Risk-Based Model for Effective Marshalling of Dangerous Goods Railway Cars

by

Morteza Bagheri

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Civil Engineering

Waterloo, Ontario, Canada, 2009

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Today, railroad companies transport many varieties of dangerous goods (DG). Train derailments, especially those involving DG, can be catastrophic in terms of loss of life and environmental damage. In North America, the transportation of DG is governed by regulations published by the Canadian and United State's governments. While the regulation is important in terms of providing overall guidelines, they do not address the problem of optimally positioning DG cars in terms of their potential for derailment and the associated risks. Currently, most rail yard operations do not consider the potential effect of the position of DG cars on the risk of derailment.

This research is concerned with the problem of how to place DG cars in a train in the train assembly process so that the overall derailment risk can be minimized. The approach considers both the probability of railway cars derailing en route by position as well as the time associated with additional operations in the rail yard. This work has resulted in a useful decision support tool for assisting rail yard operation managers to achieve an optimum trade-off between derailment risk and operating costs in assembling trains. The merits of this new car placement model are illustrated through a case study of a real railway corridor that connects Barstow Yard in California to Corwith Yard in Chicago over 2100 miles and involves a range of track features. The case study demonstrates that the proposed risk minimization strategy could be implemented with minimal rail yard operation cost.

Acknowledgements

First of all, I want to praise God for giving me the strength to complete my program. The completion of this research would not have been possible without the efforts, support, and inspiration of many individuals.

I would like to especially to acknowledge my supervisors, Dr. Frank Saccomanno and Dr. Liping Fu, for their guidance and support. It has been a privilege to work under their supervision.

Next, I would like to express sincere gratitude to Dr. Shojaeddin Chenouri for his invaluable insights and fruitful discussions with him over the last couple of years. I would also like to acknowledge the members of my Ph.D. committee, Prof.Tarek Hegazi and Prof.Bruce Hellinga, for their time and helpful comments.

Further, I am grateful to my friends, Amir Golroo, Sayyed Bashir Sadjad, Vahid Miraftab, and Hamid Zarrabi-Zadeh for their advice and time. In addition, I would like to acknowledge my colleagues from the Transportation Group, particularly David Duong, Pedram Izadpanah, Amir Hosein Ghods, Peter Park, Flavio Cunto, and Akram Nour, for their friendship and help.

I wish to thank others who also contributed their time and effort, including Mirek Wierucki, Jeff Rosenblum, Duncan Simpson, Angus Mcleod of Canadian Pacific Railway; Corrine O'Brien and Professor Peter Dederer of George Brown College in Toronto; Eva Dodsworth (Map Library), Oana Danila (Statistical Consulting Service), and Mary McPherson, for their help and support.

I also want to thank my parents for all they have done for me, especially for their patience during the past six years while I have been studying in Sweden and Canada. Finally, and most importantly, I want to thank my wife, Vahideh, for all her love and understanding.

Dedication

This is dedicated to my beloved Vahideh.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Rail accidents can be classified into three main types: derailments, collisions (including head on, rear end, and side), and highway railway grade crossing accidents. As illustrated in Figure 1.1, train derailments account for greater than 50% of all train accidents in Canada and the US. Train derailments also represent a significant share of personal injury and property damage every year [47]. For instance according to the United States Federal Railroad Administration (FRA), as many as 300 personal injuries occur and the related property damage costs can be as high as $300 million per year in the US [9].

The potential threat to the public and the environment is even more significant when derailments involve cars carrying dangerous goods (hazardous materials). This threat depends on the nature of the dangerous goods (DG) involved and their propensity for fires, explosions, and toxic impacts. According to FRA statistics, on average, $500,000 in direct property damage results from a derailment involving DG [9]. This damage is five times greater than if DG are not involved. The major threat posed by DG involvement is related to personal injury and long-term environmental damages.

Two recent train derailments involving DG illustrate the extent of personal injury and environmental impact possible. In February 2003, a freight train carrying DG derailed in Ontario, resulting in the derailment of seven cars that contained liquefied petroleum gas. The derailed cars produced an extensive fire that burned for three days. Smoke from the fire resulted in significant degradation of in the air quality in the vicinity of the derailment. More than 300 residents had to be evacuated, and two crew members sustained major injuries [46].

In January 2005, a freight train derailed in South Carolina, involving a numbers of DG cars and three bulk chlorine tankers (80 tons) whose highly toxic substances were released.

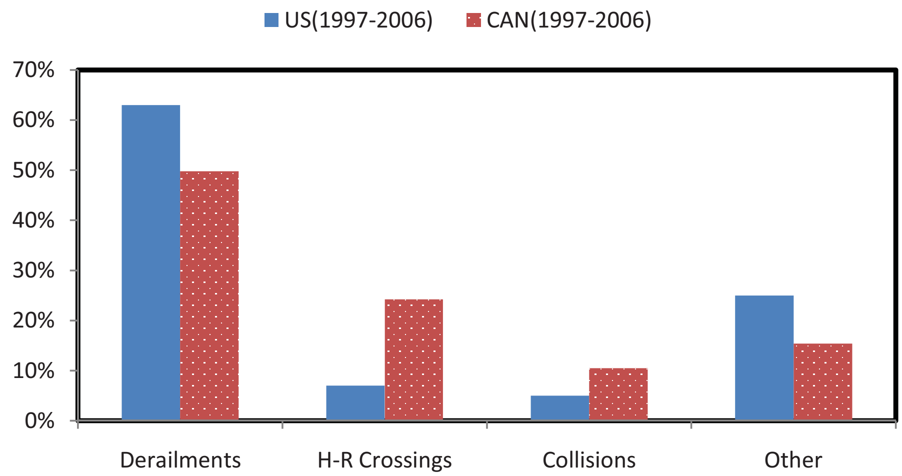

Figure 1.1: Rail accidents by type in the US and Canada (Source:[9, 48])

Nine people died, hundreds were injured, and about 6000 residents had to be evacuated. The total damage from the derailment exceeded $6.9 million [21].

US railroads carry approximately 1.8 million carloads of DG annually (approximately 5% of total freight rail traffic). Thus every day, thousands of DG are moved by rail across the US [44]. In Canada, approximately 500,000 carloads of DG are shipped annually, or 12% of total freight shipped by rail nationally [47]. Figure 1.2 shows that, based on the FRA database, DG cars were involved in half of the total number of derailments of trains carrying DG in the US in the period 1997-2006.

Recently, in response to a number of derailments involving DG in British Columbia (e.g., the release of sodium hydroxide into the Cheakamus River), Alberta and Quebec, Transport Canada announced a formal review of the Railway Safety Act to consider ways to manage and reduce the risk associated with the transport of DG [23]. This review will address all of these challenges by identifying possible changes to the Act that would strengthen Canada's regulatory regime, including train make-up issues.

In the US, the Hazardous Materials Transportation Act of 1994 recommended an investigation of the placement of DG cars in general freight trains in order to reduce the potential of DG cars involvment in derailments. Section 111 of the Hazardous Materials Transportation Authorization Act (P.L. 103-311) states that

> "The Secretary of Transportation shall conduct a study of existing practices regarding the placement of cars on trains, with particular attention to the placement of cars that carry hazardous materials. In conducting the study,

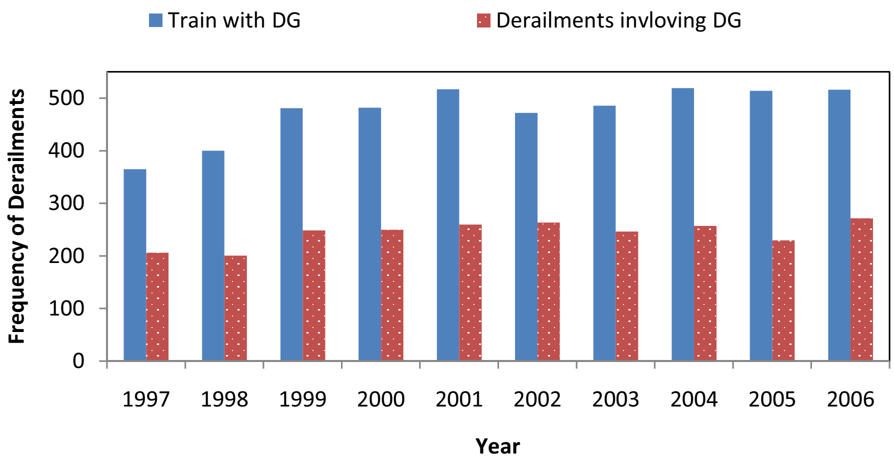

Figure 1.2: Share of car derailments involving DG from FRA 1997-2006 (Source:[9])

> the Secretary shall consider whether such placement practices increase the risk of derailment, hazardous materials spills, or tank ruptures or have any other adverse effect on safety . . ."

Recently, the Association of American Railroads [43] underlined railroad companies desire to reduce the risk of DG transportation. The association suggested that risks associated with the transport of DG could be in the billions of dollars. A number of state agencies in the US have gone even further to consider banning such rail shipments within their jurisdictions [45].

Train derailments are an important safety concern, and they become even more critical when DG are involved. There are two ways of addressing this concern:

1. Improving track infrastructure (track, switches, etc.) and rolling stock (cars, locomotives), which are major factors contributing to the occurrence of derailments.

2. Improving train operational strategies so that the total risk of derailments is reduced.

This research is concerned with the second approach, with a specific focus on mitigating derailment risk through improved DG marshalling practice in the train assembly process.

1.2 Train Assembly Process and Marshalling of DG Railway Cars

This section presents the current marshalling strategies involving DG as adopted in Canada and the US. The safety problems associated with the current situation are underscored. This review is to justify the proposed model introduced in Chapter 2.

A rail yard is a system of tracks used for making up (assembling) trains, storing cars, and other purposes[29]. In North America there are two types of rail yards, flat yards and hump yards. In hump yards, a specially built hill known as a hump is used for sorting cars, while in flat yards, yard engines sort cars by flat switching [38]. "Switching is the operation that separates two adjacent sets of cars, and sends the sets to their assigned classification tracks."[28, 38]

Figure 1.3 illustrates the four basic operational stages involved in a typical rail yard train make-up process:

1. The train arrives at the receiving track, where cars (tagged by destination) are inspected. After inspection, they are queued for humping.

2. In a conventional hump yard operation, a train is pushed to the hump for disassembly on a "first come, first served" basis. On the hump, cars are cut from the rest of the train one by one and shunted to their respective classification track, with all cars on each track sharing a common destination.

3. Once a full block of cars (all going to a common destination) is assembled on the classification track, the cars are checked for position violations such as the inadequate separation of DG cars from train personnel and unsafe distance between incompatible DG materials. In the current train assembly process, a yard engine is used to cut a specific number of cars from the classification track to ensure that these concerns are satisfied.

4. Once the full blocks of cars have been assembled in the classification yard, a yard engine shunts each block to a departure track. The order of blocks in the train on the departure track is determined by the sequence of intermediate and final destination points along the corridor. On the departure track, each train is attached to a locomotive, and a final inspection is carried out prior to departure.

In rail yard operations, the key step in train assembly takes place on the classification tracks and is sometimes referred to as the pull-down process. Pull-down crews have two main tasks: coupling cars in the classification yard and pulling them to the departure track to form the outbound train [60].

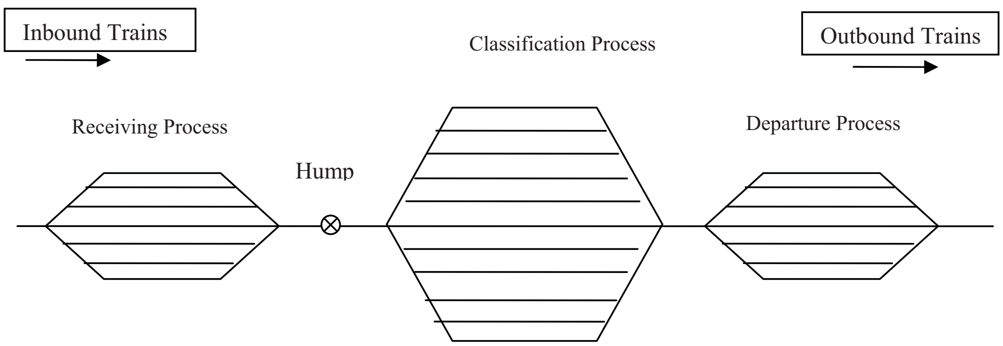

Figure 1.3: Hump yard schematic

Figure 1.4 illustrates the train assembly classification process, which consists of categorizing blocks of cars based on destination. The main problem addressed in the classification process is marshalling, which determines the specific position of individual cars in a train with a key consideration of ensuring sufficient separation of incompatible DG cars from each other. Marshalling or building a train is the process by which a yard engine (locomotive) assembles cars into a train.

The position of DG in a train has a major effect on the risk associated with derailment, more specifically,

- The involvement of DG cars in derailments increases the damages associated with these derailments. The extent of damages is a function of the type of DG material.
- For multiple DG car derailments, the location of incompatible materials near each other increases the extent of possible damage.

The probability of involving DG cars in a derailment can be decreased by systematically placing theses cars in positions that are less prone to derailing for a given route. Therefore, prior knowledge of train derailment patterns can help to improve marshalling operations involving DG. Furthermore, it is important from a practical point of view to understand the cost and efficiency implications of different marshalling strategies as the safest strategy may not be justified if train assembly costs are prohibitively high.

1.3 The Research Problem

As discussed previously, the basic problem with existing train marshalling practice is that it does not consider derailment risk by car position for a given train travelling along a

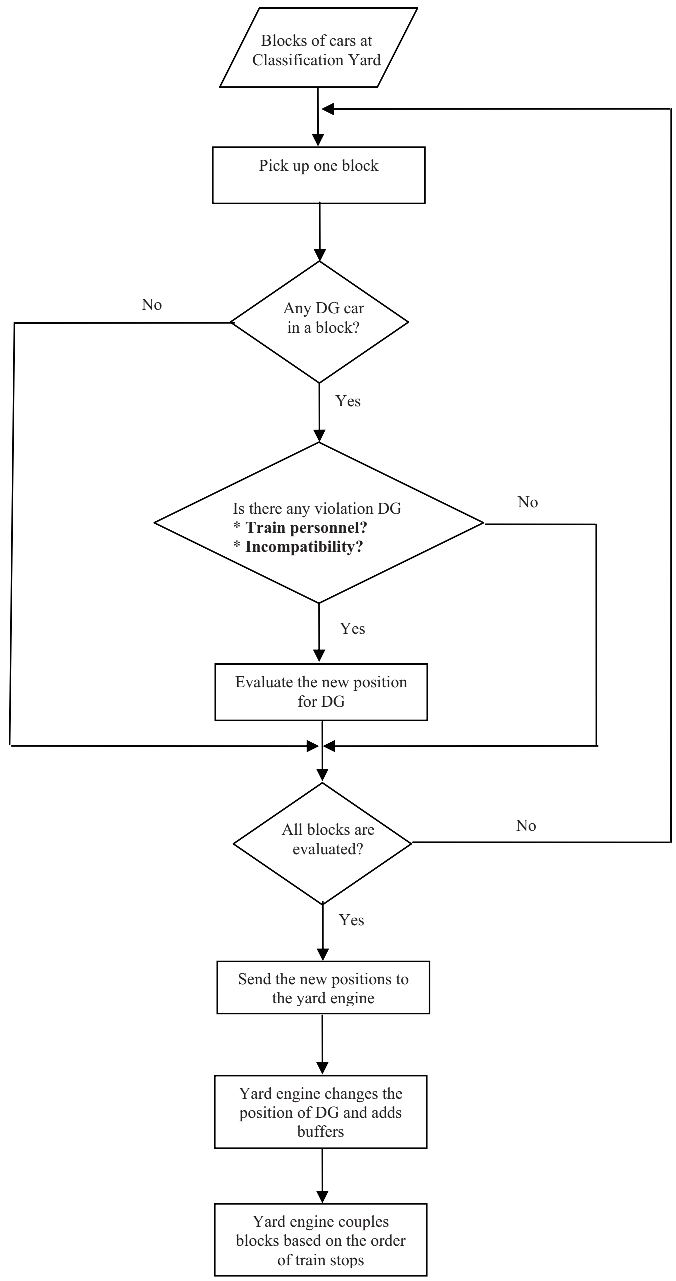

Figure 1.4: Current classification operation procedure

specific route. Hence, there is a possibility that DG cars could be placed in positions that are more prone to derailments.

Neither the Canadian or US regulations consider the relationship between DG car positions and derailment risk [22, 50]. Furthermore, although the specification advised in the International Carriage of Dangerous Goods by Rail (RID) are important in terms of providing overall guidelines, the problem of optimally positioning DG cars in terms of their potential for derailment is not addressed [35].

Over the last two decades, various studies have been conducted on train derailments. In 1979, the Volpe National Transportation Systems Center published a study [34] suggesting that the front of a train is more prone to derail under loaded conditions, thus implying that DG cars should be placed closer to the rear. A similar study by Battelle [57] divided the train into segments and then evaluated the probability of derailments for each segment. This study also provided a risk-based ranking of incompatible materials to determine the worst-case combinations of different types of DG being placed in proximity to one another. A study by FRA [40] concluded that empty cars should not be placed in the front; that is, the preferred position for loaded cars (DG and non-DG) should be the front part of the train. The Canadian Institute of Guided Ground Transport (CIGGT) [32] investigated the risk to train crews as related to position and separation distance of DG in conventional freight trains based on Canadian derailment data. This study offered no recommendations as to preferred low-risk placement of DG cars along the train.

Saccomanno and El-Hage [53] established derailment profiles by position for shipments of DG. The main focus of this research was to determine the probability of derailment for each position along a given train and develop a model to predict the number of cars derailing by train derailment cause. The study did not explicitly consider train assembly operations in the rail yard.

More recently, Anderson and Barkan [11] studied derailment probabilities at an aggregate level using recent FRA data. A recent study by English et al., [36] developed a derailment model at the disaggregate level based on the Canadian Railway Occurrence Data System (RODS). Another study by Bagheri [13] evaluated the risks associated with DG cars in sidings. These studies failed to explicitly consider DG placement risks in rail yard marshaling operations.

In summary, the research to date does not provide adequate scientific evidence on the risk implications associated with DG car placement along the train consist. In the absence of such evidence, current practice in marshaling DG cars has been guided more by rail yard assembly costs and efficiencies than by strategies to counter the underlying risks.

1.4 Research Objectives

The focus of this research is to investigate how the position of different types of DG cars in a train affects their chances of being involved in a derailment and how the relative placement of different types of materials in a train affects the consequence of a derailment. The intent is to reduce the risk of DG involvement in train derailments by reducing the probability of any car carrying DG being involved in a derailment block. This study has four specific objectives:

1. to evaluate different risk-based approaches for effective marshalling of DG cars along a train consist.
2. to develop a comprehensive risk minimization model for application that incorporates rail-yard train assembly and cost/time constraints.
3. to apply the model to several representative shipment scenarios (mix of DG, mix of destinations, different transhipment points, etc.) and investigate the effects of different marshalling strategies for typical shipment scenarios.
4. to apply the model to a case study corridor with known shipment volumes and shipment destinations and compare the case study results to existing marshalling regulations and make appropriate recommendations.

The scope of the study is limited to train assembly issues at a rail yard. Despite this focus, the model will account for the train derailment risk estimated along a given route for different DG shipments. The research is concerned only with conventional freight train assembly and does not consider unit trains, which are not subject to assembly at rail yards. Conventional freight trains normally consist of a mix of various types of freight, including different types of DG with different destination points. In this analysis, the primary safety concern is derailments as opposed to other types of train accidents, such as collisions and accidents at highway-railway crossings. Furthermore, while derailments can influence cars prior to the point of derailment, this research is limited to derailed cars situated after that point.

DG risk in this research refers to a potential derailment of cars carrying some type of DG along a given route or route segment. Subsequent events such as releases, fires,explosions, etc. are not within the scope of this analysis. Furthermore, it is assumed that the effect on total risk resulting from the interaction of incompatible DG materials derailing in proximity to one another can be ignored, such that all derailing DG cars are treated equally in terms of the potential threat they pose to the population and environment.

1.5 Organization of the Thesis

This thesis is comprised of six chapters:

Chapter 2: DG Car Placement Model Framework

Chapter 2 describes a model framework to formulate the DG car marshalling problem, focusing on introduction of two required components of the proposed model: an in-transit risk model and rail yard marshalling model.

Chapter 3: In-transit Risk Model

Chapter 3 presents the in-transit risk model as the first component of the study framework. This model provides an estimate of the risk of derailment by position along a route segment for a given mix of DG and non-DG cars making up a train consist. The following four sequential components are discussed in this chapter: causes of train derailment, probability of train derailment, point of derailment, and number of cars derailing.

Chapter 4: Rail Yard Marshalling Model

Chapter 4 introduces the rail yard marshalling model as the second component of the study framework and includes practical operations restrictions such as an estimation of the rail yard train assembly costs (or processing time) for each train and the DG placement strategy or plan.

Chapter 5: Case Study Application

Chapter 5 illustrates the merits of this new model through a case study abstracted from a real railway corridor. The corridor connects Barstow Yard in California to Corwith Yard in Chicago over 2100 miles and involves a range of track features. The case study demonstrates that the proposed risk-minimization strategy could be implemented with a minimum overhead in terms of rail yard operation costs (time).

Chapter 6: Conclusions

Chapter 6 summarizes the conclusions, with recommendations for future research.

It should be noted that this thesis does not include a separate literature review chapter. However, pervious studies are discussed at the beginning of each chapter.

Chapter 2

Methodology

This research proposes a train assembly model that differs from past models in its consideration of two important points: the risk of DG cars being involved in a derailment for a specific route; the cost of marshalling at a rail yard. This chapter presents the overall framework proposed to formulate the underlying DG-car-marshalling problem.

The focus of this study is on the classification process at rail yards where humping operations are completed and cars are separated on the classification track based on their destination. In cases involving DG cars, the yard engine is required to reassemble the order the cars in blocks that prevent any violation associated with DG. The yardmaster in the tower is responsible for sending the tasks to the yard engine. In addition, based on current regulations regarding the position of DG cars, another question should be answered in this study: How should the yard engine change the position of DG cars in the train consist in order to reduce the risk of a derailment?

This research problem has been stated as an optimization model which is the main part of this research and addresses the problem of which cars should be placed in a given block and which block should be placed in a given departing train. Minimizing total risk, taking operational limitations into consideration, is the objective of the optimization model, which calculates total risk by summing the risk of each car and then determining the combination of cars that results in the least risk and cost.

The optimization model considered in this study consists of two major models: an in-transit risk model (objective function) and a rail yard marshalling model. The in-transit risk model provides an estimate of the risk of derailment by position along a route segment for a given mix of DG and non-DG cars making up the train consist. The rail yard marshalling model introduces practical operations restrictions. For instance, obtaining the rail yard train assembly costs (or processing time) for each train and DG placement strategy is discussed in this model. The model framework is illustrated in Figure 2.1.

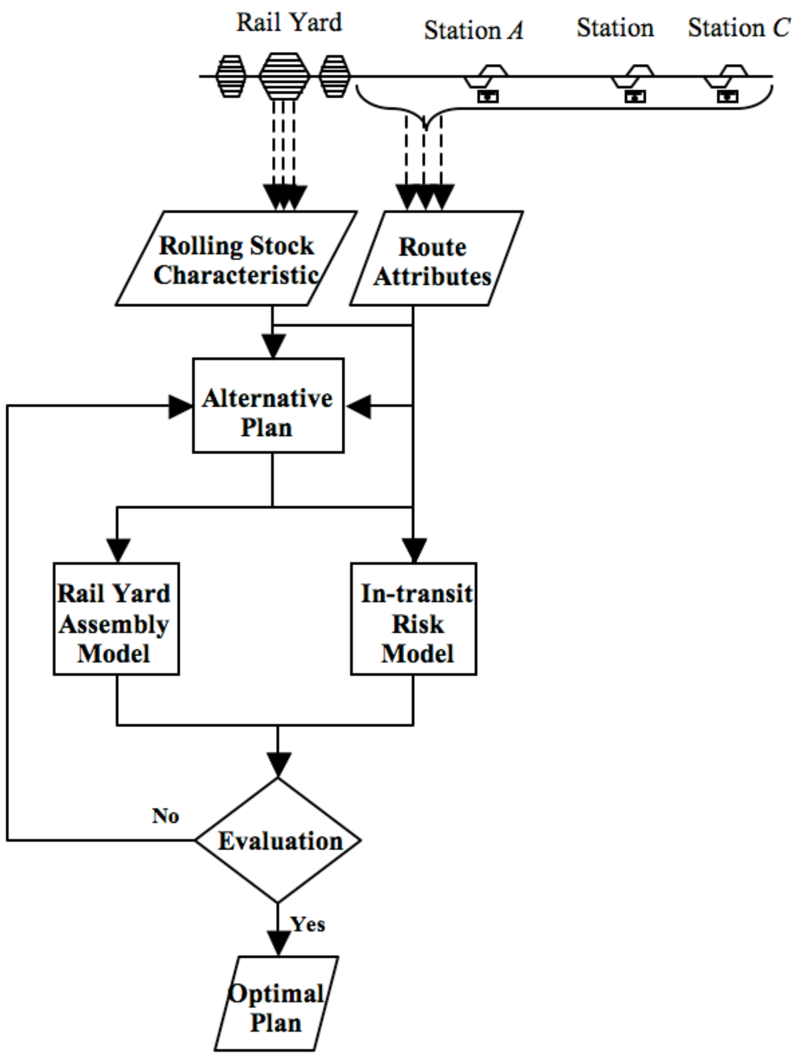

Figure 2.1: DG car placement framework

Before formulating the problem, one must know the inputs and outputs of the model. The inputs are 1) the rolling stock characteristics such as the number of cars in each block; the total number of cars, which determines the train length; the number of DG cars in each block; the types of DG in each block; and 2) the route attributes, such as the minimum radius of curvature, maximum longitudinal gradient, and maximum supperelevation. The outputs are 1) the order of cars in each block and 2) the order of blocks in each train consist. Ideally, after running the model and obtaining the optimal order of cars and blocks, the yardmaster will send the output to a yard engine to change the positions.

2.1 Model Framework

This framework (Figure 2.1) begins with the arrival of a train at the hump yard with DG and non-DG cars tagged for specific destinations. These cars are disengaged and humped to the classification track on a first-come, first-served (FCFS) basis. Each classification track contains cars that are coupled together based on a common destination point (a block) along the route. Once a block has been completed with the desired number of cars, a yard engine shunts it to the departure track, where it is combined with other blocks to make up a train, and which is subsequently inspected prior to departure. The order of blocks on a given train is set with respect to the sequence of intermediate destination points along a route, such that the block assigned to the closest destination is shunted to the front of the train, followed by the block assigned to the next closest destination, and so on until the final destination block is connected to the train.

