. European Journal of Operational Research 121, 316-329.

## Chapter 5

## Risk Analysis of Multiple-Tank-Car Release Accidents

### 5.1 Introduction

U.S. petroleum crude oil production has increased significantly in part attributed to advances in hydraulic fracturing and horizontal drilling technologies. This has resulted in a considerable increase in rail shipments of petroleum crude oil. For example, U.S. Class I railroads originated 9,500 carloads of petroleum crude oil in 2008, compared to 234,000 carloads in 2012, a 2,400\% increase (AAR 2013). Most of this traffic is shipped in unit trains, or large blocks of cars in trains. This has raised new questions about the safety implications of large, consolidated groups of tank cars transporting hazardous materials. There is particular concern regarding flammable liquids because of the potential for large-quantity releases, and possible interactive effects with other cars if there is a fire, such as occurred in the accident in Lac-Mégantic, Canada in July 2013. Consequently, there is interest in understanding the risk of transporting petroleum crude oil in unit trains.

Most previous analyses have focused on the consequences of hazardous materials releases from one or a few cars, although there are a few exceptions (Glickman and Rosenfield 1984; Glickman et al. 2007; Bagheri et al. 2013). Glickman and Rosenfield (1984) were the first to consider the risk of large, catastrophic tank car release incidents; however, their work did not account for the effect of point of derailment (POD), speed or accident cause on tank car derailment probability. Bagheri et al. (2013) considered POD, speed and accident cause in modeling tank car derailment probability, but did not account for tank car safety design and
speed in the estimation of tank car release probability. Furthermore, previous analyses have not accounted for localized differences in train derailment probability accounting for multiple infrastructure features, such as FRA track class (Nayak et al. 1983; Treichel and Barkan 1993; Anderson and Barkan 2004), method of operation and traffic density (Liu et al. 2014a).

In this chapter, I develop a new multiple-car release risk model that extends and generalizes previous models by simultaneously accounting for factors including FRA track class, method of operation, traffic density, train length, speed, point of derailment, accident cause, number and placement of tank cars, tank car safety design and population density.

This model is applicable to any train configuration, in which unit trains are a special case. The unit-train risk model is developed to evaluate the risk of mainline shipments of flammable liquids. The purpose of this study is to understand the risk of multiple-car release incidents and gain insight into the effectiveness of strategies to mitigate this risk.

The exposition of this chapter is as follows. First, I develop a methodology to estimate the risk of multiple-car release incidents. Second, I use this methodology to estimate the national risk of flammable liquid transportation. Third, I analyze the effect of tank car safety design improvement on risk reduction, especially prevention of multiple tank car releases. Finally, I highlight primary findings, discuss uncertainties in the risk analysis, and propose future research directions.

### 5.2 Methodology

Most major railroad hazardous materials release incidents occur as a result of train derailments. Derailment probability is affected by FRA track class (Nayak et al. 1983; Treichel and Barkan 1993; Anderson and Barkan 2004), method of operation and traffic density (Liu et al. 2014a). After a train is derailed (train derailment probability is denoted by $\mathrm{P}_{\mathrm{i}}(\mathrm{TD})$ ), the probability distribution of the number of tank cars releasing $\left(\mathrm{P}_{\mathrm{i}}\left(\mathrm{X}_{\mathrm{R}} \mid \mathrm{TD}\right)\right)$ can be estimated using the model developed by Liu et al. (2014b). Based on train derailment probability and the number of cars releasing per derailment, the distribution of the number of tank cars releasing per train shipment can be estimated using the Law of Total Probability (Ross 2007):

$$
\begin{equation*}
P\left(X_{R}\right)=\sum_{i=1}^{N}\left[P_{i}(T D) \times P_{i}\left(X_{R} \mid T D\right)\right] \tag{5.1}
\end{equation*}
$$

where:
$\mathrm{P}\left(\mathrm{X}_{\mathrm{R}}\right) \quad=$ probability distribution of the number of tank cars releasing per train shipment on a route
$\mathrm{P}_{\mathrm{i}}(\mathrm{TD}) \quad=$ probability of a train derailment prior to the $\mathrm{i}+1$ track segment
$\mathrm{P}_{\mathrm{i}}\left(\mathrm{X}_{\mathrm{R}} \mid \mathrm{TD}\right) \quad=$ probability distribution of the number of tank cars releasing if a train derails prior to the $\mathrm{i}+1$ track segment
$\mathrm{N} \quad=$ total number of track segments on the route

The probability that a train derails prior to the $\mathrm{i}+1$ segment is the product of the probability that there is no derailment on the prior segments and the derailment probability on the $\mathrm{i}^{\text {th }}$ segment (Erkut et al. 2007; Kawprasert 2010):
$P_{i}(T D)=\left[\prod_{j=1}^{i-1}\left(1-P_{j}\right)\right] P_{i}$
where:
$P_{i} \quad=$ probability of a train derailment on the $\mathrm{i}^{\text {th }}$ segment

Train derailment rate is sufficiently small (on the order of $10^{-6}$ per train-mile) (Nayak et al. 1983;
Treichel and Barkan 1993; Anderson and Barkan 2004; Liu et al. 2014a), thus derailment probability on a segment is approximately equal to the product of derailment rate and segment length (Kawprasert 2010):
$P_{i} \approx 1-\exp \left(-Z_{i} L_{i}\right) \approx Z_{i} L_{i}$
where:
$\mathrm{Z}_{\mathrm{i}} \quad=$ train derailment rate per train-mile
$\mathrm{L}_{\mathrm{i}} \quad=$ segment mileage

Furthermore, because segment-specific derailment probability is sufficiently small,

$$
\begin{equation*}
\prod_{j=1}^{i-1}\left(1-P_{j}\right) \approx 1 \tag{5.4}
\end{equation*}
$$

Based on Equations (5.2) to (5.4), train derailment probability is approximately equal to the product of train derailment rate and segment length:

$$
\begin{equation*}
P_{i}(T D) \approx Z_{i} L_{i} \tag{5.5}
\end{equation*}
$$

Train derailment rate ( $\mathrm{Z}_{\mathrm{i}}$ ) was developed by Liu et al. (2014a), and the number of tank cars releasing per train derailment, $\mathrm{P}\left(\mathrm{X}_{\mathrm{R}} \mid \mathrm{TD}\right)$, was developed by Liu et al. (2014b). Based on these results, the number of tank cars releasing per train shipment on a route can be estimated using the following equation:

$$
\begin{equation*}
P\left(X_{R}\right) \approx \sum_{i=1}^{N}\left\{Z_{i} \times L_{i} \times \sum_{X_{D}=0}^{\min \left(L_{r}, X\right)}\left\{P_{i}\left(X_{R} \mid X_{D}\right)\left\{\sum_{X=0}^{L-K+1} P_{i}\left(X_{D} \mid X\right)\left[\sum_{K=1}^{L} P_{i}(X \mid K) P O D_{i}(K)\right]\right\}\right\}\right\} \tag{5.6}
\end{equation*}
$$

Where:
$\mathrm{P}\left(\mathrm{X}_{\mathrm{R}}\right) \quad=$ probability distribution of the number of tank cars releasing per shipment
$\mathrm{N} \quad=$ number of track segments on the route
$Z_{i} \quad=$ train derailment rate per train-mile on the $i^{\text {th }}$ segment
$\mathrm{L}_{\mathrm{i}} \quad=$ length of the $\mathrm{i}^{\text {th }}$ segment (miles)
K = point of derailment (the position of the first car derailed)
$\mathrm{L} \quad=$ train length (total number of locomotives and railcars)
X = number of cars (tank and non-tank cars including locomotives) derailed
$X_{D} \quad=$ number of tank cars derailed
$\mathrm{X}_{\mathrm{R}} \quad=$ number of tank cars releasing
$\mathrm{L}_{\mathrm{T}} \quad=$ number of tank cars in the train
$\operatorname{POD}_{\mathrm{i}}(\mathrm{K}) \quad=$ probability distribution of point of derailment
$\mathrm{P}_{\mathrm{i}}(\mathrm{X} \mid \mathrm{K}) \quad=$ probability distribution of the number of cars derailed given point of derailment
$\mathrm{P}_{\mathrm{i}}\left(\mathrm{X}_{\mathrm{D}} \mid \mathrm{X}\right) \quad=$ probability distribution of the number of tank cars derailed given total number of all types of cars derailed
$\mathrm{P}_{\mathrm{i}}\left(\mathrm{X}_{\mathrm{R}} \mid \mathrm{X}_{\mathrm{D}}\right) \quad=$ probability distribution of the number of tank cars releasing given the number
of tank cars derailed

Figure 5.1 illustrates the analytical procedure to estimate the number of tank cars releasing per train shipment. The model accounts for a variety of track-related, rolling-stock-related and operational characteristics. The occurrence of a train derailment was assumed to follow a Poisson process (Glickman and Rosenfield 1984; Kawprasert 2010). The point of derailment (POD) was assumed to follow a Beta distribution (Anderson 2005). Given the POD, the number of all types of cars derailed was assumed to follow a truncated geometric distribution (Saccomanno and ElHage 1989, 1991). If tank cars were randomly distributed throughout a train, the number of tank cars derailed given the number of cars derailed was assumed to follow a multivariate hyper-geometric distribution (Liu et al. 2014b). Finally, given the number of different types of tank cars derailed, the number of tank cars releasing was assumed to follow a Poisson binomial distribution (Liu et al. 2014b). The integration of all these probability distributions was used to estimate the probability distribution of the number of tank cars releasing per train shipment.

The model is applicable to any train configuration on any route. A special case of this model was used for unit-train transportation risk analysis. In the next section, I develop a numerical example to illustrate the application of the methodology to nationwide flammable liquid transportation risk assessment.


Figure 5.1: Analytical procedure for estimating multiple-car release probability distribution

### 5.3 Case study

### 5.3.1 Introduction

The purpose of this case study is to assess railroad transportation risk of flammable liquids in the United States. Due to data constraints, a number of assumptions were made based on the information available to develop a high level, preliminary risk assessment. In the following subsections, I introduce the input parameters and assumptions used in this study. These assumptions are applicable to a national average risk assessment. If route-specific information is
available, it can be used to develop more accurate national, regional or route-specific estimates using this risk model.

### 5.3.2 Rail network information

U.S. Class I mainline rail network information was obtained from National Transportation Atlas Database. For each segment on this network, I identified its FRA track class and method of operation using railroad-provided infrastructure information. Approximately 83,000 route miles (41,704 segments) owned or operated by U.S. Class I railroads were analyzed in this case study (Table 5.1).

## Table 5.1

Route mileage distribution by FRA track class and method of operation, U.S. Class I mainlines

|  | FRA Track Class |  |  |  |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: | :---: |
| Method of |  |  |  |  |  | Total |  |
| Operation | 1 | 2 | 3 | 4 | 5 | $0.1 \%$ |  |
| Non-Signaled | $4.3 \%$ | $9.0 \%$ | $9.8 \%$ | $8.8 \%$ | $0 \%$ |  |  |
| Signaled | $0.9 \%$ | $3.2 \%$ | $10.4 \%$ | $40.5 \%$ | $12.9 \%$ | $68.0 \%$ |  |
| Total | $5.1 \%$ | $12.2 \%$ | $20.3 \%$ | $49.4 \%$ | $13.0 \%$ | $100.0 \%$ |  |

Note: 18,371 route miles ( $18 \%$ of national network) have unknown information regarding FRA track class and/or method of operation, thus were excluded from this analysis.

### 5.3.3 Unit train information

The makeup of the unit train analyzed in the case study was as follows. It consisted of five headend locomotives and eighty non-insulated DOT 111A100W1 tank cars. I assumed that all cars were fully loaded and optimized for the product they were transporting and has a gross rail load of 263,000 pounds each. Each locomotive was assumed to weigh 187 tons (AAR 2009). Track-class-average train derailment speed was calculated based on FRA-reportable freight-train derailment data from 2002 to 2011 (Figure 5.2).


Figure 5.2: Average derailment speed by FRA track class, U.S. Class I mainline FRA-reportable freight-train derailments, 2002 to 2011

### 5.3.4 Population density information

I assumed that each flammable liquid release would result in a fire event. The U.S. Emergency Response Guide (ERG) recommends a corresponding 0.5-mile evacuation distance. Based on the ERG recommendation, I created a 0.5 -mile spatial buffer along each segment and overlaid it on the U.S. census tract map to estimate the route mileage distribution of different population densities ranging from remote ( $<20$ persons per square mile), up to extremely high ( $>10,000$ persons per square mile) along Class I mainlines (Table 5.2 and Figure 5.3).

Table 5.2
Route mileage distribution by population density along Class I mainlines

| Population Class | Population Density <br> (persons per square mile) | Route <br> Mileage | Percentage |
| :--- | :--- | ---: | ---: |
| Remote | $<20$ | 23,565 | $28.5 \%$ |
| Rural | $20-100$ | 23,788 | $28.7 \%$ |
| Suburban | $100-1,000$ | 24,308 | $29.4 \%$ |
| Urban | $1,000-3,000$ | 7,901 | $9.5 \%$ |
| High | $3,000-10,000$ | 3,014 | $3.6 \%$ |
| Extremely High | $>10,000$ | 218 | $0.3 \%$ |
| Total |  | 82,794 | $100.0 \%$ |



Figure 5.3: Distribution of population density along Class I mainlines

### 5.4 Flammable liquid transportation risk

Based on the input information above and using Equation (5.6), I estimated the probability distribution of multiple-car release incidents per unit-train mile (Table 5.3). Detailed calculation is presented in Appendix E.

### 5.4.1 Probability distribution of multiple-car release incidents

Table 5.3 presents the probability distribution of multiple-car release incidents per unit-train mile on Class I mainlines. For example, the probability of at least 10 tank cars released is approximately 1.51 in 100 million in areas with extremely high population density (> 10,000 persons per square mile).

