RISK MANAGEMENT IN THE RAIL TRANSPORT OF HAZARDOUS MATERIALS

RISK MITIGATION AND MANAGEMENT STRATEGIES FOR ROUTING HAZARDOUS MATERIALS OVER RAILROAD NETWORK IN CANADA

By ALI VAEZI, M.Sc., B.Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy in Business Administration

Ph.D. Dissertation (2018) McMaster University
(Business Administration) Hamilton, Ontario, Canada

TITLE: Risk Mitigation and Management Strategies for Routing Hazardous Materials over Railroad Network in Canada

AUTHOR: Ali Vaezi, M.Sc., B.Sc.

SUPERVISOR: Dr. Manish Verma
COMMITTEE: Dr. Elkafi Hassini
COMMITTEE: Dr. Kai Huang

NUMBER OF PAGES: x, 185

Abstract

Railroad transportation of hazardous materials (hazmat) has grown significantly in recent years in Canada. Although rail is one of the safest modes for hazmat transport, the risk of catastrophic events such as the Lac Mégantic train disaster, does exist. In this thesis, we study a number of measures to manage and mitigate the risk associated with rail hazmat shipments. *First*, we propose a methodology that makes use of analytics to dis-aggregate national freight data to estimate hazmat traffic on rail-links and at rail-yards in Canada. Further, a focused analysis is conducted on crude oil rail shipments to develop long-term forecasts and evaluate the impact of proposed pipeline projects. *Second*, we present an emergency response planning problem, aimed at the effective and efficient response to rail hazmat incidents. A two-stage stochastic programming problem is solved over part of the Canadian railroad network, which provides recommendations on where to locate response facilities, and which equipment packages to stockpile at each facility. *Finally*, we study infrastructure investment as a strategy to mitigate the risk associated with rail hazmat shipments. This strategy is based on building new railway tracks to provide alternative routes to the riskiest parts of the network. Given the hierarchical relationship between the decisions made by regulatory agencies and railroad companies, a bilevel programming approach is used to identify the optimal set of infrastructure investment options given an allocated budget. Our computational experiments show that significant network-wide risk reduction is possible if hazardous shipments are routed using some of the proposed alternative rail tracks.

Acknowledgements

This dissertation would not have been possible without the continuous support and constructive guidance of my advisor and committee members, help from my friends, and support from my beloved family.

I would like to express my sincere gratitude and deep appreciation to my doctoral advisor, Dr. Manish Verma, for being an excellent mentor in every stage of this journey since I began the Ph.D. program in 2013. I have had his full support, and I could have not imagined having a better advisor for my Ph.D. studies. I appreciate all his contributions of time and feedback, which helped me a lot to develop skills as a researcher.

I am also grateful to the other members of my supervisory committee, Dr. Elkafi Hassini and Dr. Kai Huang, who have generously given their time and expertise to review my dissertation. Their insightful comments and recommendations have helped me to improve the quality of this dissertation.

Also, I would like to express my special appreciation to Dr. Mahmut Parlar and Dr. Prakash Abad, who have been excellent and inspirational teachers, and have taught me business courses during the past five years. I thank those who are part of the Operations Management Area and all DeGroote School of Business members for providing me with a great atmosphere for this exciting journey of discovery.

Last but certainly not least, I take this opportunity to thank my Ph.D. fellows and caring friends, and I wish them every success and happiness in life.

Contents

List of Figures

List of Tables

Chapter 1 Introduction

A significant majority of hazardous materials (hazmat) shipments, including those integral to sustain our industrial lifestyle such as crude oil, are produced and consumed at different locations, and therefore they may have to be transported over long distances. In North America, a significant portion of hazmat shipments is transported over the railroad network. Besides pipeline capacity shortfall, favorable safety statistics have made rail a growing mode for hazmat transport. For instance, Canadian railroads carried over 44 million tons of hazmat in 2015, which shows a remarkable increase from about 29 million tons in 2006 (Transport Canada, 2017). Although rail is one of the safest modes for hazmat transport, the possibility of catastrophic events resulting from multi railcar incidents does exist. For example, around 435 train accidents involving hazmat release have been reported between 2000 and 2015, which translates into an average of 27 accidents every year (FRA, 2016). The tragic event in Lac-Mégantic (Quebec), in July 2013, is a reminder of the possible catastrophe associated with rail hazmat shipments (Lavelle, 2013). In this thesis, we use relevant data and analytical tools to study a number of measures to manage and mitigate the risk associated with rail hazmat shipments.

Hazardous materials are defined as those substances that are capable of causing harm to people, property, and the environment (Erkut et al., 2007). Transport Canada divides hazardous materials into 9 classes: explosives, gases, flammable liquids, flammable solids, oxidizing substances, poisonous and infectious substances, radioactive materials, corrosives, and miscellaneous substances such as dangerous wastes

(Government of Canada, 2014). Class 2, 3, and 8 (i.e., *gases*, *flammable liquids*, and *corrosives*) account for almost 80% of the rail hazmat tonnage in Canada (Provencher, 2008). Our data extraction and analysis endeavor validates this statement and shows that the abovementioned three classes constitute the majority of hazmat transported by rail during the last decade, which also motivated us to focus on only these three classes.

1.1 Thesis Objectives and Contributions

Transportation of hazmat can pose a significant risk to human health and to the environment. However, the associated risk can be managed and mitigated by preventive and/or protective measures to reduce probabilities and/or consequences of hazmat release incidents. For example, the catastrophe from the Lac-Mégantic train disaster could have been alleviated if an appropriate emergency response network was in place. The last two decades have witnessed a number of academic and industry initiatives focused on the *assessment* and *management* (mitigation) of hazmat risk for rail shipments. We will provide further details on relevant studies and their findings in Chapter 2.

In this dissertation, we are interested in studying the problem from an analytical point of view and propose solutions, whose effective implementation would mitigate the risk associated with rail hazmat shipments. Thus, this dissertation is mainly motivated by the following question: *"How to manage and mitigate risk associated with rail hazmat shipments?"*[1] We endeavour to answer this question by first analyzing the relevant publicly

[1] While risk mitigation aims at identifying and reducing the adverse impacts before an emergency, risk management is a more general term, which considers the trade-offs involved and covers a wide range of measures to control and manage the associated risks.

available data, and then investigating a number of mitigation measures to provide useful analytical solutions and managerial insights.