The order of cars in the classification track is currently set on an FCFS basis that does not explicitly consider DG derailment risk from position along the route. In consequence, DG cars may not be excluded from those positions that are more prone to derailment. Under current regulations [22, 50], a yard engine may be used to insert a designated number of non-DG buffer cars to separate DG cars from train operating personnel in, for instance the locomotive or caboose. Current regulations may also take into account possible train instability problems caused by locating loaded and empty cars near one another.

The DG car placement framework considered in this research envisions the introduction of an in-transit risk model that assigns DG cars to those positions along the train that have the lowest probability of derailing along the different route segments. This strategy serves to modify the FCFS approach currently in use at the interface between the hump and the classification track. The modified DG placement strategy is implemented by the yard engine on instruction from the hump controller. It is proposed that this strategy be introduced at the interface between the hump and the classification track prior to these cars being shunted to the departure track.

As illustrated in Figure 2.1, the DG car placement strategy receives input from two sources: 1) DG and non-DG shipment volumes by intermediate and final destination points along a given route and 2) route attributes and operating and design features by segment. The first input is used to estimate the number of DG and non-DG cars traversing a given segment of track, while the second input is used to estimate car derailment probabilities by position for each route segment. Presumably, the yard engine plan will include the full set of marshalling instructions required to modify DG car placement based on minimizing in-transit derailment risk. The in-transit risk model in Figure 2.1 estimates three specific probabilities: train derailment, point of derailment (POD) by position, and number of cars derailing for different PODs. Within the scope of this study, the number of cars derailing reflects the basic unit of "in-transit risk", such that, the risk consequence is 1 for every DG car derailing and 0 otherwise.

2.2 Modeling the Marshaling of DG Railway Cars(MDG)

This section describes the new approach to formulating the problem: the Marshalling of DG - Two levels model (MDG2). In this model, the decision is divided into two sequential problems. The first is to find the best combination of cars within each block, while the second is to determine the order of blocks.

Finding the optimal position for DG cars along a train consist is a combinatorial problem with the objective of minimizing total risk caused by derailment and operating costs.

The model divides the marshalling problem into two levels of decision problems. The first is to find the order of cars within each block (i.e., those with the same destination), and the second attempts to find the order of blocks in the train consists. The optimal order of cars in a block depends on the probability of derailment and the incompatibility of any DG involved, and this probability in turn depends on the order of blocks. As a result, these two optimization problems are interrelated and must be solved together iteratively.

To illustrate the problem, a train with three destinations is considered: block B consisting of 30 cars, ten of which contain DG is located in the first position, block C consisting of 50 cars, 20 DG in the second, and block A consisting of 20 cars, five DG at the third. Under the assumption that there is no difference between DG in terms of the level of hazard, the possible combinations are shown in Figure 2.2. This large sized problem is fundamentally difficult to solve.

The probability of derailment for each position can be calculated. The best combination can be identified by minimizing the total risk under the incompatibility constraint. This step repeats for the next assumed order of blocks. Note that changing the sequence of the blocks will change the best corresponding combination of cars in each block. For all possible combinations of blocks (Figure 2.3), this procedure repeats.

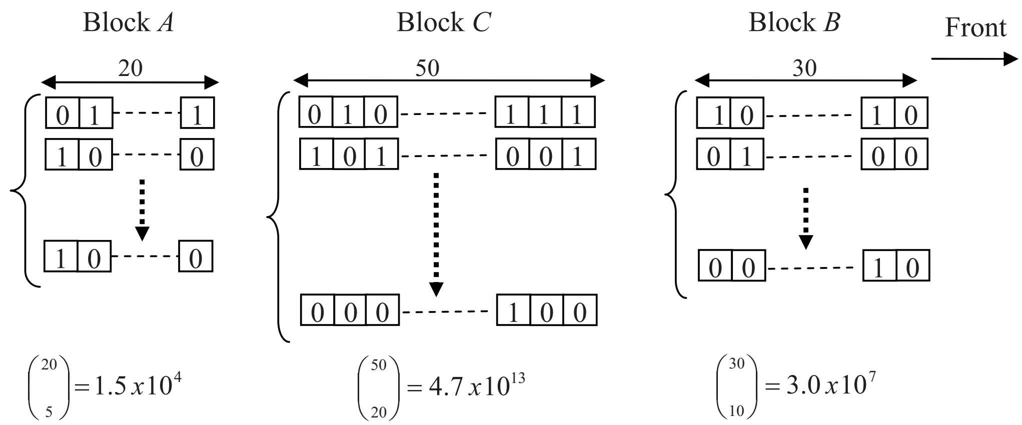

Figure 2.2: Possible combinations of cars in each block

The second level develops the order of the blocks.

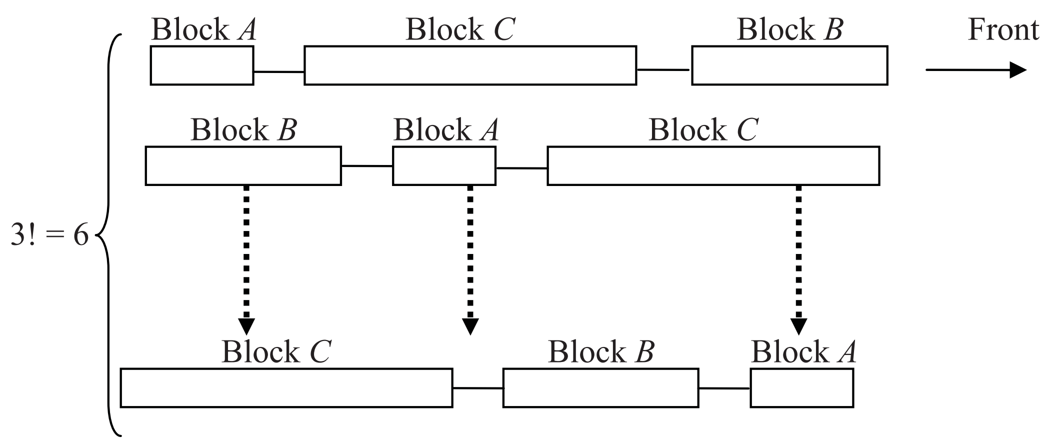

Figure 2.3: Possible combinations of the blocks

At this level, each block has a corresponding risk value calculated in the first level. Thus, the risk of a whole train consist is the summation of the risks of all blocks. The second level finds the order of blocks with the minimum accumulated risk. More details are presented in Figure 2.4.

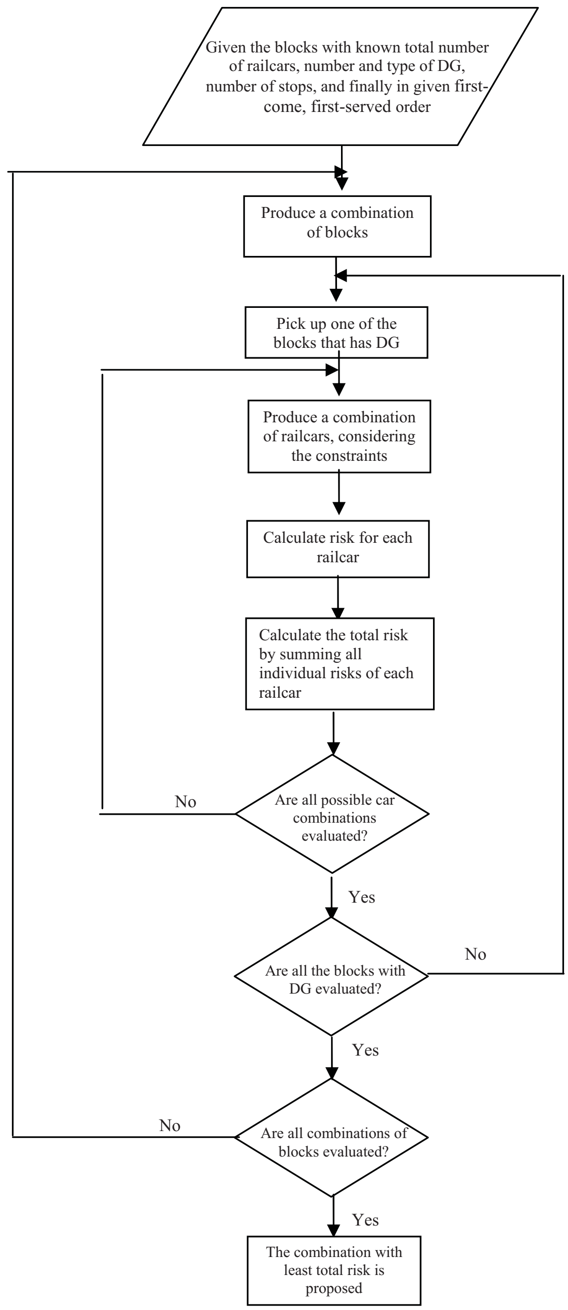

Figure 2.4: MDG2

2.3 Data Sources

The data sources for this study can be classified into two groups: the data sources required to update the previous derailment studies and the data sources needed for establishing different application scenarios and marshalling strategies. The latter required interviews and site visits to rail yards. For updating the previous models, three basic sources of data are used:

1. The Federal Railroad Administration (FRA) train accident database, published annually in the United States.
2. The Canadian Transportation Safety Board (TSB) train accident database (RODS)
3. Reports on selected derailments and collisions produced by the National Transportation Safety Board (NTSB) in the US and by the Transportation Safety Board (TSB) in Canada.

The FRA provides information on the type, causes, consequences, and mitigating circumstances of train accidents experienced annually nation-wide in the US for the period 1975-2007. These data are readily available for downloading from the FRA, Office of Safety Analysis Web Site. One problem with the FRA train derailment database is that it does not include information on all cars involved in a derailment. As a result the payload of each car in a train consist can not be determined. For Canadian derailment data, the RODS database is available for the period 1983-2005. The problem with RODS tables is that most of the cells are empty. Thus obtaining more information requires reviews of a number of NTSB and TSB accident investigation reports for the period 1993-2005. The NTSB database contained 101 investigative reports, compared to 163 reports in the TSB database.

Interviews and observations are two main ways to evaluate current operations and establish application scenarios and marshalling strategies. Various site visits were conducted to gain a general understanding of rail yard operations and included Taschereau Yard (Montreal Intermodal Terminal) in June 2006, Agincourt Yard (Canadian Pacific Railway) in Toronto in July 2007 and April 2009, Liucun South Yard (Datong-Qinhuangdao Railway in China) in June 2009, Alyth Yard (Canadian Pacific Railway) in Calgary in August 2009, and finally, the Network Management Center (NMC) of the Canadian Pacific Railway in August 2009.

In addition, interviews (by telephone, personal, email) were conducted to obtain railroad management and crew feedback. Finally, two classes were attended in the Railway Conductor Program (T151) at George Brown College in Toronto (the"Switching and Marshalling in Rail Yards" course in February 2009 and "Transportation of Dangerous Goods" in March 2009).

Chapter 3

In-Transit Risk Model

This chapter presents the In-Transit Risk Model as the first main component of the proposed model (Figure 2.1). The following four sequential components are discussed: causes of train derailment, probability of train derailment, point of derailment, and number of cars derailing.

3.1 In-transit Risk Components

There are different ways to express risk, especially as it relates to DG in transport. Some authors define risk as a function of probability and consequence of the undesirable event [39]. In addition risk has been expressed in a more comprehensive as an F-N representation of risk. Providing a more complete definition of risk overall likely covers from low frequency high consequence to high frequency low consequence. Recently researchers have explained F-N representation of risk in terms of confidence interval to refine uncertainty [54]. More common definition of risk is the product of both the probability and the consequence of the event [27, 33]. In this research, risk is defined as the probability of derailment multiplied by the consequence.

In-transit risk in this study refers to a potential derailment of cars carrying some type of DG along a given route or route segment. The in-transit risk model provides an estimate of the risk of derailment by position along a route segment for a given mix of DG and non-DG cars making up a train consist. Within the scope of this study, the number of cars derailing reflects the basic unit of “in-transit risk”; therefore, the actual severity of a car derailing is not explicitly considered.

Figure 3.1 illustrates the components involved in the in-transit risk model. To develop derailment probability profiles for trains along specific routes, four basic steps must be taken:

1. analysis of train derailments;
2. assessment of the potential causes of derailment;
3. analysis of the Point of Derailment (POD) along the train where derailments are likely to begin;
4. estimation of the number of cars derailing for each POD.

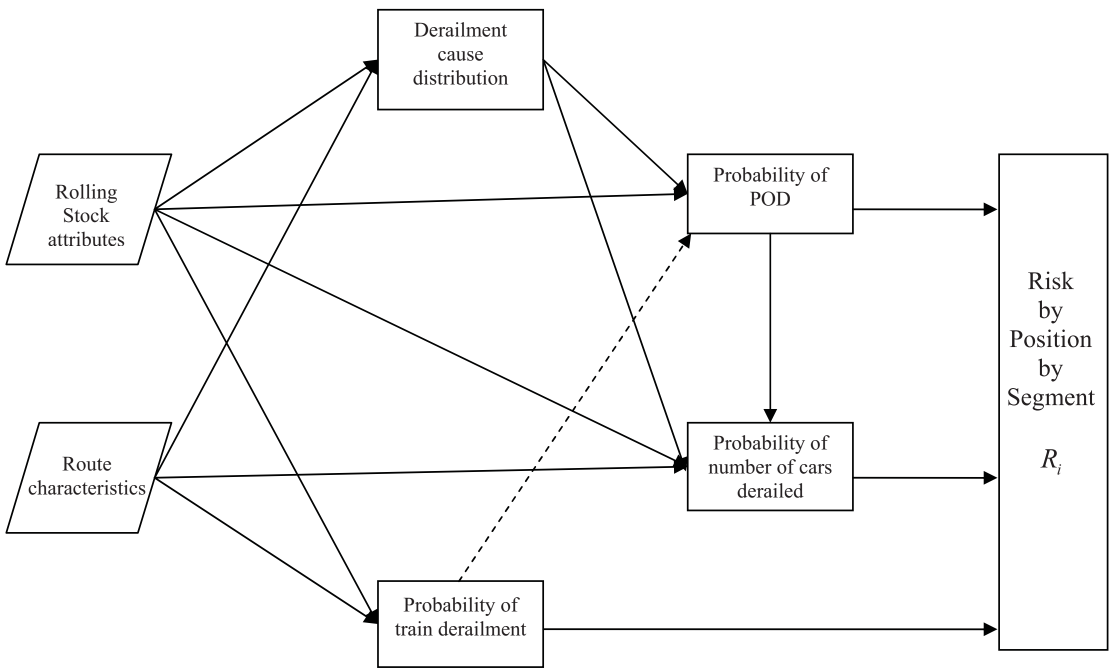

Figure 3.1: In-transit risk model components

The in-transit risk model inputs are

1. rolling stock characteristics, which include the number of cars in each block, the total number of cars (the train length), the number of DG cars in each block, and the types of DG in each block; and
2. route attributes, including the minimum radius of curvature, the maximum longitudinal gradient, and the maximum supperelevation.

The in-transit risk model output is the derailment risk for each DG car. Ideally, after running the model and obtaining the optimum order of cars and blocks, the yardmaster sends the output to a yard engine to order the cars optimally.

A comprehensive understanding of the above four elementary steps will facilitate the presentation of the model framework in this research. As a result, each of these steps is discussed after in-transit risk formulation.

3.2 In-transit risk formulation

In-transit risk requires the estimation of two constituent components, the point of derailment (POD) and the number of cars derailing. The aim of the DG placement model is to position DG cars along the train so that the route segment risk of derailment is minimized. For a given train and route segment, this risk is summed over all positions; that is, $\sum_{i=1}^{n} R_i$, where R_i is the risk of derailment at position i. R_i is defined as the probability of derailment at position i, P_i, multiplied by the consequence of derailment at position i, C_i, that is,

$$R_i = P_i \times C_i. \tag{3.1}$$

The probability of derailment at position i (P_i) can be defined by

$$P_i = P_r(TD) \times P_r(i \,|\, TD), \tag{3.2}$$

where $P_r(TD)$ is the probability of train derailment on a given route, and $P_r(i \,|\, TD)$ is the conditional probability that a car in the i^{th} position derails for a given derailment.

As noted before,

$$C_i = \begin{cases} 1 & \text{if a DG car occupies position } i \\ 0 & \text{otherwise.} \end{cases}$$

The conditional probability of a car in position i derailing given that the train has derailed can be determined by considering all possible points of derailment (POD) as follows:

$$\begin{aligned} P_r(i \,|\, TD) = \sum_{j=1}^{i} \Bigg[& P_r(\text{POD at position } j) \\ & \times \sum_{k=i-j+1}^{n-j+1} P_r(k \text{ cars derailing } \,|\, \text{POD at position } j) \Bigg]. \end{aligned} \tag{3.3}$$

P_r(POD at position j) is the probability that derailment initiated at position j along a given route segment. This probability depends on the cause of derailment and can be determined accordingly. Let $g \in G$ represents a derailment cause such as track geometry; then,

$$P_r(\text{POD at position } j) = \sum_{g \in G} \left[P_r(\text{POD at position } j \,|\, \text{Cause } g) \times P_r(\text{Cause } g) \right], \quad (3.4)$$

where $P_r(\text{Cause } g)$ is the probability of the derailment cause for the given route attributes and rolling stock characteristics, and $P_r(\text{POD at position } j \,|\, \text{Cause } g)$ is the probability of POD at position j when the derailment cause is given.

Finally, the risk associated with each position i (as in Equation (3.1)) is simply the product of Equation (3.2) and the placement of DG cars in this position, i.e.,

$$R_i = P_r(TD) \times P_r(i \,|\, TD) \times C_i \,. \quad (3.5)$$

For a given route segment and n car train, the risk associated with all positions of the train is estimated by summing Equation (3.5) for $i = 1, \ldots, n$. As noted, only positions occupied by DG cars contribute to this risk.

In this research, the total train segment risk serves as the objective function minimized with respect to each DG placement strategy considered. More details about the total train risk along a corridor are discussed in Appendix A.

3.3 Analysis of Train Derailments

Anderson and Barkan [11] show that the probability of a freight train derailing is a function of exposure (distance traveled), train length, and track class (which indicates track quality). Their resultant expression is given in Equation (3.6).

$$P_r(TD) = 1 - e^{-\text{distance} \times (RC \times (\text{train length}) + RT)}, \quad (3.6)$$

where, $P_r(TD)$ is derailment probability; RC is derailment rate per billion freight car-miles; and RT is derailment rate per million freight train-miles.

Their model is based on aggregate data for accident rates for several track classes in terms of the number of derailments per billion freight car-miles and number of derailments per million freight train-miles. The probability of a train derailment is assumed to be a function of the track and train operating characteristics along a given route segment.

The FRA classifies track into nine categories for passenger and freight trains based on various quality and speed considerations (Table 3.1). Table 3.2 summarizes several train derailment rates for different track classes. These rates are given in both train and car-mile measures of exposure. This model (Equation (3.6)) is used as the probability of a train derailment in this research.

Table 3.1: FRA Track classification based on maximum speed (Source:[5])

Track Type	Freight Train	Passenger Train
Class 1	10 mph (16 km/h)	15 mph (24 km/h)
Class 2	25 mph (40 km/h)	30 mph (48 km/h)
Class 3	40 mph (64 km/h)	60 mph (97 km/h)
Class 4	60 mph (97 km/h)	80 mph (130 km/h)
Class 5	80 mph (130 km/h)	90 mph (140 km/h)
Class 6	110 mph (180 km/h)	
Class 7	125 mph (201 km/h)	
Class 8	160 mph (260 km/h)	
Class 9	200 mph (320 km/h)	

Table 3.2: Estimated accident rates by FRA track class: 1992 - 2001(Source:[11])

FRA track class	X & 1	2	3	4	5 & 6
Derailments per million freight train miles	48.540	6.060	2.040	0.530	0.320
Derailments per billion freight car miles	720.100	92.700	31.500	7.800	4.900

3.4 Derailment Cause Analysis and Modeling

The first step in estimating P_r(POD at position j) in Equation (3.4) is obtaining the probability of derailment causes along each route segment (P_r(Cause g)). The assumption is that derailment causes are a function of route attributes and rolling stock characteristics. In this section, this assumption is validated through statistical analysis. A detailed discussion of derailment cause distribution is provided (Figure 3.1).

This section has two specific objectives:

1. to evaluate the correlation between train derailment causes and route attributes and rolling stock characteristics through log-linear analysis

2. to estimate the causes of train derailment for a given train and route

3.4.1 Previous Studies

Various studies have been conducted on train derailment causes. Arthur D. Little [7] showed that train accidents were attributable to either car miles- or train miles-related causes. Train miles-related causes are independent of train length, which essentially constitutes the number of train movements over a given track segment. For instance, "human error" is a train miles-related cause. On the other hand, causes such as "track component failures" and "equipment failures" are car miles- related causes. A similar study by Schafer and Barkan [55] evaluated this grouping of accident causes through regression analysis of US accidents for the period 1990 to 2005.

Another study by TranSys Research Ltd [36] analyzed Canadian derailment data to explore equipment and track-related causes. Barkan et al., [19] studied the US FRA mainline freight derailment records to investigate the link between various causes and release of dangerous goods from tank cars. This study carries out a similar investigation of FRA mainline fright derailment, for a different period (1997- 2006), which is presented in Table 3.3. This table arranges different causes according to their frequency.

Nicolet and Gheorghe[42] showed that a railway car can be involved in a derailment either by initiating a train derailment or by derailment due to the derailment of the cars in front of it. Thus, in estimating the probability of derailment by position along a train, two factors must be considered: (i) the point at which the train derailment begins (POD) and (ii) the number of cars derailing after the point of derailment.

Previous studies [30, 51] showed that the point at which a train derailment begins is affected by the cause of derailment. In addition, the numbers of cars involved in derailments have been found to depend on the cause of derailment and the operating speed [57]. Hence, to better understand the relationship between the POD and number of cars derailing, the distribution of causes in derailment data needs to be established.

To estimate the distribution of causes, firstly, this assumption that the derailment causes is a function of route attributes and rolling stock characteristics has been tested through log-linear analysis. Also, this log-linear model has been used to estimate the distribution of derailment causes ($P_r(\text{Cause } g)$).

According to an extensive literature review, previous studies generally neglect to consider derailment causes according to rolling stock characteristics and route attributes.

In the next section, statistical analysis is conducted to estimate probabilities of derailment causes given route attributes and rolling stock characteristics. The loglinear method is used to evaluate dependency structure between the response variable "cause" and predictors. The loglinear model is further be used to estimate conditional probability distribution of causes give other factors.

3.4.2 Derailment Data Sources

The analyses in this section use a database of freight train derailments that happened in the US from 1997-2006. The US Federal Railroad Administration (FRA) train accident database is assembled annually and provides information on the type of track, total damage, and mitigating circumstances of individual train accidents experienced nation-wide. The data are readily available for downloading from the FRA, Office of Safety Analysis Web Site [9].

For the 10-year period under discussion (1997-2006), there are 38,393 records, but this research is concerned with freight train derailments on mainlines. Moreover, another limitation had to be added to omit some records where the numbers of locomotives are more than three, so the number of records is 4,150.

The 389 distinctive causes listed in the FRA database are categorized into five major cause groups, namely track, roadbed and structure (T); signal and communication (S); mechanical and electrical failures (E); train operation-human factors (H); and miscellaneous (M).

Each cause is represented as a code of length four, where the first space in the coding scheme refers to one of the five aforementioned major cause groups. For instance, the "Broken Rail" defect represented by T201 belongs to causes group T (track or roadbed and structure), and 201 is a numeric label given by the FRA to distinguish this defect from other track-related causes.

Contrary to the FRA, the Association of American Railroads (AAR) has used a different cause grouping, which is defined by merging several FRA causes into one single cause, and resulting in only 51 distinctive accident causes instead of the FRA's 389 (Appendix B). The AAR used codes of length three to distinguish different causes. The two first spaces are numeric and the last one is one of the letters T, S, E, H, M introduced above. Following [19, 51] the current study has used the AAR cause groupings since it has better resolution for certain causes. Table 3.3 illustrates the fifteen most frequent accident cause groups and the four broad categories that combine together the remaining cause groups (Track-TO, Human Factor-HO, Equipment-EO, and Signal-01S).

3.4.3 Cause Grouping by Probability Distribution Fitting

Previous studies [53, 17, 16] showed that the point at which a train derailment begins (POD) is affected by the cause of derailment. These studies took train length into account for estimating the POD by normalizing the length in the range of 1 to 100 cars to obtain the normalized points of derailment (NPOD) for different derailment causes. For instance, for road bed defects, derailments occurred more likely in the front sections of the train than in other sections. The main focus of this section is probability distribution fitting for NPOD data.

In what follows, three steps have been used to model a dataset with a specific probability distribution: (i) determining the "best-fitting" distribution and (ii) estimating the parameters for that distribution, and (iii) testing the validity of the selected distribution. To find the best fitting distribution for all 19 causes, a variety of probability distributions are fitted using the @Risk software package. P-values for the Chi-square tests are calculated, and at level 0.05, the best fits for each accident cause is reported in Table 3.3.

As an example, for the "Broken Rails or Welds" cause (08T), five different types of distributions that have been fitted to NPOD frequencies are beta, triangle, uniform, exponential, and normal. Chi-Square tests were applied to test the validity of the aforementioned distributions, and at level 0.05 only beta distribution was selected.

In a similar fashion, best-fit distributions were obtained for all 19 causes (Table 3.3). The last column of the table shows which part of train is most likely be derailed because of the associated cause. The accident causes with a beta distribution of $\alpha < 1$ (and $\beta > 1$) tend to initiate at the front. Those with $\alpha < 1$ (and $\beta < 1$) tend to initiate at the front or rear of the train. Those accident causes with triangular distribution have a lower limit a, mode c, and upper limit b. For instance, when $b = c$ for "track geometry", the distribution suggests that the rear end of a train is more likely to be derailed than other parts.

Table 3.3: Best fit POD distributions (F=Front, R=Rear end, E=Everywhere, M=Middle)

Cause Group	Description	Frequency	Distribution with Parameters	Train
08T	Broken Rails or Welds	723	Beta (0.659, 1.139)	F
04T	Track Geometry (excl. Wide Gage)	359	Triangle(-0.174, 1.000, 1.000)	R
03T	Wide Gage	277	Beta (0.817, 0.976)	FR
10E	Bearing Failure (Car)	265	Uniform (0.005, 1.004)	E
05T	Buckled Track	180	Triangle(-0.136, 1.000, 1.000)	R
04M	Track-Train Interaction	163	Uniform (0.008, 1.006)	E
12E	Broken Wheels (Car)	145	Uniform (0.003, 1.007)	E
03M	Lading Problems	139	Uniform (0.002, 1.007)	E
09H	Train Handling (excl. Brakes)	134	Triangle(0.007, 0.007, 1.15)	F
10T	Turnout Defects - Switches	133	Triangle(0.008, 0.008, 1.091)	F
13E	Other Wheel Defects (Car)	119	Normal(0.500, 0.262)	M
05M	Other Miscellaneous	112	Beta (0.506, 0.662)	FR
09E	Sidebearing, Suspension Defects (Car)	108	Beta (1.137, 0.950)	R
11H	Use of Switches	97	Beta (0.518, 0.701)	FR
01M	Obstructions	93	Expo(0.280)	F
EO	All other Equipment Causes	505	Beta (0.730, 0.827)	FR
HO	All other Human Factor Causes	228	Beta (0.604, 0.761)	FR
TO	All other Track Causes	354	Beta (0.601, 0.902)	FR
01S	All Signal Failures	15	Beta (0.300, 0.400)	FR

Cause groups in Table 3.3 are further re-grouped into three main cause groups in the last column. Those causes that affect the front and front or rear-end of a train are categorized as group (C1). The second group (C2) includes causes that have an impact on the rear-end of a train, and finally, the last group (C3) is kept for causes that influence the middle of a train.