Table 5.3
Probability distribution of multiple-car release incidents by population density

|  |  | Number of Tank Cars Releasing per Unit-Train Mile |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Population | Population Density (persons per |  |  |  |  |  |  |  |  |  |  |
| Class | square mile) | $\geq 1$ | $\geq 2$ | $\geq 3$ | $\geq 4$ | $\geq 5$ | $\geq 6$ | $\geq 7$ | $\geq 8$ | $\geq 9$ | $\geq 10$ |
| Remote | <20 | $9.80 \mathrm{E}-07$ | 6.10E-07 | 4.10E-07 | $2.88 \mathrm{E}-07$ | 2.07E-07 | 1.52E-07 | 1.13E-07 | 8.43E-08 | 6.35E-08 | $4.82 \mathrm{E}-08$ |
| Rural | 20-100 | $1.18 \mathrm{E}-06$ | 7.02E-07 | 4.56E-07 | 3.09E-07 | 2.16E-07 | 1.54E-07 | 1.12E-07 | 8.21E-08 | 6.07E-08 | $4.51 \mathrm{E}-08$ |
| Suburban | 100-1,000 | $1.23 \mathrm{E}-06$ | 7.10E-07 | 4.53E-07 | $3.04 \mathrm{E}-07$ | $2.11 \mathrm{E}-07$ | $1.50 \mathrm{E}-07$ | $1.08 \mathrm{E}-07$ | 7.87E-08 | $5.79 \mathrm{E}-08$ | $4.29 \mathrm{E}-08$ |
| Urban | 1,000-3,000 | $1.30 \mathrm{E}-06$ | 6.86E-07 | 4.20E-07 | $2.77 \mathrm{E}-07$ | 1.90E-07 | 1.34E-07 | 9.58E-08 | 6.96E-08 | $5.10 \mathrm{E}-08$ | $3.77 \mathrm{E}-08$ |
| High | 3,000-10,000 | $1.41 \mathrm{E}-06$ | 6.87E-07 | 3.98E-07 | 2.53E-07 | 1.69E-07 | 1.16E-07 | 8.19E-08 | 5.87E-08 | 4.26E-08 | $3.13 \mathrm{E}-08$ |
| Extremely High | >10,000 | $1.81 \mathrm{E}-06$ | 6.57E-07 | 3.09E-07 | 1.72E-07 | $1.05 \mathrm{E}-07$ | 6.74E-08 | 4.49E-08 | 3.07E-08 | $2.14 \mathrm{E}-08$ | $1.51 \mathrm{E}-08$ |
| Average |  | 1.16E-06 | 6.76E-07 | 4.36E-07 | 2.96E-07 | 2.08E-07 | 1.49E-07 | $1.08 \mathrm{E}-07$ | 7.95E-08 | $5.90 \mathrm{E}-08$ | $4.40 \mathrm{E}-08$ |

Note: The "average" row represents national average probability of multiple-car release incidents, irrespective of population density

I compared the multiple-car release risk in a remote area versus an area with extremely high population density (Figure 5.4). The analysis shows that although a high population area appears to have a higher probability of releasing one or two cars, it is less likely to have a large release incident.


Figure 5.4: Risk profile in different population densities

For safety reasons, the trackage in high population areas typically has speed restrictions and corresponding lower FRA track class based on the maximum operating speed. Given all else equal, lower track classes have been found to have higher derailment rates (Nayak et al. 1983; Treichel and Barkan 1993; Anderson and Barkan 2004; Liu et al. 2014a). However, this does not mean that most track design and maintenance standards on the lower speed sections are less than the higher-speed sections. If these low-speed segments are designed and maintained to higher standards than is evident from track class alone, it may lead to over-estimation of derailment likelihood under these circumstances (Liu et al. 2014a). Although a higher population area may not actually have a higher derailment rate due to many other factors (particularly traffic density) (Liu et al. 2014a), I did not have the data needed for this case study to account for them. This uncertainty may lead to over-estimation of the risk in a high population area (similarly, the risk in a remote area may be under-estimated).

### 5.4.2 Interval between occurrences of multiple-car release incidents

Using multiple-car release probabilities, I used national traffic level data to estimate the interval between occurrences of release incidents of various magnitudes. In 2012, there were 530,000 tank carloads of petroleum crude oil and Alcohols N.O.S. originated in the United States (BOE 2013). For purposes of illustration, I assumed that all these carloads were in unit trains consisting of eighty non-insulated DOT 111A100W1 tank cars with an average shipment distance of 1,000 miles. Finally, I assumed that the distribution of route characteristics of flammable liquid corridors was the same as the overall Class I mainline network. These assumptions were made to illustrate a preliminary nationwide risk assessment methodology and preliminary risk estimates. A more refined analysis would use detailed route information.

I estimated a total of 6.6 million unit-train miles on Class I mainlines in 2012. The traffic exposure (measured in unit-train miles) for each population density was estimated based on route mileages in different population densities (Table 5.4). For example, remote areas have $28.5 \%$ of route miles in the national network. The estimated unit-train miles in remote areas were 1.89 million ( $6.6 \times 28.5 \%$ ).

## Table 5.4

Traffic exposure by population density on Class I mainlines

| Population Class | Population Density <br> (persons per square mile) | Traffic Exposure (million <br> unit-train miles) |
| :--- | :--- | :---: |
| Remote | $<20$ | 1.89 |
| Rural | $20-100$ | 1.90 |
| Suburban | $100-1,000$ | 1.95 |
| Urban | $1,000-3,000$ | 0.63 |
| High | $3,000-10,000$ | 0.24 |
| Extremely High | $>10,000$ | 0.02 |
| Total |  | 6.63 |

I next estimated the expected interval (measured in years) between occurrences of multiple-car release incidents (Table 5.5 and Figure 5.5). For example, I estimated an average of 42 years (based on 2012 traffic) between ten-or-more-car-release events in urban areas ( 1,000 to 3,000 persons per square mile). By comparison, the interval in remote areas for a similar magnitude event is 11 years. The occurrence interval for multiple-car release incidents is longer in high population areas than in low population areas. Evaluation of the risk analysis indicates that this is due to relatively low exposure to urban or higher population density levels (13\%) and the lower operating speeds in these areas, resulting in the lower probability of tank car derailment and release.

## Table 5.5

Estimated interval between occurrences of multiple-car-release incidents on U.S. Class I mainlines, assuming 80 non-insulated, DOT 111A100W1 tank cars in a unit train

|  |  | Number of Tank Cars Releasing |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Population Class | Density (persons per square mile) | $\geq 1$ | $\geq 2$ | $\geq 3$ | $\geq 4$ | $\geq 5$ | $\geq 6$ | $\geq 7$ | $\geq 8$ | $\geq 9$ | $\geq 10$ |
| Remote | <20 | 0.5 | 0.9 | 1.3 | 1.8 | 2.6 | 3.5 | 4.7 | 6.3 | 8.3 | 11.0 |
| Rural | 20-100 | 0.4 | 0.7 | 1.2 | 1.7 | 2.4 | 3.4 | 4.7 | 6.4 | 8.7 | 11.6 |
| Suburban | 100-1,000 | 0.4 | 0.7 | 1.1 | 1.7 | 2.4 | 3.4 | 4.8 | 6.5 | 8.9 | 12.0 |
| Urban | 1,000-3,000 | 1.2 | 2.3 | 3.8 | 5.7 | 8.3 | 11.8 | 16.5 | 22.7 | 31.0 | 42.0 |
| High | 3,000-10,000 | 2.9 | 6.0 | 10.4 | 16.4 | 24.6 | 35.7 | 50.6 | 70.6 | 97.2 | 132.7 |
| Extremely High | >10,000 | 31.6 | 87.2 | 185.2 | 333.2 | 546.7 | 849.9 | 1,275.9 | 1,867.3 | 2,679.8 | 3,787.7 |
| Average |  | 0.1 | 0.2 | 0.3 | 0.5 | 0.7 | 1.0 | 1.4 | 1.9 | 2.6 | 3.4 |

Note: The "average" row represents expected interval between releases, irrespective of population

(a) Remote area ( $<20$ persons per square mile)

(b) Rural area (20 to 100 persons per square mile)

(c) Suburban area (100 to 1,000 persons per square mile)

Figure 5.5 (cont.)

(d) Urban area ( 1,000 to 3,000 persons per square mile)

(e) High population area ( 3,000 to 10,000 persons per square mile)

(f) Extremely high population area ( $>10,000$ persons per square mile)

Figure 5.5: Estimated interval between occurrences of multiple-car release incidents by population density

### 5.5 Tank car safety design improvement for risk reduction

The conditional probability of release of a derailed tank car is dependent on its safety design (Treichel et al. 2006; Barkan 2008; Saat 2009; Saat and Barkan 2011) and derailment speed (Treichel et al. 2006; Kawprasert and Barkan 2010). The DOT 111A100W1 is the most common type of tank car used for transporting flammable liquids in the United States (BOE 2013). Of interest is understanding the possible reduction in risk when certain design modifications are made to this type of tank car.

In 2011, the Association of American Railroads (AAR) petitioned the Pipeline and Hazardous Materials Safety Administration to adopt more stringent requirements for DOT-111 tank cars used to transport packing group I and II materials. That petition, P-1577, proposed a series of safety improvement options, such as top fittings protection, reclosing pressure relief valve, thicker tank car shell and head, and head shields (AAR 2011). Subsequently, AAR adopted an interchange standard (CPC-1232) with the same requirements applicable to tank cars ordered after October 1, 2011.

The CPC-1232 standard requires either a thicker tank shell and head; or a jacket. Both types of CPC-1232 compliant cars are required to be equipped with top fittings protection and a minimum half-inch, half height head shield (AAR 2011). In this section, I analyze the effectiveness of the following four types of tank cars in reducing the risk of multiple-car release incidents (Table 5.6).

Table 5.6
Minimum design specifications of tank cars

|  | Conventional Non-Jacketed 1111A100W1 | Conventional Jacketed <br> 111A100W1 | Non-Jacketed CPC-1232 Compliant Car | Jacketed CPC-1232 <br> Compliant Car |
| :---: | :---: | :---: | :---: | :---: |
| Head Thickness (inch) | 0.4375 | 0.4375 | 0.5 | 0.4375 |
| Shell Thickness (inch) | 0.4375 | 0.4375 | 0.5 | 0.4375 |
| Jacket | No | Yes | No | Yes |
| Head Shields | None | None | Half Height | Full Height |
| Top Fittings Protection | No | No | Yes | Yes |
| Bottom Fittings | Yes | Yes | Yes | Yes |
| Average Conditional Probability of Release | 0.35 | 0.21 | 0.24 | 0.15 |

Note: Conditional probability of release was estimated based on the method described in Treichel et al. (2006).

Treichel et al. (2006) developed a series of regression analyses to estimate CPR for most common designs of tank cars. Based on their work, Kawprasert and Barkan (2010) subsequently developed an approach to estimate speed-dependent CPR of a derailed tank car. Using Kawprasert and Barkan's results, I analyzed the relationship between CPR and speed by tank car safety design (Figure 5.6).


Figure 5.6: Speed-dependent conditional probability of release by tank car design

The CPR of a conventional non-jacketed, DOT-111 tank car is expected to decrease by 31 percent if it is replaced by a non-jacketed CPC-1232 car, and by 57 percent if it is changed to a jacketed CPC-1232 car.

Using the risk analysis model I developed, I estimated the interval between occurrences of multiple-car release incidents if different tank car designs are used (Table 5.7 and Figure 5.7). The analysis shows that use of jacketed CPC-1232 compliant cars can substantially reduce the likelihood of large, multiple-car release incidents. For example, in urban areas, there is an average of 10 -or-more-cars release incident every 40 years using the conventional non-jacketed, DOT 111A100W1 tank car. The interval for a similar magnitude event is 800 years if all the conventional, non-jacketed cars were replaced by jacketed CPC-1232 cars.

Table 5.7
Estimated interval between occurrence of multiple-car-release incidents on U.S. Class I mainlines, enhanced tank car safety designs, assuming 80 tank cars in a unit train
(a) conventional jacketed DOT 111A100W1

|  |  | Number of Tank Cars Releasing |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Population Class | Population Density (persons per square mile) | $\geq 1$ | $\geq 2$ | $\geq 3$ | $\geq 4$ | $\geq 5$ | $\geq 6$ | $\geq 7$ | $\geq 8$ | $\geq 9$ | $\geq 10$ |
| Remote | <20 | 0.7 | 1.2 | 2.1 | 3.4 | 5.4 | 8.4 | 13.0 | 19.8 | 30.2 | 46.1 |
| Rural | 20-100 | 0.5 | 1.1 | 1.9 | 3.3 | 5.4 | 8.7 | 13.7 | 21.5 | 33.6 | 52.7 |
| Suburban | 100-1,000 | 0.5 | 1.1 | 1.9 | 3.3 | 5.5 | 8.9 | 14.1 | 22.2 | 34.8 | 54.7 |
| Urban | 1,000-3,000 | 1.5 | 3.4 | 6.5 | 11.4 | 19.0 | 30.8 | 49.3 | 78.0 | 122.8 | 193.5 |
| High | 3,000-10,000 | 3.8 | 9.3 | 18.7 | 33.8 | 57.9 | 95.8 | 155.0 | 247.1 | 391.0 | 617.0 |
| Extremely High | >10,000 | 43.2 | 149.4 | 371.0 | 765.9 | 1,436.8 | 2,548.2 | 4,356.6 | 7,265.7 | 11,924.6 | 19,395.2 |
| Average |  | 0.2 | 0.3 | 0.6 | 1.0 | 1.6 | 2.6 | 4.0 | 6.3 | 9.8 | 15.2 |

(b) non-jacketed CPC-1232 compliant cars

|  |  | Number of Tank Cars Releasing |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Population Class | Population Density (persons per square mile) | $\geq 1$ | $\geq 2$ | $\geq 3$ | $\geq 4$ | $\geq 5$ | $\geq 6$ | $\geq 7$ | $\geq 8$ | $\geq 9$ | $\geq 10$ |
| Remote | <20 | 0.6 | 1.1 | 1.8 | 2.9 | 4.4 | 6.5 | 9.6 | 14.1 | 20.7 | 30.1 |
| Rural | 20-100 | 0.5 | 1.0 | 1.7 | 2.7 | 4.3 | 6.6 | 10.0 | 15.1 | 22.6 | 33.7 |
| Suburban | 100-1,000 | 0.5 | 0.9 | 1.7 | 2.7 | 4.3 | 6.7 | 10.3 | 15.5 | 23.3 | 34.9 |
| Urban | 1,000-3,000 | 1.4 | 3.1 | 5.6 | 9.3 | 15.0 | 23.4 | 35.9 | 54.4 | 82.0 | 123.2 |
| High | 3,000-10,000 | 3.5 | 8.2 | 15.8 | 27.5 | 45.3 | 72.2 | 112.2 | 171.9 | 260.5 | 392.9 |
| Extremely High | >10,000 | 39.5 | 128.7 | 305.7 | 606.7 | 1,094.7 | 1,868.7 | 3,074.8 | 4,930.5 | 7,765.5 | 12,088.0 |
| Average |  | 0.2 | 0.3 | 0.5 | 0.8 | 1.3 | 2.0 | 3.0 | 4.4 | 6.6 | 9.8 |

(c) jacketed CPC-1232 compliant cars

|  | Population Density (persons per square mile) | Number of Tank Cars Releasing |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Population Class |  | $\geq 1$ | $\geq 2$ | $\geq 3$ | $\geq 4$ | $\geq 5$ | $\geq 6$ | $\geq 7$ | $\geq 8$ | $\geq 9$ | $\geq 10$ |
| Remote | <20 | 0.8 | 1.7 | 3.2 | 5.9 | 10.6 | 18.7 | 32.7 | 57.2 | 101.0 | 180.7 |
| Rural | 20-100 | 0.6 | 1.5 | 3.0 | 5.9 | 10.9 | 19.9 | 36.0 | 65.2 | 118.9 | 220.0 |
| Suburban | 100-1,000 | 0.6 | 1.5 | 3.0 | 5.9 | 11.1 | 20.5 | 37.2 | 67.6 | 123.8 | 229.5 |
| Urban | 1,000-3,000 | 1.8 | 4.8 | 10.4 | 20.5 | 38.8 | 71.7 | 131.0 | 238.8 | 437.8 | 812.5 |
| High | 3,000-10,000 | 4.6 | 13.3 | 30.3 | 62.2 | 120.4 | 225.5 | 415.2 | 759.4 | 1,392.1 | 2,575.4 |
| Extremely High | >10,000 | 54.9 | 226.8 | 643.8 | 1,504.6 | 3,188.6 | 6,391.1 | 12,391.7 | 23,596.7 | 44,639.6 | 84,656.1 |
| Average |  | 0.2 | 0.4 | 0.9 | 1.7 | 3.2 | 5.8 | 10.5 | 18.8 | 34.0 | 62.2 |

Note: The "average" row represents expected interval between releases, irrespective of population

(a) Remote area ( $<20$ persons per square mile)

(b) Rural area ( 20 to 100 persons per square mile)

(c) Suburban area (100 to 1,000 persons per square mile)

Figure 5.7 (cont.)