More specifically, the *first* research piece (Chapter 3) is concerned with the use of both *descriptive* and *predictive analytics* (Evans, 2012) to analyze historical rail hazmat traffic flows in Canada, identify the associated patterns, and forecast future volumes and trends at the link-level. It is important that such link (yard) level data is not publicly available, and extremely challenging to procure from the railroad companies because of the proprietary nature of the industry and the unwillingness to share. Hence, we propose a methodology that uses relevant historical data, predictive modeling, and optimization, collectively to estimate rail hazmat traffic and also to provide a basis to assess hazmat risk across the Canadian railroad network. Furthermore, this chapter includes a focused analysis on crude oil rail shipments in Canada, where long-term forecasts are developed, different routing approaches and their risk management implications are discussed, and the impact of proposed pipeline projects is evaluated at the rail-link level.

Then, we propose two risk management techniques, i.e., *emergency response planning* (Chapter 4) and *rail track infrastructure investment* (Chapter 5), each of which is applied to the Canadian railroad network. In particular, we will investigate how and to what extent each measure can be useful in terms of risk reduction.

In Chapter 4, we propose a mathematical model, which seeks to optimally locate emergency response facilities across the railroad network and to identify specialized equipment packages to be stockpiled at each facility. A *two-stage stochastic programming* *with recourse* technique is used to integrate strategic and tactical level decisions as well as to account for uncertainties in some model parameters, e.g., incident location. This optimization problem explicitly accounts for the probability of hazmat incident as a model parameter, which is partially calculated based on the outcome of Chapter 3, i.e., link-level hazmat traffic estimates. Due to complexities involved in the problem, a *decomposition heuristic methodology* is combined with *Lagrangian relaxation* to efficiently solve the problem.

In Chapter 5, a country-wide risk assessment of the Canadian network is conducted to select candidate areas for investment. Information about origins and destinations of hazmat shipments and their supply/demand shares is extracted from the related analyses in Chapter 3. A *bilevel programming* approach is then proposed to identify the optimal set of infrastructure investment options given an allocated budget. Employing the special structure of the model, a *heuristic* solution methodology is used to solve the realistic size problem.

1.2 Thesis Layout Description

Chapter 2 provides a general literature review on *risk assessment and management* in the *railroad transportation of hazmat*. Additionally, this chapter reviews the relevant literature specific to each of the following three chapters. In the end, we will elaborate on how our research, including our data analyses and the proposed risk management techniques, would fit within the existing literature.

Chapter 3 outlines a data analysis approach to better understand rail hazmat traffic flow in Canada, and also contains risk management recommendations for an important hazmat commodity, i.e., crude oil.

Chapter 4 investigates emergency response planning as a risk management technique, aimed at alleviating the harmful consequences associated with rail hazmat incidents.

Chapter 5 studies an infrastructure investment strategy to mitigate the risk associated with rail hazmat shipments. This strategy is based on building new railway tracks to provide alternative routes to the riskiest parts of the network.

Finally, in Chapter 6, we summarize the thesis and its findings. We will also present a number of future research opportunities for the three research questions introduced in this thesis.

Chapter 2 Literature Review

We review the relevant literature under four themes: the first part will provide an overview of *risk assessment and management* efforts specific to rail hazmat shipments; the second part will briefly review works related to *freight traffic forecasting* (Chapter 3); the third part will be related to *emergency response facility location* (Chapter 4); and, the fourth will review the relevant literature on *infrastructure investment* as a risk mitigation strategy (Chapter 5).

2.1 Risk Assessment and Management

Risk is defined as a "measure of probability and severity of adverse effects" (Lowrance, 1976). In hazmat transport literature, this is referred to as *traditional risk* (TR), which is more precisely defined as the product of the *probability* and the *consequence* of an undesirable event. This expected consequence approach was originally developed to evaluate the risk associated with highway hazmat shipments and fixed facilities, but later was adapted to study rail hazmat shipments. Verma and Verter (2013) specify a number of features that differentiate rail hazmat transport from highway hazmat transport, such as the possibility of hazmat release from multiple railcars; taking such features into account would result in a more accurate risk analysis.

Alternate measures such as *incident probability* (IP) and *population exposure* (PE) have been developed, mainly due to the dearth of relevant data and/or limitations associated with the expected consequence approach. The former neglects the incident consequences, and therefore is appropriate for hazmat with relatively small danger zones (Abkowitz et al., 1992). On the other hand, the latter is focused on incident consequences only, and indicate the total number of people exposed to the possibility of harmful consequences due to a hazmat shipment (Batta and Chiu, 1988; ReVelle et al., 1991). Finally, in an effort to both incorporate the catastrophic nature of hazmat incidents and provide routing flexibility to the decision makers, Hosseini and Verma have proposed *value-at-risk* and *conditional value-at-risk* models for rail hazmat shipments (2017a, b). It is important to note that the efforts made in applying these measures to the railroad transportation domain (Verma, 2011; Bagheri, 2010; Verma and Verter, 2007) have been only at a network level. Access to hazmat volume estimates at a rail-link level could lead to more accurate risk analyses (Vaezi and Verma, 2017, 2018). A brief review of the related literature on freight traffic estimation is provided in Section 2.2.

The railroad industry has adopted various measures to manage hazmat risk by reducing the frequency of tank car accidents as well as the likelihood of release in case of an accident. A number of efforts have focused on mitigating the risk through safer *tank-car designs* (Barkan et al. ,2007; Barkan, 2008; Saat & Barkan, 2011; Raj and Pritchard, 2000). Another measure studied to mitigate or control hazmat risk is a proper *placement of hazmat railcars* (Fang and Reed, 1979; Thompson et al., 1992; Verma, 2011; Cheng et al., 2017). Also, optimal *routing of hazmat railcars* has been studied in the literature (Glickman, 1983; Glickman et al., 2007), which is primarily concerned with reducing the number of people exposed to hazmat shipments. The perspective of competing stakeholders (i.e., regulators and railroad companies) in routing decisions was captured in Verma (2009) and Verma et al. (2011).

Finally, we briefly mention other risk management techniques studied in hazmat transport literature. In the road transportation domain, *signal control policy* (Chiou, 2016) and *toll setting policy* (Marcotte et al., 2009; Bianco et al., 2015) have been presented as regulatory measures to mitigate hazmat risk. Xie et al. (2012) studied a joint *facility location* and routing problem on a multimodal transportation network. Purdy (1993) and Mazzarotta (2002) studied *modal choice* by analyzing and comparing the risk associated with road and rail hazmat shipments. *Emergency response planning*, which is aimed at reducing the consequences associated with potential hazmat incidents, has been studied in the road and marine transportation domains (Berman et al., 2007; Verma et al., 2013). To the best of our knowledge, rail-link *infrastructure investment* has not been studied in hazmat transport literature as a risk mitigation measure. Reviews of the literature specific to emergency response and infrastructure investment are provided in Sections 2.3 and 2.4, respectively.