3.4.4 Factors Affecting Derailment Cause

Ideally, geometric attributes of rail corridors, such as minimum radius (meter), maximum longitudinal gradient, and super elevation (millimeter), should be used as important measures for evaluating the dependency of cause on route attributes. However, because of lack of available information, topographical features of the rail segment (area) have been used instead. The Arc Map Version 9.2 package has been used to separate the rail network in the US into two areas (levels): flat and mountain. The flat area, which covers mostly the central and eastern parts of the country, has an elevation difference of less than 1500 meters, while the mountain area is defined as a region with an elevation difference higher than that of the flat area.

The US Federal Railroad Administration classifies track for freight trains into nine categories, based on various qualities and speed considerations (Table 3.1). However, in this study, track class is categorized as either poor (1-4)(G1) or good (5-9)(G2)to ensure sufficient data for the analysis.

In addition, train length is used as a source of rolling stock attributes (car information). Following a previous study [15], train length has been categorized into three possibilities: short (less than 40 cars), medium, and long (more than 120 cars).

3.4.5 Relating Train Derailment Causes to Route Attributes and Rolling Stock Characteristics

The FRA data (1997-2006) can be summarized in a 4-way contingency table as below (Table 3.4). Log-linear analysis has been conducted to uncover the relationship structure among the factors: Cause groups, Train length, Area, and Track class. For simplicity, these factors are labeled 1, 2, 3, 4, respectively, in the analysis below. Note that the factors Track class and Area are route attributes, while Train length is a rolling stock characteristic. It should be noted that in Table 3.4, a very small number (1.00E-12) is used instead of zero (sampling zero).

The Statistica software package provides as the platform for this analysis. The results of fitting hierarchical loglinear models are reported in Table 3.5. The model showing only the main effects is referred to by K-factor =1, and 2-way, 3-way, and 4-way interactions models are referred to by K-factor=2,3,4 respectively. This table shows that the largest but least-complex model that best fits the data is the one with all 2-way interactions

Table 3.4: Contingency table based on cause, train length, area, and track class

Track Class	Area	Train Length	Cause Group		
			Cause Group 1	Cause Group 2	Cause Group 3
Track Class G1	Flat	Short	325	116	4
		Medium	488	193	23
		Long	58	9	3
	Mountain	Short	7	3	1.00E-12
		Medium	18	9	2
		Long	5	1	2
Track Class G2	Flat	Short	71	32	3
		Medium	402	206	47
		Long	80	28	18
	Mountain	Short	7	2	1.00E-12
		Medium	41	13	7
		Long	5	2	2

(associations). It is worth noting that the null hypothesis in each row of Table 3.5 is that all K-factor interactions are simultaneously zero.

Table 3.5: Results of fitting all K-factor interactions

K-factors	df	LR statistic	P-value
1	6	4686.016	0
2	13	290.502	0
3	12	13.351	0.344
4	4	0.709	0.950

Table 3.6 reports the results of the partial and marginal analyses of individual effects. In the partial association test of a specific effect for K-way association or interaction, the model that includes all K-way interactions is compared with the one without that specific effect. For example, effect 12 in Table 3.6 represents the association between factors 1 (Cause) and 2 (Train Length). When this effect is removed from the model with all 2-way associations, the value of the test statistic, which is the difference between the log-likelihood values, is equal to 30.94, with 4 degrees of freedom. Examination of the corresponding P-value, makes clear that the model fit becomes significantly worse when this two-way interaction is excluded from the model.

In the marginal association test of a specific effect for K-way association, the model without any K-way interactions is compared with the model that includes this specific effect. For example, when effect 12 in Table 3.6 is added to the model showing all main effects (1-way interaction), the value of the test statistic that is the difference between the log-likelihood values is equal to 45.85, with 4 degrees of freedom. The corresponding P-value reveal the model fit becomes significantly better when this two-way interaction is included in the model.

Table 3.6: Test of marginal and partial association

Effect	df	Statistic (partial)	P-value (partial)	Statistic (marginal)	P-value (marginal)
1	2	1449.983	0	1449.983	0
2	2	1090.096	0	1090.096	0
3	1	2106.317	0	2106.317	0
4	1	39.620	0	39.620	0
12	4	30.946	0	45.851	0
13	2	6.284	0.043	10.185	0.006
14	2	21.464	0.00002	37.457	0
23	2	1.607	0.448	7.791	0.021
24	2	172.983	0	191.188	0
34	1	14.301	0.0001	21.502	0.000004
123	4	1.536	0.820	1.254	0.869
124	4	1.982	0.739	1.837	0.765
134	2	3.965	0.138	4.923	0.085
234	2	5.306	0.070	5.402	0.067

Choosing the effects for the model require a review of the "tests of partial and marginal association" in Table 3.6. Several strategies exist for combining the results of these tests. A factor is retained when both partial and marginal tests are in agreement; otherwise, it is removed from the model [26]. Applying this strategy, all two-way interactions except the 23 association should be included in the model.

In summary, it is concluded that the log-linear model that fits the observed table (Table 3.4) is given by

$$\begin{aligned} \log(m_{ijkl}) &= \lambda + \lambda_i^1 + \lambda_j^2 + \lambda_k^3 + \lambda_l^4 + \lambda_{i\,j}^{12} + \lambda_{i\,k}^{13} + \lambda_{i\,l}^{14} + \lambda_{j\,l}^{24} + \lambda_{k\,l}^{34} \\ & i,\ j = 1,2,3 \quad k,\ l = 1,2, \end{aligned} \tag{3.7}$$

where m_{ijkl} is the expected frequency of cell $i\,j\,k\,l$ in Table 3.4. In addition λ_i^1, λ_j^2, λ_k^3, λ_l^4 are main effects at different levels, and $\lambda_{i\,j}^{12}$, $\lambda_{i\,k}^{13}$, $\lambda_{i\,l}^{14}$, $\lambda_{j\,l}^{24}$, $\lambda_{k\,l}^{34}$ are 2-way interaction effects. Note that it is necessary to impose some technical constraints on the effects in order to estimate them. These constraints are not discussed here. The reader can consult with Christensen [26] for more details on log-linear models. The p-value for testing the model given by Equation (3.7) versus the saturated model is 0.61569 with $df = 18$. Therefore, it can be concluded that the specified model is sufficient to explain the frequencies in Table 3.4.

The model given by Equation (3.7) shows that, given factors 1 (cause) and 4 (track class), factors 2 (train length) and 3 (area) are independent. In addition, it can be concluded that the factor "cause" is associated with all three other factors.

Table 3.7 provides the estimates for λs in the log-linear Equation (3.7), which was in the SPSS software package. The last two columns report 95% confidence intervals for the parameters' λs. It can be seen that all main effects (parameters) are significant, except for level 2 of the factor "cause". In addition, all 2-way associations are also significant at the level 5% except λ_{11}^{13} which refers to the combination of the front of a train and a flat area.

Looking at Table 3.7, and the sign of the estimates for λs, the data using Equation (3.7) can be further investigated. Note that the positive and negative signs for the estimates in Table 3.8 refer to positive and negative effects of the corresponding levels of the factors on the probability of derailment in that factor level. Below are some of our findings:

- The model shows that cause group one, which affects the front of trains, is the main reason for more derailments compared to other types of causes. Similar results have been reported in a previous study [51].
- The highest number of derailments occur with trains of 40 -120 cars (medium), while the lowest number occurs with short trains(less than 40) as expected.
- Track quality improvement from poor to high results in corresponding decrease in the number of derailment. However, this issue has been justified in a previous study [19] as follows: In the US, the right to use mainlines with good quality track is mainly reserved for Class I railroads (with a high volume of traffic); Class II and III railroads (with less traffic) are confined to lines of with lower quality track.
- Short trains are more prone to derailment than longer trains as a result of cause groups 1 or 2.
  - Derailments for short trains happen for cause group 1 (affect front) more than for other types of causes, and fewer are expected for longer trains
  - The derailment of short trains happen for cause group 2 (affect rear end) for other causes, and fewer are expected for longer trains
- Trains in flat areas are more prone to be derailed compared to those in mountain areas as a result of cause group 2.

Table 3.7: Estimates of λ's in Equation (3.7)

Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Constatnt	1.026	0.33	3.115	0.002	0.381	1.672
(Cause=1)	0.932	0.358	2.604	0.009	0.231	1.634
(Cause=2)	-0.413	0.406	-1.017	0.309	-1.209	0.383
(Cause=3)	0	.	.	.	.	.
(TL=1)	-2.033	0.443	-4.584	0	-2.902	-1.164
(TL=2)	1.004	0.233	4.307	0	0.547	1.461
(TL=3)	0	.	.	.	.	.
(AREA=1)	1.816	0.299	6.073	0	1.23	2.402
(AREA=2)	0	.	.	.	.	.
(TRK=1)	-2.133	0.304	-7.012	0	-2.729	-1.537
(TRK=2)	0	.	.	.	.	.
(CAUSE=1)-(TL=1)	1.897	0.449	4.228	0	1.018	2.777
(CAUSE=1)-(TL=2)	0.6	0.249	2.409	0.016	0.112	1.088
(CAUSE=1)-(TL=3)	0	.	.	.	.	.
(CAUSE=2)-(TL=1)	2.312	0.473	4.886	0	1.385	3.24
(CAUSE=2)-(TL=2)	0.117	0.258	3.922	0	0.559	1.675
(CAUSE=2)-(TL=3)	0	.	.	.	.	.
(CAUSE=3)-(TL=1)	0	.	.	.	.	.
(CAUSE=3)-(TL=2)	0	.	.	.	.	.
(CAUSE=3)-(TL=3)	0	.	.	.	.	.
(CAUSE=1)-(AREA=1)	0.625	0.321	1.946	0.052	-0.005	1.255
(CAUSE=1)-(AREA=2)	0	.	.	.	.	.
(CAUSE=2)-(AREA=1)	0.778	0.353	2.204	0.028	0.086	1.469
(CAUSE=2)-(AREA=2)	0	.	.	.	.	.
(CAUSE=3)-(AREA=1)	0	.	.	.	.	.
(CAUSE=3)-(AREA=2)	0	.	.	.	.	.
(CAUSE=1)-(TRK=1)	0.905	0.219	4.126	0	0.475	1.334
(CAUSE=1)-(TRK=2)	0	0	.	.	.	.
(CAUSE=2)-(TRK=1)	0.65	0.229	2.841	0.004	0.202	1.099
(CAUSE=2)-(TRK=2)	0	.	.	.	.	.
(CAUSE=3)-(TRK=1)	0	.	.	.	.	.
(CAUSE=3)-(TRK=2)	0	.	.	.	.	.
(TL=1)-(TRK=1)	1.891	0.179	10.549	0	1.54	2.243
(TL=1)-(TRK=2)	0	.	.	.	.	.
(TL=2)-(TRK=1)	0.576	0.154	3.731	0	0.274	0.879
(TL=2)-(TRK=2)	0	.	.	.	.	.
(TL=3)-(TRK=1)	0	.	.	.	.	.
(TL=3)-(TRK=2)	0	.	.	.	.	.
(AREA=1)-(TRK=1)	0.843	0.194	4.347	0	0.463	1.224
(AREA=1)-(TRK=2)	0	.	.	.	.	.
(AREA=2)-(TRK=1)	0	.	.	.	.	.
(AREA=2)-(TRK=2)	0	.	.	.	.	.

- Trains on poor quality track are derailed more than ones on good quality track for either cause group 1 or 2.
- Short trains on poor track are derailed more than short trains on good track. This finding is true for medium trains too
- As the quality of track in flat areas improves, the number of derailments decreases.

Table 3.8 is the conditional probability distribution of cause given different combinations of "train length", "area", and "track classes". These probabilities have been estimated using the model given by Equation (3.7), and the estimates of λs reported in Table 3.8. To be precise, the joint probability distribution of all four factors is

$$p_{ijkl} = \frac{m_{ijkl}}{2230} = \frac{e^{\lambda+\lambda_i^1+\lambda_j^2+\lambda_k^3+\lambda_l^4+\lambda_{i\,j}^{12}+\lambda_{i\,k}^{13}+\lambda_{i\,l}^{14}+\lambda_{j\,l}^{24}+\lambda_{k\,l}^{34}}}{2230}.$$

The joint probability distribution of the factors "train length", "area", and "track class" is now $p_{.\,jkl} = p_{1jkl} + p_{2jkl} + p_{3jkl}$, and, thus, the conditional probability of cause being at level i given that the other factors are at levels train length $= j$, area $= k$, track class $= l$ is

$$p_{i\,|\,jkl} = \frac{p_{ijkl}}{p_{.\,jkl}}.$$

Table 3.8: Estimated conditional probability distributions of derailment causes given the factors "train length", "area", "track class"

(train length,area,track class)	C1	C2	C3
(1,1,1)	0.730	0.260	0.009
(1,1,2)	0.670	0.310	0.021
(2,1,1)	0.694	0.274	0.032
(2,1,2)	0.615	0.313	0.071
(3,1,1)	0.757	0.178	0.064
(3,1,2)	0.660	0.200	0.139
(1,2,1)	0.752	0.229	0.018
(1,2,2)	0.688	0.271	0.041
(2,2,1)	0.701	0.237	0.061
(2,2,2)	0.605	0.264	0.131
(3,2,1)	0.735	0.148	0.117
(3,2,2)	0.604	0.157	0.238

3.5 Point of Derailment Analyses and Modeling

A rail car can be involved in a derailment either by initiating a derailment or by being derailed itself due to the derailment of the cars in front of it [42].Thus, in estimating the probability of derailment by position along a train, two factors need to be considered: the point at which the train derailment begins (POD) and the number of cars derailing after the point of derailment.

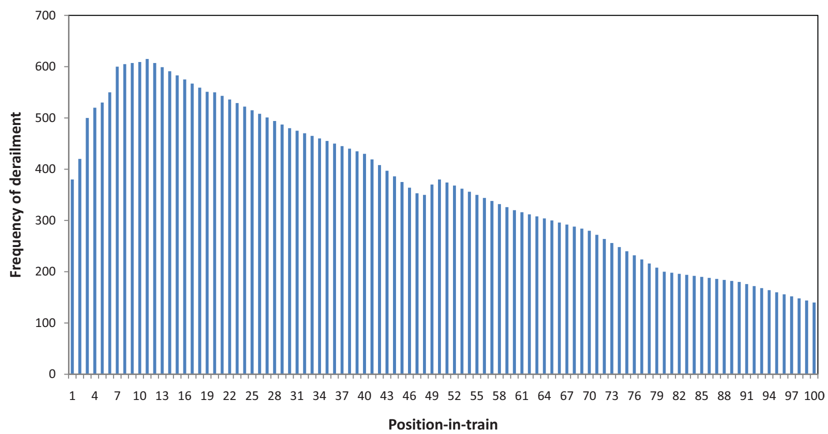

Figure 3.2: Frequency of derailment by position (Source:[12])

Figure 3.2 illustrates the frequency of derailment by position for a standard 100-car train, based on historical derailment data from the US for the period 1992-2001. This figure clearly shows a strong relationship between position and derailment. For instance, under the same conditions and for the same train, a car in position 11 is more likely to derail than a car in position 80.

This section provides a detailed discussion of the point of derailment (POD), and the number of cars derailing is discussed in the next section.

Previous work by El-Hage [30] suggests a link between the causes of derailment and the position of the train where the derailment begins (POD). He took train length into account for estimating the POD by normalizing the length in the range 1 to 100 cars; then he generated probability tables and histograms of the normalized points of derailment for the different derailment causes. For instance, derailments caused by road bed defects occurred in the front sections of trains rather than in other sections. Anderson and Barkan [12] modeled normalized POD probabilities by regression of data using beta distribution with parameters α and β. They considered all accident causes simultaneously. Figure 3.3 clearly shows that derailments are more likely to initiate in front sections of trains.

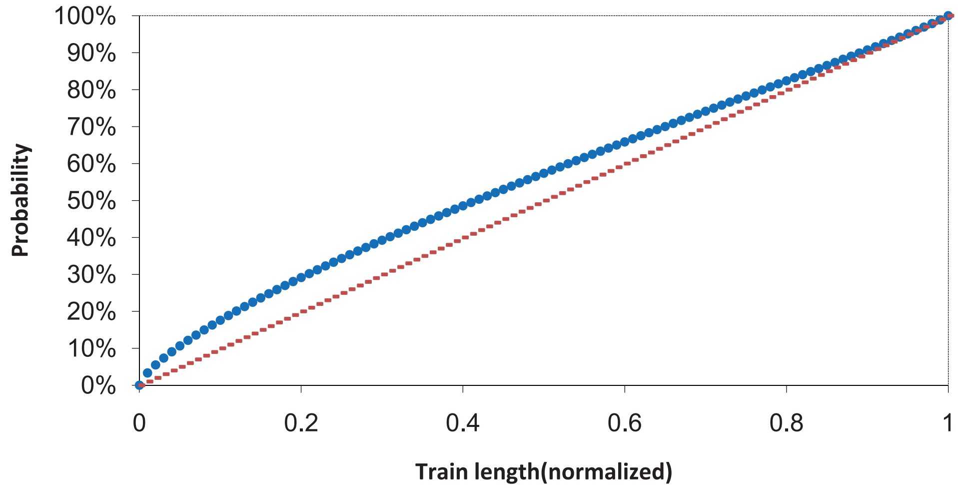

Figure 3.3: Likelihood of points of derailment (Source: [12])

The link between train derailment and POD requires further specification of the primary cause of each train derailment. The underlying assumption is that different causes are likely to lead to different POD profiles. For example, it is expected that the front of a train would be more likely to derail if a track-related problem causes the derailment, since it is the front of a train that first encounters any track faults as the train traverses a segment. Rolling stock defects, on the other hand, can take place at different points. A previous study [53] used ANOVA to confirm the relationship between cause and POD empirically by analyzing derailment data from Canada (years 1983-1985). Another study [51] presented a methodology to estimate the probability of POD for different causes and train lengths. The first step in estimating POD for each track segment is obtaining the probability of different derailment causes along each route segment, as has been discussed in the previous section. In addition, the previous section shows that the probability of derailment causes is a function of route attributes and rolling stock characteristics.

3.5.1 Factors Affecting POD

This section explores several factors that explain POD, based on historical train derailments reported by the FRA for the period 1997-2006. As illustrated in Figure 3.4, out of 4148 derailments in this database (after data cleaning), over 18% were found to take place at the front of trains, and only 3% were found to take place beyond the 100^{th} car position.

One of the major problems with using these observations to establish POD rates is that they do not take train-length distribution into account. The distribution of train lengths

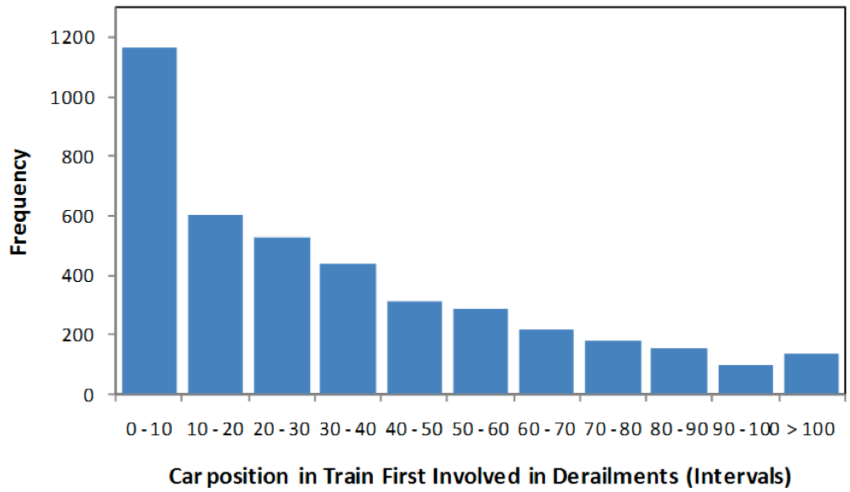

Figure 3.4: POD distribution from FRA database 1997-2006 (Source:[9])

in the FRA database is shown in Figure 3.5, and indicates that over 30% of trains are in the 0-50 car-length range. Hence, there appears to be an over-representation of front-end positions in the distribution, which leads to an over-representation of front-of-train POD in the database.

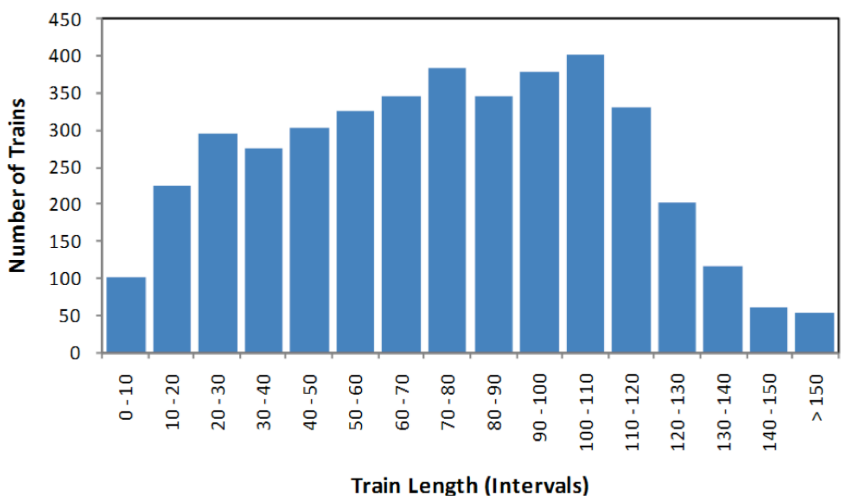

Figure 3.5: Train length distribution from FRA database 1997-2006(Source:[9])

To adjust for train length, a number of studies [30, 51] have expressed POD in a normalized form (NPOD) that reflects a standard 100-car train. While the NPOD accounts for absolute train length, it fails to reflect dynamic forces acting on the train with respect to POD that cause car-track instability leading to derailments. This failure requires specification of actual positions along a train in lieu of standardized measures. For example the 50$^{\text{th}}$ percentile position in a 10-car train is subjected to different dynamic forces along a sharp curve than those for the 50$^{\text{th}}$ percentile position in a 100-car train.

To account for limitations in NPOD, FRA train derailments in this study have been classified into short (< 40 cars), medium (40 to 120 cars), and long (> 120 cars) trains, and analyzed separately with respect to their NPOD. The nonparametric Kruskal-Wallis test, a one-way analysis of variance by ranks, was applied to determine whether train length (short, medium, and long) provides a statistically significant explanation for median POD. The results of this analysis are summarized in Table 3.9.

Table 3.9: Kruskal-Wallis test for comparing NPOD in 3 different train length groups

Train length	Number of records	Median
Short	2537	0.326
Medium	1090	0.574
Long	2160	0.473

Using the Minitab 14.0 software package, the Kruskal-Wallis test examines the validity of the null hypothesis that the median NPOD does not differ according to train length, against its alternative assertion that median NPOD differs for at least two out of the three train length groups. The sample NPOD medians for the three groups were estimated to be 0.33 for short trains, 0.57 for medium trains, and 0.47 for long trains. The Kruskal-Wallis test yielded a test statistic value of 239.5 well above its critical value of $\chi^2_{0.05,2} = 5.99$. Hence, there is sufficient evidence to reject the null hypothesis H_o that there is no difference in NPOD medians for different groupins of train length. It can also be stated that the absolute train length has a significant effect on POD. The next factor possibly affecting POD is derailment cause. Using Canadian derailment data for the period 1983-1985, [30] established a statistical expression linking POD to five primary causes: roadbed defect, track geometry, railbar, switches, and general car. Using more recent US-FRA derailment data, [51] also found that derailments are more likely to begin at the front of a train when the primary cause is track-related, and the middle of the train when the cause is rolling stock-related.

Figure 3.6 illustrates the distribution of NPOD for the FRA database for the period 1997- 2006. These distributions are controlled for three underlying and mutually-exclusive causes of derailment: rail, joint bar, and rail anchoring defects; track geometry and brake failures. In this figure, train position was aggregated into quartiles to ensure that sufficient observations were obtained from the data for each train length category. For "rail, joint bar, and rail anchoring", over 35% of all derailments were found to take place at the front of the train (first quarter), while for derailments caused by "brake failures", the highest frequency NPOD intervals were observed in the middle. For "track geometry," the highest frequency intervals occurred in the rear section of the train.

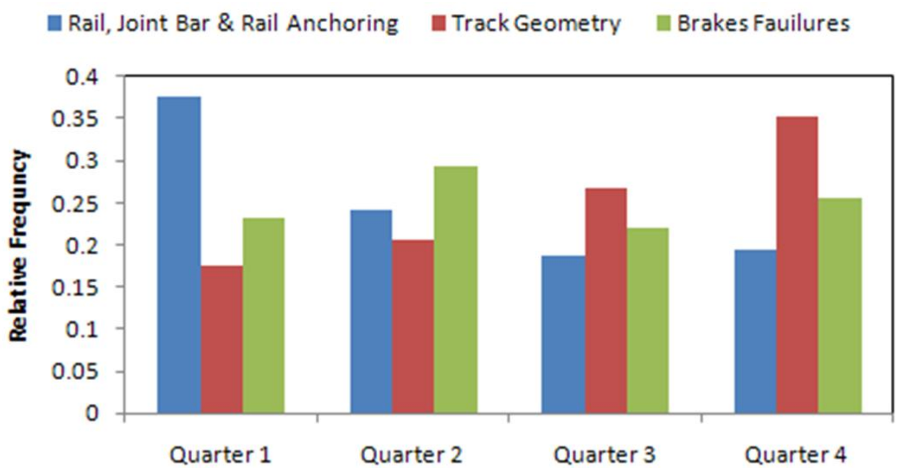

Figure 3.6: Comparison of point of derailment probability for different causes

Table 3.10 summarizes the results of Kruskal-Wallis testing comparing median NPOD to cause of derailment for the FRA (1997-2006) data. The sample medians for the three cause groups in Table 3.10 were estimated to be 0.40, 0.64, and 0.50, respectively. The value of the test statistic, 130.27, exceeded its critical value of $\chi^2_{0.05,\,2} = 5.99$ at the 5% level of significance. Hence, there is sufficient evidence to reject the null hypothesis H_o (no cause effect) and conclude that the causes have a statistically significant effect on POD when trains are adjusted for length. The above test provides some statistical evidence that the probability of POD along a given route segment depends on train length and primary cause of derailment.