(d) Urban area (1,000 to 3,000 persons per square mile)

(e) High population area ( 3,000 to 10,000 persons per square mile)

(f) Extremely high population area ( $>10,000$ persons per square mile)

Figure 5.7: Estimated interval between occurrence of multiple-car release incidents by tank car safety design

Based on the estimated interval between occurrences of multiple-car release incidents for the conventional, non-jacketed DOT 111A100W1 tank car (Table 5.5) and the intervals for enhanced tank car designs (Table 5.7), I calculated the increase in the interval for enhanced tank car designs compared to the baseline case (Table 5.8).

The analysis shows that tank car safety design improvement has a substantial, multiplicative effect on prevention of large, multiple-car release incidents. For example, if conventional, non-jacketed DOT 111A100W1 tank cars are replaced by jacketed, CPC-1232 cars, the expected interval between occurrences of one-or-more-car-release incidents increases by $50 \%$. However, the same tank car upgrade strategy results in an 18 -fold increase in the interval between occurrences of 10-or-more-car release incidents. Tank car safety design improvement reduces the occurrence of hazardous materials release incidents of all magnitudes, but this effect is more substantial for large, multiple-car release incidents.

It should be pointed out that an assumption in this analysis is that the release probability of each tank car in a given accident is independent, but this may be violated due to circumstancespecific factors (e.g., the interaction between tank cars and infrastructure at the accident location) that are not accounted for in the CPR estimation. Further research is needed to better understand the validity of the assumption of independence of different cars' CPRs in multiple-car release incidents.

## Table 5.8

Increase in the interval between occurrences of multiple-car release incidents for various enhanced tank car designs compared to the baseline case
(a) Conventional jacketed car versus conventional non-jacketed, DOT 111A100W1

|  | Number of Tank Cars Releasing |  |  |  |  |  |  |  |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | :---: |
| Population Density <br> (persons per square | $\geq 1$ | $\geq 2$ | $\geq 3$ | $\geq 4$ | $\geq 5$ | $\geq 6$ | $\geq 7$ | $\geq 8$ | $\geq 9$ | $\geq 10$ |  |
| mile) | 1.2 | 1.4 | 1.6 | 1.9 | 2.1 | 2.4 | 2.8 | 3.1 | 3.6 | 4.2 |  |
| $<20$ | 1.2 | 1.4 | 1.7 | 1.9 | 2.2 | 2.5 | 2.9 | 3.4 | 3.9 | 4.5 |  |
| $20-100$ | 1.2 | 1.5 | 1.7 | 2.0 | 2.3 | 2.6 | 3.0 | 3.4 | 3.9 | 4.6 |  |
| $100-1,000$ | 1.3 | 1.5 | 1.7 | 2.0 | 2.3 | 2.6 | 3.0 | 3.4 | 4.0 | 4.6 |  |
| $1,000-3,000$ | 1.3 | 1.5 | 1.8 | 2.1 | 2.4 | 2.7 | 3.1 | 3.5 | 4.0 | 4.7 |  |
| $3,000-10,000$ | 1.4 | 1.7 | 2.0 | 2.3 | 2.6 | 3.0 | 3.4 | 3.9 | 4.4 | 5.1 |  |
| $>10,000$ | 1.2 | 1.5 | 1.7 | 1.9 | 2.2 | 2.5 | 2.9 | 3.3 | 3.8 | 4.4 |  |

(b) Non-jacketed CPC-1232 car versus conventional non-jacketed, DOT 111A100W1

|  | Number of Tank Cars Releasing |  |  |  |  |  |  |  |  |  |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :---: |
| Population Density <br> (persons per square | $\geq 1$ | $\geq 2$ | $\geq 3$ | $\geq 4$ | $\geq 5$ | $\geq 6$ | $\geq 7$ | $\geq 8$ | $\geq 9$ | $\geq 10$ |  |
| mile) | 1.1 | 1.3 | 1.4 | 1.6 | 1.7 | 1.9 | 2.0 | 2.2 | 2.5 | 2.7 |  |
| $<20$ | 1.1 | 1.3 | 1.4 | 1.6 | 1.8 | 1.9 | 2.1 | 2.4 | 2.6 | 2.9 |  |
| $20-100$ | 1.2 | 1.3 | 1.5 | 1.6 | 1.8 | 2.0 | 2.2 | 2.4 | 2.6 | 2.9 |  |
| $100-1,000$ | 1.2 | 1.3 | 1.5 | 1.6 | 1.8 | 2.0 | 2.2 | 2.4 | 2.6 | 2.9 |  |
| $1,000-3,000$ | 1.2 | 1.4 | 1.5 | 1.7 | 1.8 | 2.0 | 2.2 | 2.4 | 2.7 | 3.0 |  |
| $3,000-10,000$ | 1.2 | 1.5 | 1.7 | 1.8 | 2.0 | 2.2 | 2.4 | 2.6 | 2.9 | 3.2 |  |
| $>10,000$ | 1.2 | 1.3 | 1.4 | 1.6 | 1.8 | 1.9 | 2.1 | 2.3 | 2.6 | 2.9 |  |

(c) Jacketed CPC-1232 car versus conventional non-jacketed, DOT 111A100W1

|  | Number of Tank Cars Releasing |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Population Density (persons per square mile) | $\geq 1$ | $\geq 2$ | $\geq 3$ | $\geq 4$ | $\geq 5$ | $\geq 6$ | $\geq 7$ | $\geq 8$ | $\geq 9$ | $\geq 10$ |
| <20 | 1.4 | 1.9 | 2.5 | 3.2 | 4.1 | 5.3 | 6.9 | 9.1 | 12.1 | 16.4 |
| 20-100 | 1.4 | 2.0 | 2.6 | 3.5 | 4.5 | 5.9 | 7.7 | 10.2 | 13.7 | 18.9 |
| 100-1,000 | 1.5 | 2.0 | 2.7 | 3.5 | 4.6 | 6.0 | 7.8 | 10.4 | 13.9 | 19.2 |
| 1,000-3,000 | 1.5 | 2.1 | 2.8 | 3.6 | 4.7 | 6.1 | 7.9 | 10.5 | 14.1 | 19.3 |
| 3,000-10,000 | 1.6 | 2.2 | 2.9 | 3.8 | 4.9 | 6.3 | 8.2 | 10.8 | 14.3 | 19.4 |
| >10,000 | 1.7 | 2.6 | 3.5 | 4.5 | 5.8 | 7.5 | 9.7 | 12.6 | 16.7 | 22.4 |
| Average | 1.5 | 2.0 | 2.6 | 3.4 | 4.4 | 5.8 | 7.5 | 9.9 | 13.3 | 18.2 |

### 5.6 Discussion

### 5.6.1 Risk uncertainties

The parameters in this model represent the "average" accident circumstances. Therefore, there may be uncertainty using this model to analyze an extraordinary release incident such as the LacMégantic accident. That accident was extremely rare in terms of both the kinetic energy of the derailed train itself, and the quantity of flammable and combustible product released. In that accident, 63 tank cars out of 72 tank cars and one boxcar were derailed. Figure 5.8 is an analysis of derailment severity as measured by number of cars derailed per accident in freight train derailments on U.S. Class I railroads in the interval 2002 to 2011. The Lac-Mégantic derailment is in the 99.9th percentile in terms of number of cars derailed.

In addition, the parameters in the risk model I used were estimated based on all accident causes combined. Previous studies have found that train derailment probability, POD and derailment severities may vary by accident cause (Saccomanno and El-Hage 1989, 1991; Anderson 2005; Bagheri 2009; Bagheri et al. 2011; Liu et al. 2011, 2012, 2013b). The next stage in this research is to quantify multiple-car release risk associated with major accident causes, such as a broken rail or bearing failure. This would enable a better understanding of the causal effects in hazardous materials transportation risk analysis, as well as the reduction in risk associated with different approaches to derailment prevention.


Figure 5.8: Number of cars (including locomotives) derailed per FRA-reportable freight-train derailment, all accident causes combined on U.S. Class I railroad mainlines, 2002 to 2011

### 5.6.2 Risk reduction strategies

In addition to tank car safety improvement, it is also important to evaluate other potential risk reduction strategies, particularly approaches to derailment prevention. Detailed cost-benefit analysis of these risk reduction strategies is beyond the scope of this chapter; however, in the following subsections, I identify certain promising approaches.

### 5.6.2.1 Accident prevention

It is important to identify major accident causes and prioritize accident prevention initiatives.
Track and equipment-related accident causes accounted for over 70 percent of Class I mainline, freight-train derailments (Barkan et al. 2003; Liu et al. 2012). In particular, broken rails caused more hazardous materials car incidents and the corresponding releases than any other individual cause group (Liu et al. 2013a; Liu et al. 2014c). Therefore, broken rail prevention appears to be a promising risk reduction strategy. Liu et al (2014c)'s study developed a single-car release risk
model (CCPS 1995; Kawprasert and Barkan 2008, 2010; Saat 2009), but I am unaware of any study analyzing the effect of broken rail prevention on prevention of multiple-car release incidents.

Another approach to derailment prevention is use of wayside defect detectors to prevent equipment-related accident causes (Ouyang et al. 2009; Schlake and Barkan 2011). However, no prior research quantified the effect of wayside defect detector implementation on reduction of hazardous materials transportation risk. Further research is needed to evaluate how effectively wayside defect detectors could prevent train accidents, and the corresponding impact on risk.

### 5.6.2.2 Speed reduction

Train speed has a two-fold effect on the number of tank cars releasing. First, lower speed derailments may result in fewer cars derailed (Saccomanno and El-Hage 1989, 1991; Anderson 2005; Bagheri et al. 2011; Liu et al. 2011). Second, lower derailment speed results in a lower average tank car CPR (Kawprasert and Barkan 2010). The model developed in this chapter can be used to simultaneously account for both of these effects of derailment speed on multiple-car release incidents.

### 5.7 Conclusion

In this chapter, I develop a preliminary nationwide risk analysis of flammable liquid transportation on U.S. railroads. The analysis shows that the interval between occurrences of multiple-car release incidents is lower in high population areas than in low population areas. Tank car safety design improvement has a more substantial, multiplicative effect on prevention
of large, multiple-car release incidents. Use of CPC-1232 standards can significantly reduce the probability of a large, catastrophic release incident.

## References

Anderson, R.T., Barkan, C.P.L. (2004). Railroad accident rates for use in transportation risk analysis. Transportation Research Record 1863, 88-98.

Center for Chemical Process Safety (CCPS) (1995). Guidelines for chemical transportation risk analysis. Center for Chemical Process Safety of the American Institute of Chemical Engineers, New York.

Association of American Railroads (AAR) (2009). Analysis of Class I railroads. Association of American Railroads, Washington D.C.

Association of American Railroads (AAR) (2011). Petition for rulemaking: tank car standards for DOT class 111 tank cars used for packing group I and II materials. Association of American Railroads, Washington D.C.

Association of American Railroads (AAR). (2013). Moving crude oil by rail. https://www.aar.org/keyissues/Documents/Background-Papers/Crude-oil-by-rail.pdf. Accessed on October 19, 2013.

Bagheri, M., Saccomanno, F.F., Chenouri, S., Fu, L.P. (2011). Reducing the threat of in-transit derailments involving dangerous goods through effective placement along the train consist. Accident Analysis and Prevention 43(3), 613-620.

Bagheri, M., Verma, M., Verter, V. (2013). Transport mode selection for toxic gases: rail or road? Risk Analysis (in press) DOI: 10.1111/risa. 12063.

Barkan, C.P.L. (2008). Improving the design of higher-capacity railway tank cars for hazardous materials transport: optimizing the trade-off between weight and safety. Journal of Hazardous Materials, 122-134.

Bureau of Explosives (BOE) (2013). Annual report of hazardous materials transported by rail: calendar year 2012. Bureau of Explosives, Association of American Railroads, Washington D.C.

Erkut, E., Tjandra, S.A., Verter, V. (2007). Hazardous materials transportation, in handbooks in operations research and management science, volume 14 Transportation, C. Barnhart and G. Laporte (Eds.).

Federal Railroad Administration (FRA) (2003). Track safety standards. U.S. Department of Transportation, Washington D.C.

Glickman, T.S., Rosenfield, D.B. (1984). Risks of catastrophic derailments involving the release of hazardous materials. Management Science 30(4), 503-511.

Glickman, T.S., Erkut, E., Zschocke, M.S. (2007). The cost and risk impacts of rerouting railroad shipments of hazardous materials. Accident Analysis and Prevention 39(5), 1015-1025.

Hong, Y. (2013). On computing the distribution function for the Poisson binomial distribution. Computational Statistics and Data Analysis 59, 41-51.

Kawprasert, A., Barkan, C.P.L. (2008). Effects of route rationalization on hazardous materials transportation risk. Transportation Research Record 2043, 65-72.

Kawprasert, A., Barkan, C.P.L. (2010). Effect of train speed on risk analysis of transporting hazardous materials by rail. Transportation Research Record 2159, 59-68.

Kawprasert, A. (2010). Quantitative analysis of options to reduce risk of hazardous materials transportation by railroad. Ph.D. Dissertation, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Kletz, T. (1990). Critical aspects of safety and loss prevention. Butterworth-Heinemann, London.

Liu, X., Barkan, C.P.L., Saat, M.R. (2011). Analysis of derailments by accident cause: evaluating railroad track upgrades to reduce transportation risk. Transportation Research Record 2261, 178-185.

Liu, X., Saat, M.R., Barkan, C.P.L. (2012). Analysis of causes of major train derailment and their effect on accident rates. Transportation Research Record 2289, 154-163.

Liu, X., Saat, M.R., Barkan, C.P.L. (2013a). Safety effectiveness of integrated risk reduction strategies for rail transport of hazardous materials. Transportation Research Record 2374: 102110.