2.2 Freight Traffic Forecasting

As alluded to earlier, having link (yard) level traffic data could significantly impact the precision of the aforementioned risk assessment and management efforts. As a result, access to relevant and reliable data is crucial in analyzing and mitigating the risk associated with rail hazmat shipments. Although new technologies are collecting more data than ever before, many organizations are still looking for better ways to achieve value from their data

(LaValle et al., 2011). Such a value is obtained from collecting, extracting, and analyzing relevant data and linking the analytics-driven insights to business strategy, which is known as *analytics* (Davenport and Harris, 2007). Liberatore and Luo (2010) presented a process view of analytics, which is made up of visualization (*descriptive*), prediction (*predictive*), and optimization (*prescriptive*). Analytical approaches to freight traffic forecasting generally involve descriptive, predictive, and/or prescriptive analytics to gain insight and support decision making in freight transportation. Though we review a number of relevant works, studies conducted in urban contexts (Chrobok et al., 2004), as well as those on passenger traffic modeling (Owen & Phillips, 1987), are not covered.

One of the earliest works on freight transportation forecasting resulted from the endeavour to develop long-term forecasts for energy resources, including crude oil, in France (Chateau & Lapillonne, 1978). In general, freight traffic models can be classified into aggregate and disaggregate categories, wherein the former has a cost-minimizing focus whereas the latter could be tailored to accommodate the behavioral realities of the decision maker (Winston, 1983). A nonlinear optimization framework was proposed by Crainic et al. (1990) to minimize the total generalized system costs in freight traffic assignment to railroad network. They presented a case study in Brazil and modeled the flow of products that represent high volumes and are of national importance, e.g., agricultural products, petroleum, etc. However, their model neglected the risk associated with hazardous commodities and required extensive commodity-specific data for each origin and destination pair. De Jong et al. (2004) reviewed freight transport models at national and international levels, including those for production and attraction of goods, their distribution, their modal split, and their assignment. Chow et al. (2010) studied existing freight forecasting model classes and analytical approaches in traffic dis-aggregation, but also recognized the possibility of using hybrid and ad-hoc methodologies. Finally, some transport geography researchers have opined the need for empirical investigations involving the entire value chain (Hesse and Rodrigue, 2004).

To sum, to the best of our knowledge, the peer reviewed efforts in rail hazmat transport have primarily focused on risk assessment and risk management, and assumes the existence of precise data for thorough analysis and resulting decision making. The true potential benefits resulting from these efforts could not be demonstrated in the absence of rail-link level information about the type and volume of hazmat. Furthermore, existing analytical approaches in freight traffic modeling/forecasting cannot necessarily be used in other studies because of the differences in planning level, transportation mode, data requirement, decision criteria, etc.

2.3 Emergency Response Facility Location

Emergency management, in general, involves all activities and risk management measures for *prevention/mitigation* of, *preparedness* for, *response* to, and *recovery* from emergencies (Public Safety Canada, 2010). A hazmat *incident* is defined as an accident that results in a hazmat *release* (Erkut et al., 2007). While operational readiness (*ex-ante*) and timely response (*ex-post*) are crucial in alleviating the harmful consequences resulting from rail hazmat incidents, longer-term planning approaches are also needed to assure the availability of specialized response resources at the incident site. In other words, *strategic* and *tactical* decision levels interact with one another, which is why an integrated framework is recommended to be used to locate response facilities and allocate specialized equipment to hazmat incident situations (Iakovou et al., 1997).

One of the earliest works in emergency response facility location was done by Toregas et al. (1971), who proposed a *set covering* problem to locate emergency service facilities. This view is based on the idea that each potential demand point in the area of interest must be reached from its allocated emergency response facility within a specified time in case of emergency. Given a specified coverage distance, this covering problem minimizes the number of facilities or an equivalent budget required to cover every demand over the network. On the other hand, *maximal covering location* problem (Church & ReVelle, 1974) aimed at covering as many demand points as possible within a predetermined critical distance or time, given a finite budget or a finite number of facilities. As argued by Berman et al. (2007), in hazmat transportation, emergency response needs to cover a set of *arcs* that make the transport network rather than a set of points. They introduced a mixed-integer nonlinear formulation for the *maximal arc-covering* problem, which involves locating a finite number of facilities so as to maximize the total weighted arc length covered.

The question of locating emergency response facilities with *stochastic* inputs has not received much attention in the literature. *Stochastic programming with recourse* was first introduced and studied by Beale (1955) and Dantzig (1955). *Two-stage stochastic programming with recourse* (Birge and Louveaux, 2011) was used by Louveaux (1986) to solve *plant location* and *p-median* problems. In these problems, decisions on location and size of facilities are considered in the first stage, and decisions on allocation of production resources are considered in the second stage. Owen and Daskin (1998) conducted an extensive review of strategic facility location literature, which also includes works related to stochastic location theory and locating emergency service facilities.

A *partial set covering* model (Daskin and Owen, 1999) was applied by Belardo et al. (1984) to the problem of locating oil spill response resources on Long Island Sound. They used multiple objective functions, to evaluate strategies based on both probability of occurrence and the impact of occurrence. Their model also accounted for resource constraints, but equipment needs were determined on the basis of a single spill volume, which is not the case in oil spill events. Psaraftis et al. (1986) proposed a mixed integer programming problem to (1) locate appropriate levels and types of response resources, and (2) allocate such capability among oil spill potential locations. Their model took into account a number of relevant elements, such as frequency of spill occurrence and variability of spill volumes. A framework integrating strategic and tactical level decisions is suggested by Iakovou et al. (1997) to locate emergency response facilities and allocate response resources. A linear integer programming model was used to formulate such a problem, which was then applied to two examples related to the east coast of Florida. However, the probabilistic nature of oil spills was ignored in their emergency response planning formulation.

Finally, Verma et al. (2013) studied the problem of location and capability of oil-spill response facilities for the south coast of Newfoundland. They divided this area into several zones, recognizing each of which as a likely location for oil spill events characterized by oil type, weather conditions, and spill volume. The authors then proposed an emergency response problem using a two-stage stochastic programming structure, where facility location and equipment acquisition decisions are made in the first stage, and equipment dispatching decisions are made in the second stage.