3.5.2 Method for Estimating POD

The primary aim of this section is to develop expressions for estimating POD probability for different derailment causes and train lengths. A number of distributions were explored to explain NPOD by train length and cause (C). The best-fit distribution for each of the nine groupings of train length (short, medium, and long) and derailment cause (C1: track,

Table 3.10: Kruskal-Wallis test for comparing NPOD in 3 different cause groups

Cause	Number of records	Median
Track related (excluding track geometry)	1573	0.397
Track geometry	831	0.639
Car related	1799	0.504

roadbed, and structure, excluding track geometry, C2: track-geometry related and C3: causes related to each car (such as mechanical and electrical causes) was established for the FRA data using BestFit software developed by Palisade [1]. FRA derailment data by position was initially classified into three train length groups (short, medium and long) and subsequently into three causes. An example of the distribution for track-related causes for medium length trains is illustrated in Figure 3.7. Four different types of distribution (beta, triangle, uniform, and exponential) have been fitted to POD frequency for medium train length and track-related causes.

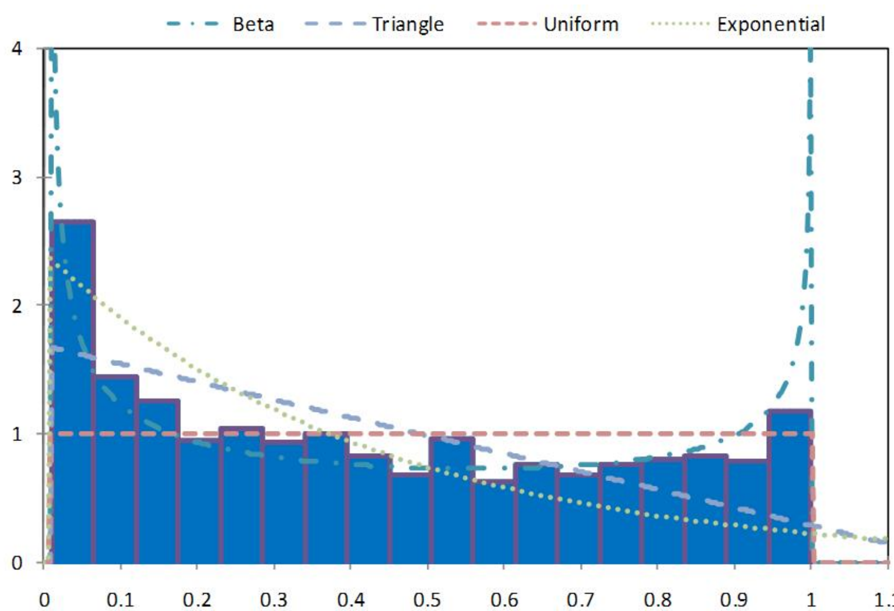

Figure 3.7: Four distributions for track related derailments involving medium trains

Comparison of observed POD frequency from the FRA data with expected values from the underlying distribution provided χ^2 statistic. Table 3.11 summarizes these Chi-Square values of the four distributions considered. For this combination of derailment cause and train length, the Beta distribution yielded the lowest Chi-Square value and, hence, the best-fit result.

In a similar fashion, best-fit distributions were obtained for all nine train length/cause combinations or groupings (Table 3.12).

Table 3.11: Values of the χ^2 goodness of fit statistic of distributions for track-related derailments involving medium length trains

Distribution	χ^2 statistic
Beta	75.150
Triangle	154.600
Uniform	243.400
Exponential	334.700

Table 3.12: Best fit POD distributions (U=Uniform, T=Triangle, Beta=General Beta) for all derailment causes and train lengths

	Short train	Medium train	Long train
Cause Group 1	U(0.03,1)	Beta(0.575,0.6579)	Beta(0.602,0.745)
Cause Group 2	T(-0.094,1, 1)	Beta(0.782,0.504)	Beta(0.646,0.59)
Cause Group 3	U(0.031,1.005)	U(0.008,1)	Beta(0.763,0.799)

To illustrate how these results can be used, we consider the probability of derailment for the 10^{th} position of a 39-car train (at the top of the short train class). The best-fit distribution for track-related causes is $U(0.03, 1)$, and this finding yields a probability of 0.10 for the 10^{th} position for a short-train membership. On the other hand, for the same 10^{th} position and cause on a 41 car train, a $Beta(0.58, 0.66)$ distribution is used for a medium-train length membership. This distribution yields a probability of 0.09 for the same 10^{th}-car position. The difference between these estimates can be explained by the uncertainty associated with assigned train length membership of between 39 and 41 cars (i.e., short versus medium).

To incorporate this uncertainty, it is considered a membership expression [58] of the form

$$P_r(\text{POD at position } j) = \frac{f_1(j) \cdot m_1(e) + f_2(j) \cdot m_2(e)}{m_1(e) + m_2(e)}\,, \tag{3.8}$$

where $f_1(j)$ and $f_2(j)$ are estimated from probabilities of NPOD from Table 3.12 for a given position j, train length, and cause group. The membership values m_1 and m_2 are obtained for a given train length e, from Figure 3.8.

As shown in Figure 3.8, assuming a 41-car train over a segment of track subject to cause group 2 (track-geometry related), it is obtained a membership factor of $m_1 = 0.68$ for a short-train classification and $m_2 = 0.32$ for a medium-train classification. Since a 41-car train is considerably shorter than the 80-car minimum for long train classification, the membership function associated with long trains is assumed to be zero.

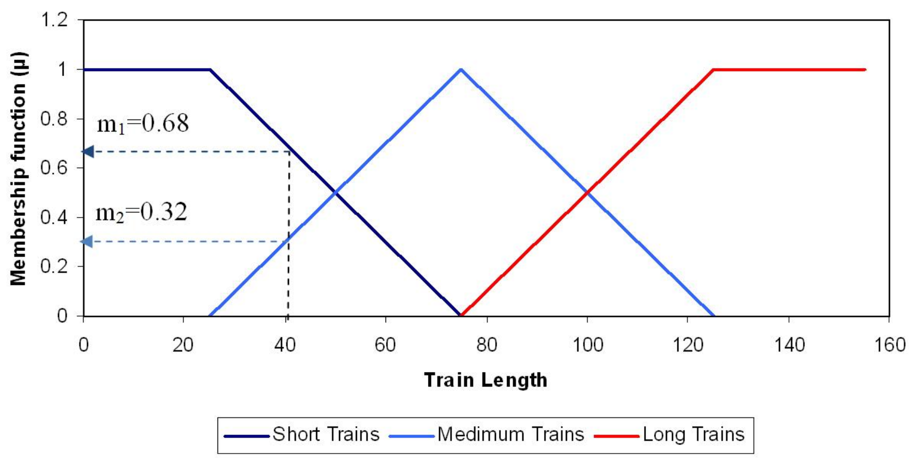

Figure 3.8: Membership function for different classes of train length

The probability of derailment for the 10$^{\text{th}}$ position (in a train with 41 cars) with a cause-2 derailment (track- geometry related) is obtained by applying the following four steps:

i. For the 41-car train, Figure 3.8 suggests the membership factors $m_1 = 0.68$ and $m_2 = 0.32$ for short and medium trains, respectively.

ii. The ratio of NPOD is calculated as $10/41 = 0.244$.

iii. From Table 3.12, the corresponding derailment probabilities are obtianed: $f_1 = 0.033$ and $f_2 = 0.07$.

   Assuming a short train membership, f_1 is established from the Triangle distribution from Table 3.12, $T(-0.094, 1, 1)$. The resultant derailment probability for position 10 on this 41 car train is estimated to be 0.0133.

Assuming a medium train membership, f_2 is established from the Beta distribution, $Beta(0.782, 0.504)$, the derailment probability for position 10 becomes 0.0165, a difference of about 20%.

iv. The probability of initiating a derailment per position is obtained by combining the fuzzy function memberships with the underlying NPOD distribution values as per the above membership expression, such that

$$P_r(\text{POD at position 10}) = \frac{f_1(10) \cdot m_1(e) + f_2(10) \cdot m_2(e)}{m_1(e) + m_2(e)} = 0.014.$$

In a similar fashion, the probabilities for each position on the 41-car train can be obtained, with the results given in Table 3.13 for track geometry-related derailments. A similar set of values can be obtained for the other cause groups for the 41-car train.

As discussed previously (Equation (3.4)), the probability of POD along a given route segment also depends on the cause of derailment. Thus, the probability of point of derailment can be expressed as

$$P_r(\text{POD at position } j) = \sum_{g=1}^{3} \left[P_r(\text{POD at position } j \mid \text{Cause } g) \times P_r(\text{Cause } g) \right].$$

using Causes 1, 2, and 3 introduced in Section 3.4. Note that $P_r(\text{Cause } g)$ is obtained from Table 3.8. For the above example, to calculate the probability of derailing for the 10th position (in a train with 41 cars) with a cause-2 derailment (track geometry related), it is assumed that $P_r(\text{Cause 2}) = 1$ while $P_r(\text{Cause 1}) = 0$ and $P_r(\text{Cause 3}) = 0$. However, for the case study (Chapter 5), this assumption is not be followed.

3.6 Analysis of Number of Cars Derailed

A.D. Little Inc. [41] has suggested a non-linear relationship between the mean number of cars derailing and train speed (in mph), such that

$$M_{nd} = 1.7\,(\text{Speed})^{0.5}\,,. \tag{3.9}$$

where M_{nd} is the mean number of cars derailing and the train speed is in miles per hour.

A number of studies [52, 12] have suggested that the number of cars derailing is also affected by residual train length and cause of derailment. The longer the residual length of the train (post POD), the higher the potential number of cars derailing. The link between cars derailing and train length is affected by the cause of derailment, such that,

Table 3.13: POD probability for a 41 car train (m_1=0.68, m_2=0.32)

POS j	NPOD	$f_1(j)$	$f_2(j)$	P_r(POD at position j)
1	0.0244	0.0043	0.0304	0.0127
2	0.0488	0.0053	0.0222	0.0107
3	0.0732	0.0063	0.0200	0.0107
4	0.0976	0.0073	0.0188	0.0110
5	0.1220	0.0083	0.0181	0.0114
6	0.1463	0.0093	0.0175	0.0119
7	0.1707	0.0103	0.0171	0.0125
8	0.1951	0.0113	0.0169	0.0131
9	0.2195	0.0123	0.0167	0.0137
10	0.2439	0.0133	0.0165	0.0143
11	0.2683	0.0143	0.0164	0.0150
12	0.2927	0.0153	0.0164	0.0156
13	0.3171	0.0163	0.0163	0.0163
14	0.3415	0.0173	0.0164	0.0170
15	0.3659	0.0182	0.0164	0.0177
16	0.3902	0.0192	0.0165	0.0184
17	0.4146	0.0202	0.0166	0.0191
18	0.4390	0.0212	0.0167	0.0198
19	0.4634	0.0222	0.0169	0.0205
20	0.4878	0.0232	0.0171	0.0212
21	0.5122	0.0242	0.0173	0.0220
22	0.5366	0.0252	0.0175	0.0227
23	0.5610	0.0262	0.0178	0.0235
24	0.5854	0.0272	0.0181	0.0243
25	0.6098	0.0282	0.0185	0.0251
26	0.6341	0.0292	0.0189	0.0259
27	0.6585	0.0302	0.0194	0.0267
28	0.6829	0.0312	0.0199	0.0276
29	0.7073	0.0322	0.0205	0.0284
30	0.7317	0.0332	0.0213	0.0293
31	0.7561	0.0342	0.0221	0.0303
32	0.7805	0.0351	0.0230	0.0313
33	0.8049	0.0361	0.0242	0.0323
34	0.8293	0.0371	0.0256	0.0334
35	0.8537	0.0381	0.0273	0.0347
36	0.8780	0.0391	0.0295	0.0360
37	0.9024	0.0401	0.0324	0.0376
38	0.9268	0.0411	0.0365	0.0396
39	0.9512	0.0421	0.0430	0.0424
40	0.9756	0.0431	0.0555	0.0471
41	1	0.0441	0.1320	0.0722

for example, causes that have a more pronounced effect on lateral instability result in more cars jumping the track following the initial derailment.

Using data reported by the Canadian Transport Commission (CTC) for the period 1983 to 1985, [52] explored the relationship between the probability of a specific number of cars derailing and the residual train length (L_r), speed and cause of derailment. For a train length L, and position j, they introduced a truncated geometric distribution for the probability of k cars derailing, such that:

$$P_r(k \text{ cars derailing} \,|\, \text{POD at position } j) = \begin{cases} \dfrac{p\,(1-p)^{(k-1)}}{1-(1-p)^{L_r}} & \text{if } k = 1, \ldots, L_r \\ 0 \, \text{otherwise} \end{cases} \tag{3.10}$$

where, $L_r = L - j + 1$ is the residual length or simply number of cars after POD, and $1-p$ is the probability of derailment for a position after POD. The probability p, is assumed to be related to the factors/covariates through the logit link function

$$p = \frac{e^z}{1+e^z}\,,$$

where z is a linear function of speed, L_r, and causes. As a standard approach for categorical variables, considering the cause "railbar" as a baseline, the other causes are entered to this model as binary variables (i.e., 0 or 1 for absence or presence of a specific cause, respectively). Thus for a typical train we have

$$z = \beta_o + \beta_1\,(\text{speed}) + \beta_2\,L_r + \beta_3\,(\text{roadbed}) + \cdots + \,\beta_8\,(\text{all other})\,.$$

The results from fitting a logit expression to derailment data from the CTC are summarized in Table 3.14. With the exception of switching causes, all factors have a statistically significant effect in explaining mean number of cars derailed.

Similar to the previous study, in this research, the distribution in Equation (3.10) is linked to the covariates through the logit function. This truncated geometric logistic model is fitted to the FRA data (instead of the CTC data) for the period of 1997-2006. Parameter values were estimated by maximizing a likelihood function of the form

$$L(\beta_o,\, \beta_1,\, \ldots,\, \beta_8) = \prod_{i=1}^{n} \frac{p_i\,(1-p_i)^{(x_i-1)}}{1-(1-p_i)^{L_r}} \tag{3.11}$$

using statistical software R (`http://cran.r-project.org/`). The results of this calibration are summarized in Table 3.15. The calibration procedure is presented in Appendix C.

With the possible exception of residual train length and track geometry, all input factors were found to provide a significant explanation for mean number of cars derailing at the

Table 3.14: Summary statistics for POD logit expression from CTC derailment data

Parameters	Estimates	Std.Error	Student T-test	Lower (95% interval)	Upper (95% interval)
Intercept	1.674	0.334	5.01	1.017	2.331
Residual length	-0.638	0.053	11.855	-0.744	-0.532
Speed effect	-0.575	0.082	7.036	-0.736	-0.414
Roadbed	0.648	0.143	4.505	0.365	0.931
Track geometry	0.382	0.094	4.060	1.197	0.577
Railbar	NA				
Switches	0.470	1.425	0.330	-2.33	3.270
General car	1.672	0.323	5.181	1.308	2.306
Axles/wheels	1.510	0.128	11.771	1.258	1.763
All other	1.329	0.261	5.091	0.816	1.842

Table 3.15: Summary statistics for estimates for FRA database (1997-2006)

Parameters	Estimates	Std.Error	Z statistics	Lower (95% interval)	Upper (95% interval)
Intercept	-2.013	0.082	-24.465	-1.850	-2.170
Residual length	0.001	0.001	1.266	0.002	-0.001
Speed effect	-0.032	0.002	-17.075	-0.029	-0.036
Roadbed	0.419	0.018	2.367	0.766	0.072
Track geometry	0.171	0.089	1.921	0.346	-0.003
Railbar	NA				
Switches	0.715	0.119	6.013	0.949	0.482
General car	0.841	0.085	10.132	1.030	0.697
Axles/wheels	1.108	0.077	14.404	1.260	0.958
All other	0.444	0.073	6.056	0.587	0.300

5% level. Track geometry and residual train length were found to be significant at the 10% level. Table 3.15 also gives the 95% confidence interval for each of the factors considered. Note that since the cause railbar is considered as a baseline, there is no parameter associated with in above tables. This table also shows that increasing the train speed would reduce the z value and subsequently increase the derailment probability. In addition, the effect of axles/wheels defects on the probability of derailment would be less compare to track geometry defect.

Chapter 4

Rail Yard Marshalling Model

According to the framework presented in Figure 2.1, the rail yard marshalling model (assembly model) is the second component of this study which describes the practical operations restrictions to the proposed model.

This research problem has been stated as an optimization problem. The Chapter 3 discussed how to obtain the components of the objective function. The missing part of the optimization model is its constraints which address realistic operating conditions. This chapter introduces the three constraints including DG marshalling violations (crew safety, and incompatibility) and yard operation cost (time). The final mathematical formulation and solution methods are then presented.

It should be noted that since 1980, many studies have been conducted about rail yard operations; however, these models mainly focus on minimizing overall rail yard operation time through hump sequence [61], blocking plan [20, 10] and train scheduling [37]. Previous works have failed to suggest any model for the yard engine operation in the classification yard. The algorithm in Appendix E has been developed to model the yard engine operation in case of presence of a DG car in a block.

4.1 DG Marshalling Regulatory Constraints

Both Canada and the US have specific regulations for the marshalling of DG cars. These regulations serve two basic purposes: to keep DG separate from personnel and to keep incompatible DG materials separate from each other[22, 50]. For example, most regulations prevent locating any DG car next to an operating engine, occupied rail vehicle, and caboose to increase the safety of rail personnel. In addition, it is prohibited to assign a DG car to next to a car with a source of ignition or next to a flatcar with protruding lading, to reduce the likelihood of DG material being released. Furthermore, DG cars must be separated by at least one car.

Canadian Transport Dangerous Goods (TDG) classifies DG into nine classes :

Class 1 (1.1, 1.2, 1.3 and 1.4) explosives

Class 2 (2.1, 2.2 and 2.3) gases (flammable, toxic and oxidizing)

Class 3 flammable liquids

Class 4 (4.1, 4.2 and 4.3) flammable solids (liable to spontaneous combustion and water reactive subtances)

Class 5 (5.1 and 5.2) oxidizing substances and organic peroxides

Class 6 (6.1 and 6.2) poisonous (toxic) and infectious substances

Class 7 radioactive materials

Class 8 corrosives

Class 9 miscellaneous products or substances

In addition, each DG has a unique number (UN number) that must be displayed on a car placard (More details are provided in Appendix D).

Following the 1979 Mississauga train derailment, the Canadian Transport Commission (CTC) regulated the marshaling of DG cars. The main intent of the CTC regulations is to provide sufficient distance to separate train personnel from DG cars. Table 4.1 summarizes the current regulations concerned with the placement of DG cars along a train consist. Based on these regulations, it is prohibited that DG cars described in column 1 be placed next to cars described in column 2.

U.S. regulations on DG car positions are similar to Canadian regulations, as shown in Table 4.2. Nine DG classes reclassified into four placard groups and the cars that carry specific groups are not allowed to be located next to each other. For instance, Placard Group 1, which includes explosive materials, can not be adjacent to Placard Group 2, which includes compressed gas.

The number of buffers between the crew and first DG car is not the same in the US and Canada. Recently, US Department of Transportation, Pipeline and Hazardous Materials Safety Administration reviewed the importance of buffer size [56]. This review states interestingly, in Canada, there is not any report shows a major incident as the result of having fewer of buffers.

Column 1 Dangerous Goods	Column 2 Railway Vehicle
Any class of dangerous goods	(a) an operating engine or an engine tender unless all the railway vehicles in the train, other than engines, tenders and cabooses, have placards displayed on them; (b) an occupied railway vehicle unless all the other railway vehicles in the train, other than engines, tenders and cabooses, are occupied or have placards displayed on them; (c) a railway vehicle that has a continual source of ignition; or (d) a railway vehicle that is a flat car from which part of the lading protrudes.
Dangerous goods included in Class 1.1 or Class 1.2	Any railway vehicle that is required to have a placard displayed on it for Class 2, 3, 4 or 5.
UN1008, BORON TRIFLUORIDE COMPRESSED UN1026, CYANOGEN UN1051, HYDROGEN CYANIDE, STABILIZED UN1067, DINITROGEN TETROXIDE or NITROGEN DIOXIDE UN1076, PHOSGENE UN1589, CYANOGEN CHLORIDE, STABILIZED UN1614, HYDROGEN CYANIDE, STABILIZED	Any railway vehicle that is required to have a placard displayed on it for Class 1, 2, 3, 4 or 5, unless the railway vehicle next to it contains the same dangerous goods.
UN1660, NITRIC OXIDE, COMPRESSED UN1911, DIBORANE, COMPRESSED UN1975, NITRIC OXIDE AND DINITROGEN TETROXIDE MIXTURE or NITRIC OXIDE AND NITROGEN DIOXIDE MIXTURE UN2188, ARSINE UN2199, PHOSPHINE UN2204, CARBONYL SULPHIDE or CARBONYL SULFIDE UN3294, HYDROGEN CYANIDE, SOLUTION IN ALCOHOL	Any railway vehicle that is required to have a placard displayed on it for Class 1, 2, 3, 4 or 5, unless the railway vehicle next to it contains the same dangerous goods.

Figure 4.1: Regulations of transport of dangerous goods in Canada (Source: [22])

Restrictions	Placard Group1	Placard Group2		Placard Group3		Placard Group4
	Railcar	Tank Car	Rail Car	Tank Car	Rail Car	Rail Car
1-When train permits, placarded car may not be nearer than the sixth car from the engine or occupied caboose.	X	X		X		
2-When train length does not permit, placarded car must be placed near the middle of the train, but not nearer than the second car from an engine or occupied caboose.	X	X		X		
3-A placarded car may not be placed next to an open-top car when any of the lading in the open top car protrudes beyond the car ends, or if the lading shifted, would protrude beyond the car ends.	X	X		X		
4-A placarded car may not be placed next to a loaded flat car, except closed TOFC/COFC equipment, auto carriers, and other especially equipped cars with tie-down devices for securing vehicles. Permanent bulk head flat cars are considered the same as open-top cars.	X	X		X		
5-A placarded car may not be placed next to any transport vehicle or freight container having an internal combustion engine or an open-flame device in operation.	X	X		X		
6- Placarded cars may not be placed next to each other based on the following:						
Placard Group 1………………………………		X	X	X	X	X
Placard Group 2………………………………	X			X	X	X
Placard Group 3………………………………	X	X	X			X
Placard Group 4………………………………	X	X	X	X	X	

PLACARD GROUP:
Group1-Divisions 1.1 and 1.2 (explosive) materials.
Group2 - Divisions 1. 3, 1. 4, 1. 5 (explosives). Class 2 (compressed gas: other than Div 2. 3, PG I, Zone A). Class (flammable liquid), Class 4 (flammable solid), Class 5 (oxidizing). Class 6 (poisonous liquid; other than Div 6.1, PG I, Zone A), and Class 8 (corrosive) materials.
Group3 – Divisions 2. 3 (Zone A: poisonous gas) and 6.1 (PG I, Zone A; poisonous liquid) materials.
Group 4 – Class 7 (radioactive) materials.

Figure 4.2: Regulations of placarded cars in the US (Source: [50])

While the regulation provides overall guidelines, they do not address the problem of optimally positioning DG cars in terms of their potential for derailment and operating costs.

This section shows how the model accommodates DG marshalling violations in train assembly. Canadian Pacific Railway General Operating Instructions (GOI) [49] summarize the regulations for DG marshalling so that DG are adequately separated from

- occupied equipment, such as an operating locomotive;
- other incompatible DG cars on the train;
- a car equipped with a source of ignition, a mechanical heating source, or cooling device;
- an open top car with lading protrudes beyond the car or lading above the car end that is liable to shift lengthwise;
- a loaded flat car.

The above five restrictions are included in the proposed model as follows: five cars for the US, one car for Canada acts as buffers at the front of each block; incompatible DG cars must be kept separate from one another, and the last three restrictions (related to source of ignition, open top car and loaded flat car) are combined to the incompatibility constraint.

This section introduces the above restrictions as new constraints in the optimization model. It should be noted that the marshalling restrictions apply to all DG cars, whether they are loaded or contain only the residue of their DG load (residues).

Table 4.3 shows the incompatibility restriction for DG cars in the train consist. The nine classes of DG materials are categorized into four groups (A, B, C, and D). For instance, based on the regulations, it is prohibited for group A DG cars to be placed next to cars of group B and C. There should be at least one buffer (non-DG) car between incompatible DG cars (Notes 1 and 2 discussed in Table 4.3 are not considered for simplicity). The last three restrictions grouped into a hypothetical group-Group E, which includes a car equipped with a source of ignition, a mechanical heating source or cooling device; an open top car from which lading protrudes beyond the car or lading rising above the car end is liable to shift lengthwise; and a loaded flat car. Group F (hypothetical group) accommodates cars with non-DG material.

The alternative combination of cars (X) is defined as

$$X = \left\{x_1, x_2, x_3, \ldots, x_n\right\}$$

Dangerous Goods Cars in Group/Class:	Must not be placed next to: Group A	Group B	Group C
Group A: Explosives Classes 1.1 & 1.2 		X	X(1)
Group B: (Infrequently handled. See list below.)	X	X(2)	X
Group C: Explosives Classes 1.3 to 1.6,  Classes 2, 3, 4, and 5. 	X(1)	X	
Group D: Classes 6, 7, 8, 9, and mixed loads 	Only general marshalling restrictions (item 6.3) apply.		

Notes:
X "must not be next to" restriction
(1) not applicable to explosives in Classes 1.3 to 1.6.
(2) not applicable if the next car has the same UN number.

Group B Dangerous Goods (Infrequently handled)

UN 1008, CLASS 2.3	UN 1660, CLASS 2.3
UN 1026, CLASS 2.3	UN 1911, CLASS 2.3
UN 1051, CLASS 6.1	UN 1975, CLASS 2.3
UN 1067, CLASS 2.3	UN 2188, CLASS 2.3
UN 1076, CLASS 2.3	UN 2199, CLASS 2.3
UN 1589, CLASS 2.3	UN 2204, CLASS 2.3
UN 1614, CLASS 6.1	UN 3294, CLASS 6.1

Figure 4.3: Incompatibility restrictions among DG cars (Source:[49])

where, x_i is ID of the car assigned to position i and n is the total number of poitions in the train consist.