Liu, X., Saat, M.R., Barkan, C.P.L. (2013b). Analysis of U.S. freight-train derailment severity using zero-truncated negative binomial regression and quantile regression. Accident Analysis and Prevention 59, 87-93.

Liu, X., Saat, M.R., Barkan, C.P.L. (2014a). Analysis of freight-train derailment rates. Accident Analysis and Prevention (under review).

Liu, X., Saat, M.R., Barkan, C.P.L. (2014b). Probability analysis of multiple-tank-car release incidents in railway hazardous materials transportation. Journal of Hazardous Materials (under review).

Liu, X., Dick, C.T., Saat, M.R., Barkan. C.P.L. (2014c). Optimization of rail defect inspection frequency to reduce railroad hazardous materials transportation risk. Proceedings of 93nd Annual Meeting of Transportation Research Board, Washington D.C.

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

Ouyang, Y., Li, X., Lai, Y.C., Barkan, C.P.L., Kawprasert, A. (2009). Optimal locations of railroad wayside defect detection installations. Computer-Aided Civil and Infrastructure Engineering 24, 1-11.

Ross, S. (2007). Introduction to probability models, Academic Press.
Saat, M.R. (2009). Optimizing railroad tank car safety design to reduce hazardous materials transportation risk. Ph.D. Dissertation, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Saat, M.R., Barkan, C.P.L. (2011). Generalized railway tank car safety design optimization for hazardous materials transport: addressing the trade-off between transportation efficiency and safety. Journal of Hazardous Materials 189, 62-68.

Saccomanno, F.F., El-Hage, S.M. (1989). Minimizing derailments of railcars carrying dangerous commodities through effective marshaling strategies. Transportation Research Record 1245, 3451.

Saccomanno, F.F., El-Hage, S.M. (1991). Establishing derailment profile by position for corridor shipments of dangerous goods. Canadian Journal of Civil Engineering 18(1), 67-75.

Schlake, B.W., Barkan, C.P.L., Edwards, J.R. (2011). Train delay and economic impact of inservice failures of railroad rolling stock. Transportation Research Record 2261, 124-133.

Treichel, T.T., Barkan, C.P.L. (1993). Working paper on mainline freight train accident rates. Research and Test Department. Association of American Railroads, Washington D.C.

Treichel, T.T., Hughes, J.P., Barkan, C.P.L., Sims, R.D., Phillips, E.A., Saat, M.R. (2006). Safety performance of tank cars in accidents: probability of lading loss. Report RA-05-02, RSIAAR railroad tank car safety research and test project. Association of American Railroads, Washington, D.C.

## Chapter 6

# Integrated Risk Reduction Framework to Improve Railway Hazardous Materials Transportation Safety 

Adapted from
Liu, X., Saat, M.R., Barkan, C.P.L. (2013). Integrated risk reduction framework to improve railway hazardous materials transportation safety. Journal of Hazardous Materials 260, 131-140.

### 6.1 Introduction

Approximately 2 million carloads of hazardous materials are shipped by rail in North America each year (BOE 2013). Although the majority of these shipments (over 99.99\%) safely reach their destinations, the risk of a hazardous materials release incident remains a major safety concern to the rail industry, government and public. For example, the release of chlorine gas from a train collision in Graniteville, South Carolina in January 2005 resulted in 9 fatalities, hundreds of injuries, an evacuation of about 5,400 people and economic loss exceeding $\$ 6.9$ million (NTSB 2005). There has been growing interest and intensifying regulatory pressure to further improve the safety of rail transport of hazardous materials. Improvements have focused on enhancing packaging and tank car safety design (Barkan et al. 1991; Saat and Barkan 2005; Barkan et al. 2007; Tyrell et al. 2007; Barkan 2008; Saat 2009; Saat and Barkan 2011), deploying wayside inspection technologies (Resor and Zarembski 2004; Tournay and Cummings 2005; Ouyang et al. 2009; Kalay et al. 2011), upgrading track infrastructure (Kawprasert and Barkan 2010; Liu et al. 2011a, 2011b), routing (Glickman 1983; Glickman and Rosenfield 1984; Abkowitz et al. 1989; Erkut et al. 2007; Kawprasert and Barkan 2008), reducing train speed (Kawprasert 2010) and improving emergency response practices (CCPS 2008). Each strategy has a direct effect on hazardous materials transportation risk, and different strategies may also have interactive effects.

However, how to optimize the integration of different risk reduction strategies in the most cost-efficient manner is not well understood. In this chapter, I develop an integrated risk reduction framework, accounting for the cost-effectiveness of individual risk reduction strategies, their interactive effects and optimal integration. The chapter is structured as follows. First, I formulate hazardous materials transportation risk management as a multi-attribute decision analysis problem. Then, I develop a Pareto-optimality approach to determine the lowest risk that can be achieved at a specific level of investment. Understanding the risk-and-cost relationship can lead to development of a decision model to determine the "optimal" investment. To illustrate the methodology, I analyze the cost-effectiveness of broken rail prevention, tank car safety design enhancement and their optimal combination given a budgetary constraint. Although the model implementation was based on U.S. data, the methodology can be adapted to the rail systems in other regions.

### 6.2 Multi-attribute model for hazardous materials transportation risk management

Hazardous materials transportation risk management can be formulated as a multi-attribute decision problem. It is assumed that certain risk reduction strategies are implemented to reduce the baseline risk $\left(\mathrm{R}_{0}\right)$ to a lower level (R). The associated implementation cost is I. Given a specific investment $\mathrm{I}_{\mathbf{i}}$, there is an optimal combination of risk reduction strategies so as to achieve the lowest risk. Let $R^{*}\left(I_{i}\right)$ represent the lowest risk given investment $I_{i}$. For rational decision making, additional investment should not reduce safety, that is:

If $\mathrm{I}_{\mathrm{j}}>\mathrm{I}_{\mathrm{i}}, \quad$ Then $\mathrm{R}^{*}\left(\mathrm{I}_{\mathrm{j}}\right) \leqslant \mathrm{R}^{*}\left(\mathrm{I}_{\mathrm{i}}\right)$

In Equation (6.1), the equality holds when the additional investment $\left(\mathrm{I}_{\mathrm{j}}-\mathrm{I}_{\mathrm{i}}\right)$ does not result in additional safety benefit. Figure 6.1 illustrates the relationship between $\mathrm{R}^{*}(\mathrm{I})$ and I. This relationship is called Pareto-optimality (Fudenberg and Tirole 1983). In the context of hazardous materials transportation risk management, Pareto-optimality means that the safety cannot be further improved without additional investment. When Pareto-optimality is used in a multiattribute decision analysis model, the "optimal" investment ( $\mathrm{I}^{* *}$ ) and the corresponding risk ( $\mathrm{R}^{* *}$ ) can be determined.


Figure 6.1: Schematic illustration of Pareto-optimality in hazardous materials transportation risk management

In decision analysis, the value function is a general approach to account for the decision maker's preference and trade-off between multiple attributes (such as the risk and cost) (Dyer and Sarin 1979). The linear value function has been used in prior research (Glickman et al. 2007). A value function $V(R, I)$ is defined based on the risk and the investment to reduce the baseline risk to a lower level of risk.
$V(R, I)=W_{R} R+W_{I} I$

Where:
$\mathrm{V}(\mathrm{R}, \mathrm{I})=$ value function of risk and investment
R = proportion of baseline risk ( $\mathrm{R}=100 \%$ for baseline risk)
I $\quad=$ investment for reducing the baseline risk $\left(\mathrm{R}_{0}\right)$ to a lower level of risk R
$\mathrm{W}_{\mathrm{R}} \quad=$ first-order partial derivative of the value function with respect to risk
$\mathrm{W}_{\mathrm{I}}$ = first-order partial derivative of the value function with respect to investment

On the Pareto-optimality frontier, the minimum risk can be estimated as a function of investment, denoted as $\mathrm{R}=\mathrm{g}(\mathrm{I})$. Equation (6.2) can be re-written as:
$V(R, I)=W_{R} g(I)+W_{I} I$

The optimal investment $\left(\mathrm{I}^{* *}\right)$ is determined by solving the following equation:

$$
\begin{equation*}
\frac{\partial V\left(R, I^{* *}\right)}{\partial I^{* *}}=W_{R} \frac{\partial g\left(I^{* * *}\right)}{\partial I^{* *}}+W_{I}=0 \tag{6.4}
\end{equation*}
$$

Equation (6.4) can be re-written as :

$$
\begin{equation*}
\frac{\partial g\left(I^{* * *}\right)}{\partial I^{* *}}=-\frac{W_{I}}{W_{R}} \tag{6.5}
\end{equation*}
$$

If the optimal investment $\left(\mathrm{I}^{* *}\right)$ exceeds the budget $\left(\mathrm{I}_{\max }\right)$, the optimal decision may be either no investment $\left(I^{* *}=0\right)$ or using all the budget $\left(I^{* *}=I_{\max }\right)$, depending on the value function. In order to optimize the allocation of investment, we need to estimate the safety effectiveness and cost of a risk reduction strategy. In the next section, I introduced a railroad hazardous materials transportation risk model.

### 6.3 Railroad hazardous materials transportation risk analysis model

Railroad hazardous materials transportation risk is defined as a multiplication of derailment rate of a hazardous materials car, traffic exposure, conditional probability of release (CPR) of a derailed hazardous materials car and the consequence of a release (Anand et al. 2005;

Kawprasert and Barkan 2008, 2010; Saat 2009; Kawprasert 2010; Liu et al. 2011b, 2013)
$\mathrm{R}=\mathrm{Z} \times \mathrm{M} \times \mathrm{P} \times \mathrm{C}$
Where:

R = hazardous materials transportation risk
Z = hazardous materials car derailment rate per billion car-miles
M = traffic exposure (e.g., billion car-miles)
P $\quad=\mathrm{CPR}$ of a derailed hazardous materials car
C = consequences of a car release (e.g., number of people affected)

Hazardous materials car derailment rate is defined as the number of cars derailed per unit of traffic exposure. Car derailment rates vary by track characteristics (Nayak et al. 1983; Treichel and Barkan 1993; Anderson and Barkan 2004). The CPR of a hazardous materials car reflects its safety performance. The majority of railroad hazardous materials shipments (72\%) and the largest quantities are in tank cars (BOE 2013), thus tank car safety design improvement has been a priority for the U.S. rail industry and government. Treichel et al. (2006) developed logistic regression models to estimate the CPR of a derailed tank car given its configuration. Kawprasert and Barkan (2010) extended Treichel et al's model by accounting for derailment speed (Kawprasert and Barkan 2010). The consequences of a release can be measured by several
metrics, including property damage, disruption of service, environmental impact, human impact (e.g., number of people potentially affected by a release), litigation or other types of impacts (Barkan et al. 1991; Saat 2009; Kawprasert 2010). Among the consequence measures, population in the affected area of a release incident is often used (CCPS 2008; Saat 2009; Kawprasert 2010). The hazard exposure model provided in the U.S. Department of Transportation (U.S. DOT) Emergency Response Guidebook (ERG) can be used to estimate the affected area based on the material and scenario of release (fire, spill, daytime, nighttime) (PHMSA 2012). Once the affected area is determined, the number of people affected can be estimated by multiplying the affected area of each segment by the corresponding average population density. The assessment of release consequence can be conducted using a Geographical Information System (GIS) platform (Kawprasert 2010).

Figure 6.2 illustrates two basic strategies to reduce tank car release risk: (1) reduce the likelihood of a hazardous materials release incident; (2) reduce release consequences. This study focuses on the former - reducing the likelihood of a release incident. Reducing the likelihood of a hazardous materials car release incident (tank cars are used as an example in this chapter) can be achieved by reducing tank car derailment probability and/or reducing the CPR of a derailed tank car. For illustration, I consider two potential risk reduction strategies, broken rail prevention to reduce tank car derailment probability and tank car safety design enhancement to reduce its release probability.


Figure 6.2: Hierarchy of risk reduction strategies in this chapter

In the next section, I analyze the cost-effectiveness of broken rail prevention and tank car safety design enhancement, respectively, for risk reduction.

### 6.4 Cost-effectiveness of risk reduction strategies

### 6.4.1 Broken rail prevention

### 6.4.1.1 Broken-rail-caused car derailment rate

Broken rails are among the most common and severe accident causes on U.S. railroads (Barkan et al. 2003; Liu et al. 2012, 2013a). I use the number of all types of railcar derailment as a proxy to evaluate the safety effectiveness of broken rail prevention. This approach assumes that the proportional reduction of hazardous materials car derailments is equal to other types of railcars due to broken rail prevention (Liu et al. 2013b), since a broken rail can derail any type of railcar,
regardless of whether it is transporting hazardous materials. I assumed that the number of broken-rail-caused cars derailed follows a Poisson distribution:
$\mathrm{P}(\mathrm{Y}=\mathrm{k})=\frac{\lambda^{\mathrm{k}}}{\mathrm{k}!} \mathrm{e}^{-\lambda}$
Further, the Poisson mean $(\lambda)$ is assumed to follow a gamma distribution (Hilbe 2007):
$p(\lambda=m)=\frac{\left(\frac{\eta}{\mu}\right)^{\eta}}{\Gamma(\eta)} m^{\eta-1} e^{-\left(\frac{\eta}{\mu}\right)^{m}}$

It can be shown that the marginal distribution of broken-rail-caused derailment count follows a negative binomial distribution (Hilbe 2007):
$\int \operatorname{Poi}(y \mid \lambda) \operatorname{Gamma}(\lambda \mid \eta, \mu) d \lambda=\frac{\Gamma(y+\eta)}{y!\Gamma(\eta)}\left(\frac{\eta}{\eta+\mu}\right)^{\eta}\left(\frac{\mu}{\eta+\mu}\right)^{y}$
$\mu=\exp \left(\sum_{\mathrm{p}=0}^{\mathrm{k}} \beta_{\mathrm{p}} \mathrm{X}_{\mathrm{p}}\right) \mathrm{M}$

Where:
$\mu \quad=$ expected broken-rail-caused car derailment count
$\beta_{\mathrm{p}} \quad=\mathrm{p}^{\text {th }}$ parameter coefficient
$\mathrm{X}_{\mathrm{p}} \quad=\mathrm{p}^{\text {th }}$ explanatory variable
$\mathrm{M} \quad=$ traffic exposure (gross ton-miles)
$\eta \quad=$ gamma parameter (also called inverse dispersion parameter)

Equations (6.7) to (6.10) represent the widely used Poisson-gamma (negative binomial) regression model for estimating accident rates (Hilbe 2007). In this chapter, I use annual rail
maintenance (including rail repair and renewal) cost per track mile (adjusted to dollars in year 2008) as an explanatory variable to estimate the rate of broken-rail-caused car derailments. The accident data were from the Rail Equipment Accident (REA) database from the Federal Railroad Administration (FRA) of U.S. Department of Transportation (DOT). This database contains information regarding all accidents that exceed a monetary threshold of damages to on-track equipment, signals, track, structures, and roadbed (FRA 2011). In this chapter, I focus on Class I freight railroads (Class I railroads are defined as those having operating revenue that exceeded $\$ 378.8$ million in 2009). They accounted for approximately $69 \%$ of route miles and $88 \%$ of carloads transported on U.S. railroads in 2012 (AAR 2013).