While first responders are typically able to reduce damages to some extent, the timely presence of specialized personnel and equipment is necessary to effectively respond to hazmat incidents. To the best of our knowledge, there is no peer-reviewed study in the literature that incorporates the differentiating aspects of rail hazmat transport into a planning framework for response to rail hazmat incidents. It is pertinent that the initiatives taken by railroad companies are not publicly available, and thus difficult to ascertain.

2.4 Infrastructure Investment

There are numerous research threads in the broad area of infrastructure investment, such as methods to assess infrastructure projects and the roles of corporate/regulatory players. In this section, we only cite the most relevant works in the literature.

Cost benefit analysis (CBA) is cited as a general method to assess infrastructure investment projects. Quah and Toh (2011) list three general steps in conducting a CBA: deciding which items to include as costs and benefits; assigning values to the included items; and, arriving at a conclusion to support decision making. Salling and Banister (2009) conduct CBA to assess large transport infrastructure projects, in which they account for uncertainties through probability distribution functions. A well-known example of CBA for regulatory decision making is the Superfund program by the US government to clean up some of the most contaminated lands in the country to mitigate the risk to human health and to the environment (U.S. DHHS, 2017). This analysis involved identifying the most dangerous sites (US EPA, 1992), and evaluating the associated cleanup benefits in terms of health effects or other surrogate measures. Using a similar risk-mitigating framework, Flammini et al. (2008) present a CBA approach to assessing the extent to which physical protection measures cost but on the other hand reduce the risk of intentional threats in the railroad domain. However, they only account for monetary values of expected damages and exclude the risk to human lives. Finally, Špačková and Straub (2015) use CBA to find optimal risk protection strategies against natural hazards under a limited budget, which is shared among a number of subsystems. They formulate a *discrete hierarchical optimization* problem to minimize the sum of risk and cost at the system level under budget constraints.

Infrastructure investment as a measure to mitigate the risk associated with rail hazmat shipments has only been briefly mentioned in the existing literature. For instance, Milazzo et al. (2002) recommend a number of strategies for mitigating the risk associated with the land transport of hazmat through the downtown of Messina in Italy. In particular, they argue that it is necessary to take into account *volumes* of hazmat commodities and their *production sites* when making decisions about long-term railroad investments. Van der Vlies and Suddle (2008) study the case of rail hazmat transport in the Netherlands, and list a number of *structural safety measures* while considering the interdependencies between investment on such measures and future urban developments.

Infrastructure investment projects typically involve an *interactive* decision making environment with both *corporate* and *regulatory* players. Depending on the type of infrastructure project, the role of regulators may vary. Sometimes they even provide the infrastructure. In particular, governments/regulators play a key role in some infrastructure investments since they are beneficial to the society or they can promote or induce growth in a region, e.g., investments on airports or railroads (Smit, 2003). In a given infrastructure investment project, there may be several players, each with its own perspective of the situation and hence its own objective. For instance, in railroad transportation of hazmat, the regulator is usually concerned about transport risk whereas the carrier seeks to minimize routing costs while satisfying the safety requirements set by the regulator (Kara and Verter, 2004). Therefore, the presence of a *bilevel* setting is conceivable.

A *bilevel programming problem* (BLPP) represents a hierarchical relationship between two autonomous decision makers who may have conflictual objectives (Colson et al., 2007). It can be viewed as a *static* version of the *non-cooperative*, *two-person leader-follower game*, which was initially introduced by Von Stackelberg (2010). In a bilevel setting, *perfect information* is assumed, which means each player knows the objective and feasible strategies available to the other player (Bard, 2013).

In general, route planning for regular freight simply involves selecting the shortest path between an origin-destination pair. By contrast, route planning decisions for hazmat freight should also account for the risk associated with such shipments. Bianco et al. (2013) review hazmat route planning models on a road network, in which they recognize the presence of government authorities in addition to carriers. In particular, they suggest that a carrier solves a single-commodity problem for each hazmat shipment, i.e., *local route planning*, whereas a government authority is mostly concerned about the associated societal and environmental risks, and therefore solves a *global route planning* problem by means of regulatory tools such as *network design* and *toll setting*. More detailed review of each tool is provided in the following paragraphs.

Kara and Verter (2004) propose a bilevel programming approach to *network design* for highway hazmat transportation, which is then applied to Western Ontario in Canada. In such a setting, the government has the authority to close certain roads to vehicles carrying hazmat. In their bilevel formulation, the *outer* problem involves minimizing the total risk resulting from the carrier's route choices, while the *inner* problem involves minimizing the total cost over the network designed by the regulator. Since for every designated network, the inner problem is a *minimum cost flow* model which is *unimodular*, they use the *Karush-Kuhn-Tucker (KKT)* conditions of its *linear relaxation* to obtain an *equivalent feasibility* problem. As a result, the original bilevel program is transformed into a *single-level Mixed Integer Programming (MIP)* problem, which is then solved using a commercial optimization software. Several extensions of this model have been proposed, e.g., by Erkut and Gzara (2008), Verter and Kara (2008), Marcotte et al. (2009), and Fontaine and Minner (2018).

In the highway domain, *toll setting* is considered as an alternative policy tool to mitigate the risk associated with hazmat shipments. This policy is based on discouraging (rather than preventing) hazmat carriers from using certain road segments by imposing tolls on risky links. Hence, it is a more flexible regulation compared to the network design policy as it would maintain the carriers' freedom in using the entire network. Marcotte et al. (2009) propose a bilevel model to minimize hazmat transport risk. They show that using *inverse optimization*, the regulator can adopt toll setting policies that would induce the carriers to choose the routing plan corresponding to a minimum risk solution. Their work was later extended by Bianco et al. (2015) and Esfandeh et al. (2016).

Bilevel programs are intrinsically difficult, which is why most research conducted on solution algorithms has focused on the simplest cases such as those with linear problems at both inner and outer levels. Even such a *linear-linear BLPP* has been shown to be *NP*-hard (Jeroslow, 1985). In fact, many *combinatorial optimization* problems can be reformulated as bilevel programs; based on such an equivalency, Colson et al. (2007) suggest that employing the combinatorial structure of bilevel programs could lead to the development of efficient solution strategies. The interested reader is referred to Colson et al. (2007), Bard (2013), and Talbi (2013) for a comprehensive survey of existing methods including *heuristic* and *metaheuristic* techniques.