For any given combination of cars, the types of materials, either DG or non-DG are inputs to the model. Thus for any initial combination of cars the types of loaded material for a train, is given as follows

$$
\begin{aligned}
H & \text{As Vector} \\
H &= \{h(x_i)\} \\
h(x_i) &= \{1, 2 \ldots, w\} \\
i &\in \{1, \ldots, n\}
\end{aligned}
$$

where $h(x_i)$ is the type of car x_i, n is the total number of positions in the train consist, and w represents the number of types of loaded materials. The first restriction (1) can be introduced as a constraint in the model so:

$$h(x_i) = 6, \text{ for } i = 1, \ldots, m \tag{4.1}$$

where m, is the minimum size of buffers ($m < n$), determined by the regulation; for instance, in the US, m equals to five cars while in Canada it is only one car.

Table 4.3 and GOI restrictions (discussed before) can be represented as a matrix:

$$
\begin{aligned}
\pi & \text{ As Matrix} \\
\pi &= \{a_{df}\} \\
d &\in \{1, \ldots, w\} \\
f &\in \{1, \ldots, w\}
\end{aligned}
$$

where $a_{df} = 1$ when the car involving DG material d must not be assigend next to a car with DG material f or $a_{ij} = 0$ otherwise. Each cell in the matrix shows the interaction of two types of materials. For the case $w = 6$, $a_{13} = 1$ since, based on Table 4.3, a car with full of TNT (Group A or h=1) must not be located next to a car loaded with chlorine (Group C or h=3).

$$
\pi = \begin{pmatrix}
0 & 1 & 1 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

For a given train, the incompatibility restriction and remaining restrictions can be shown as in the following constraint:

$$\pi_{x_i,x_{i+1}} = 0,\ i \in \{1, \ldots, n-1\} \tag{4.2}$$

4.2 Yard Operations Cost Constraint

Yard operation cost is one of the main factors influencing marshalling. A marshalling solution that minimizes risk could mean additional costs because rail crews have to cut off and couple cars to set a safer train consist. This study hence investigates the costs associated with additional tasks to change the order of cars within blocks at a rail yard.

Currently, the order of blocks on a given train is generally set based on the sequence of intermediate destination points along a route such that the block assigned to the closest destination is shunted to the front of the train, followed by the block assigned to the next closest destination, etc., until the final destination block is connected. Positioning the blocks according to railway station sequence reduces the time for setting off procedures at intermediate destinations. However, in some circumstances, this style of block positioning is not followed. For instance, sometimes the order of blocks is changed because of loading or unloading restrictions.

In June 2005, the US Federal Railroad Administration released a report that reviewed several previous studies on the safe placement of train cars [8]. Because of the costs associated with the additional switching of cars, the report rejected a suggestion made by the Hazardous Materials Transportation Authorization Act of 1994 (P.L. 103-311) for regulatory changes to increase the safe shipment of DG cars. Cars switching at rail yards does increase the time of operation and thus means higher immediate cost. However, this is a very short-sighted view of the situation, one that completely neglects the huge environmental, safety, and financial costs of derailments involving the transport of DG goods.

There are two options for taking into account the cost of yard operations: option 1, adding a cost constraint to the optimization model; or option 2 incorporating the incremental time required for marshalling into the objective function.

In this study, rail yard operation costs are considered as an additional constraint in the model (option 1). The rail yard marshalling constraint considers that the rail yard train assembly costs (or processing time) for each train and DG placement strategy should be less than the threshold time (yard master's decision). In this study, it is assumed that the additional time required at intermediate destinations can be ignored for simplification.

Two cost constraints are required: one for cutting off and marshaling cars and a second for changing the order of blocks, which may increase the cost of cutting off the blocks during the trip. The second option considers the incremental time required for marshalling. In this case, once a train consist is determined, the cost associated with achieving it is reviewed to balance safety and cost. By considering the cost of additional marshalling, a mutually agreeable point can be found to show the most cost-effective yet safe strategy for a given scenario.

Figure 4.4 illustrates an example of a given combination (X_0) of cars in a block. It is assumed that there is no difference between the X_0 combination and the alternative one (X) from position 1 to position i; thus, the first part of the train (between 1 to i) is ignored in this analysis. For the second part of the train (between i to j), there is at least one difference between the (X_0) and (X) in that car y should replace car x in the i position.

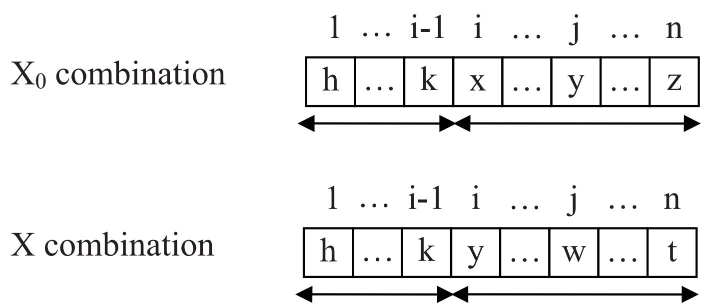

Figure 4.4: General representation for a given and alternative combination of cars for one block in a classification track

The proposed model begins once blocks have been formed with the desired number of cars but before the blocks are shunted by a yard engine to the departure track. According to information collected during interviews and site visits, the total time required for one position change for a car within a block (t_f) can be assigned to the time required for switching (t_{sw}) and the time for additional marshalling (t_m). In addition, the total time (T) required to complete all the changes (F) is as follows:

$$T(X_0, X) = \sum_{f=1}^{F} t_f = \sum_{f=1}^{F} (t_{sw} + t_m)$$

The total time (T) to change from the X_0 to the X order of cars is the summation of the time for each individual replacement, which should be less than the threshold time

(T_{Max}) determined by the yard master.

$$T(X_0, X) \leq T_{Max} \tag{4.3}$$

The following procedure details the operational steps involved in moving car y in position j to the designed position i where x is located (Figure 4.5):

i. tie yard engine to car z and set off the train consist from position i

ii. move the cars $(n - i + 1)$ to the side

iii. back the cars$(n - i + 1)$ to the empty track

iv. set off the cars from position j, and move the cars ($n - j + 1$) to the side

v. back the cars $(n - j + 1)$ to the track 1, and couple the cars to car k

vi. move the rest of the cars $(i - 1)$ from track 2 to the side and back to track 1

The total time for one replacement includes the time for four switching operations and four trips of $(n - i + 1)$ cars back and forth.

The time for additional switching is a product of the number of times switching takes place (c) and the fixed time required to switch from one line to another (t_0 minute). The total number of switches required to change the order of cars from a given combination to an alternative one is the summation of all individual switches to change one position in a given combination to an alternative.

$$t_{sw} = c \times t_0$$

Figure 4.5 shows one way to change one car position assuming that one empty track is available in the classification track.

The time for additional marshaling can be obtained based on the number of cars marshaled from one track to another to reach the alternative combination. To estimate this time, the rule of $time = distance/velocity$ is used, where, the distance (mile) is the product of the number of cars marshaled (l) to an average length of a car (d_0=15 meters) as

$$t_m = \frac{4ld_0}{v}$$

where v is the average speed of the yard engine in the classification yard used for marshalling cars. v is usually in the range of 7 mph (11 km/h).

To achieve the alternative combination, the algorithm in Appendix E has been developed to estimate the number of switches (c) and number of cars being marshaled (l).

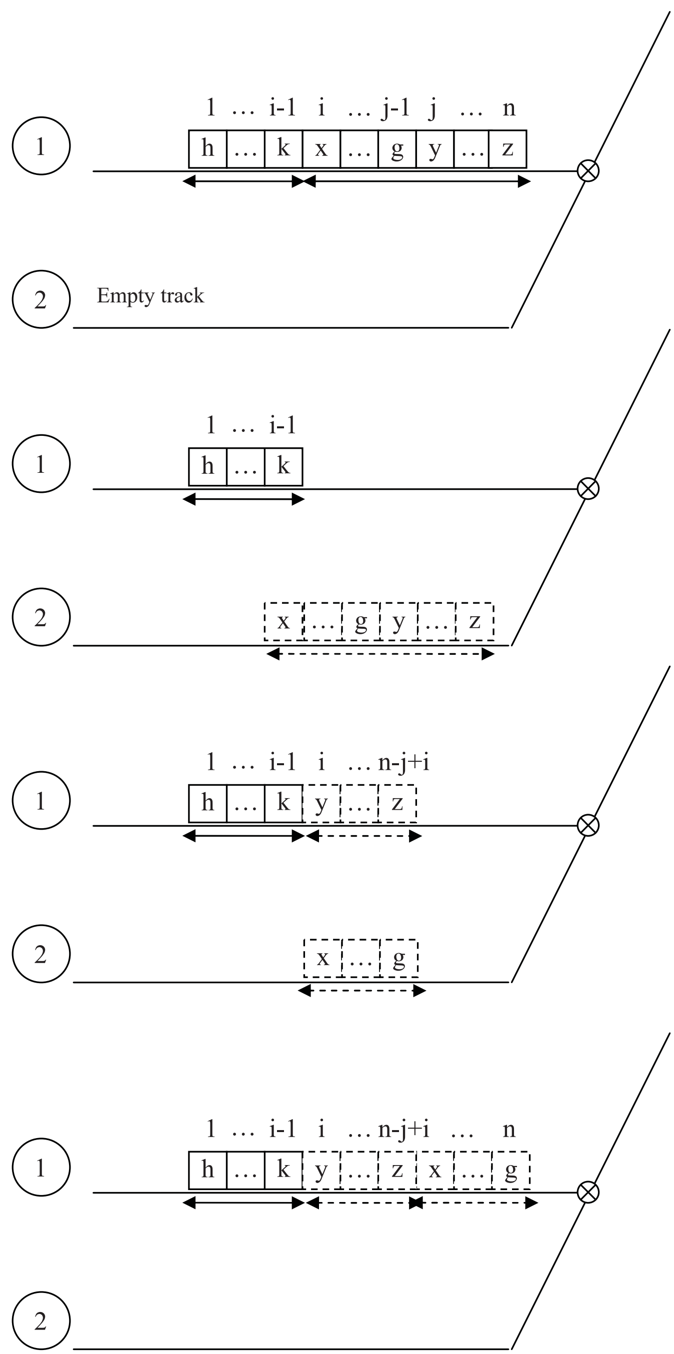

Figure 4.5: Changing one car position when one empty track is available

4.3 Minimizing Derailment Risk Subject to Rail Yard Constraint

As discussed before, for a given route segment and an n car train, the risk associated with all positions of the train is estimated by summing in Equation (3.5) for $i = 1, \ldots, n$.

$$\sum_{i=1}^{n}\left[P_r(TD) \times \sum_{j=1}^{i}\left[P_r^{POD}(j) \times \sum_{k=i-j+1}^{n-j+1} P_r(k \mid j)\right] \times C(x_i)\right]$$

Finally, combining all the above equations ((3.5),(4.1),(4.2),(4.3)) results in the following formulation as the optimization model with three main constraints.

$$min \qquad P_r(TD) \times \sum_{i=1}^{n}\left[\sum_{j=1}^{i}\left[P_r^{POD}(j) \times \sum_{k=i-j+1}^{n-j+1} P_r(k \mid j)\right] \times C(x_i)\right]$$

s.t.

1. $h(x_i) = 6, \; i = 1, \ldots, m$
2. $\pi_{x_i, x_{i+1}} = 0, \; i \in \{1, \ldots, n\}$
3. $T(X_0, X) \leq T_{Max}$

where the decision variable is a column vector: $X = \{x_i\}$ and x_i is ID of the car assigned to position i.

4.4 Solution Methods

This problem formulated previously is an integer programming (IP) problem, which is fundamentally difficult to solve for large sized problems. For instance, consider a train with 100 cars and half of them carrying DG materials. Assuming the train has only one destination and all the DG have the same level of risk, the number of possible combinations is more than 10^{29}.

In this research, two solution approaches are considered: (1) enumeration; (2) Genetic Algorithm (GA). The first approach is only viable for solving small sized problems and therefore the second approach is selected to solve the research problem.

The Genetic Algorithm approach can find near optimal solutions to problems that are "unsolvable" for standard linear and non-linear optimizers. A solution delivered from a genetic algorithm is called a “chromosome”, which is illustrated as a string including a set of components called “genes”. In this study, a block of cars can be defined as a chromosome and each car can be defined as a gene. Each chromosome is evaluated against an objective function. The best chromosomes produce offspring chromosomes. Each time, the offspring chromosome is evaluated and if it provides better results it replaces weaker members. This process continues for a number of generations to obtain a near optimum solution [31]. This research uses existing software called Evolver Palisade which provides powerful genetic algorithm-based optimization techniques to solve the formulated optimization problems.

To implement the proposed model several software packages have been used (Figure 4.6). The in-transit risk model was encoded in Matlab (Appendix F) while the optimization model has been implemented in Evolver (Palisade). The three restrictions have been added as the constraints to Evolver. To calculate the time required for additional marshaling operation in classification track, the algorithm (Appendix E) is coded in Visual Basic (Appendix G). In Evolver, Visual Basic macros can be run at different times during an optimization for each trial solution. This allows the development of custom calculations that will be invoked during an optimization. This feature allows calculations which only can be performed through the use of a macro to be made during an optimization. The developed macro runs after recalculation of each trial that is executed (After Recalculation of Each Trial).

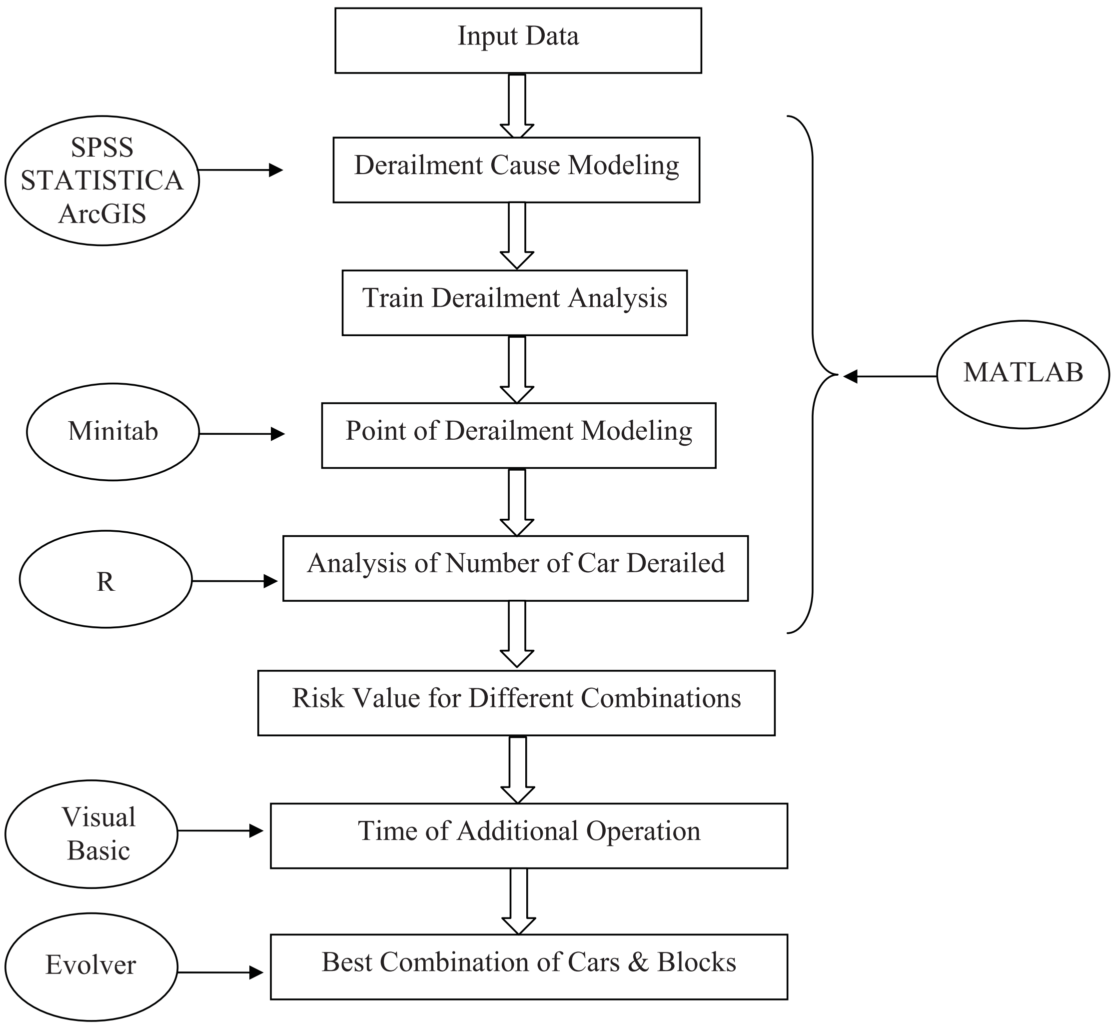

Figure 4.6: Different software packages and techniques used in this study

Chapter 5

Case Study Application

The merits of a new car placements model lie in its ability to achieve maximum risk reduction at a minimum cost. The proposed model must be able to provide insights about changes in overall risk for the optimal scenario compared to the current operation scenario. For instance, how much of the risk of DG car derailment will be decreased if the order of cars and blocks are changed based on the result of proposed model? What would be the processing cost (time) for these changes? Are the results sensitive to changes in operation conditions such as the proportion of DG?

In this chapter, the proposed risk-based model is applied to a rail corridor to demonstrate how overall derailment risks can be effectively reduced through the marshaling of DG along a train.

In addition, solutions from the proposed model are compared with results from other suggested strategies. A sensitivity analysis has been carried out to determine how changes in the proportion of DG, buffer size, and route attributes affect the model.

5.1 Description of Corridor and Shipment Characteristics

The model is applied to a railway corridor (Figure 5.1) to illustrate how risks along a route can be used to develop cost effective DG placement strategies at the rail yard. The corridor connects Barstow (train origin point), California to Chicago (Corwith Station), Illinois through two intermediate stations (Albuquerque and Kansas City) over 2100 miles and involves a range of track geometry features for six segments. The Albuquerque, Kansas City, and Corwith stations are situated respectively, at a distance of 768, 1663, and 2100 miles from the Barstow Hump Yard.

Figure 5.1: The route map with six segments (Source:[6])

There are two main reasons for using this corridor. First, according to the US Commodity Flow Survey (CFS) in 2002 and Freight Analysis Framework (FAF), California and Illinois are DG shipment origin and destination by rail, respectively [3, 4]. Second, the corridor includes several segments with various geometric characteristics, including track class and elevation differences (Table 5.1). The corridor is one of the few mainlines with high-quality track (track class 5). Thus, a train along the corridor experiences poor and high quality track from California to Illinois. In addition, the train traverses several areas with different elevation differences.

Table 5.1: Corridor attributes

Station	Seg.	Distance (mile)	Track Class	Track Type Poor (1) Good (2)	Max Speed (mph)	Elevation Difference (m)	Topology Flat (1) Mount.(2)
Albuquerque	1	169	4	1	60	4,507(CA)	2
	2	599	5	2	80	3,830 (AZ)	2
Kansas City	3	243	4	1	60	3,147(NM)	2
	4	157	5	2	80	3,391(CO)	2
	5	495	5	2	80	1,024 (KS)	1
Corwith	6	437	5	2	80	470 (MO), 291(IL)	1

To understand the track class of each segment, the timetable speeds of passenger train has been used. The Southwest Chief passenger train operates along 2260 miles to connect Los Angles, California to Chicago, Illinois. For each segment, the average speed is calculated using the distance and the time of operation. Then the track class is obtained on the basis of FRA's track classification (Table 3.1).

The assumed car shipments to each intermediate station, consisting of a given mix of DG and non-DG cars, is summarized in Table 5.2. The train length 110 cars is selected because, according to train length distributions from FRA database 1997-2006 (Figure 3.5), the most frequent train derailments occurred for trains with 100-110 cars. Usually 10-15% of a train consists of DG cars (Marshaling Switching course at George Brown College). Thus, in this study, a train is assumed to consist of ten DG cars.

Table 5.2: Car shipments along the case study corridor

Station	Distance from rail yard (mile)	Total number of cars	Number of DG cars
Albuquerque	768	25	2
Kansas City	1663	30	3
Corwith	2100	55	5

In this application, information is needed concerning train derailment rates and causes along each route segment. The route considered consists of six uniform segments: two segments between the Barstow Hump Yard, and Albuquerque of lengths 169 (Barstow-Needles), and 599 (Needles-Albuquerque) miles; three segments between Albuquerque and Kansas City of lengths 243 (Albuquerque-Raton), 157 (Raton-Lamar), and 495 (Lamar-Kansas City) miles; and one segment of 437 mile between Kansas City and Chicago.

5.2 Estimating Probability of Derailment by Position

Table 5.3 summarizes the derailment rates and the probability of train derailments for the different route segments obtained by the derailment rates(per billion freight car-miles and per million freight train-miles) from Table 3.2 and assuming that approximately 25% of all derailments can be classified as train-mile caused and the remaining as car-mile caused. For example, for segment 1, RC and RT are calculated using values from Table 3.2,as follows:

$$RC = 7.8 \times 0.75 = 5.85, \qquad \text{and} \qquad RT = 0.53 \times 0.25 = 0.132\,.$$

The values in the last column of Table 5.3 are calculated by utilizing the formula in

Table 5.3: Probability of train derailment along the case study corridor

Station	Seg.	Distance (mile)	Track class	Speed (mph)	Train derailment rate		Train length	Expectation of train derailment
					RC($\times 10^{-9}$)	RT($\times 10^{-6}$)		
Albuquerque	1	169	4	60	5.850	0.132	110	0.131
	2	599	5	80	3.670	0.080	110	0.290
Kansas City	3	243	4	60	5.850	0.132	85	0.153
	4	157	5	80	3.670	0.080	85	0.061
	5	495	5	80	3.670	0.080	85	0.194
Corwith	6	437	5	80	3.670	0.080	55	0.123

Equation 3.6. These values are basic estimates of the probability of train derailment along different route segments which provide estimates of $P_r(TD)$ in Equation (3.2).

$P_r(i \,|\, TD)$ for a given segment is obtained from Equation (3.4) based on the POD and the probability of a number of cars derailing. As discussed in Chapter 3, this calculation requires information concerning the expected causes of derailment along the segment and various train operating characteristics. According to Table 3.14, the distribution of derailment causes associated with individual/specific segments can be predicted based on route attributes and rolling stock characteristics(Table 5.4).

Table 5.4: Probability of causes of derailment for different segments along the case study corridor

Station	Segments	(train length,area,track class)	Cause groups		
			C1	C2	C3
Albuquerque	1	(2,2,1)	0.701	0.237	0.061
	2	(2,2,2)	0.605	0.264	0.131
Kansas City	3	(2,2,2)	0.605	0.264	0.131
	4	(2,2,1)	0.701	0.237	0.061
	5	(2,1,2)	0.615	0.313	0.071
Corwith	6	(2,1,2)	0.615	0.313	0.071

To estimate P_r(POD at position j) in Equation (3.8), Table 3.12 is used. The parameters m_1 and m_2 can be obtained from Figure 3.8 for each segment (Table 5.5).

To estimate $P_r(k$ cars derailing $|$ POD at position $j)$, Equation (3.10) and Table 3.15 are used. Finally, by applying Equation (3.2) the probability of car derailment by position for a given train and track segment (P_i) can be obtained.

Table 5.5: Membership functions of distributions based on the train length

Station	Segment	m_1	m_2
Albuquerque	1 2	0.3	0.7
Kansas City	3 4 5	0.8	0.2
Corwith	6	0.4	0.6

5.3 Discussion of Alternative Solutions

The involvement of DG cars in derailment depends on the DG car placement strategy. As discussed before, a block consists of a set of cars for same destination. As discussed in Chapter 2, the proposed solution method starts with a given order of blocks and then the best combination of cars with in each block can be identified by minimizing the total risk. This process repeats for each order of blocks as discussed for MDG2 in Section 2.2. In this analysis, all six possible block sequence strategies are considered. Each block sequence is labelled using the first letters of the destination cities. For instance, CKA means that the first block after the engine is block A (Albuquerque), then block K (Kansas City), and finally the last block is C (Chicago). In addition, the current sequence strategy (determined based on destination, i.e., CKA^0) is considered as the base do-noting case. As a result, the following seven solutions are presented :

CKA^0 do nothing, current operation when the order of blocks is based on intermediate and final destinations, and the order of cars is based on FCFS considering DG marshalling violation (transportation of DG regulations).

CKA^* the same as CKA^0 but with the order of cars in each block determined from the application of the proposed model.

CAK, ACK, AKC, KAC, and KCA are five other block arrangements along a train with the order of cars determined from the application of the proposed model.

For each block placement, the risk of DG placement is obtained using Equation (3.1) for all train positions over the entire route. From Table 5.2, only those positions that are relevant for a given route segment with respect to the assumed shipment profile are considered in the estimation of risk. For example, cars destined for Albuquerque (the first station along the route) will not be considered for segments beyond this station. Thus, the train length varies over the route, from 110 cars traveling from the Barstow Hump Yard to Albuquerque, 85 cars to Kansas City, and 55 cars to Chicago. More details can be found in Appendix A.

Using Evolver (genetic algorithm-based optimization software) the objective function is to minimize the total risk considering the three following restrictions: a) total additional time required for marshalling and switching is less than 300 minutes (five hours), b) incompatibility between DG cars is addressed based on the regulations, and c) the first five cars behind the locomotive should not carry DG materials, according to US DG regulation. It should be noted that the incompatibility between DG cars is considered within blocks and between blocks.

Table 5.6 summarizes the case study corridor application in terms of train-wide DG car derailments along the constituent segments from Barstow to Chicago. To estimate the total risk value for current operation CKA^0, the order of DG and non-DG cars are allocated randomly to reflect an FCFS operation.

Table 5.6: Comparison of total risk values for optimum placement of cars in all the sequences of blocks

Label	Risk		Time	
	Value($\times 10^{-6}$)	Difference(%)	(min)	Difference(%)
CKA^0	1658	0	142	0
CKA^*	1106	33	182	28
KCA	1172	29	225	58
CAK	1082	35	184	29
ACK	1239	25	195	37
KAC	1122	32	234	65
AKC	1176	29	227	60

The above results can be represented in terms of percent reduction in risk considering the additional rail yard train assembly time with respect to the Do-nothing strategy (CKA^0). For the case study, the results can be illustrated as in Figure 5.2.

The results suggest that more than 33% reduction in risk can be realized over the entire route if the existing block order is maintained (i.e., CKA), but individual DG cars within each block are positioned to reflect the in-transit risk minimizing principle. The strategy that yields the lowest route risk corresponds to the CAK block order. The benefit of the risk-minimum strategy (CAK) over the base case is 35% reduction in risk.