Annual rail maintenance cost and traffic exposure data were from Class I railroad's annual report to the U.S. Surface Transportation Board (STB 2002-2008). Based on the negative binomial regression model, the expected broken-rail-caused car derailment rate on Class I mainlines $\left(\mathrm{Z}_{\mathrm{br}}\right)$ is estimated as follows:

$$
\begin{equation*}
Z_{b r}=\exp (-0.1868-0.3356 C) \tag{6.11}
\end{equation*}
$$

where:
$\mathrm{Z}_{\mathrm{br}} \quad=$ expected broken-rail-caused car derailment rate per billion gross ton-miles

C = annual rail maintenance cost per track-mile (thousand dollars)

The goodness-of-fit of the regression model is evaluated by a statistic called Deviance, which asymptotically follows a Chi-squared distribution (Hilbe 2007). Based on this criterion, the model exhibits an overall acceptable fit $($ Deviance $=37.2$, degrees of freedom $=33, \mathrm{P}=0.28$ $>0.05)$. Appendix F presents the original data for the statistical analysis. In addition to
normalizing annual rail maintenance cost per mile, some railroads may have information regarding annual rail maintenance cost per tonnage. The two variables (cost per mile and cost per ton) are highly correlated (Pearson correlation coefficient is 0.96 ), therefore using either explanatory variable could fit the data well. Appendix G presents the regression results when using rail maintenance cost per ton as the predictor variable in lieu of the cost per mile as shown in Table 6.1.

## Table 6.1

Regression analysis of broken-rail-caused car derailment rate, U.S. class I mainlines, 2002 to 2008

| Parameter | Estimate | Standard Error | 95\% Confidence Limits | Pr $>$ ChiSq |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Intercept | -0.1868 | 0.3053 | -0.7852 | 0.4115 | 0.5405 |
| Annual Rail Maintenance Cost per |  |  |  |  |  |
| Track Mile (Thousand Dollars) | -0.3356 | 0.1101 | -0.5514 | -0.1198 | 0.0023 |
| Dispersion Parameter $(1 / \eta)$ | 0.3682 | 0.0857 | 0.2333 | 0.5811 |  |

Deviance $=37.2($ Degrees of Freedom $=33)$
$P=0.28>0.05$

Table 6.1 shows that the expected broken-rail-caused car derailment rate declines when rail maintenance cost increases, all else being equal. The probability distribution of the number of broken-rail-caused cars derailed given traffic exposure and rail maintenance cost can be estimated using a negative binomial distribution:
$\mathrm{P}(\mathrm{y})=\frac{\Gamma(\mathrm{y}+2.7159)}{\mathrm{y}!\Gamma(2.7159)}\left(\frac{2.7159}{2.7159+\exp (-0.1868-0.3356 \mathrm{C}) \mathrm{M}}\right)^{2.7159}\left(\frac{\exp (-0.1868-0.3356 \mathrm{C}) \mathrm{M}}{2.7159+\exp (-0.1868-0.3356 \mathrm{C}) \mathrm{M}}\right)^{\mathrm{y}}$
$(2.7159 \approx 1 / 0.3682)$

Where:
$\mathrm{P}(\mathrm{y}) \quad=$ probability that there are y broken-rail-caused cars derailed given traffic M
$\mathrm{M} \quad=$ traffic exposure (billion gross ton-miles)

C = annual rail maintenance cost per track-mile (thousand dollars)

Figure 6.3 shows the distribution of the annual number of broken-rail-caused cars derailed assuming that annual rail maintenance cost (C) is $\$ 2,000$ per track-mile or $\$ 4,000$ per track-mile, respectively, on the 160,240 track miles of U.S. Class I railroad mainlines (AAR 2011). Annual traffic exposure (M) is 3,446 billion gross ton-miles (AAR 2011). The analysis shows that the higher the rail maintenance cost per track-mile, the smaller the expected number of broken-rail-caused cars derailed given traffic exposure. This may be because improved rail maintenance reduces rail failure rate, thereby reducing the probability of broken-rail-caused derailments (Dick et al. 2003).


Total Number of Car Derailments

Figure 6.3: Probability distribution of annual number of broken-rail-caused car derailments by annual rail maintenance cost, U.S. Class I mainlines, annual traffic exposure is 3,446 billon gross ton-miles (only a portion of the entire distribution is displayed)

The percent reduction in broken-rail-caused car derailment rate is an exponential function of additional rail maintenance cost (proof is in Appendix H ):
$\Delta Z_{b r}=1-\exp (-0.3356 \Delta C)$
Where:
$\Delta Z_{b r} \quad=$ percent reduction in broken-rail-caused car derailment rate per billion ton-miles
$\Delta \mathrm{C} \quad=$ increment of annual rail maintenance cost per track-mile (thousand dollars)

The relationship between $\Delta \mathrm{Z}_{\mathrm{br}}$ and $\Delta \mathrm{C}$ is illustrated in Figure 6.4 using Equation (6.13).
It shows that when the investment for broken rail prevention increases to a certain level, the percent reduction of car derailment rate levels off.


Figure 6.4: Percent reduction of broken-rail-caused car derailment rate by increasing annual rail maintenance cost per track-mile

This chapter focuses on the overall effect of broken rail prevention, without accounting for a specific broken rail prevention technique. Further analysis is needed to analyze the variability in cost and effectiveness of different broken rail prevention techniques, such as rail grinding or advanced rail defect detection technologies.

### 6.4.1.2 Proportion of broken-rail-caused car derailment

The analysis above is related to the percent reduction of broken-rail-caused car derailments by increasing rail maintenance cost. In order to assess the effect of broken rail prevention on the overall accident rate, I estimate the proportion of car derailments caused by broken rails versus other causes.

Figure 6.5 shows the total number of freight cars derailed due to primary accident cause for FRA-reportable freight-train derailments on Class I mainlines from 2001 to 2010. It shows that broken rails or welds accounted for $23 \%$ of railcars derailed.


Figure 6.5: Railcar derailment frequency by accident cause, Class I mainline freight-train derailments, 2001 to 2010 (Liu et al. 2013b)

The proportion of broken-rail-caused car derailments may vary by track characteristics.
In this analysis, each car derailed is assumed to be caused by either a broken rail or another cause. The probability that a car derailment is caused by a broken rail can be estimated using a logistic regression model (Liu et al. 2013b):
$P_{b \mid a}=\frac{\exp \left(\sum_{k} \beta_{k} X_{k}\right)}{1+\exp \left(\sum_{k} \beta_{k} X_{k}\right)}$
Where:
$\mathrm{P}_{\mathrm{b} \mid \mathrm{a}} \quad=$ probability that a car derailment is due to broken rails
$\beta_{\mathrm{k}} \quad=$ parameter coefficient
$\mathrm{X}_{\mathrm{k}} \quad=$ explanatory variables

The following factors are considered. The selection of these variables is based on prior research (Dick et al. 2003; Schaffer 2008; Liu et al. 2013b), and communication with railway engineering professionals.

- TrkClas (ordinal variable)

This variable represents FRA track class (1, 2, 3, 4 and 5). FRA track class is a proxy of railroad track quality by U.S. federal regulations. The higher the FRA track class, the greater the allowable maximum operating speed, and correspondingly more stringent engineering standards apply.

- MOO (dummy variable)
$\mathrm{MOO}=1$ indicates a signaled track territory; 0 otherwise.
- MGT (dummy variable)

It represents annual traffic density level, measured in million gross tons (MGT). "MGT $=1$ " represents annual traffic density greater than 20 MGT (threshold for high-density track defined by the Associated of American Railroads (AAR 2011)), 0 otherwise.

- Season (dummy variable)
"Season $=1 "$ represents a train derailment occurred in colder period between September and February, 0 otherwise.

Table 6.2 presents the significance of each variable using likelihood ratio (LR) test (Hilbe 2007). It shows that FRA track class (TrkClas) is not significant given the other variables. Traffic density (MGT) has a minor effect on the probability that a car derailment is caused by
broken rails. Whereas, both method of operation (MOO) and climate (Season) are significant.
Table 6.3 presents $95 \%$ confidence interval of the estimated proportion of broken-rail-caused car derailments by track characteristics. It shows that non-signaled track has a greater proportion of broken-rail-caused car derailments than signaled track, probably due to the absence of track circuits to detect broken rails in non-signaled track territories. In addition, a car derailment is more likely to be caused by broken rails during colder seasons (September to February), probably because of thermal stress in rails (Dick et al. 2003).

## Table 6.2

Statistical significance of explanatory variables

| Variable | Chi-Square | Pr $>$ ChiSq |
| :--- | ---: | ---: |
| TrkClas | 1.21 | 0.8773 |
| MGT | 3.79 | 0.0516 |
| MOO | 24.38 | $<.0001$ |
| Season | 28.4 | $<.0001$ |

## Table 6.3

Estimated probability that a car derailment is due to broken rails, by method of operation, annual traffic density and season, Class I mainlines, 2001 to 2010

|  |  |  | $95 \%$ Confidence Interval |  |  |
| :--- | :--- | :--- | :---: | :---: | :---: |
| Method of | Annual Traffic |  | Mean | Lower Bound | Upper Bound |
| Operation | Density | Season | March to August | 0.2862 | 0.2281 |
| Non-Signaled | <20MGT | Sen | 0.3523 |  |  |
| Non-Signaled | <20MGT | September to February | 0.4779 | 0.4088 | 0.5478 |
| Non-Signaled | $\geq 20 M G T$ | March to August | 0.2254 | 0.1689 | 0.2943 |
| Non-Signaled | $\geq 20 M G T$ | September to February | 0.3992 | 0.3245 | 0.4789 |
| Signaled | <20MGT | March to August | 0.1455 | 0.1063 | 0.1961 |
| Signaled | <20MGT | September to February | 0.2799 | 0.2169 | 0.3531 |
| Signaled | $\geq 20 M G T$ | March to August | 0.1100 | 0.0858 | 0.1400 |
| Signaled | $\geq 20 M G T$ | September to February | 0.2201 | 0.1833 | 0.2619 |

*Sample contains 6,383 broken-rail-caused car derailments out of 27,516 cars derailed by all causes (including all types of empty or loaded railcars)

### 6.4.2 Tank car safety design improvement

An alternative risk reduction strategy is to reduce the release probability of a derailed tank car. This can be achieved by improving tank car safety design (Barkan et al. 1991; Saat and Barkan 2005; Barkan et al. 2007; Tyrell et al. 2007; Barkan 2008; Saat 2009; Saat and Barkan 2011), Treichel et al. (2006) used a logistic regression model to estimate the CPR for common tank car designs as well as hypothetical configurations based on design elements currently in use. This model has been used by Barkan (2008) and Saat (2009) to develop optimized tank car safety design models. Table 6.4 shows the CPR of two types of tank cars transporting ethanol, a common hazardous materials transported by rail in North America (BOE 2013). The enhanced tank car is thicker and has head protection. Thus, it is expected to have a lower release probability ( 0.2947 for the baseline tank car versus 0.2412 for the enhanced car).

Table 6.4
Conditional probability of release by tank car safety design

| Design | Head Thickness <br> (inch) | Shell Thickness <br> (inch) | Headshield | Capacity <br> (gallon) | Conditional Probability <br> of Release (CPR) |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Baseline | 0.4375 | 0.4375 | None | 32,000 | 0.2947 |
| Enhanced | 0.5000 | 0.5000 | 0.5 inch Half Height | 31,186 | 0.2412 |

Saat (2009) developed a financial model to estimate the capital and operational cost for upgrading the baseline tank car fleet. In this chapter, I used his model to quantify the cost-effectiveness of tank car design enhancement for risk reduction. The cost for upgrading the baseline tank car fleet was estimated based on the following information:

- 29,013 ethanol baseline tank cars in the fleet in the U.S. (BOE 2012)
- 317,370 annual tank car shipments (BOE 2012)
- Average shipment distance is 926 miles (STB 2009)
- The capital cost is $\$ 80,485$ and $\$ 84,494$ (dollars in year 2008) for a baseline tank car and an enhanced tank car, respectively using Saat's model (Saat 2009)
- Life cycle of a tank car is assumed to be 40 years (Saat 2009)
- Annual discount rate is 5\% (OMB 2008)

Using Saat's model, the annualized cost over a 40-year period is approximately $\$ 9.7$ million if the entire ethanol baseline tank car fleet is replaced by the enhanced tank cars at once. Correspondingly, the average CPR is expected to be reduced by $18.2 \%$ ([0.2947-0.2412]/0.2947=18.2\%). Thus, on average, $\$ 1$ million annual investment for tank car fleet upgrade is expected to yield $1.9 \%$ reduction of release probability, all else being equal $(18.2 \% \div 9.7=1.9 \%)$. Therefore, the percent reduction of average release probability of baseline tank car fleet is a function of annual cost:
$\Delta_{\mathrm{CPR}}=\operatorname{minimum}(1.9 \% \times \mathrm{C}, 18.2 \%)$
Where:
$\Delta_{\mathrm{CPR}}=$ percent reduction of average CPR of baseline tank car fleet
C = annual cost for upgrading baseline tank car fleet (million dollars)

Using Equation (6.15), the relationship between percent CPR reduction and annual cost for tank car fleet upgrade is shown in Figure 6.6.


Figure 6.6: Percent reduction of average CPR by annual cost for upgrading tank car fleet

### 6.5 Pareto-optimality between risk and cost

The analysis above shows the cost-effectiveness of different risk reduction strategies. The following optimization model is developed to determine the "optimal" portfolio of risk reduction strategies given budgetary constraint (derived in Appendix I).

Objective function:
$\operatorname{maximize}\left\{-(1-\mathrm{em})\left[(1-\mathrm{b})+\frac{\mathrm{P}_{\mathrm{u}}}{\mathrm{P}_{\mathrm{b}}} b\right]\right\}$

Subject to:
$-\frac{\ln (1-e)}{0.3356} \times 160.240+9.7 b \leq$ Budget
$0 \leq e \leq 1$
$0 \leq b \leq 1$

The objective function is to maximize the national average percent risk reduction. The decision variables are the proportion of broken-rail-caused car derailments to prevent (e), and the proportion of the tank car fleet to upgrade (b). The first constraint is that total implementation cost cannot exceed the budget. The other constraints require that the decision variables are between 0 and 1 .

I applied this optimization model to the 160,240 track miles on the U.S. Class I mainline rail network (AAR 2011). On average, $23 \%$ of cars derailed were due to broken rails $(\mathrm{m}=0.23)$ (Liu et al. 2012). The CPR of a baseline and an enhanced tank car are 0.2947 and 0.2412 , respectively. A sensitivity analysis was conducted varying annual budget from $\$ 0$ to $\$ 200$ million, in increments of $\$ 2$ million. Given this budget, the optimization model was solved using the software called General Algebraic Modeling System (GAMS) with its built-in solver MINOS. The relationship between the minimum risk level and annual budget constructs a Pareto-optimality frontier (Figure 6.7).