Finally, it should be noted that the vast majority of published works that adopt a bilevel modeling approach to hazmat transportation belong to the highway transportation domain. To our knowledge, there are only a few such applications to the railroad domain. For instance, Assadipour et al. (2016) propose a *bilevel bi-objective toll setting policy* model to discourage carriers from using certain rail intermodal terminals. In another recent study, Fontaine et al. (2016) extend the hazmat transport road network design problem to a multimodal network setting, where they also address the issue of *risk equity*.

2.5 Research Gaps

In this chapter, we reviewed the most relevant literature on the railroad transportation of hazmat. While Section 2.1 provided a general overview of risk assessment and management efforts in the rail hazmat transport domain, the subsequent sections reviewed the existing literature on freight traffic forecasting (Chapter 3), emergency response facility location (Chapter 4), and infrastructure investment (Chapter 5).

We identify three research gaps in the literature. *First*, we observe a lack of integrated analytical approaches to understanding the current and future rail hazmat transport in Canada. More specifically, a crucial input for risk assessment and management efforts, i.e., link-level hazmat volume across the railroad network, is missing from the existing literature. *Second*, we are not aware of any optimization-based strategic and tactical planning framework for response to rail hazmat incidents. Further, no emergency response planning problem is studied that accounts for specific features of rail hazmat transport. *Third*, to the best of our knowledge, infrastructure investment has not been studied in hazmat transport literature as a risk mitigation measure.

Chapter 3 addresses the first research gap using an integrated approach to estimate hazmat traffic over the railroad network in Canada. Chapter 4 addresses the second research gap by presenting a mathematical modeling approach toward planning and responding to rail hazmat incidents. Finally, Chapter 5 addresses the third research gap through studying infrastructure investment as a risk mitigation strategy in the rail hazmat transport domain.

Chapter 3 An Analytics Approach to Estimating Hazardous Materials Traffic on Railroad Network in Canada

3.1 Introduction

Given the significance of rail hazmat shipments and the possibility to more accurately determine hazmat risk mitigation, we make a first attempt to estimate hazmat traffic on various links and yards across the railroad network in Canada. To that end, we propose a methodology that applies popular analytics techniques to country-level commodity flow data made available through Statistics Canada (2016a). More specifically, the proposed methodology makes use of the following steps: estimating hazmat traffic from historical data; forecasting hazmat traffic; routing hazmat traffic; and, estimating hazmat volume on links and yards.

The proposed methodology is applied to estimate the volumes of three hazmat classes responsible for around 80% of the rail hazmat shipments in Canada. In addition, developing prediction intervals around each forecast incorporates the impact of uncertainties, emanating from different sources. A number of problem instances are solved and analyzed to identify the high-traffic points in the railroad network in Canada, and thus the possible inputs for developing emergency response plan; and, to discern the impact of shifts in supply locations, and the relative volume-based ranking of different rail yards and rail links from 2016 to 2030. Furthermore, this chapter contains a focused analysis on crude oil rail shipments, which facilitates the development of long-term forecasts and evaluation of the impact of proposed pipeline projects.

The rest of this chapter is organized as follows. Section 3.2 provides an overview of the proposed methodology, each step of which is then elaborated upon in Section 3.3. Relevant discussion and analysis following the application of the proposed technique to the railroad network in Canada are provided in Section 3.4. Section 3.5 contains a focused analysis on crude oil rail shipments in Canada, which is followed by the conclusion in Section 3.6.

3.2 Proposed Dis-aggregation Methodology[2]

In this section, we provide an outline of the proposed dis-aggregation methodology to estimate hazmat traffic at rail-yard and rail-link levels, which is then applied to the Canadian setting in Section 3.3. As alluded to earlier, we are making the first effort at developing a technique that could enable researchers (and interested practitioners) to circumvent the challenges associated with procuring dis-aggregated data on hazmat shipments. To accomplish our objective, we make use of analytics (Davenport and Harris, 2007). More specifically, we use descriptive analytics to understand the country-level rail traffic, predictive analytics to forecast future volumes, and prescriptive analytics to route shipments over the railroad network.

[2] Sections 3.2 to 3.4 have appeared in a journal article with following citation details:
Vaezi, A., & Verma, M. (2017). An analytics approach to dis-aggregate national freight data to estimate hazmat traffic on rail-links and at rail-yards in Canada. *Journal of Rail Transport Planning & Management*, *7* (4), 291-307.