Given the rather modest risk safety gains associated with the CAK block order over the CKA^* order with restricted DG placement, it may be advisable for this corridor application

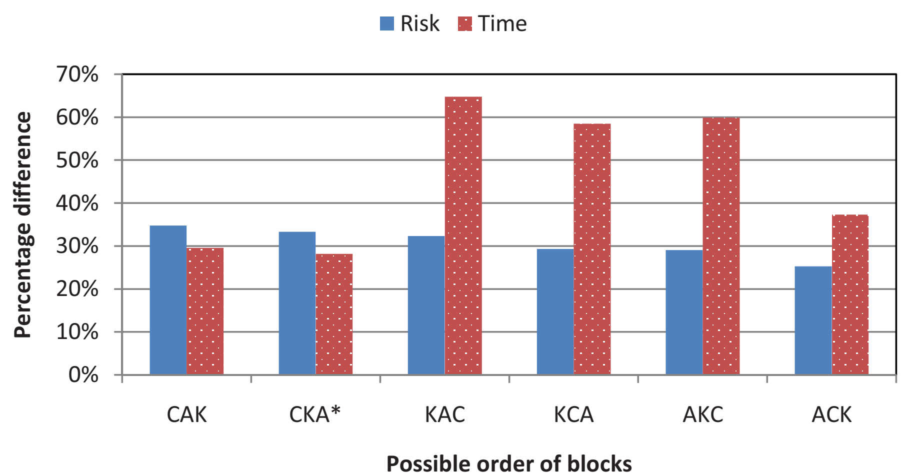

Figure 5.2: Comparison of risk and time percentage difference for six possible sequences of blocks

to continue with the existing blocking order (CKA) and only restrict DG placement within their respective blocks to the lowest risk positions.

The results obtained from optimization (Table 5.6) are compared in Figure 5.3. This figure is quite revealing in several ways. Comparison of CKA^0 and CKA^* shows that the latter is preferable as it reduces the risk value significantly while experiencing small increase in train assembly time. For better understanding of these strategies, the upper left corner of Figure 5.3 is amplified in Figure 5.4.

On the other hand, the evaluation of CKA^* and CAK (Figure 5.4) illustrates that allocating more time for additional operations could reduce the risk further. Obviously, KAC, KCA, AKC, and ACK are not interesting cases as they are high-risk, high-cost scenarios. However, it does not mean that the assessment of different possible combinations of blocks is unnecessary as other examples have shown that in some cases they achieve better results.

Table 5.7 illustrates the positions of the DG cars for various combinations of blocks. These results are illusterated in Figure 5.5. The front and rear of the train is the safer place for allocating DG cars for segments 1- 5. However, for the last segment, it seems that the rear of the train is preferred for positing DG cars.

The Evolver setting has a runtime feature that determines the total time of the optimization. There are stopping conditions that specify how and when Evolver stops during an optimization. Once an optimization starts, the Evolver continuously runs, and searches for better solutions and running simulations until the selected stopping criteria are met.

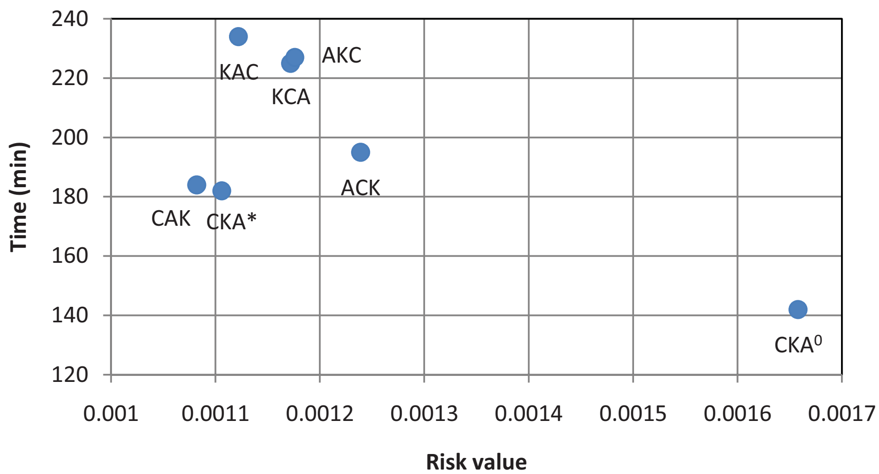

Figure 5.3: Risk and time graph of sequences of blocks

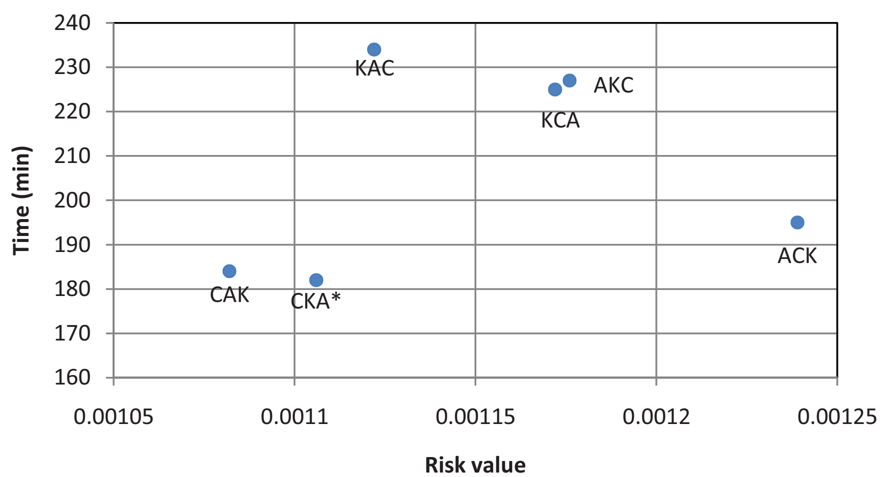

Figure 5.4: Risk and time graph of sequences of blocks excluding CKA^0

Optimization runtime options includes an option when set, stops the Evolver from simulating scenarios when the improvement in the target cell is less than the specified amount (change criterion). In this case, Evolver's answers have not improved at least 0.0001% over the last 1000 simulations, so it assumes there is little more improvement to be found and it stops the search. Figure 5.6 provides more details for each individual sequence of blocks. The time spent for each optimization is reported in Table 5.8. As can be seen the average

Table 5.7: Position of DG cars for different sequence of blocks

DG Description			Position of DG Cars							
DG Name	UN	Group	Given	CKA^0	CKA*	KCA	CAK	ACK	KAC	AKC
Nitrogen dioxide	1067	2	22	22	8	6	34	110	80	110
ANFO	0082	1	23	25	6	9	32	108	78	108
Cholorine	1017	3	26	42	35	110	10	10	108	85
Carbonyl sulphide	2204	2	28	32	33	109	8	6	105	83
TNT	0209	1	30	54	31	108	6	8	110	80
Nitroglycerine	0144	1	56	106	108	35	104	85	12	9
Diborane	1911	2	57	83	104	31	106	71	9	13
Hydrochloric acid	1789	4	58	74	109	32	110	84	6	6
Arsine	2188	2	60	104	106	80	107	82	10	12
Ethanol	1170	3	62	62	110	33	109	81	7	7

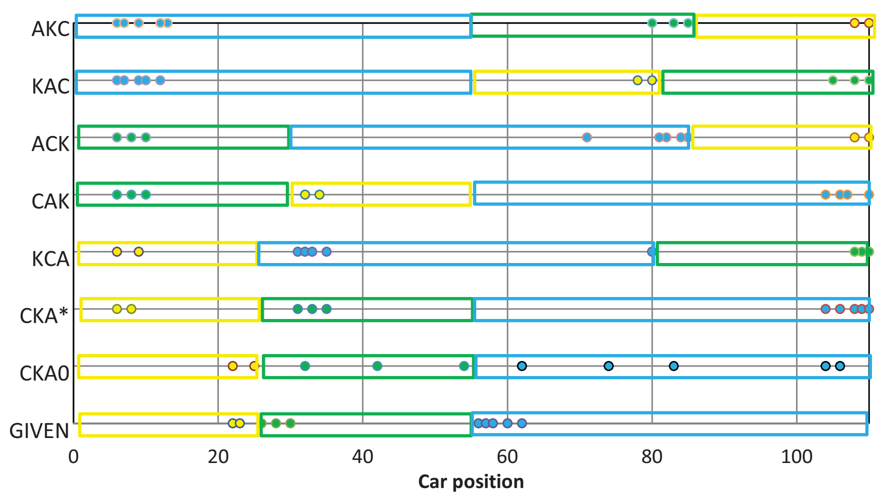

Figure 5.5: Position of DG cars for different sequence of blocks

running time for each optimization is approximately 10 minutes.

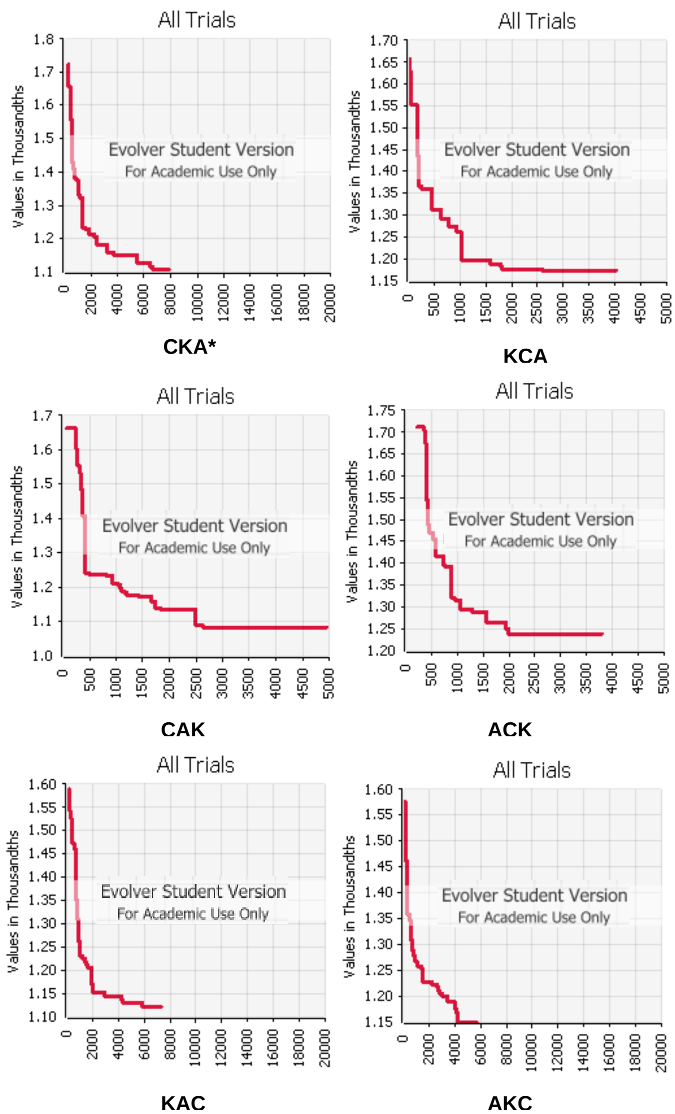

Figure 5.6: Variation of objective fuction for six block sequences

Table 5.8: Run time for the optimization of each sequence of blocks

Scenario	Time(min)
CKA*	00:13:08
KCA	00:07:16
CAK	00:08:44
ACK	00:07:40
KAC	00:14:26
AKC	00:09:39

5.3.1 A Comparison to Other DG Placement Strategies

The following analysis is conducted to compare the results of proposed model to the other strategies suggested by previous studies. A number of DG combination strategies are possible. In addition to the first-come first-serve (random) option (CKA^0) and CKA^*, three other possible CKA block combinations include

1. allocating DG cars to the rear end of each block (CKA^1).
2. putting DG cars at front of each block (CKA^2).
3. allocating DG cars to the lowest risk slots while ignoring all the constraints (CKA^3).

The first alternative is based on two previous studies [34, 57] which state that DG cars should be placed closer to the rear. A study by FRA [40] supports the second strategy that the preferred position for loaded cars (DG and non-DG) should be the front part of the train. The third strategy simply ignores all the constraints and only minimizes the risk of DG derailment.

The data in Figure 5.7 suggest that the risk would be reduced around 42% if no constraints are considered (CKA^3). Allocating DG cars to the front of the train (CKA^2) is not recommended as it would result in the higher risk, higher cost (time). A comparison of the strategy CKA^1 and CKA^* with CKA^0 shows clearly that the CKA^* is preferred as it has almost half of the risk with only a small increase in operational time (40 min).

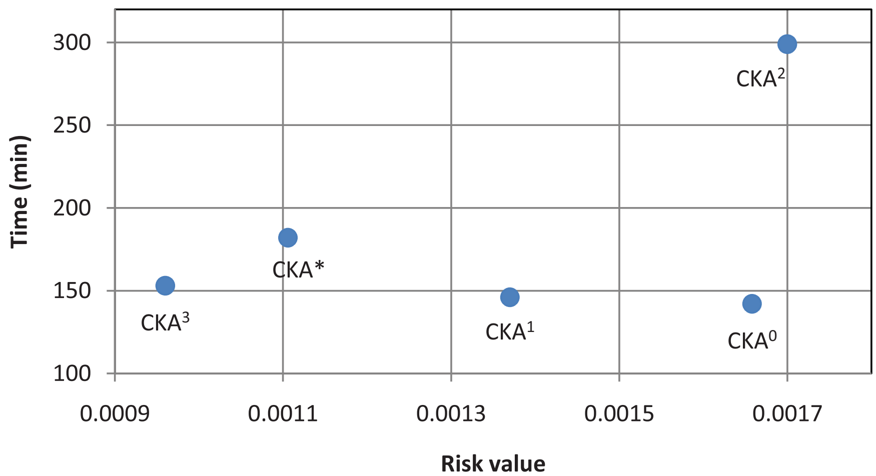

Figure 5.7: Comparison of five different marshaling strategies

5.4 Sensitivity Analysis

Many factors may affect the findings described in the previous section. The main objective of this section is to investigate the potential effect of several important condition factors on the benefits of the proposed model.

Initial Sequence of Cars

The relative benefit of the proposed model depends on the initial sequence of cars which is random under the operation rue of the FCFS. To evaluate the impact of the initial sequence of cars on the result of the model, in addition to the case study scenario with a given random sequence of cars CKA^0, fifteen sequences have been produced randomly. Note that in all of these sequences, the existing block order is maintained (i.e., CKA), but individual DG cars within each block are positioned randomly. Figure 5.8 shows the risk and time for fifteen random sequences of cars. Each of the individual sequences in this figure can be seen as an alternative for CKA^0. There are some cases where the random placement of DG cars do not break DG marshalling violations (crew safety, and incompatibility) and thus in these scenarios, the blocks leave the classification yard right a way without spending any additional time. Furthermore, there are two sequences where their associated times are higher than the case study scenario. Relative benefit of reduced risk with regard to higher yard costs less pronounced for trains that differ in profit from the one used in the case study.

Figure 5.8 could be divided into two areas upper and lower areas. It would be easy to justify the proposed model for scenarios in the upper area as they are close to the proposed strategy (CKA*) in terms of time. However, it would be difficult to justify the additional marshalling for scenarios with zero time. This analysis clearly shows that the sequence of cars when they arrive can affect the results; therefore, implication of the proposed model depends on the initial sequence of cars.

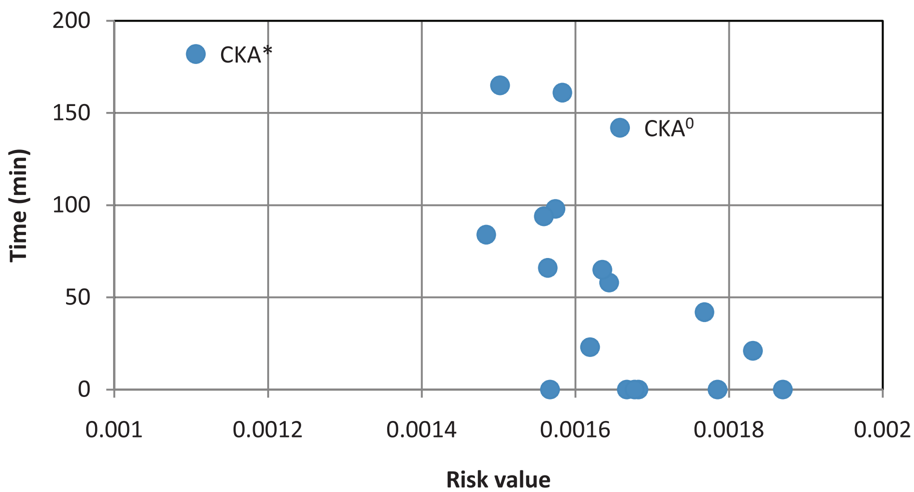

Figure 5.8: Risk and time for fifteen random sequences of cars (CKA^0)

DG Proportion

To evaluate the impact of DG proportion on the model, in addition to the case study scenario with ten DG cars, two other scenarios are introduced with five and fifteen DG cars, respectively. The five DG car scenario is similar to the case study scenario with ten DG cars Table (5.7) excluding ANFO, TNT, carbonyl sulphide, arsine, and ethanol. The last scenario includes five additional DG cars: sulfuric acid (within block A); sodium hydroxide (within block C); and phosphoric acid, maleic anhydride, and hydrodic acid (within block K).

The results of the proposed model for different sequences of blocks are illustrated in Figures 5.9 and 5.10. As can be seen in 5.9, in all three scenarios, risk can be reduced at least by 30%. The trend shows that risk can be mitigated more when there are less DG cars in the train consist due to more available slots for placement of DG cars.

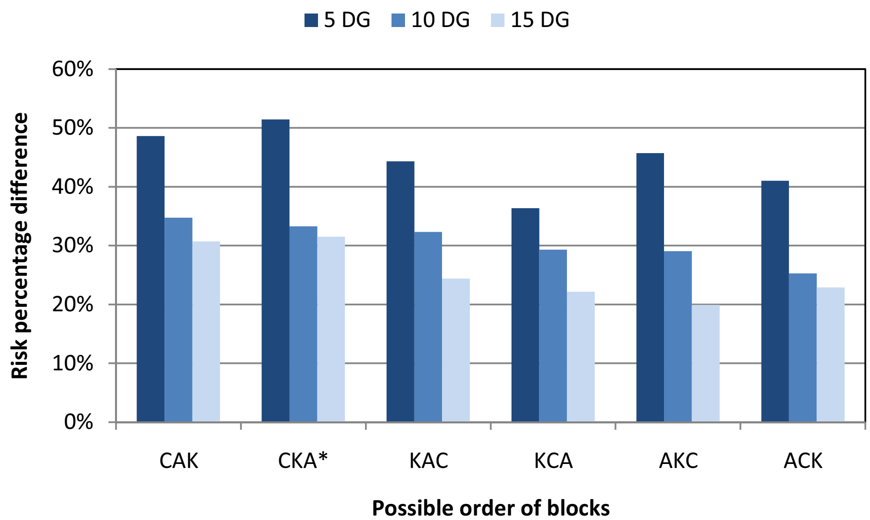

Figure 5.9: Comparison of risk percentage difference for three different DG porportion scenarios (base strategy:CKA^0: Do nothing)

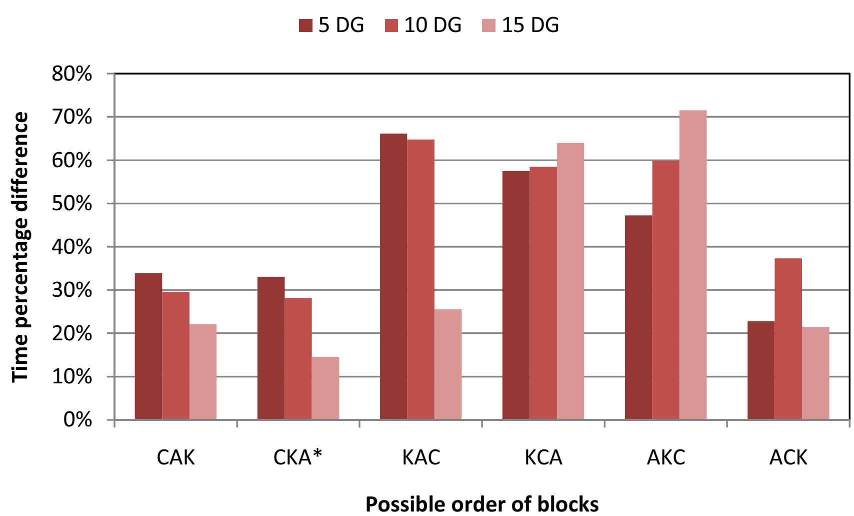

Figure 5.10: Comparison of time percentage difference for three DG porportion scenarios

Size of Buffers

As discussed in Equation 4.1, DG regulations do not permit the placement of DG cars next to an occupied locomotive. In the previous case study analysis, it is assumed that at least five buffer cars are placed after locomotive, following the US regulation. To evaluate the impact of this constraint on the results of proposed model, the Canadian regulation is evaluated, which requires only one buffer car.

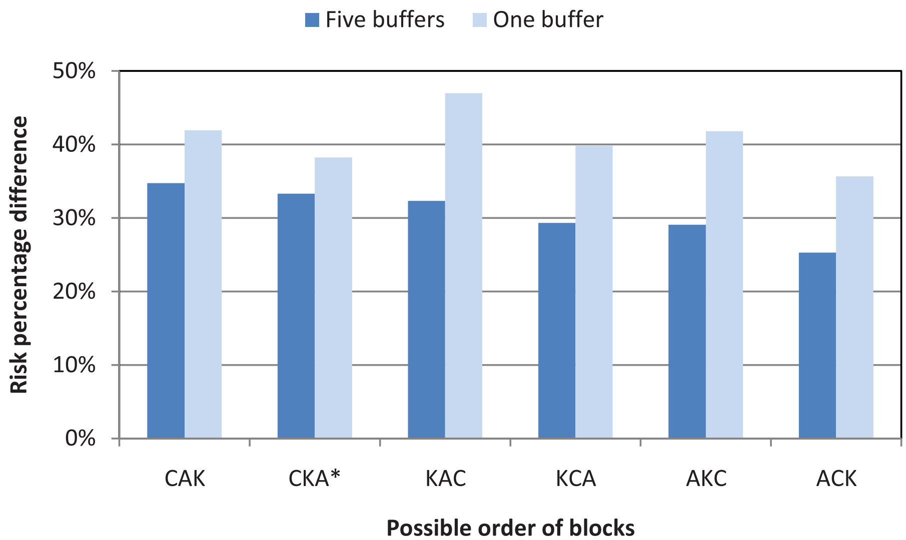

Figure 5.11: Comparison of risk percentage difference for US and Canada

According to Figures 5.11, 5.12, KAC would be a better sequence of blocks to choose if there is only one buffer. Note that, under the constraint of five buffers, the best risk reduction strategy was CAK.

Route Attributes

The route attributes affect the results of the proposed model. To evaluate the impact, the track class is selected for further investigation. The details of route attributes for the case study is given in Table 5.3. Two other scenarios are generated by changing the track class of each segment to one level higher or lower from the base scenario.

The interpretation of this result (Figure 5.13) is not easy since, for the higher track class, the risk of derailment is supposed to be less. However, the higher track class means the higher speed and subsequently higher derailment probability. This complexity can

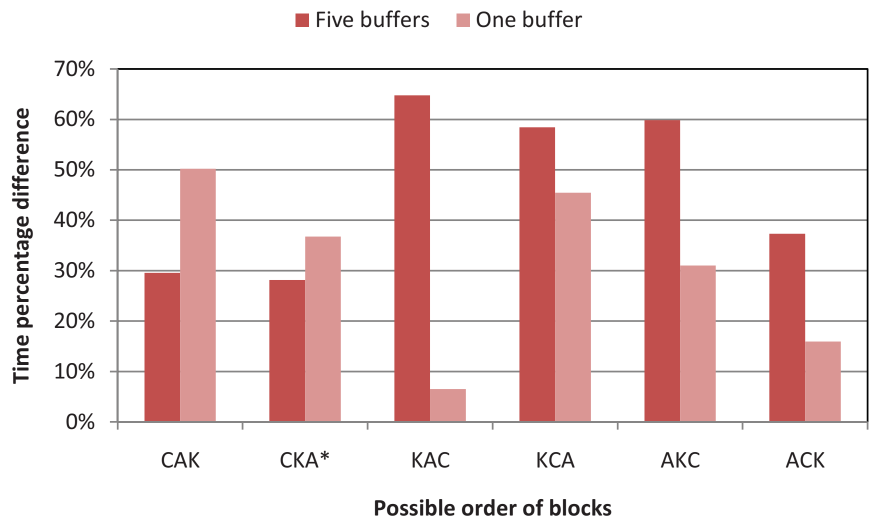

Figure 5.12: Comparison of time percentage difference for US and Canada

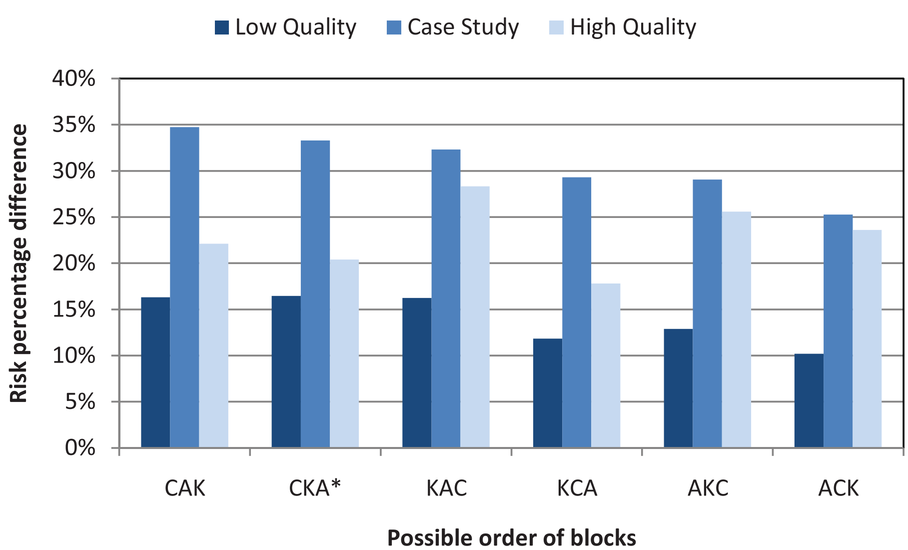

Figure 5.13: Comparison of time percentage difference between US and Canada regulations

explain the mixed pattern shown in the Figure 5.13. Nevertheless, the potential of the proposed risk mitigation strategy persists across all track classes.