Figure 6.7: Pareto-optimality of implementation of integrated risk reduction strategies

This chapter aims to address "what strategies to invest" rather than "how to implement". For example, I estimate that $23 \%$ of broken-rail-caused car derailments should be prevented without specifying how to achieve this objective. The next step of this research is to develop an operational model to determine what technologies to implement and extent of implementation to prevent major accident causes.

### 6.6 Optimal investment

The risk-and-investment relationship in Figure 6.7 is fitted by the following step-wise function:

$$
\begin{array}{rll}
\mathrm{R}=\mathrm{g}(\mathrm{I})=\left\{\begin{array}{ll}
1-0.0184 \times \mathrm{I} & \text { If } \mathrm{I} \leqslant 9.7 \\
& 0.8198 \exp (-0.0004 \times \mathrm{I})
\end{array}\right\} & \text { Otherwise }
\end{array}
$$

Recall that the value function is $V(R, I)=W_{R} R+W_{I} I$. If maximum annual budget does not exceed $\$ 9.7$ million ( $\mathrm{I}_{\max } \leqslant 9.7$ ), the optimal investment $\left(\mathrm{I}^{* *}\right)$ is either $\$ 0$ or $\$ 9.7$ million, depending on the ratio of $\mathrm{W}_{\mathrm{I}}$ to $\mathrm{W}_{\mathrm{R}}$. Specifically,

If $W_{\mathrm{I}} / \mathrm{W}_{\mathrm{R}}>0.0184, \quad \mathrm{I}^{* *}=9.7$
If $\mathrm{W}_{\mathrm{I}} / \mathrm{W}_{\mathrm{R}} \leqslant 0.0184, \quad \mathrm{I}^{* *}=0$

If annual investment exceeds $\$ 9.7$ million, the risk decreases as an exponential function of investment. The optimal investment is determined by Equation (6.5):
$\frac{\partial g\left(I^{* *}\right)}{\partial I^{* *}}=-\frac{W_{I}}{W_{R}}$
$g(I)=0.8198 \exp (-0.0004 \times I)$ for $\mathrm{I}>9.7$, thus,
$\frac{\partial g\left(I^{* *}\right)}{\partial I^{* *}}=0.8198 \exp \left(-0.0004 I^{* *}\right)(-0.0004)=-\frac{W_{I}}{W_{R}}$
Therefore, $I^{* * *}=\frac{\ln \left(\frac{W_{I}}{0.8198 \times 0.0004 W_{R}}\right)}{-0.0004}$

If $\mathrm{I}^{* *}$ exceeds the budget ( $\mathrm{I}_{\max }$ ), the optimal investment is either $\$ 0, \$ 9.7$ million or the maximum budget, depending on the value function. In the multi-attribute value function, $W_{I} / W_{R}$ is numerically equal to the monetary value of one percent risk reduction. For example, if the decision maker is willing to invest 32 million dollars annually to achieve $1 \%$ baseline risk reduction, $\mathrm{W}_{\mathrm{I}} / \mathrm{W}_{\mathrm{R}}$ is equal to $0.000313(1 \% \div 32)$. Using Equation (6.21), the optimal investment
$\left(\mathrm{I}^{* *}\right)$ is 120 million dollars. It means that the decision maker is willing to invest $\$ 120$ million annually in the rail network for safety improvement. Figure 6.7 has shown that this investment, if used in a cost-efficient manner, is expected to reduce the baseline risk by approximately $22 \%$. Figure 6.8 shows that the optimal investment varies by the monetary value of risk reduction. It illustrates that the greater monetary value per unit of risk reduction, the more investment the decision maker is willing to spend for risk reduction.


Figure 6.8: Illustrative example of optimal investment by monetary value of one percent risk reduction

### 6.7 Conclusion

This chapter develops a framework model to optimize the integration of multiple risk reduction strategies to reduce the risk in the most cost-efficient manner. Broken rail prevention represents an accident prevention strategy to reduce tank car derailment probability, while tank car safety design enhancement affects the probability that a derailed car releases. The interactive effects among multiple risk reduction strategies are taken into account in evaluating the combined safety
effectiveness of these strategies. The methodology presented in this chapter is the first step towards a larger, integrated risk management framework under development. The method can be further developed and applied to a broader set of risk reduction strategies.

### 6.8 Future research

The intent of this research is to identify the strategic options for risk reduction, without detailing specific techniques or practices to achieve the safety goal of a specific risk reduction strategy. Further research is needed to better understand the cost-effectiveness of various risk reduction measures within the same strategy (for example, rail grinding or rail inspection for broken rail prevention). Additionally, railroads may need to prioritize the track segments for safety improvement based on the trade-off between risk and cost, resource available, and operational characteristics. Implementation and prioritization of integrated risk reduction framework are based on extensive data. Each railroad should use their enterprise-specific data to optimize the allocation of risk reduction resources.

The next stage in this research is to implement integrated risk reduction strategies on representative hazardous materials routes. When extending the analysis to a rail network involving multiple routes, network-specific optimization models can be developed to simultaneously determine the optimal routing and cost-efficient infrastructure and rolling stock improvement (Lai et al. 2011).

In addition to the risk while traversing main lines between terminals, the risk in classification yards is also important (Glickman and Erkut 2007; Cozzani et al. 2007). Other risk
reduction strategies, such as hazardous materials car marshaling, can be incorporated into an integrated risk reduction framework (Bagheri et al. 2011). Ultimately, the risks of different transportation modes will be evaluated and optimized to facilitate a risk-informed decision related to multimodal transportation of hazardous materials (Paltrinieri et al. 2009).

In addition, some researchers considered risk aversion in hazardous materials transportation risk management accounting for the decision maker's avoidance toward catastrophic consequences (Erkut and Ingolfsson 2000). Use of utility (or disutility) model (Pratt 1964) might be useful in the group decision making involving multiple stakeholders.

## References

Abkowitz, M., List, G., Radwan, A.E. (1989). Critical issues in safe transport of hazardous materials. Journal of Transportation Engineering 115(6), 608-629.

Anand, P., Barkan, C.P.L., Schaeffer, D.J., Werth, C.J., Minsker, B.S. (2005). Environmental risk analysis of chemicals transported in railroad tank cars. Proceedings of the 8th International Heavy Haul Conference, Rio de Janiero, 395-403.

Anderson, R.T., Barkan, C.P.L. (2004). Railroad accident rates for use in transportation risk analysis. Transportation Research Record 1863, 88-98.

Association of American Railroads (AAR) (2011). Class I railroad statistics. http://www.aar.org/~/media/aar/Industry\ Info/AAR\ Stats\ 2010\ 1123.ash. Association of American Railroads, Washington D.C.

Association of American Railroads (AAR) (2013). Class I railroad statistics. https://www.aar.org/StatisticsAndPublications/Documents/AAR-Stats-2013-01-10.pdf. Association of American Railroads, Washington D.C.

Bagheri, M., Saccomanno, F.F., Chenouri, S., Fu, L.P. (2011). Reducing the threat of in-transit derailments involving dangerous goods through effective placement along the train consist. Accident Analysis and prevention 43, 613-620.

Barkan, C.P.L., Glickman, T.S., Harvey, A.E. (1991). Benefit-cost evaluation of using different specification tank cars to reduce the risk of transporting environmentally sensitive chemicals. Transportation Research Record 1313, 33-43.

Barkan, C.P.L., Dick, C.T., Anderson, R.T. (2003). Analysis of railroad derailment factors affecting hazardous materials transportation risk. Transportation Research Record 1825, 64-74.

Barkan, C.P.L., Ukkusuri, S., Waller, S.T. (2007). Optimizing the design of railway tank cars to minimize accident-caused releases. Computers \& Operations Research 34, 1266-1286.

Barkan, C.P.L. (2008). Improving the design of higher-capacity railway tank cars for hazardous materials transport: optimizing the trade-off between weight and safety. Journal of Hazardous Materials 160, 122-134.

Bureau of Explosives (BOE) (2013). Annual report of hazardous materials transported by rail calendar year 2012, report 12-1. Bureau of Explosives, Transportation Technology Center, Inc., Pueblo, CO.

Center for Chemical Process Safety (CCPS) (2008). Guidelines for chemical transportation safety, security, and risk management (second edition). American Institute of Chemical Engineers, Hoboken, NJ.

Cozzani, V., Bonvicini, S., Spadoni, G., Zanelli, S. (2007). Hazmat transport: a methodological framework for the risk analysis of marshalling yards. Journal of Hazardous Materials 147(1-2), 412-423.

Dick, C.T., Barkan, C.P.L., Chapman, E.D., Stehly, M.P. (2003). Multivariate statistical model for predicting occurrence and location of broken rails. Transportation Research Record 1825, 4855.

Dyer, J.S., Sarin, R.K. (1979). Measurable multiattribute value functions. Operations Research 27(4), 810-822.

Erkut, E., Ingolfsson, A. (2000). Catastrophe avoidance models for hazardous materials route planning. Transportation Science 34(2), 165-179.

Erkut, E., Tjandra, S.A., Verter, V. (2007). Hazardous materials transportation. Handbooks in operations research and management science, volume 14 Transportation, C. Barnhart and G. Laporte (Eds.), North-Holland, Amsterdam, The Netherlands.
Federal Railroad Administration (FRA) (2011). FRA guide for preparing accident/incident Reports. FRA, U.S. Department of Transportation, Washington D.C.

Fudenberg, D., Tirole, J. (1983). Game theory, MIT Press, Cambridge, Massachusetts.
Glickman, T.S. (1983). Rerouting railroad shipments of hazardous materials to avoid populated areas. Accident Analysis and prevention 15(5) 329-335.

Glickman, T.S., Rosenfield, D.B. (1984). Risks of catastrophic derailments involving the release of hazardous materials. Management Science 30(4), 503-511.

Glickman, T.S., Erkut, E., Zschocke, M.S. (2007). The cost and risk impacts of rerouting railroad shipments of hazardous materials. Accident Analysis and Prevention 39(5), 1015-1025.

Glickman, T.S., Erkut, E. (2007). Assessment of hazardous material risks for rail yard safety. Safety Science 45(7), 813-822.

Hilbe, J.M. (2007). Negative Binomial Regression. Cambridge University Press, Cambridge, England.

Kalay, S., French, P., Tournay, H.M. (2011). The safety impact of wagon health monitoring in North America. Proceedings of the World Congress on Railway Research, Lille, France.

Kawprasert, A., Barkan, C.P.L. (2008). Effects of route rationalization on hazardous materials transportation risk. Transportation Research Record 2043, 65-72.

Kawprasert, A., Barkan, C.P.L. (2010). Effect of train speed on risk analysis of transporting hazardous materials by rail. Transportation Research Record 2159, 59-68.

Kawprasert, A. (2010). Quantitative analysis of options to reduce risk of hazardous materials transportation by railroad. Ph.D. Dissertation, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Lai, Y.C., Kawprasert, A., Lin, C.Y., Saat, M.R., Liang, C.H., Barkan, C.P.L. (2011). Integrated optimization model to manage the risk of transporting hazardous materials on railroad networks. Transportation Research Record 2261, 115-123.

Liu, X., Saat, M.R., Barkan, C.P.L. (2011a). Benefit-cost analysis of heavy haul railway track upgrade for safety and efficiency. Proceedings of the International Heavy Haul Association Conference, Calgary, Canada.

Liu, X., Barkan, C.P.L., Saat, M.R. (2011b). Analysis of derailments by accident cause: evaluating railroad track upgrades to reduce transportation risk. Transportation Research Record 2261, 178-185.

Liu, X., Saat, M.R., Barkan, C.P.L. (2012). Analysis of causes of major train derailment and their effect on accident rates. Transportation Research Record 2289, 154-163.

Liu, X., Saat, M.R., Barkan, C.P.L. (2013a). Analysis of major derailment causes on heavy haul railways in the United States. International Heavy Haul Association Conference, New Delhi, India.

Liu, X., Saat, M.R., Barkan, C.P.L. (2013b). Safety effectiveness of integrated risk reduction strategies for rail transport of hazardous materials. Transportation Research Record 2374: 102110.

National Transportation Safety Board (NTSB) (2005). Railroad accident report, RAR-05-04, United States.

Nayak, P.R., Rosenfield, D.B., Hagopian, J.H. (1983). Event probabilities and impact zones for hazardous materials accidents on railroads. Report DOT/FRA/ORD-83/20, FRA, U.S.
Department of Transportation, Washington D.C.
Office of Management and Budget (OMB) (2008). Guidelines and discount rates for benefit-cost analysis of federal programs, Washington D.C.

Ouyang, Y., Li, X., Lai, Y.C., Barkan, C.P.L., Kawprasert, A. (2009). Optimal locations of railroad wayside defect detection installations. Computer-Aided Civil and Infrastructure Engineering 24, 1-11.

Paltrinieri, N., Landucci, G., Molag, M., Bonvicini, S., Spadoni, G., Cozzani, V. (2009). Risk reduction in road and rail LPG transportation by passive fire protection. Journal of Hazardous Materials 167(1-3), 332-344.

Pipeline and Hazardous Materials Safety Administration (PHMSA) (2012). Emergency Response Guidebook 2012. http://www.phmsa.dot.gov/staticfiles/PHMSA/DownloadableFiles/Files/Hazmat/ERG2012.pdf.

Pratt, J.W. (1964). Risk aversion in the small and in the large. Econometrica 32(1/2), 122-136.
Resor, R.R., Zarembski, A.M. (2004). Factors determining the economics of wayside defect detectors. Proceedings of Transportation Research Board of the National Academies, Washington, D.C.

Saat, M.R., Barkan, C.P.L. (2005). Release risk and optimization of railroad tank car safety design. Transportation Research Record 1916, 78-87.

Saat, M.R. (2009). Optimizing railroad tank car safety design to reduce hazardous materials transportation risk. Ph.D. Dissertation, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Saat, M.R., Barkan, C.P.L. (2011). Generalized railway tank car safety design optimization for hazardous materials transport: addressing the trade-off between transportation efficiency and safety. Journal of Hazardous Materials 189(1-2), 62-68.

Schafer, D.H. (2008). Effect of train length on railroad accidents and a quantitative analysis of factors affecting broken rails. M.S. Thesis, Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Surface Transportation Board (STB) (2002-2008). Class I railroad annual report 2002 to 2008. STB, Washington D.C.

Surface Transportation Board (STB) (2009). Public use waybill sample. STB, Washington D.C.
Tournay, H.M., Cummings, S. (2005). Monitoring the performance of railroad cars by means of wayside detectors in support of predictive maintenance. Proceedings of the 8th International Heavy Haul Conference, Rio De Janeiro, Brazil.