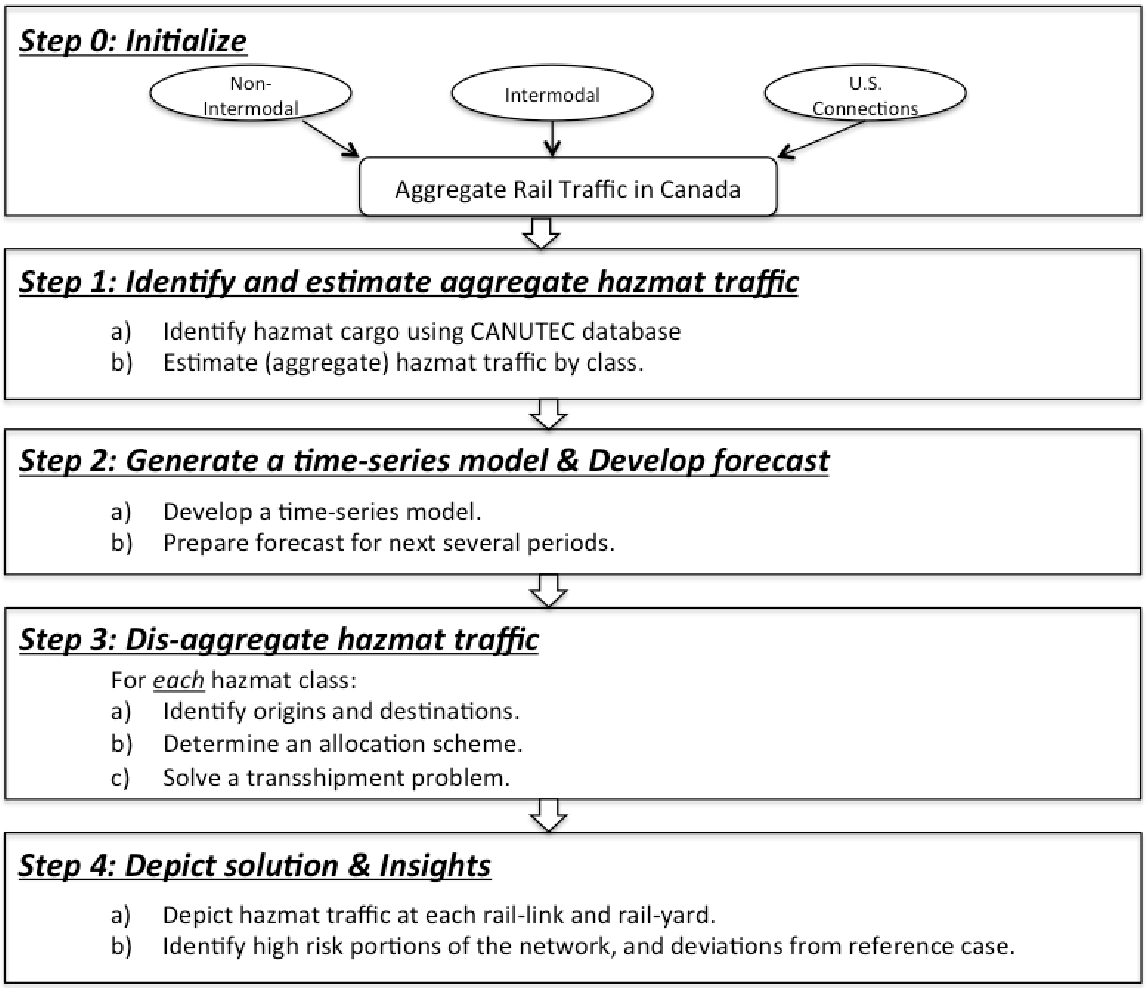


Figure 3.1. Flowchart of the proposed dis-aggregation methodology

The proposed methodology, summarized in Figure 3.1, consists of 5 steps. *Step 0* depicts the three revenue-generating sources of rail freight, which collectively determine the aggregate country-level rail traffic in Canada (Statistics Canada, 2012). We provide all the pertinent details about initialization in Section 3.3.1. *Step 1* takes the aggregate rail traffic output from the initialization step, and then applies both descriptive and diagnostic analytics to parse hazmat cargo. This is done in two steps: *first*, for each commodity transported, the emergency response database is consulted to ascertain if the freight is hazardous or not; and *second*, if hazardous, then based on characteristics, it is placed under one of the nine classes. At the end of this step, we have the country-level rail hazmat traffic by class. We outline the details for the Canadian context in Section 3.3.2.

Given our objective of providing better estimates now and in the near future, *Step 2* does two things: *first*, based on the historical dataset, it determines the best forecasting model; and *second*, develops forecast for several periods ahead taking into account some anticipated sources of uncertainty (i.e., predictive analytics). We provide the relevant details about the context in question in Section 3.3.3.

The resulting output, for each traffic class and each simulated scenario, is dis-aggregated in *Step 3* via a three-stage processing: *first*, based on publicly available information, identify the origins and destinations; *second*, determine an allocation scheme to meet demand at destinations; and *third*, solve a transhipment model to move shipments from origins to destinations. Thus, this step makes use of prescriptive analytics, and we illustrate the application in Canada in Section 3.3.4.

Finally, *Step 4* could be used to develop a pictorial and/ or tabular depiction of the solution, and conduct relevant analyses. Section 3.3.5 outlines the discussion for the Canadian context.

Before closing this section, it is important that the proposed methodology is general enough to be applied to any country, however, the individual components might have to be customized to account for the kind of data and/or restrictions specific to that jurisdiction.

3.3 Application of the Dis-aggregation Methodology to Canada

3.3.1 Initialize

Commodity-based railway car loading information from 1999 to 2015 was procured from Statistics Canada (2016a), which as indicated earlier is collected only for revenue-generating freight (Statistics Canada, 2012). The total traffic carried across the railroad network in Canada is made up of three components: non-intermodal traffic; intermodal traffic; and, traffic received from connections to the United States.

The non-intermodal traffic component is the largest of the three, and is organized into two groups: *first*, specified commodity, which is grouped into 60 categories; and *second*, non-specified commodity, which is grouped into 4 categories. The intermodal traffic is divided into container-on-flat-car and trailer-on-flat-car, while the U.S. connections are split into non-intermodal and intermodal groups. It is pertinent to bear in mind two things: *first*, the above dataset provides commodity-based information only for the 60 categories with specified commodity, while the other components contain only aggregated volume information. *Second*, each entry in the dataset corresponds to a monthly time-horizon, and thus we consolidated them to generate annual numbers for each commodity. Although visual inspection of the parsed monthly data for hazmat cargo did convey the absence of seasonality, we ran a time series regression analysis to test to validate it. Figure 3.2 depicts the contribution of the three components to the rail traffic in Canada from 1999 to 2015. It is clear that the total tonnage started an upward trend in 2009, in large part driven by the increase in the non-intermodal traffic that accounts for 80% of the rail traffic. On the other hand, the intermodal traffic increased marginally over the same period, while the U.S. connections dropped to 35 million tons after peaking at 40 in 2013. It is possible to surmise that the relative share of the three sources of freight is rather stable over the most recent years. For instance, in 2015, a total rail traffic of 355.47 million tons was registered, of which around 287.33 million tons was carried by non-intermodal – including 1.45 million tons by the 4 categories without specified commodity, 32.8 million tons by intermodal, and 35.33 million tons via U.S. connections.