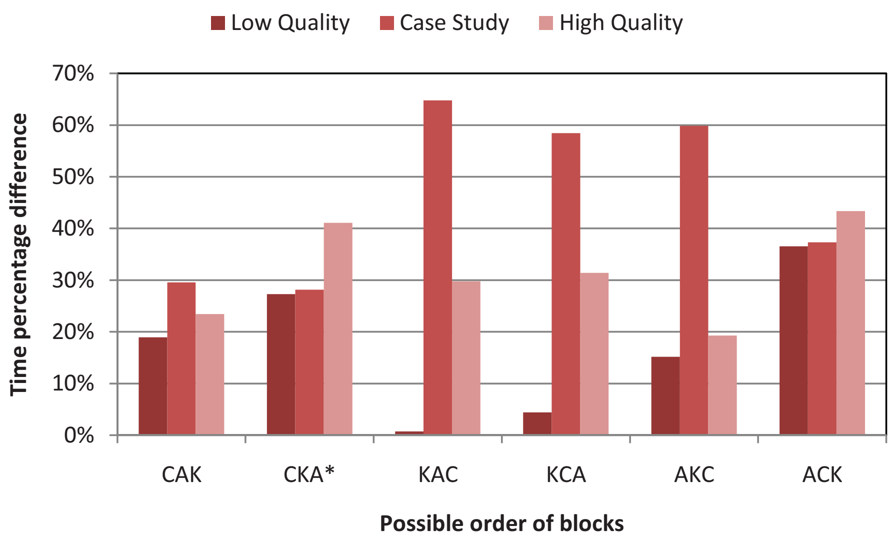

Figure 5.14: Comparison of time percentage difference for US and Canada

The risk associated with DG cars is sensitive to different factors, including the proportion of DG cars, size of buffers, and track geometry attributes. As the number of DG cars increase in the train consist, the potential to reduce the risk decreases. Similarly, as the size of buffer behind the rail crew and first DG car increases, the potential to reduce the risk decreases. These two results are likely due to the restriction of the number of available slots for placement of DG cars. The relationship between route attributes (track class) and risk of derailment is not straightforward because of presence of train speed (confounding factor).

In this chapter, the proposed model is applied to a railway corridor to illustrate how risks along a route can be used to develop cost-effective DG placement strategies at the rail yard. The results indicate that current first-come first-serve marshaling strategies potentially produce risks that may be significantly higher than the minimum risk DG placement strategy for the particular corridor under consideration. In addition, the proposed result has been compared with results from other suggested strategies. A sensitivity analysis has been carried out to determine how changes in proportion of DG, buffer size, and route attributes affect the model.

Chapter 6

Conclusions

This Chapter highlights the main conclusions and contributions of this thesis research and presents directions for future work. Two basic types of conclusions are attained from this work, academic and practical.

From an academic perspective, this research provides a unique and systematic risk-based approach for marshalling DG cars in a train consist. The approach considers both the probabilities of cars derailing en route by position as well as the time that is required to change the position of cars at the rail yard.

Current regulations do not formally consider the risks of derailment by position. Hence, there is chance that certain DG cars are located at high-risk positions along each train. While current regulations do consider the incompatibility of materials located near each other, they do not provide a systematic method to position DG cars in terms of their potential for derailment and operating costs.

This research presents a unique opportunity to investigate a very important safety problem in the rail transport sector. Based on an extensive literature review, previous work has failed to consider the risk of DG involvement in marshalling trains carrying different types of DG. It is expected that this work will result in a tool that can assist rail yard operation managers to achieve an optimum trade-off between derailment risk and operations costs in assembling trains.

Based on this research, a number of major conclusions are thus made :

- This study demonstrates that it is possible to develop a model that can suggest risk minimization strategy for DG car placements along a train consist.
- The development of such a model requires linking derailment profiles, route attributes, and rolling stock characteristics. The FRA data used in this research to examine and validate this relationship.

- Risk of derailment by position is found to be affected by speed, train length, derailment cause and track class.
- To determine the probability of derailment by position, it is necessary to include a complex chain of elements that include the train derailment, the cause of derailment, the point of derailment, number of cars derailing, placement of DG cars, and sequence of blocks.
- The effective marshalling of DG cars depends not only on the sequence of DG cars within each block but also on the position of the blocks.
- The implementation of the proposed strategy can lead to the significant reduction in risk for shipments of DG cars along a rail corridor.
- The benefit of reducing risk with the proposed marshalling strategy is accompanied by increasing rail yard marshalling time. However, this additional time is modest compared to the risk reduction. Therefore, this approach is both practical and cost effective.
- Current DG marshalling strategy may lead to solutions with significantly higher risk as compared to the proposed strategy. This includes one, five buffers between an occupied locomotive and the first DG car in Canada and the US, respectively.
- The risk associated with DG cars is sensitive to different factors, including the proportion of DG cars, size of buffers, and track geometry attributes. As the number of DG cars increase in the train consist, the potential to reduce the risk is decreased. Similarly, as the size of buffer between the rail crew and first DG car increases, the potential to reduce the risk is decreased. These two results are likely due to the restriction of the number of available slots for the placement of DG cars.

6.1 Major Contributions

Four important contributions are emphasized : 1) the development of a new point of derailment model that takes into account train length and primary cause of derailment, 2) the development of a new algorithm to calculate the yard operation time that takes into account the number of switches and number of cars marshaled, 3) the development of an optimization model which minimize the total risk considering DG marshalling constraint and yard operation time, and 4) the comprehensive analysis and evaluation of various marshalling strategies.

6.1.1 Development of Point of Derailment Model

This study introduces a new procedure for obtaining the probability of POD. A nonparametric statistical test was applied to determine whether train length and cause of derailment affect POD. The test provides some basic statistical evidence that the probability of POD along a given route segment depends on train length and the primary cause of derailment. Then the best-fit distributions are obtained for all train length/cause combinations or groupings. The uncertainty associated with distribution estimates explained by train length membership function through fuzzy logic approach.

Finally, the probabilities for each position at which the train derailment begins can be obtained by combinaing the probabilities estimates and membership functions.

6.1.2 Development of Algorithm to Calculate Rail Yard Operation Time

It is important from a practical point of view to understand the cost and efficiency implications of different marshalling strategies as the safest strategy may not be justified if train assembly costs are prohibitively high. The classification operation time is one of the main factors that influence marshalling. A marshalling solution that minimizes risk could mean additional costs as the rail crew have to cut off and couple cars to set a specific train consist. To consider this issue, this study developes an algorithm to calculate the time associated with additional tasks at rail yard, considering one empty track. Obviously, the time would be reduced in case of availability two or more empty track for the operation. This algorithm calculates the number of switching and number of cars marshalling for each alternative. The total time can be obtained by calculating the time required for each individual switching and movement of cars back and forth to the empty track.

6.1.3 Development Risk Based Model for Sterategic Marshalling DG Cars

This research introduces a risk-based model for considering placement of DG railway cars along a train. The model makes use of derailment probabilities for different railway car positions along the train. These probabilities are affected by the speed and length of the train and the causes of derailment for given track segments. The causes of derailment depend on track and train operating characteristics. This study presents a procedure for obtaining the probability of derailment by position for a given derailment cause. The model estimates the overall risks of different DG railway car marshaling strategies subject to destination block constraints.

The research problem is formulated as a non-linear integer programming problem. The resulting formula is not as a closed form as the risk term, and operational time can only be evaluated for a given solution. A heuristic genetic algorithm is used to obtain near optimal results.

6.1.4 Comprehensive Analysis and Evaluation of Various Marshalling Strategies

An application of the model to a rail corridor is presented. The main restrictions have been addressed in the model for the case study. According to regulations, the incompatible DG cars must be separated by at least one buffer car. Moreover, the regulation asked for five buffers between the locomotive and the first DG car. Furthermore, the operating time associated with marshaling and switching was considered in placement optimization to identify truly optimal solutions.

The results indicate that current first-come, first-serve marshaling strategies potentially produce risks that may be significantly higher than the minimum risk DG placement strategy for the particular corridor under consideration.

Finally, it should be noted that a number of assumptions are introduced in the model. For example, all DG cars are assumed to impose the same level of hazard, while, in reality, different DG could result in significantly different damages. Moreover, in current operation, empty cars are assigned to the rear end of train and opposite for loaded cars to consider the mechanical forces among cars. The underlying issue is not addressed in this study. Future research should therefore investigate the implications of these assumptions and develop improved models to address more realistic operating conditions.

6.2 Future Research

Train Derailment Cause Classification Model

It is recommended to conduct a study to investigate several approaches for classifying freight train derailment causes using rolling stock attributes and route characteristics. Combining the results from classification and log-linear analyses will improve understanding of causal relationships. Predicting the cause responsible for an accident at a given track segment and for a given train can help to estimate the probability of derailment.

It would also be valuable to investigate classification approaches (including Adaboost, LDA, Logistic, ANN, Random Forrest, SVM, Bagging, and KNN) and evaluate their misclassification rates (Training Error and Testing Error) to identify the best classifier to perform the cause prediction of a given train and track segment. This issue is covered in the study by Bagheri et al. [14].

Modeling Point of Derailment

Concerning the treatment of train lengths in POD modelling, a number of studies express POD in normalized form (NPOD) to address an over-representation of front-of-train POD in the database. While the NPOD accounts for relative train length, it fails to reflect dynamic forces acting on the train with respect to POD that cause car-track instability and derailment. In this study, to account for limitations in NPOD, FRA train derailments are classified into short, medium, and long trains, and they are analyzed separately with respect to their NPOD. To avoid using NPOD, it is recommended to develop a model that explicitly includes train length as a predictor. This issue will be covered in the study by Chenouri and Bagheri [25].

Modelling Number of Cars Derailed

Modeling the number of cars derailed using a truncated geometric distribution does not seem to be realistic. Two assumptions are behind gemoteric distribution : indepency of each individual event and equality of probability for each event. In reality, each car derailment depends on derailing the car ahaead. In addition, the probability of derailment for cars after POD decreases as derailment propagates due to dissipating kinetic energy. These two concerns will be addressed in the study by Chenouri and Bagheri [24].

In-transit Risk Consequence Modelling

DG risk in this research refers to a potential derailment of cars carrying some type of DG along a given route or route segment. Subsequent events such as releases, fires, explosions are not considered in this analysis. Furthermore, it is assumed that the effect on total risk resulting from the interaction of incompatible DG materials derailing in proximity to one another can be ignored. Therefore, such that all derailing DG cars are treated equally in terms of the potential threat they pose to population and property.

Future research should examine the implementations of all these assumptions and if necessary extend the model to consider the subsequent events and their associated costs. It would highlight the importance of DG cars derailments and justify the costs (time) associated with additional switching and marshalling cars at classification tracks. In this case, the in-transit risk and additional cost would have the similar measuring unit (for instance, US dollar) and can be minimized together as part of the objective function.

Modeling an Effective Buffering DG Cars

The placement of DG cars carrying different types of materials (compatible and incompatible) does not formally consider all possible combinations and their associated risk. Hence, if a derailment occurs, incompatible materials may be involved in the same derailment, so they may compound the consequences.

In a future study, the types of DG materials and DG classes could be inputs to the model. Pervious studies such as the chemical incompatibility guide by the American Society for Testing and Materials (ASTM) [2] which provides the chart that illustrates the consequence of mixing different types of DG, such as heat generation, fire, explosion, etc can be used for considering this issue. The model will assign different weights to different types of DG based on their hazardous consequences. For example, if the consequence of mixing two DG cars is explosion, they should be separated with a greater buffer compared to another case that results in only heat generation. In addition, a consequence-based ranking approach, such as Thompson et al. [57], can help compare the combination of incompatible DG and determine the worst case. This extension will investigate the effects of buffer size on safety benefits and costs.

Train Derailment Mechanics

In current operations, railroad companies allocate loaded cars to the front of the train and the empty cars for the end. This practice shows the important role of mechanical forces in train derailment. Currently, the TRAM program originated in Canada simulates the mechanical forces for a given train consist. This program can be integrated with the proposed model to position the DG cars in position which has less derailment likelihood.

APPENDICES

Appendix A

Estimation of Corridor Risks

Consider a railway corridor consisting of a rail yard, three stations and six segments, as shown in Figure A.1. Trains originate at a rail yard and are destined for Stations 1, 2 and 3. At each destination station, a block of cars is set off from the train. Therefore, each train consists of three blocks, labelled 1, 2, and 3, with Block 1 set off at Station 1, Block 2 at Station 2, and finally Block 3 at the final destination - Station 3.

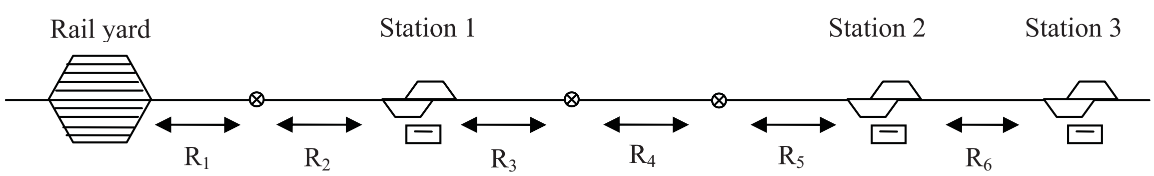

Figure A.1: A rail corridor with three destinations and six segments

As mentioned, the objective of the marshaling operations is to minimize the total risk (R), such that:

$$R = R^1 + \cdots + R^6$$

where $R^1, \ldots, R^6$ is the total risk in six segments.

Block 1 only traverses segment one and two (s =1 and 2), while Block 2 traverses the first five segments (s =1,...,5). Block 3 traverses all six segments (s =1,...,6) of the corridor. This can be expressed as

$$
\begin{aligned}
R^1 &= R^{11} + R^{12} + R^{13} \\
R^2 &= R^{21} + R^{22} + R^{23} \\
R^3 &= R^{32} + R^{33} \\
R^4 &= R^{42} + R^{43} \\
R^5 &= R^{52} + R^{53} \\
R^6 &= R^{63}
\end{aligned}
$$

Thus, the above equations can be expressed as

$$R = \sum_{s=1}^{S} \sum_{b=1}^{B} R^{sb}$$

where R^{sb} is the total risk of derailment of block b in segment s. For above example R_{31}, R_{41}, R_{51}, R_{62}, and R_{63} is zero.

Total risk of block b in segment s is a summation of each car risk in the block. For instance, for n_b cars in block b, the total risk over segment s is expressed as

$$R^{sb} = \sum_{i=1}^{n_b} R_i^{sb} = R_1^{sb} + R_2^{sb} + \cdots + R_{n_b}^{sb}$$

where R_i^{sb} is the risk for the car at position i. As mentioned in the previous section, the risk associated with the car at position i can be calculated from Equation (3.5)(R_i^{sb} in this equation is replaced by R_i).

Appendix B

Cause Code Groups

Train-mile related causes (TM) are independent of train length but depend essentially on the number of train movements traversing a given track segment. For instance, the cause “human error” is a train-miles related cause. On the other hand, causes such as “track component failures” and “equipment failures” are car-miles (CM) related causes. It should be noted that Figure B.1 is borrowed from another study [51].

Cause Group	Description	CM/TM	FRA Cause Codes										
01T	Roadbed Defects	CM	T001	T099									
02T	Non-Traffic, Weather Causes	TM	T002	T401	T402	T403							
03T	Wide Gauge	CM	T110	T111	T112	T113							
04T	Track Geometry (excl. Wide Gauge)	CM	T101	T102	T103	T104	T105	T106	T107	T108	T199		
05T	Buckled Track	CM	T109										
06T	Rail Defects at Bolted Joint	CM	T201	T211									
07T	Joint Bar Defects	CM	T213	T214	T215	T216							
08T	Broken Rails or Welds	CM	T202	T203	T204	T207	T208	T210	T212	T218	T219	T220	T221
09T	Other Rail and Joint Defects	CM	T299										
10T	Turnout Defects - Switches	CM	T307	T308	T309	T310	T311	T312	T313	T314	T315	T319	
11T	Turnout Defects - Frogs	CM	T304	T316	T317	T318							
12T	Misc. Track and Structure Defects	CM	T205	T206	T217	T222	T301	T302	T303	T305	T306	T399	T499
			S001	S002	S003	S004	S005	S006	S007	S008	S009	S010	S011
01S	Signal Failures	TM	S012	S013	S099								
01E	Air Hose Defect (Car)	CM	E00C										
02E	Brake Rigging Defect (Car)	CM	E07C										
03E	Handbrake Defects (Car)	CM	E08C	E0HC									
04E	UDE (Car or Loco)	CM	E05C	E05L									
05E	Other Brake Defect (Car)	CM	E01C	E02C	E03C	E04C	E06C	E09C					
06E	Centerplate/Carbody Defects (Car)	CM	E20C	E21C	E22C	E23C	E24C	E25C	E26C	E27C	E29C		
07E	Coupler Defects (Car)	CM	E30C	E31C	E32C	E33C	E34C	E35C	E36C	E37C	E39C		
08E	Truck Structure Defects (Car)	CM	E44C	E45C									
09E	Sidebearing, Suspension Defects (Car)	CM	E40C	E41C	E42C	E43C	E47C	E48C					
10E	Bearing Failure (Car)	CM	E52C	E53C									
11E	Other Axle/Journal Defects (Car)	CM	E51C	E54C	E55C	E59C							
12E	Broken Wheels (Car)	CM	E60C	E61C	E62C	E63C	E6AC						
13E	Other Wheel Defects (Car)	CM	E64C	E65C	E66C	E67C	E68C	E69C					
14E	TOFC/COFC Defects	CM	E11C	E12C	E13C	E19C							
			E07L	E40L	E41L	E42L	E43L	E44L	E45L	E46L	E47L	E48L	E4TL
			E49L	E51L	E52L	E53L	E54L	E55L	E59L	E60L	E61L	E62L	E63L
15E	Loco Trucks/Bearings/Wheels	CM	E64L	E65L	E66L	E67L	E68L	E6AL	E69L	E70L	E77L		
16E	Loco Electrical and Fires	CM	E71L	E72L	E73L	E74L	E76L						
			E00L	E01L	E02L	E03L	E04L	E06L	E08L	E0HL	E09L	E20L	E21L
			E22L	E23L	E24L	E25L	E26L	E27L	E29L	E30L	E31L	E32L	E33L
17E	All Other Locomotive Defects	CM	E34L	E35L	E36L	E37L	E39L	E79L	E99L				
18E	All Other Car Defects	CM	E49C	E80C	E81C	E82C	E83C	E84C	E85C	E86C	E89C	E99C	
19E	Stiff Truck (Car)	CM	E46C										
20E	Track/Train Interaction (Hunting) (Car)	CM	E4TC										
21E	Current Collection Equipment (Loco)	CM	E75L										
			H510	H511	H512	H513	H514	H515	H516	H517	H518	H519	H520
01H	Brake Operation (Main Line)	TM	H521	H525	H526								
02H	Handbrake Operations	TM	H017	H018	H019	H020	H021	H022	H025	M504			
03H	Brake Operations (Other)	TM	H008	H099									
04H	Employee Physical Condition	TM	H101	H102	H103	H104	H199						
			H201	H202	H203	H204	H205	H206	H207	H208	H209	H215	H216
05H	Failure to Obey/Display Signals	TM	H217	H299									
06H	Radio Communications Error	TM	H210	H211	H212	H405							
			H301	H302	H303	H304	H305	H306	H307	H308	H309	H310	H311
07H	Switching Rules	TM	H312	H313	H314	H315	H399						
08H	Mainline Rules	TM	H401	H402	H403	H404	H406	H499					
			H501	H502	H503	H504	H505	H506	H507	H508	H509	H522	H523
09H	Train Handling (excl. Brakes)	TM	H524	H599									
10H	Train Speed	TM	H601	H602	H603	H604	H605	H606	H699				
11H	Use of Switches	TM	H701	H702	H703	H704	H705	H799					
12H	Misc. Human Factors	TM	H821	H822	H823	H824	H899	H991	H992	H993	H994	H995	H999
01M	Obstructions	TM	M101	M102	M103	M104	M105	M199	M402	M403	M404		
02M	Grade Crossing Collisions	TM	M301	M302	M303	M304	M305	M306	M307	M399			
03M	Lading Problems	CM	M201	M202	M203	M204	M205	M206	M207	M299	M409	M410	
04M	Track-Train Interaction	CM	M405										
05M	Other Miscellaneous	TM	M401	M406	M407	M408	M501	M502	M503	M505	M599		

Figure B.1: Comparison of FRA and AAR cause code groups

Appendix C

Calibration Procedure to Estimate the Parameters

The geometric distribution gives the probability of "*x*" failures before the first success when the probability of success at each trail is a constant probability "*p*". In other words, the probability of "*x*" cars derailing before the first non-derailing car is a geometric distribution with the probability of a car to derail (given that a derailment has occurred), equal to *(1-p)*. The general equation of the geometric distribution is

$$P(X = x) = p(1-p)^{(x-1)} \qquad x = 1 \ldots \infty$$

In this case there is a need for truncating the range of the geometric distribution at a certain value before infinity. For a given derailment on a given train , the number of cars that could derail is restricted by the number of cars on the train. Assuming that only cars behind the point of derailment could derail, then the number of cars that could derail is actually restricted by the number of cars behind the point of derailment (L_r = residual length). The theoretical geometric distribution assigns probabilities for values of "x" as large as infinity and thus the equation of the geometric distribution should be modified.

Let the random variable X be the number of cars derailing in a train accident. Given the number of cars after POD (L_r), we have

$$P(X = x) = cp(1-p)^{(x-1)} \qquad x = 1,\ldots L_r$$

$$P(X = x) = cp(1-p)^{(x-1)} \qquad x = 1,\ldots L_r\,,$$

$$\sum_{x=1}^{L_r} P(X = x) = 1$$

$$\sum_{x=1}^{L_r} cp(1-p)^{(x-1)} = 1$$

$$cp\sum_{x=1}^{L_r} (1-p)^{(x-1)} = 1$$

$$if \;\; q = 1-p$$

$$cp\sum_{x=1}^{L_r} q^{(x-1)} = 1$$

$$cp[\sum_{x=1}^{\infty} q^{(x-1)} \text{ - } \sum_{x=L_r+1}^{\infty} q^{(x-1)}] = 1$$

We know $\sum_{x=0}^{\infty} k^x = \dfrac{1}{1-k}$ thus we have

$$cp[\frac{1}{1-q} - q^{L_r}(\frac{1}{1-q})] = 1$$

$$\sin ce \;\; q = 1-p$$

$$cp[\frac{1}{p} - (1-p)^{L_r}(\frac{1}{p})] = 1$$

$$cp[\frac{1}{p}(1-(1-p)^{L_r})] = 1$$

$$c[(1-(1-p)^{L_r})] = 1$$

$$c = \frac{1}{[1-(1-p)^{L_r}]}$$

To determine constant c we utilize that fact that $\sum_{x=1}^{L_r} P(X=x)=1$. This yields

$$c=\frac{1}{[1-(1-p)^{L_r}]}$$

Thus the distribution of X given L_r is

$$P(X=x)=\frac{p(1-p)^{(x-1)}}{[1-(1-p)^{L_r}]} \qquad x=1,\ldots L_r$$

To estimate parameters, we consider logit link functions of the forms:

$$\log it(p_i)=\log\frac{p_i}{1-p_i}=x_i\beta$$

where, $x_i\beta=\beta_0+(\beta_1\times Speed)+(\beta_2\times L_r)+(parameters\ for\ causes:\beta_3,\beta_4,\ldots\beta_9)$

Train causes in the above expression were entered as dummy variables (0, 1). Speed (mph) and train length were given scalar values.

The likelihood function of the model based on the sample is

$$L(\beta)=\prod_{i=1}^{n}\frac{p_i(1-p_i)^{(x_i-1)}}{[1-(1-p_i)^{L}]}$$

The log-likelihood is then given by

$$\begin{aligned}
\ell(\beta)&=\sum_{i=1}^{n}[\log(p_i(1-p_i)^{(x_i-1)}-\log(1-(1-p_i)^{L_i})]\\
&=\sum_{i=1}^{n}[\log p_i+(x_i-1)\log(1-p_i)-\log(1-(1-p_i)^{L_i})]\\
&=\sum_{i=1}^{n}[\log p_i-\log(1-p_i)+x_i\log(1-p_i)-\log(1-(1-p_i)^{L_i})]\\
&=\sum_{i=1}^{n}[\log\frac{p_i}{(1-p_i)}+x_i\log(1-p_i)-\log(1-(1-p_i)^{L_i})]\\
&=\sum_{i=1}^{n}[\log\frac{p_i}{(1-p_i)}+x_i\log(1-p_i)-\log(1-(1-p_i)^{L_i})]
\end{aligned}$$

As we know $\log it(p_i) = \log \frac{p_i}{1-p_i} = x_i \beta$

Thus we have

$$\ell(\beta) = \sum_{i=1}^{n} [x_i \beta + x_i \log(1+e^{x_i \beta})^{-1} - \log(1-(1+e^{x_i \beta})^{-L_i})]$$

$$\ell(B) = \sum_{i=1}^{n} [x_i \beta - x_i \log(1+e^{x_i \beta}) - \log(1-(1+e^{x_i \beta})^{-L_i})]$$

The above log-likelihood function can be maximized by calibrating values for the parameters in the response function $x_i \beta$ while minimizing the negative log-likelihood function. The calibration is conducted using R Software (version 2.7.2). Since the above probability function is truncated, thus R classified the log-likelihood as user-defined likelihood function. Below commands are used in R to introduce above log-likelihood function.

```
TGlogit<-function(beta){
Train<-as.data.frame(Train)
y<-Train$NDC
L<-Train$LR
n<-length(y)
intercept<-rep(1,n)
x<-cbind(intercept,Train[,c(1:8)])
x<-as.matrix(x)
l.lik.i<-x%*%beta-y*log(1+exp(x%*%beta))-log(1-
(1+exp(x%*%beta))^(-L))
l.lik<--sum(l.lik.i)
return(l.lik)
}
```

To calibrate above log-likelihood function different algorithms such as "Nelder-Mead", "BFGS", "CG", "L-BFGS-B", "SANN" were applied with different initials. After substantial try and error most of the solutions are converged to specific value.

The summary statistics for the maximum likelihood calibration exercise are given in Table 3.15 for the 1997-2006 FRA accident data base.

Mean Number of Cars Derailed

The below general formula has been used to obtain the mean number of cars derailing

$$P(X=x)=\begin{cases} \dfrac{p(1-p)^{(x-1)}}{[1-(1-p)^{L_r}]} & if \quad x=1,...L_r \\ \\ 0 & o.w \end{cases}$$

$$E[X]=\sum_{x=1}^{L_r} x.P(X=x)=\sum_{x=1}^{L_r} x.\ \frac{p(1-p)^{(x-1)}}{[1-(1-p)^{L_r}]} \ =\frac{p}{[1-(1-p)^{L_r}]}\sum_{x=1}^{L_r} x.(1-p)^{(x-1)}$$

If $q=1-p$

$$E[X]=\frac{1-q}{[1-q^{L_r}]}\sum_{x=1}^{L_r} x.q^{(x-1)} \ =\frac{1-q}{[1-q^{L_r}]}\sum_{x=1}^{L_r}(\frac{\partial}{\partial q}q^x)=\ \frac{1-q}{[1-q^{L_r}]}.\frac{\partial}{\partial q}[\sum_{x=1}^{L_r}q^x]$$

We know $\frac{\partial}{\partial q}[\sum_{x=1}^{L_r}q^x] \ =\frac{\partial}{\partial q}[\sum_{x=0}^{L_r}q^x]$ thus we have

$$E[X]=\frac{1-q}{[1-q^{L_r}]}.\frac{\partial}{\partial q}[\sum_{x=0}^{L_r}q^x] \ =\frac{1-q}{[1-q^{L_r}]}.\frac{\partial}{\partial q}[\sum_{x=0}^{\infty}q^x-\sum_{x=L_r+1}^{\infty}q^x]$$

We know $\sum_{x=0}^{\infty}k^x=\frac{1}{1-k}$ thus we have

$$E[X]=\frac{1-q}{[1-q^{L_r}]}.\frac{\partial}{\partial q}[\frac{1}{1-q}-q^{L_r+1}.(\frac{1}{1-q})]=\frac{1-q}{[1-q^{L_r}]}.[\frac{1}{(1-q)^2}-(\frac{(L_r+1)q^{L_r}}{1-q}+\frac{q^{L_r+1}}{(1-q)^2})]$$

We know $q=1-p$

$$E[X]=\frac{p}{[1-(1-p)^{L_r}]}.[\frac{1}{p^2}-(\frac{(L_r+1)(1-p)^{L_r}}{p}+\frac{(1-p)^{L_r+1}}{p^2})]$$

$$E[X]=\frac{1}{p}-\frac{L_r(1-P)^{L_r}}{1-(1-p)^{L_r}}$$

A recent investigation [11] resulted in the similar model for the mean number of cars derailing (M_{nd}).