Treichel, T.T., Barkan, C.P.L. (1993). Working paper on mainline freight train accident rates. Research and Test Department, Association of American Railroads, Washington, D.C.

Treichel, T.T., Hughes, J.P., Barkan, C.P.L., Sims, R.D., Phillips, E.A., Saat, M.R. (2006). Safety performance of tank cars in accidents: probability of lading loss, report RA-05-02. RSIAAR railroad tank car safety research and test project, Association of American Railroads, Washington, D.C.

Tyrell, D., Jeong, D.Y., Jacobsen, K. (2007). Improved tank car safety research. Proceedings of the ASME Rail Transportation Division Fall Technical Conference, Chicago, IL.

## Chapter 7

## Future Research

### 7.1 Introduction

In this chapter I highlight the needs for future research related to the topics in this dissertation. More specific points related to future research are also discussed in chapters 4, 5 and 6.

### 7.2 Future research

### 7.2.1 Trend of accident rates

The analysis of train derailment probability and severity in this dissertation was based on aggregated accident data within the study period. This approach enabled a larger sample size, but may also introduce uncertainty if there are substantial temporal effects. The next step will be to investigate possible trends in train and car derailment rates on U.S. railroads using appropriate statistical models (Evans 2007, 2010; Silla and Kallber 2012). This would provide insights regarding possible changes in rail safety performance in different periods.

### 7.2.2 Consequence of a multiple-tank-car release incident

This dissertation analyzes the probability of a multiple-tank-car release incident. Future work is needed to analyze the consequences (e.g., population affected) of such releases. Statistical and engineering models may be developed to analyze the relationship between release consequences and the number of tank cars releasing using a geographical information system (GIS) platform.

### 7.2.3 Route-specific flammable liquid transportation risk analysis

In chapter 5, I focused on nationwide risk analysis of unit-train shipments of hazardous materials. The next stage of this work is to perform risk analysis on key corridors transporting flammable liquids accounting for localized differences in infrastructure and operating practices, and their consequent effect on tank car derailment and release probabilities and exposed population densities. In addition, the effectiveness of various risk reduction strategies, individually and in combination needs to be analyzed.

### 7.2.4 Risk comparison under parameter uncertainty

Single-point risk estimates have often been used to evaluate and compare route-specific risk. These do not account for the uncertainty in the estimation of input parameters in the risk analysis, such as accident rate, tank car release probability and affected population. Future research needs to develop an approach to assess the variance in risk estimates and to analyze whether the differences between them on different routes are statistically significant. This method would facilitate a better understanding of the reliability of risk estimation and could aid decision makers in identifying the most appropriate hazardous materials routes.

### 7.2.5 Disutility function in risk analysis

The majority of traditional hazardous materials transportation risk models assume that decision makers are risk-neutral. i.e., they are indifferent between different probability distributions of release consequences, if expected value of consequences are the same. However, previous studies suggest that human beings often exhibit risk aversion when dealing with low-frequency-high-consequence (LFHC) events (Evans and Erlanger 1997; Erkut and Ingolfsson 2000). This
risk-neutrality assumption will be relaxed in the future research by accounting for a decision maker's risk attitudes. A disutility-based model should be developed to understand how different risk attitudes may change decisions related to hazardous materials transportation risk management.

### 7.2.6 Other risk reduction strategies

This dissertation focused on broken rail prevention and tank car safety design improvement as potential risk reduction strategies. The next step is to consider additional risk reduction strategies, such as implementation of wayside inspection technologies, track quality upgrade (increased track inspection and maintenance standards), traffic control system upgrades, high performance rolling stock components (wheels, axles, bearings) and rerouting. The preliminary integrated risk reduction framework developed in this dissertation can be modified to account for the safety benefits and costs of these risk reduction strategies, individually and in combination, in order to address the trade-off between transportation safety and efficiency.

## References

Erkut, E., Ingolfsson, A. (2000). Catastrophe avoidance models for hazardous materials route planning. Transportation Science 34(2), 165-179.

Evans, A.W. (2007). Rail safety and rail privatisation in Great Britain. Accident Analysis and Prevention 39(3), 510-523.

Evans, A.W. (2010). Rail safety and rail privatization in Japan. Accident Analysis and Prevention 42(4), 1296-1310.

Evans, A.W., Verlander, N.Q. (1997). What is wrong with criterion FN-lines for judging the tolerability of risk? Risk Analysis 17(2), 157-168.

Silla, A., Kallber, V.P. (2012). The development of railway safety in Finland. Accident Analysis and Prevention 45, 737-744.

## Appendix A Variance in car derailment rate

Average car derailment rate can be estimated by multiplying train derailment rate by average number of cars derailed per derailment (Anderson and Barkan 2004).
$C^{*}=\frac{m^{*} \times D^{*}}{M}$
Where:
C* = estimated car derailment rate per traffic exposure
$\mathrm{m}^{*} \quad=$ estimated train derailment count given traffic exposure
D* = average number of cars derailed per train derailment
M = traffic exposure

The variance in estimated car derailment rate is denoted by $\operatorname{Var}\left(\mathrm{C}^{*}\right)$. Assuming that estimated train derailment count and estimated derailment severity are independent, $\operatorname{Var}\left(\mathrm{C}^{*}\right)$ is calculated using the model developed by Goodman (1962):
$\operatorname{Var}\left(C^{*}\right)=\frac{\left[E\left(m^{*}\right)\right]^{2} \operatorname{Var}\left(D^{*}\right)+\left[E\left(D^{*}\right)\right]^{2} \operatorname{Var}\left(m^{*}\right)+\operatorname{Var}\left(m^{*}\right) \operatorname{Var}\left(D^{*}\right)}{M^{2}}$

The higher the FRA track class, the greater the average number of cars derailed. It is probably due to the greater maximum allowable operating speeds on higher track classes (Table A.1). Track-class-specific average number of cars derailed per derailment was calculated and used to estimate car derailment rate, measured by number of cars derailed per billion gross tonmiles.

## Table A. 1

Average number of cars derailed per freight-train derailment on Class I railroad mainlines, 2005 to 2009

|  | FRA Track Class |  |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: |
|  | 1 | 2 | 3 | 4 | 5 |
| Average Number of Cars <br> Derailed per Derailment | 5.3 | 7.3 | 8.5 | 9.3 | 10.0 |
| Standard Error | 0.3 | 0.7 | 0.9 | 1.2 | 1.4 |
| Maximum Operating Speed <br> $(m p h)$ of Freight-Trains | 10 | 25 | 40 | 60 | 80 |

The traffic volumes provided by railroads are in gross ton-miles (GTM). A railroadspecific conversion factor was developed to project car-mile data, for converting car derailment rate per billion gross ton-miles to car derailment rate per billion car-miles. The conversion factor (91.61) was developed based on the gross ton-miles and car-miles statistics on Class I mainlines (AAR 2005-2009). The results of estimated car derailment rate per billion car-miles were presented in Table 2.5.

## Appendix B <br> Variable selection of zero-truncated negative binomial model

## Table B. 1

Hierarchical variable selection based on deviance

| Model | Predictors | log- <br> likelihood | Number of parameters <br> (including intercpt) | Model <br> Comparison | Deviance <br> Difference |
| :--- | :--- | :---: | :---: | :---: | :---: |
| 1 | RL+DS+LO+PO | -1367.4 | 5 | $(2)-(1)$ |  |
| 2 | $R L+D S+L O$ | -1367.4 | 4 | $0.002(d f=1)$ |  |
| 3 | $R L+D S+L O+(R L)^{2}+(D S)^{2}+(L O)^{2}$ | -1361.8 | 7 | $(3)-(2)$ | $11.187(d f=3)$ |
| 4 | $R L+D S+L O+(R L)^{2}$ | -1363.1 | 5 | $(4)-(3)$ | $2.61(d f=2)$ |
| 5 | $R L+D S+L O+(R L)^{2}+R L \times D S+R L \times L O+D S \times L O$ | -1348.2 | 8 | $(5)-(4)$ | $29.9(d f=3)$ |
| 6 | $R L+D S+L O+(R L)^{2}+R L \times D S+R L \times L O$ | -1348.2 | 7 | $(6)-(5)$ | $0.006(d f=1)$ |
| 7 | $R L+D S+L O+(R L)^{2}+R L \times D S+R L \times L O+(R L)^{2} \times D S+(R L)^{2} \times L O$ | -1348.1 | 9 | $(7)-(6)$ | $0.20(d f=2)$ |

The final model is as below (Table 3.3)

| Variable | Coefficient | Standard Error | P-value |
| :--- | :---: | :---: | :---: |
| Intercept | 1.38 | 0.71 | 0.05 |
| RL | -0.03 | 0.23 | 0.91 |
| DS | -0.26 | 0.17 | 0.13 |
| LO | -0.57 | 0.34 | 0.09 |
| $(\text { RL })^{2}$ | -0.06 | 0.02 | 0.01 |
| RL $\times$ DS | 0.24 | 0.04 | 0.00 |
| RL $\times$ LO | 0.21 | 0.08 | 0.01 |

Notes:
RL = logarithmic residual train length
DS = logarithmic derailment speed
LO = loading factor

## Appendix C <br> Truncated geometric model for derailment severity

The model assumes that the number of cars derailed follows a truncated geometric distribution (Saccomanno and El-Hage 1989, 1991)
$P N(x)=\frac{p(1-p)^{x-1}}{1-(1-p)^{L_{r}}}$
Where, $\mathrm{PN}(\mathrm{x})$ represents the probability of derailing x cars (including locomotives) in a train derailment. The probability of each car derailment $(P)$ is affected by derailment speed and residual train length (number of cars following the POD, including the POD):
$p=\frac{e^{Z}}{\left(1+e^{Z}\right)}$
$Z=a+b \times \ln (S)+c \times \ln \left(L_{r}\right)$
The mean number of cars derailed per derailment is:
$E(x)=\sum_{x=1}^{L_{r}} P N(x) x=\frac{1}{p}-\frac{L_{r}(1-p)^{L_{r}}}{1-(1-p)^{L_{r}}}$

Where:
$\mathrm{PN}(\mathrm{x}) \quad=$ probability of derailing x cars in a train derailment
$x \quad=$ number of cars derailed
$\mathrm{L}_{\mathrm{r}}=\mathrm{L}-\mathrm{POD}+1=$ residual train length (number of cars following the POD)
$\mathrm{L} \quad=$ train length (total number of cars in a train)
$\mathrm{S} \quad=$ derailment speed $(\mathrm{mph})$
$\mathrm{a}, \mathrm{b}, \mathrm{c} \quad=$ parameter estimates (dependent on accident cause)
In addition to the two variables (speed and residual train length) considered in previous studies, I considered a new factor - loading status of the first car derailed. The loading status of POD (variable name $L_{d}$ ) was treated as a dummy variable (loading $=1$ represents a loaded POD, empty otherwise). Parameter coefficients were estimated using a nonlinear regression on the mean number of cars derailed (Anderson 2005). Table C. 1 presents the parameter estimates for modeling broken-rail-caused number of cars derailed.

Table C. 1
Parameter estimates for modeling the number of cars derailed per broken-rail-caused freighttrain derailment, Class I mainlines, 2002 to 2011

| Parameter | Estimate | Standard Error | $95 \%$ Confidence Interval |  |
| :--- | ---: | :---: | :---: | :---: |
| a (intercept) | 1.2153 | 0.3762 | 0.4761 | 1.9545 |
| b (speed) | -1.2064 | 0.0689 | -1.3417 | -1.0711 |
| c (residual train length) | 0.0039 | 0.0856 | -0.1643 | 0.1721 |
| d (loading status of POD) | -0.3124 | 0.0668 | -0.4436 | -0.1813 |

Figure C. 1 compares the empirical and predicted distributions of train derailment severity. The Kolmogorov-Smirnov (K-S) test shows that the two distributions are statistically identical (Ksa $=$ $0.93 ; \mathrm{P}=0.35$ ). It indicates that the truncated geometric model has a reasonable fit for modeling broken-rail-caused train derailment severity.


Figure C.1: Comparison of empirical and predicted train derailment severity distribution, broken-rail-caused Class I railroad freight-train mainline derailments, 2002 to 2011

## Appendix D <br> Tank car derailment probability by position-in-train

I used the worst-case placement scenarios of all five DOT 111A100W1 tank cars as an example. The conditional derailment probability of these cars are $0.29295,0.29284,0.29283,0.29252$ and 0.29250 , respectively (Figure 4.10). The derailment probability of each car in a specific train position is a Bernoulli variable. The sum of Bernoulli variables with heterogeneous probabilities follows a Poisson binomial distribution (Wang 1993). Therefore, the probability of derailing a certain number of 111 A 100 W 1 tank car is calculated using the algorithm implemented in R (Hong 2013) (Figure D.1).


Figure D.1: Probability distribution of the number of 111A100W1 tank cars derailed per train derailment

## R codes

install.packages('poibin', repos='http://cran.us.r-project.org')
library(poibin)
kk1=0:5
ppl=c(0.29295,0.29284,0.29283,0.29252,0.29250)
$\mathrm{CDF}=$ ppoibin(kk=kk1, $\mathrm{pp}=\mathrm{pp} 1$, method $=$ "RF",wts=c( $1,1,1,1,1$ )

## Appendix E <br> Probability of multiple-car release incidents

Using Equation (5.6), I estimated the probability distribution of the number of tank cars releasing per unit-train mile given FRA track class and method of operation (Table E.1). For example, in signaled, track class 4 , the estimated probability of releasing 10 tank cars is 1.17 in 100 million per unit-train mile.