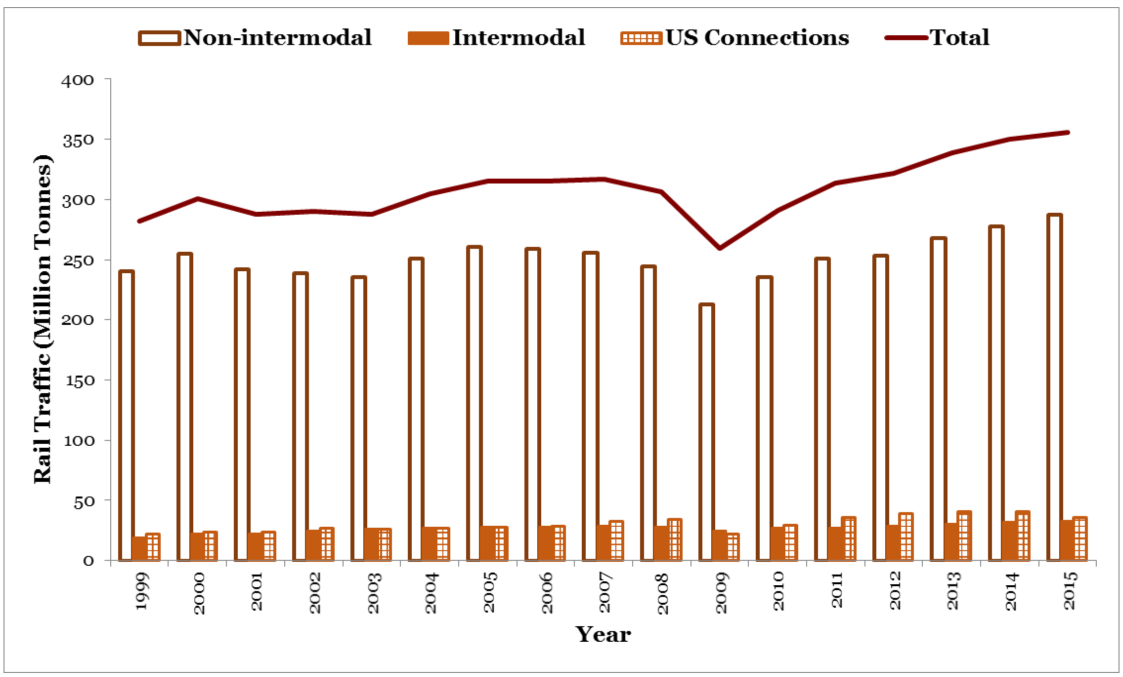


Figure 3.2. Rail traffic in Canada

3.3.2 Identify and Estimate Aggregate Hazmat Traffic

For expositional reasons, we illustrate the estimation process using the rail traffic data for 2015, and note that the equivalent information for other years could be determined similarly.

Revenue Generating Freight Sources		Hazmat				
		Class	Commodity Groups[3]	Proportion	Volume (tonnes)	Aggregate Percent
Non-intermodal traffic (specified commodity groups)	*"with"*	***2*** (Gases)	*Gaseous HC*	1.0000	7,049,767	2.47%
			Other products	0.2808		
			Fertilizers	0.0631		
		3 (Flammable Liquids)	*Beverages*	0.0668	16,412,613	5.74%
			Gasoline	1.0000		
			Fuel & crude oil			
			Other products	0.2808		
			Plastic & rubber	0.0391		
		4 (Flammable Solids)	*Sulphur*	1.0000	5,844,763	2.04%
			Plastic & rubber	0.2842		
		5 (Oxidizing substances)	*Fertilizers and materials*	0.0138	56,156	0.02%
		8 (Corrosives)	*Sulphuric Acid*	1.0000	2,827,641	0.99%
			Potash	0.0100		
		Unidentified	*Basic chemicals*	1.0000	11,199,109	3.92%
			Other chemicals			
				Sub-Total	***43,390,069***	***15.18%***
	"without"				219,883	15.18%
Intermodal traffic					4,920,541	
US connections					5,299,710	

Table 3.1. Country-level breakdown of hazmat traffic by class for 2015

[3] The names of some commodity-groups have been shortened to fit the table.

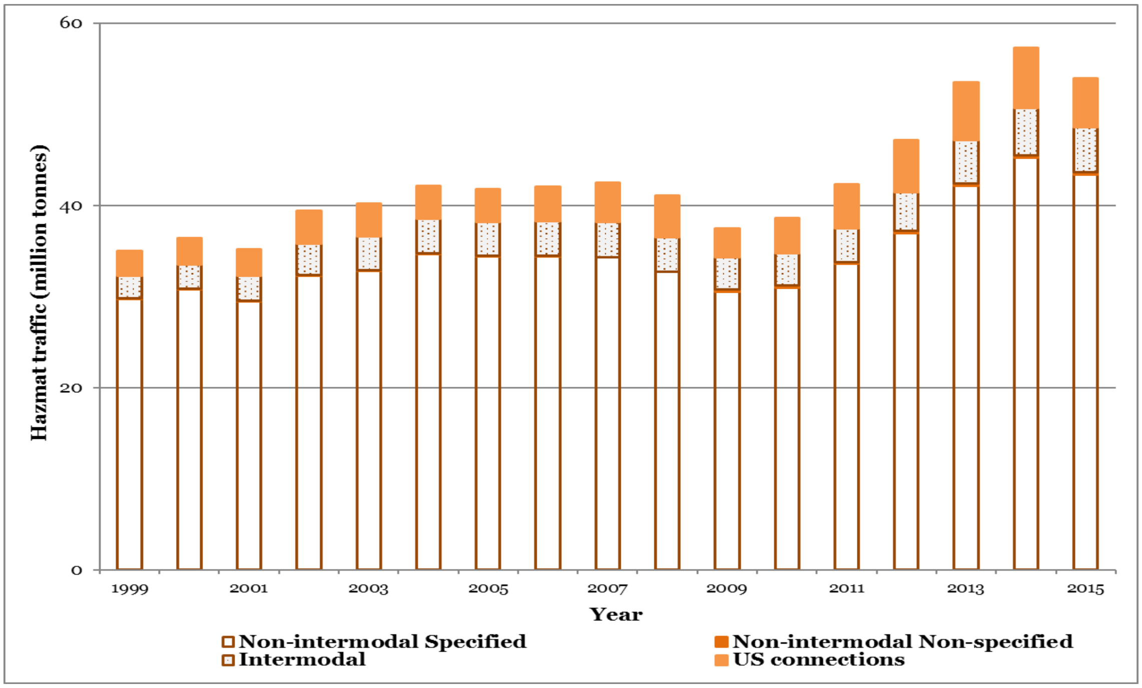
As indicated earlier, Statistics Canada (2016a) provides commodity-based information for 60 categories of products for non-intermodal traffic, but not sufficient details for the remaining two revenue generating freight sources. It is pertinent that any information about commodity facilitates consulting the emergency response database to ascertain the chemical properties of the product. Hence, each of the sixty products was entered into the Canadian Transport Emergency Centre (CANUTEC) database (Transport Canada, 2012a), of which one-fifth contained hazardous properties. Table 3.1 depicts these twelve hazmat commodity groups and the hazmat class, respectively, in the 3rd and 2nd columns. Table 3.1 provides a snapshot of the various hazmat attributes associated with the three revenue generating freight sources, and we will elaborate on each in the paragraphs to follow. Note that for two commodity-groups (i.e., basic chemicals and other chemicals) no additional information was available, and thus to be conservative we assumed the entire freight to be hazardous and indicated them against “unidentified”.