Appendix D

Dangerous Goods Classification

Canadian Transport Dangerous Goods (TDG) classified DG to 9 classes:

Figure D.1: Class 1 explosives

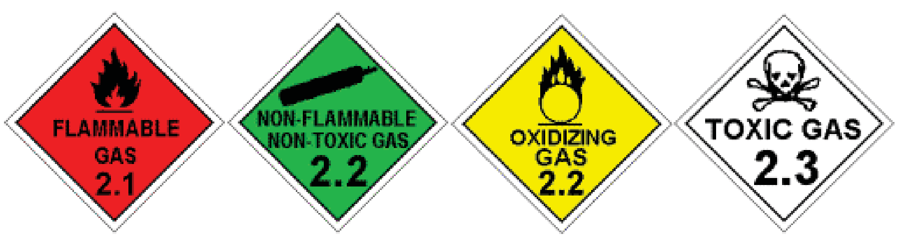

Figure D.2: Class 2 gases

Figure D.3: Class 3 flammable liquids

Figure D.4: Class 4 flammable solids, spontaneously combustibles and substances that, on contact with water, emit flammable gases

Figure D.5: Class 5 oxidizing substances and organic peroxides

Figure D.6: Class 6 poisonous (toxic) and infectious substances

Figure D.7: Class 7 radioactive materials

Figure D.8: Class 8 corrosives

Figure D.9: Class 9 miscellaneous products or substances

Appendix E

Time Calculation Algorithm

For this algorithm, the given and alternative combinations are assumed. This algorithm has been coded with Visual Basic.NET. As can be seen in Figure E.1, the algorithm includes four functions. The given and alternative combinations have been defined as two arrays in addition to Temp as a temporary array. Two outputs, c and l, will be used to estimate the switching and marshalling times.

```
Count ()
        for i =1 to n-1
                if G[i] ≠ A[i]
                c = c + 4
                Fix (i)
        next i

Fix(i)
        for j = i +1 to n
                if G[j] = A[i]
                Swap (i,j)
                end if
        next j

Swap (i,j)
        Move(G, i, n, Temp, 1)
        Move (Temp, j-i+1, n-i+1,G, i)
        Move (Temp, 1, j-I, G, i+n-j+1)

Move (A, p, q, B, r)
        for k = p to q
        B[r + k-p] =A[k]
        next k
        l=l+ (q-p+1)
```

Figure E.1: The algorithm to calculate the time required for changing car positions

Appendix F

Matlab Code

F.1 Matlab Code for Calculating Probability of Derailment by Position

```
clear all;
clc;

%%% This Program is for P(i|TD)%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%This part of program related to Cause disteribution (CPM)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

constant=1.026;
cause=[0.932 -0.413 0];
TL= [-2.033 1.004 0];
AREA=[1.816 0];
TRK=[-2.133 0];
cause_TL = [1.897 0.6 0; 2.312 1.117 0; 0 0 0];
cause_AREA = [0.625 0; 0.778 0; 0 0];
cause_TRK = [0.905 0; 0.65 0; 0 0];
TL_TRK=[1.891 0; 0.576 0; 0 0];
AREA_TRK=[0.843 0; 0 0];

NR=2230;
p=0;
```

```
z=1;

for l = 1:2
    for k = 1:2
        for j = 1:3
            for i = 1:3
                p=(exp((constant + cause(i)+ TL(j)+AREA(k)+TRK(l)+ cause_TL(i,j)+
                cause_AREA(i,k) + cause_TRK(i,l)+ TL_TRK(j,l)+ AREA_TRK(k,l))))/NR;
                a(z)=p;
                z=z+1;

            end
        end
    end
end

y=1;
for y=0:11
   b(y+1)=(a(1+(3*y))+a(2+(3*y))+a(3+(3*y)));
end

x=1;
for w=1:12
  for v=1:3
  c(x)=a(x)/b(w);
  x=x+1;
   end
end
%%%%%%%%%%%

t_t=input('Train type (short=1, medium=2, long=3)=');
r_a=input('Area (flat=1, mountain=2)=');
t_c=input('Track type(poor=1,good=2)=');


if  t_t ==1 & r_a==1 & t_c==1
   W1=c(1);W2=c(2);W3=c(3);
elseif t_t ==2 & r_a==1 & t_c==1
    W1=c(4);W2=c(5);W3=c(6);
elseif t_t ==3 & r_a==1 & t_c==1
```

```
    W1=c(7);W2=c(8);W3=c(9);
elseif t_t ==1 & r_a==2 & t_c==1
    W1=c(10);W2=c(11);W3=c(12);
elseif t_t ==2 & r_a==2 & t_c==1
    W1=c(13);W2=c(14);W3=c(15);
elseif t_t ==3 & r_a==2 & t_c==1
    W1=c(16);W2=c(17);W3=c(18);
elseif t_t ==1 & r_a==1 & t_c==2
    W1=c(19);W2=c(20);W3=c(21);
elseif t_t ==2 & r_a==1 & t_c==2
    W1=c(22);W2=c(23);W3=c(24);
elseif t_t ==3 & r_a==1 & t_c==2
    W1=c(25);W2=c(26);W3=c(27);
elseif t_t ==1 & r_a==2 & t_c==2
    W1=c(28);W2=c(29);W3=c(30);
elseif t_t ==2 & r_a==2 & t_c==2
    W1=c(31);W2=c(32);W3=c(33);
else
    W1=c(34);W2=c(35);W3=c(36);
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%This part of program related to Probability of train derailment (PTD)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

dis=input('Distance=');
t_l=input('Train length=');
t_c2=input('Track class(1-5)=');
RT=[12.135e-6 1.515e-6 0.51e-6 0.1325e-6 0.08e-6];
RC=[540.075e-9 69.525e-9 23.625e-9 5.85e-9 3.675e-9];
P_TD=1-exp (-dis*((RC(t_c2)*t_l)+RT(t_c2)));

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%This part of program related to Probability of Point of derailment (POD)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

if t_l<=25
    m_1=1;
    m_2=0;
end
```

```
if t_l==75
    m_1=0;
    m_2=1;
end
if t_l>=125
    m_1=0;
    m_2=1;
end
if 25<t_l & t_l<75
        m_1=(-0.02*t_l)+1.5;
        m_2=(0.02*t_l)-0.5;
end
if 75<t_l & t_l<125
       m_1=(-0.02*t_l)+2.5;
       m_2=(0.02*t_l)-1.5;
end

%%%%%%%%%%%%%%%%%%%%%%%%

aU=0.03;
bU=1;
out_U = U(aU,bU,t_l);
out_U_Sh_G1=out_U;

aT=-0.094;
bT=1;
cT=1;
out_T = T(aT,bT,cT,t_l);
out_T_Sh_G2=out_T;

aU=0.031;
bU=1.005;
out_U = U(aU,bU,t_l);
out_U_Sh_G3=out_U;

Short_Matx=[];
for i=1:t_l
Short_Matx(i, :)= [out_U_Sh_G1(i) out_T_Sh_G2(i) out_U_Sh_G3(i)];
end
```

```
%%%%%%%%%%%%%%%%%%%%%%%%

aB=0.575;
bB=0.6579;
out_B = B(aB,bB,t_l);
out_B_Me_G1=out_B;

aB=0.782;
bB=0.504;
out_B = B(aB,bB,t_l);
out_B_Me_G2=out_B;

aU=0.008;
bU=1;
out_U = U(aU,bU,t_l);
out_U_Me_G3=out_U;

Medium_Matx=[];
for i=1:t_l
Medium_Matx(i, :)= [out_B_Me_G1(i) out_B_Me_G2(i) out_U_Me_G3(i)];
end

%%%%%%%%%%%%%%%%%%%%%

aB=0.602;
bB=0.745;
out_B = B(aB,bB,t_l);
out_B_Lo_G1=out_B;

aB=0.646;
bB=0.59;
out_B = B(aB,bB,t_l);
out_B_Lo_G2=out_B;

aB=0.763;
bB=0.799;
out_B = B(aB,bB,t_l);
out_B_Lo_G3=out_B;

Long_Matx=[];
```

```
for i=1:t_l
Long_Matx(i, :)= [out_B_Lo_G1(i) out_B_Lo_G2(i) out_B_Lo_G3(i)];
end

%%%%%%%%%%%%%%%%%%%

if t_l<75
A=Short_Matx;
B=Medium_Matx;
end

if t_l>=75
A=Medium_Matx;
B=Long_Matx;
end

for i=1:size(A,1)
    for j=1:size(A,2)
        C(i,j)=(A(i,j)*m_1)+(B(i,j)*m_2);
    end
end

for i=1:size(C,1)
    D(i)=C(i,1)*W1 + C(i,2)*W2 + C(i,3)*W3;
end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%This part of program related to Probability of No. of cars derailing (PN)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

t_s=input('Train Speed=');

B_Int=-2.013;
%B_Lr=0.001;
B_Sp=-0.032;
B_Rb=0.419;
B_Tg=0.171;
%B_Sw=0.715;
B_Gc=0.841;
%B_Aw=1.108;
```

```
%B_Ao=0.444;
z_G1=B_Int+(B_Sp*t_s)+B_Rb;
z_G2=B_Int+(B_Sp*t_s)+B_Tg;
z_G3=B_Int+(B_Sp*t_s)+B_Gc;
pN_G1=((exp(z_G1))/(1+exp(z_G1)));
pN_G2=((exp(z_G2))/(1+exp(z_G2)));
pN_G3=((exp(z_G3))/(1+exp(z_G3)));

p= (W1*pN_G1) + (W2*pN_G2) + (W3*pN_G3);

q={};
for ic = 1:t_l
    for ir = 1:t_l - (ic - 1)
         q{ic}(ir)=(p * ((1 - p) ^ ir)) / (1 - ((1 - p) ^ (t_l - ic + 1)));
    end
end
%%%%%%%%%%%%%%%%%%%%%
for i = 1:t_l
tt = 0;
pp = 0;
total = 0;

    for j = 1:i
        for x = i - j + 1:t_l - j + 1
            total = total + q{j}(x);
        end
        pp = D(j) * total;
        tt = tt + pp;
        pp = 0;
        total = 0;
    end
     P_final(i) = tt;
end
%%%%%%%%%%%%%%%%%%%
for i=1:t_l
    P(i)=P_TD*P_final(i);
end
%%%%%%%%%%%%%%%%%%
```

Appendix G

Visual Basic Code

G.1 Visual Basic Code for Calculating Time

The following code was integrated with Evolver Software Package. The first part of the code solves the illegal combinations (infeasible route in TSP problem) that Genetic Algorithm (GA) produces for each iteration. The second part of the code is to calculate the time required for switching and marshalling to make-up a train based on alternative combination of cars.

```
Public n As Integer
Public c As Integer
Public l As Integer
Public g As Integer
Public z As Integer  ' output list index

Sub Macro1()
'
' Macro1 Macro
' Macro recorded 22/05/2009 by Admin
'
Range(Cells(1, 2), Cells(400, 2)).Value = Range(Cells(1, 1), Cells(400, 1)).Value

For ii = 1 To 23 Step 2
aa = Cells(Cells(ii, 55), 2)
bb = Cells(Cells(ii + 1, 55), 2)
Cells(Cells(ii + 1, 55), 2) = aa
```

```
Cells(Cells(ii, 55), 2) = bb
Next
For ii = 26 To 54 Step 2
aa = Cells(Cells(ii, 55), 2)
bb = Cells(Cells(ii + 1, 55), 2)
Cells(Cells(ii + 1, 55), 2) = aa
Cells(Cells(ii, 55), 2) = bb
Next
For ii = 56 To 108 Step 2
aa = Cells(Cells(ii, 55), 2)
bb = Cells(Cells(ii + 1, 55), 2)
Cells(Cells(ii + 1, 55), 2) = aa
Cells(Cells(ii, 55), 2) = bb
Next

ThisWorkbook.Sheets("Processing").Select

Dim i As Integer
Dim s As Integer
Dim t As Integer

s = ThisWorkbook.Sheets("Input").Cells(13, 3).Value

g = 0
z = 1
n = 0

Range(Cells(1, 3), Cells(400, 7)).ClearContents

For t = 1 To s

n = ThisWorkbook.Sheets("Input").Cells(12 + t, 2).Value
Range(Cells(1, 7), Cells(400, 7)).Value = Range(Cells(1, 1), Cells(400, 1)).Value

c = 0
l = 0

For i = 1 To n - 1
    If Cells(i + g, 7).Value <> Cells(i + g, 2).Value Then
        Call fix(i)
```

```
        c = c + 4
    End If
Next i

Cells((2 * t - 1), 4).Value = c
Cells((2 * t), 4).Value = l
Cells((2 * t - 1), 5).Value = c * 1
Cells((2 * t), 5).Value = l * 0.2

Cells(z + g, 6).Value = "END"

g = n + g
Next t
End Sub

Public Sub fix(i)

    For j = i + 1 To n
        If Cells(j + g, 7).Value = Cells(i + g, 2).Value Then
            Call swap(i + g, j + g)
        End If
    Next j

End Sub

Public Sub swap(i, j)
    Call move(7, i, n + g, 3, 1)
    Call move(3, j - i + 1, n + g - i + 1, 7, i)
    Call move(3, 1, j - i, 7, i + n + g - j + 1)
End Sub

Public Sub move(a, p, q, b, r)
    For k = p To q
        Cells(r + k - p, b).Value = Cells(k, a).Value
    Next k
    l = l + (q - p + 1)
    Cells(z + g, 6).Value = Str(a) + "[" + Str(p) + "," + Str(q) + "] => " +
    Str(b) + "[" + Str(r) + "," + Str(r + q - p) + "]"
    z = z + 1
End Sub
```

Appendix H

Glossary

AAR
Association of American Railroads.

CTC
Canadian Transport Commission.

DG
Dangerous Goods.

DOT
Department of Transportation (US).

FRA
Federal Railroad Administration (US).

MDG
Marshalling of Dangerous Goods.

NTSB
National Transportation Safety Board (US).

POD
Point of Derailment.

RODS
Railway Occurrence Data System (CANADA).

TC
Transport Canada.

TDG

Dangerous Goods.

TSB

Transportation Safety Board of Canada.

Bibliography

[1] Palisade corporation, @risk package. http://www.palisade.com. Last accessed on October 31, 2007. 36

[2] Proposed guide for estimating the incompatibility selected wastes based on binary chemical reactions. Technical report, American Society for Testing Materials (ASTM), D-34, Proposal P 168, 1987. 79

[3] Commodity flow survey (cfs),research and innovative technology administration (rita), u.s. department of transportation (us dot). http://www.bts.gov/programs/commodity_flow_survey/, 2002. Last accessed on November 06, 2009. 58

[4] Freight analysis framework (faf), fhwa office of freight management and operations. http://ops.fhwa.dot.gov/freight/freight_analysis/faf/faf2_com.htm, 2002. Last accessed on November 06, 2009. 58

[5] Fra guide for preparing accident/incident reports. Technical report, Federal Railroad Administration, Office of Safety, 2003. ix, 21

[6] Southwest chief map, los angeles to chicago route map. http://www.vacationsbyrail.com/usa/luxury_rail/grandluxe_limited/los_angeles_chicago/index.html, 2004. Last accessed on November 06, 2009. xii, 58

[7] Risk assessment for the transportation of hazardous materials by rail. Technical report, Arthur D. Little, Inc., April 1996. 21

[8] Safe placement of train cars: A report. Technical report, Federal Railroad Administration, June 2005. 50

[9] Federal Railroad Administration. Accident data on demand. http://safetydata.fra.dot.gov/officeofsafety, 1997-2006. Last accessed on October 31, 2007. xi, 1, 2, 3, 23, 33

[10] R.K. Ahuja, K.C. Jha, and J. Liu. Solving real-life railroad blocking problems. *Interfaces*, 37(5):404, 2007. 43

[11] R.T. Anderson and C.P.L. Barkan. Railroad accident rates for use in transportation risk analysis. *Transportation Research Record: Journal of the Transportation Research Board*, 1863(1):88–98, 2004. ix, 7, 20, 21

[12] RT Anderson and CPL Barkan. Derailment probability analysis and modeling of mainline freight trains. In *Proceedings of the 8th International Heavy Haul Conference*, pages 491–497, 2005. xi, 31, 32, 39

[13] M. Bagheri. Risk analysis of stationary dangerous goods railway cars: a case study. *Journal of Transportation Security*, 2(3):77–89, 2009. 7

[14] M. Bagheri, K. Blass, S. Chenouri, F. F. Saccomanno, and L. Fu. Cause prediction analysis for railroad derailments. *for submission to Journal of Rail and Rapid Transit.* 78

[15] M. Bagheri, F.F. Saccomanno, S. Chenouri, and L. Fu. Reducing the risk of dangerous goods car derailments through effective positioning along the train. *Journal of Accident Analysis & Prevention*, (AAP-D-09-00209), Submitted on 18 May, 2009. 25

[16] M. Bagheri, F.F. Saccomanno, and L. Fu. Dangerous goods railway car placement model. In *9th International Heavy Haul Conference in Shanghai*, pages 863–871, 2009. 23

[17] M. Bagheri, F.F. Saccomanno, and L. Fu. Risk based model for marshalling dangerous goods railway cars in rail yards. In *Proceedings of the 88th Annual Meeting of Transportation Research Board (TRB) in Washington DC*, 2009. 23

[18] M. Bagheri, F.F. Saccomanno, and L. Fu. Effective placement of dangerous goods cars in rail yard marshaling operation. *Canadian Journal of Civil Engineering*, (09-159), Accepted on 02 October, 2009.

[19] C.P.L. Barkan, C. Tyler Dick, and R. Anderson. Railroad derailment factors affecting hazardous materials transportation risk. *Transportation Research Record: Journal of the Transportation Research Board*, 1825(-1):64–74, 2003. 22, 23, 28

[20] C. Barnhart, H. Jin, and P.H. Vance. Railroad blocking: A network design application. *Operations Research*, 48(4):603–614, 2000. 43

[21] National Transportation Safety Board. Railroad accident report, pb2005-916304. `http://www.ntsb.gov/publictn/2005/RAR0504.pdf`, 2005. Last accessed on October 31, 2007. 2

[22] Transport Canada. Transportation of dangerous goods regulations, part 10.6 "location of placarded railway vehicles in a train". http://www.tc.gc.ca/tdg/clear/part10.htm#sec106. Last accessed on October 31, 2007. xi, 7, 12, 43, 45

[23] Transport Canada. Transport dangerous goods newsletter. http://www.tc.gc.ca/tdg/newsletter/spring2007.pdf, Spring 2007. Last accessed on October 31, 2007. 2

[24] S. Chenouri and M. Bagheri. Derailment probability analysis using markov chain. *for submission to Transportation Research Part B*. 78

[25] S. Chenouri and M. Bagheri. Estimating point of derailment probability by position. *for submission to ASCE Journal of Transportation Engineering*. 78

[26] R. Christensen. *Log-linear models and logistic regression*. Springer Verlag, 1997. 27

[27] V.T. Covello and M.W. Merkhofer. *Risk assessment methods: approaches for assessing health and environmental risks*. Plenum Pub Corp, 1993. 17

[28] C.F. Daganzo, R.G. Dowling, and R.W. Hall. Railroad Classification Yard Throughput:The Case of Multistage Triangular Sorting. *Transportation Research-Part A*, 17A(2):95–106, 1983. 4

[29] J.A. Droege. *Freight Terminals and Trains*. McGraw-Hill, New York, 1925. 4

[30] S.M. El-Hage. Evaluating train marshaling regulations for railcars carrying dangerous commodities. Master's thesis, University of Waterloo, 1988. 22, 31, 33, 34

[31] E. Elbeltagi, T. Hegazy, and D. Grierson. Comparison among five evolutionary-based optimization algorithms. *Advanced Engineering Informatics*, 19(1):43–53, 2005. 55

[32] G.W. English, T.K. Cattani, and C. Schweir. Assessment of dangerous goods regulations in railway train marshalling. Technical report, Canadian Institute of Guided Ground Transport, March 1991. 7

[33] E. Erkut and V. Verter. Modeling of transport risk for hazardous materials. *Operations Research*, pages 625–642, 1998. 17

[34] P. Fang and H.D. Reed. Strategic positioning of railroad cars to reduce their risk of derailment. Technical report, Volpe Transportation Systems Center (DOTITSC), 1979. 7, 67

[35] Intergovernmental Organization for International Carriage by Rail(OTIF). Regulations concerning the international carriage of dangerous goods by rail(rid). `http://www.mtc.government.bg/upload/docs/RID_en01.pdf`, 2007. Last accessed on June 28, 2008. 7

[36] English G.W., Higham G., Bagheri M., Moynihan T.W., and Saccomanno F.F. Evaluation of risk associated with stationary dangerous goods railway cars. Technical report, TranSys Research Limited, 2007. 7, 22

[37] S. He, R. Song, and S.S. Chaudhry. An integrated dispatching model for rail yards operations. *Computers and Operations Research*, 30(7):939–966, 2003. 43

[38] Dirnberger J.R. Development and application of lean railroad engineering to improve classification terminal performance. Master's thesis, University of Illinois at Urbana-Champaign, 2006. 4

[39] S. Kaplan and B.J. Garrick. On the quantitative definition of risk. *Risk analysis*, 1(1):11–27, 1981. 17

[40] P.R. Nayak and D.W. Palmer. Issues and dimensions of freight car size: A compendium. Technical report, Federal Railroad Administration, January 1980. 7, 67

[41] PR Nayak, DB Rosenfield, and JH Hagopian. Event probabilities and impact zones for hazardous material accidents on railroads. Technical report, Arthur D. Little, Inc., 1983. 39

[42] M. Nicolet-Monnier and A.V. Gheorghe. *Quantitative risk assessment of hazardous materials transport systems: rail, road, pipelines, and ship*. Springer, 1996. 22, 31

[43] Association of American Railroads. Current rail hazmat conditions called "untenable". `http://www.aar.org/ViewContent.asp?Content_ID=3763`, 2006. Last accessed on October 31, 2007. 3

[44] Association of American Railroads. Railroads: the safe way to move, january 2006. `http://www.aar.org/GetFile.asp?File_ID=161`, 2006. Last accessed on October 31, 2007. 2

[45] Association of American Railroads. Mandatory hazmat rerouting. `http://www.aar.org/rail_safety/HazmatRerouting.pdf`, April 2007. Last accessed on October 31, 2007. 3

[46] Transportation Safety Board of Canada. Railway investigation report, r03t0080. `http://www.tsb.gc.ca/en/reports/rail/2003/r03t0080/r03t0080.pdf`, 2003. Last accessed on October 31, 2007. 1

[47] Transportation Safety Board of Canada. Statistical summary railway occurrences 2004. http://www.tsb.gc.ca/ENG/stats/rail/2004/ss04-sec1.asp, 2004. Last accessed on October 31, 2007. 1, 2

[48] Transportation Safety Board of Canada. Statistical summary railway occurrences 2006. http://www.tsb.gc.ca/ENG/stats/rail/2006/ss04-sec1.asp, 2004. Last accessed on October 31, 2007. xi, 2

[49] Transportation Safety Board of Canada. General operating instructions(goi),canadian pacific railway. http://www.tcrc76.com/Legislative/GOI/GOInal.pdf, 2009. Last accessed on November 06, 2009. xi, 47, 48

[50] US Department of Transportation. Transportation of hazardous material by rail in the us, us dot 49 cfr 174.85. http://a257.g.akamaitech.net/7/257/2422/12feb20041500/edocket.access.gpo.gov/cfr_2004/octqtr/pdf/49cfr174.85.pdf. Last accessed on October 31, 2007. xi, 7, 12, 43, 46

[51] Anderson R.T. Quantitative analysis of factors affecting railroad accident probability and severity. Master's thesis, University of Illinois at Urbana-Champaign, 2005. 22, 23, 28, 32, 33, 34, 83

[52] FF Saccomanno and SM El-Hage. Minimizing Derailments of Railcars Carrying Dangerous Commodities through Effective Marshalling Strategies. *Transportation Research Record*, 1245:34–51, 1989. 39, 41

[53] FF Saccomanno and SM El-Hage. Establishing derailment profiles by position for corridor shipments of dangerous goods. *Canadian Journal of Civil Engineering*, 19(1):194–195, 1992. 7, 23, 32

[54] FF Saccomanno and JH Shortreed. Hazmat transport risks: societal and individual perspectives. *Journal of Transportation Engineering*, 119:177, 1993. 17

[55] D.H. Schafer and C.P.L. Barkan. Relationship between train length and accident causes and rates. *Transportation Research Record*, 2043:73–82, 2008. 22

[56] Gerard S.L. Us department of transportation, pipeline and hazardous materials safety administration. http://www.phmsa.dot.gov/staticfiles/PHMSA/DownloadableFiles/Files/NTSB_R_08_13_08-18-08.pdf, 2008. Last accessed on October 31, 2009. 44

[57] R.E. Thompson, E.R. Zamejc, and D.R. Ahlbeck. Hazardous materials car placement in a train consist. Technical report, Battelle Columbus Division, 1989. 7, 22, 67, 79

[58] L.H. Tsoukalas and R.E. Uhrig. *Fuzzy and neural approaches in engineering.* John Wiley & Sons, Inc. New York, NY, USA, 1996. 37

[59] R. Wegner. Map of the month: major yards. *Trains*, pages 42–43, April 2003.

[60] PJ Wong, M. Sakasita, and W. Stock. Railroad Classification Yard Technology Manual. Volume I. Yard Design Methods. *NTIS, SPRINGFIELD, VA,272*, 1981. 4

[61] S. Yagar, F. F. Saccomanno, and Q. Shi. An efficient sequencing model for humping in a rail yard. *Transportation Research A*, 17(4):251, 1983. 43