## Table E. 1

Probability distribution of the number of tank cars releasing per unit-train mile by FRA track class and method of operation

| Method of Operation | FRA Track Class | Number of Tank Cars Releasing per Unit-Train Mile |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Non-Signaled | 1 | 2.00E-06 | 4.02E-07 | 8.47E-08 | 1.83E-08 | 4.04E-09 | 9.00E-10 | $2.02 \mathrm{E}-10$ | $4.57 \mathrm{E}-11$ | $1.04 \mathrm{E}-11$ | $2.34 \mathrm{E}-12$ |
| Non-Signaled | 2 | $1.67 \mathrm{E}-06$ | 7.93E-07 | 3.93E-07 | 1.99E-07 | 1.02E-07 | 5.27E-08 | $2.74 \mathrm{E}-08$ | 1.43E-08 | 7.47E-09 | 3.90E-09 |
| Non-Signaled | 3 | $7.36 \mathrm{E}-07$ | $4.55 \mathrm{E}-07$ | 2.95E-07 | $1.95 \mathrm{E}-07$ | 1.30E-07 | 8.79E-08 | 5.96E-08 | 4.06E-08 | $2.77 \mathrm{E}-08$ | 1.89E-08 |
| Non-Signaled | 4 | $2.98 \mathrm{E}-07$ | 2.07E-07 | 1.51E-07 | 1.13E-07 | 8.52E-08 | 6.49E-08 | $4.97 \mathrm{E}-08$ | 3.82E-08 | $2.95 \mathrm{E}-08$ | $2.28 \mathrm{E}-08$ |
| Non-Signaled | 5 | $9.15 \mathrm{E}-08$ | 6.64E-08 | 5.09E-08 | 3.98E-08 | $3.15 \mathrm{E}-08$ | $2.52 \mathrm{E}-08$ | 2.02E-08 | 1.63E-08 | 1.32E-08 | 1.07E-08 |
| Signaled | 1 | 1.09E-06 | 2.18E-07 | 4.60E-08 | 9.96E-09 | 2.20E-09 | $4.89 \mathrm{E}-10$ | 1.10E-10 | $2.48 \mathrm{E}-11$ | 5.62E-12 | 1.27E-12 |
| Signaled | 2 | 8.33E-07 | 3.95E-07 | 1.95E-07 | 9.88E-08 | 5.07E-08 | 2.62E-08 | 1.36E-08 | 7.12E-09 | 3.72E-09 | 1.94E-09 |
| Signaled | 3 | $3.77 \mathrm{E}-07$ | 2.33E-07 | 1.51E-07 | 9.96E-08 | $6.66 \mathrm{E}-08$ | 4.50E-08 | 3.05E-08 | $2.08 \mathrm{E}-08$ | 1.42E-08 | 9.69E-09 |
| Signaled | 4 | $1.53 \mathrm{E}-07$ | 1.06E-07 | $7.75 \mathrm{E}-08$ | 5.78E-08 | $4.37 \mathrm{E}-08$ | $3.32 \mathrm{E}-08$ | $2.54 \mathrm{E}-08$ | 1.96E-08 | $1.51 \mathrm{E}-08$ | 1.17E-08 |
| Signaled | 5 | $6.60 \mathrm{E}-08$ | $4.79 \mathrm{E}-08$ | 3.67E-08 | $2.87 \mathrm{E}-08$ | $2.27 \mathrm{E}-08$ | 1.82E-08 | 1.46E-08 | 1.18E-08 | 9.53E-09 | 7.73E-09 |

Note: the probabilities of 11-or-more-car release incidents were also estimated but not displayed in this table

Each population density (e.g., remote areas) has its distribution of FRA track class and method of operation (Table E.2).

## Table E. 2

Route miles in remote areas by FRA track class and method of operation

| Method ofOperation | FRA Track Class |  |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 1 | 2 | 3 | 4 | 5 |  |
| Non-Signaled | 558 | 1,601 | 2,267 | 2,721 | 35 | 7,181 |
| Signaled | 41 | 229 | 1,351 | 8,359 | 6,404 | 16,383 |
| Total | 598 | 1,830 | 3,618 | 11,080 | 6,439 | 23,565 |

The average probability distribution of the number of tank cars releasing per unit-train mile in a specific population density can be calculated using the following equation:

$$
\begin{equation*}
P_{i}=\sum_{j k}\left[P_{i j k} \times \frac{M_{j k}}{\sum_{j k} M_{j k}}\right] \tag{E.1}
\end{equation*}
$$

Where:
$P_{i} \quad=$ probability of itank cars releasing per unit-train mile
$P_{i j k} \quad=$ probability of i tank cars releasing per unit-train mile given FRA track class j and method of operation $k$
$\mathrm{M}_{\mathrm{jk}} \quad=$ route miles for trackage with FRA track class j and method of operation k

Based on Tables E. 1 and E. 2 and using Equation (E.1), I calculated the weighted average probability distribution of the number of tank cars releasing in remote areas and other population densities (Table E.3).

Table E. 3
Probability distribution of the number of tank cars releasing per unit-train mile by population density

|  |  | Number of Tank Cars Releasing per Unit-Train Mile |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Population Class | Density (persons per square mile) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Remote | <20 | 3.70E-07 | $1.99 \mathrm{E}-07$ | 1.23E-07 | 8.07E-08 | 5.54E-08 | 3.91E-08 | 2.82E-08 | $2.07 \mathrm{E}-08$ | $1.54 \mathrm{E}-08$ | $1.15 \mathrm{E}-08$ |
| Rural | 20-100 | $4.76 \mathrm{E}-07$ | 2.47E-07 | 1.46E-07 | 9.30E-08 | 6.19E-08 | $4.25 \mathrm{E}-08$ | $2.99 \mathrm{E}-08$ | $2.14 \mathrm{E}-08$ | $1.56 \mathrm{E}-08$ | $1.14 \mathrm{E}-08$ |
| Suburban | 100-1,000 | $5.15 \mathrm{E}-07$ | 2.57E-07 | $1.49 \mathrm{E}-07$ | 9.33E-08 | $6.14 \mathrm{E}-08$ | 4.18E-08 | 2.92E-08 | $2.08 \mathrm{E}-08$ | $1.50 \mathrm{E}-08$ | $1.10 \mathrm{E}-08$ |
| Urban | 1,000-3,000 | 6.13E-07 | 2.66E-07 | 1.43E-07 | 8.70E-08 | 5.63E-08 | $3.79 \mathrm{E}-08$ | 2.62E-08 | $1.86 \mathrm{E}-08$ | $1.33 \mathrm{E}-08$ | 9.72E-09 |
| High | 3,000-10,000 | 7.22E-07 | $2.88 \mathrm{E}-07$ | 1.46E-07 | $8.41 \mathrm{E}-08$ | $5.24 \mathrm{E}-08$ | 3.43E-08 | $2.32 \mathrm{E}-08$ | $1.61 \mathrm{E}-08$ | $1.14 \mathrm{E}-08$ | 8.19E-09 |
| Extremely High | >10,000 | $1.16 \mathrm{E}-06$ | $3.47 \mathrm{E}-07$ | $1.37 \mathrm{E}-07$ | $6.71 \mathrm{E}-08$ | $3.74 \mathrm{E}-08$ | 2.25E-08 | 1.42E-08 | $9.31 \mathrm{E}-09$ | $6.26 \mathrm{E}-09$ | $4.30 \mathrm{E}-09$ |
| Average |  | $4.81 \mathrm{E}-07$ | $2.40 \mathrm{E}-07$ | $1.40 \mathrm{E}-07$ | 8.86E-08 | 5.90E-08 | $4.05 \mathrm{E}-08$ | $2.86 \mathrm{E}-08$ | $2.05 \mathrm{E}-08$ | $1.50 \mathrm{E}-08$ | $1.10 \mathrm{E}-08$ |

Note: the probabilities of 11-or-more-car release incidents were also estimated but not displayed in this table

Based on Table E.3, I estimated the probability of at least a certain number of cars releasing per unit-train mile for each population density. The results were presented in Table 5.3.

## Appendix F

## Data for modeling broken-rail-caused car derailment rate

## Table F. 1

Input data for regression analysis

| Number of Broken-RailCaused Car Derailments | Railroad Indicator | Year | Annual Rail Maintenance \& Repair Cost per Track-Mile (\$) | Billion Gross Ton-Miles |
| :---: | :---: | :---: | :---: | :---: |
| 53 | A | 2002 | 2,194 | 959 |
| 334 | B | 2002 | 1,729 | 469 |
| 23 | C | 2002 | 1,405 | 38 |
| 35 | D | 2002 | 2,443 | 373 |
| 277 | E | 2002 | 2,560 | 1,086 |
| 155 | A | 2003 | 2,331 | 1,003 |
| 310 | B | 2003 | 2,179 | 488 |
| 37 | C | 2003 | 1,019 | 37 |
| 126 | D | 2003 | 2,443 | 380 |
| 449 | E | 2003 | 2,395 | 1,109 |
| 270 | A | 2004 | 2,614 | 1,110 |
| 165 | B | 2004 | 3,636 | 510 |
| 34 | C | 2004 | 711 | 44 |
| 59 | D | 2004 | 2,915 | 408 |
| 280 | E | 2004 | 2,434 | 1,130 |
| 200 | A | 2005 | 2,609 | 1,162 |
| 148 | B | 2005 | 3,421 | 504 |
| 36 | C | 2005 | 535 | 54 |
| 160 | D | 2005 | 3,080 | 417 |
| 352 | E | 2005 | 2,751 | 1,141 |
| 316 | A | 2006 | 2,986 | 1,228 |
| 237 | B | 2006 | 3,560 | 513 |
| 20 | C | 2006 | 880 | 56 |
| 133 | D | 2006 | 3,424 | 419 |
| 134 | E | 2006 | 2,991 | 1,175 |
| 77 | A | 2007 | 3,031 | 1,231 |
| 219 | B | 2007 | 3,618 | 496 |
| 52 | C | 2007 | 1,388 | 55 |
| 153 | D | 2007 | 3,397 | 401 |
| 234 | E | 2007 | 3,265 | 1,155 |
| 83 | A | 2008 | 3,432 | 1,237 |
| 150 | B | 2008 | 4,153 | 487 |
| 53 | C | 2008 | 2,773 | 54 |
| 93 | D | 2008 | 2,956 | 393 |
| 155 | E | 2008 | 3,224 | 1,119 |

## Appendix G <br> Broken-rail-caused car derailment rate by maintenance cost per ton

## Table G. 1

Broken-rail-caused car derailment rate on U.S. Class I mainlines, 2002 to 2008

| Parameter | Estimate | Standard Error | $95 \%$ Confidence Limits |  | Pr $>$ ChiSq |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Intercept | -0.0676 | 0.2859 | -0.6279 | 0.4928 | 0.8132 |
| Annual Rail Maintenance Cost per |  |  |  |  |  |
| Million Gross Tons (Thousand Dollars) | -0.0056 | 0.0015 | -0.0086 | -0.0027 | 0.0002 |
| Dispersion Parameter $(1 / \eta)$ | 0.3307 | 0.0772 | 0.2093 | 0.5225 |  |
| Derian |  |  |  |  |  |

Deviance $=36.7$ (Degrees of Freedom $=33$ )
$\mathrm{P}=0.30>0.05$
Compared to Table 6.1, using either rail maintenance cost per mile or rail maintenance cost per ton as the predictor variable would fit the data well, because of the significant correlation between these two variables (Pearson correlation coefficient is $0.96, \mathrm{P}<0.001$ ). Depending on the data available to the railroads, they may use the most appropriate predictor to estimate broken-rail-caused car derailment rate on their routes.

## Appendix H Car derailment rate reduction by increased maintenance cost

Let $\mathrm{C}_{1}$ and $\mathrm{C}_{2}$ represent annual rail maintenance cost per track mile (thousand dollars), respectively. It is assumed that $\mathrm{C}_{2}>\mathrm{C}_{1}$. Let $\Delta \mathrm{C}=\mathrm{C}_{2}-\mathrm{C}_{1}$. The percent broken-rail-caused car derailment rate reduction is:

$$
\begin{equation*}
\Delta Z_{b r}=\frac{Z_{b r}\left(C_{1}\right)-Z_{b r}\left(C_{2}\right)}{Z_{b r}\left(C_{1}\right)} \tag{H.1}
\end{equation*}
$$

From Equation (6.11), broken-rail-caused car derailment rate is an exponential function of rail maintenance cost C :

$$
\begin{align*}
& Z_{b r}\left(C_{1}\right)=\exp \left(-0.1868-0.3356 \mathrm{C}_{1}\right)  \tag{H.2}\\
& Z_{b r}\left(C_{2}\right)=\exp \left(-0.1868-0.3356 \mathrm{C}_{2}\right) \tag{H.3}
\end{align*}
$$

From Equations (H.1) to (H.3),
$\Delta Z_{b r}=1-\exp \left[-0.3356\left(C_{2}-C_{1}\right)\right]=1-\exp (-0.3356 \Delta C)$
So proof of Equation (6.13) is completed.

## Appendix I

## Derivation of optimization model

Let $\Delta \mathrm{R}$ denote the percent risk reduction by preventing broken rails and enhancing tank car safety design. Let $Z, M, C$ represent the national average car derailment rate per traffic exposure, traffic exposure and average release consequence, respectively. Let $\mathrm{P}_{\mathrm{b}}$ and $\mathrm{P}_{\mathrm{u}}$ represent the CPR of the baseline tank car and the enhanced tank car, respectively.

Broken rail prevention and tank car design enhancement affect tank car derailment rate and conditional probability of release, respectively. By definition, we have
$\Delta R=\frac{Z M P_{b} C-Z_{\text {new }} M P_{\text {new }} C}{Z M P_{b} C}$
It is assumed that the proportion of the baseline tank car fleet to upgrade is $b$, so the weighted average of CPR in the mixed tank car fleet (baseline tank car and enhanced tank car) is:

$$
\begin{equation*}
P_{\text {new }}=P_{b}(1-b)+P_{u} b \tag{I.2}
\end{equation*}
$$

It is assumed that the proportion of broken-rail-caused car derailments to prevent is $e$. In addition, the proportion of broken-rail-caused car derailments among all accident causes is $m$. Therefore, the overall tank car derailment rate after implementation of broken rail prevention is estimated as:
$\mathrm{Z}_{\text {new }}=\mathrm{Z}(1-\mathrm{em})$
From Equations (I.1) to (I.3),
$\Delta R=\frac{Z M P_{b} C-Z(1-e m)\left[P_{u} b+P_{b}(1-b)\right] M C}{Z M P_{b} C}$
Equation (I.4) is simplified as:
$\Delta R=1-(1-e m)\left[(1-b)+\frac{P_{u}}{P_{b}} b\right]$
To maximize percent risk reduction,
Maximize $\left\{1-(1-e m)\left[(1-b)+\frac{P_{u}}{P_{b}} b\right]\right\} \Leftrightarrow$ Maximize $\left\{-(1-e m)\left[(1-b)+\frac{P_{u}}{P_{b}} b\right]\right\}$
From Equation (6.13), $\Delta Z_{b r}=1-\exp (-0.3356 \Delta C)$
Let $\Delta \mathrm{Z}_{\mathrm{br}}=\mathrm{e}$ (percent reduction of broken-rail-caused car derailment rate):
$\Delta C=\frac{\ln (1-e)}{-0.3356}$
$\Delta \mathrm{C}$ is the increment of annual rail maintenance cost per track-mile, measured by thousand dollars. So total increment of annual rail maintenance cost over 160,240 track miles of U.S. Class I rail network is:

$$
\begin{equation*}
C_{\text {rail }}=\frac{\Delta C \times 160240}{1000} \text { (measured by million dollars) } \tag{I.8}
\end{equation*}
$$

Also, if the entire baseline tank car fleet is upgraded, the total annual cost is $\$ 9.7$ million.
Assuming that the proportion of baseline tank cars to upgrade is $b$, the corresponding cost is $\mathrm{C}_{\operatorname{tank}}$ $=9.7 \mathrm{~b}$.

The total annual cost for broken rail prevention and tank car upgrade does not exceed the budget, that is:
$\mathrm{C}_{\text {rail }}+\mathrm{C}_{\text {tank }} \leqslant$ Budget
Equation (I.9) is re-written as:
$\frac{\frac{\ln (1-e)}{-0.3356} \times 160240}{1000}+9.7 b \leq$ Budget

The optimization model is derived.