For a given commodity-group, the 4th column indicates the proportion of hazmat freight. For example, 6.68% of beverages transported in Canada has alcoholic component that is classified as flammable liquids and thus is a hazmat. It is important that rail traffic data does not distinguish between alcoholic and non-alcoholic beverage, and hence we relied on published works to develop the estimates. More specifically, we went through the following three steps to estimate 6.68%: *first*, the aggregate beverage consumption (in grams) statistics by age group and gender provided was extracted (Garriguet, 2008a; Garriguet, 2008b); *second*, the organized consumption statistics were combined with the population statistics provided by Statistics Canada (2016b) to compute the population-weighted consumption of different beverages; and *third*, the population-weighted alcoholic beverage consumption of adults was divided by the total population-weighted beverage consumption. For brevity, we note that relevant published works were consulted and appropriate estimation techniques were developed to ascertain hazmat proportions for other commodity-groups, but are not reported here.

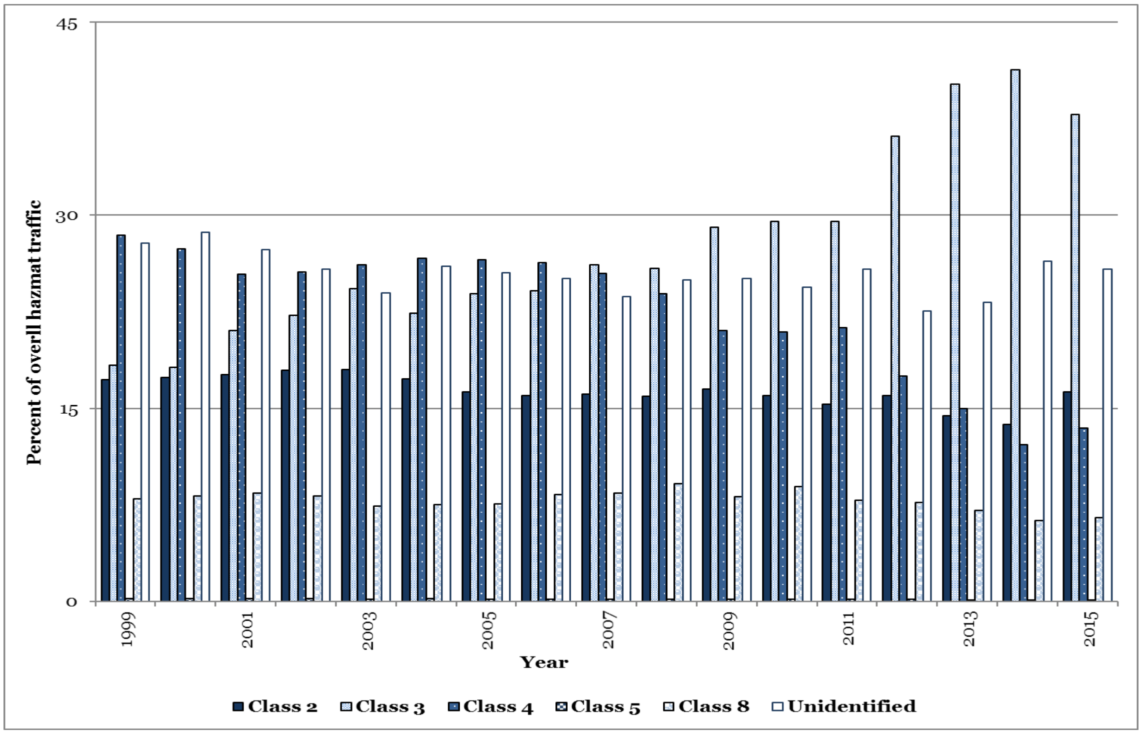
The 5th column reports the volume of hazmat cargo, while the 6th indicates the percentage of total freight that could be categorized as hazardous. A total of 285.89 million tons of non-intermodal with specified commodity was transported in 2015, which together with 1.45 million tons for non-intermodal without specified commodity is depicted in Figure 3.2. As indicated in Table 3.1, three of the twelve commodity groups belong to Class 2, i.e., *Gaseous hydrocarbons (HC)*, *Other products*, and *Fertilizers*. The volume (in million tons) and the hazmat proportion, respectively, of the three groups are: 5.54, 100%; 4.44, 28.08%; and, 4.07, 6.31%. Thus, for Class 2, the weighted average volume is $\{(5{,}547{,}149 \times 1.00) + (4{,}436{,}784 \times 0.2808) + (4{,}069{,}247 \times 0.0631)\} = 7{,}049{,}767$, and the weighted average percent is: $\frac{7{,}049{,}767}{285{,}881{,}951} = 2.47\%$. Computations for other classes can be performed similarly, which when aggregated would yield a hazmat volume of 43.39 million tons, and a percent of 15.18%.

As noted earlier, the aggregate rail traffic data does not provide any information about the type of commodity for the other two revenue generating freight sources, and also for the "without specified commodity" segment of non-intermodal traffic. In an effort to be conservative, and consistent, we assume that the hazmat percentage for each of these categories is 15.18% (i.e., same as that for commodity-based groups). The corresponding hazmat volumes are depicted in the last three rows of Table 3.1, and are shaded to reflect the above extrapolation.

The aforementioned hazmat estimation technique was applied to each of the seventeen years depicted in Figure 3.2, which in turn yields Figure 3.3 (a). It is clear from Table 3.1 and Figure 3.3 (a) that, based on our conservative estimate, over 50 million tonnes of hazmat were transported on the Canadian railroad network in 2015, and the same information is depicted by various hazmat classes in Figure 3.3 (b). Note that the relative share of Class 3 has increased substantially in recent years and it accounts for more than one-third of total hazmat freight in Canada, i.e., 5.74% of the 15.18% of hazmat traffic. Concurrently, the relative share of Class 4 has decreased from 28.45% in 1999 to 13.47% in 2015, and that for Class 5 is too small to be even visible on the relative scale. Finally, though 25% of the hazmat freight is classified unidentified, it is reasonable to assume the presence of hazmat from various classes in this group. Hence, either one could divide the unidentified based on the relative share of identified hazmat classes, or discard that part of the dataset. Consequently, we focus on only three hazmat classes, *viz.*, 2, 3, and 8. It is clear from Figure 3.4 (a) that these three classes have accounted for at least 60%, and over 80% in the most recent years. The upward trend is also underlined in Figure 3.4 (b), which exhibits a more than doubling of the volume over the seventeen-year period. It is important that these numbers are going to go up further in 2016, and beyond, primarily driven by the higher volume of Class 3 hazmat, i.e., crude oil and refined petroleum products. A more detailed analysis of crude oil rails shipments in Canada is provided in Section 3.5.

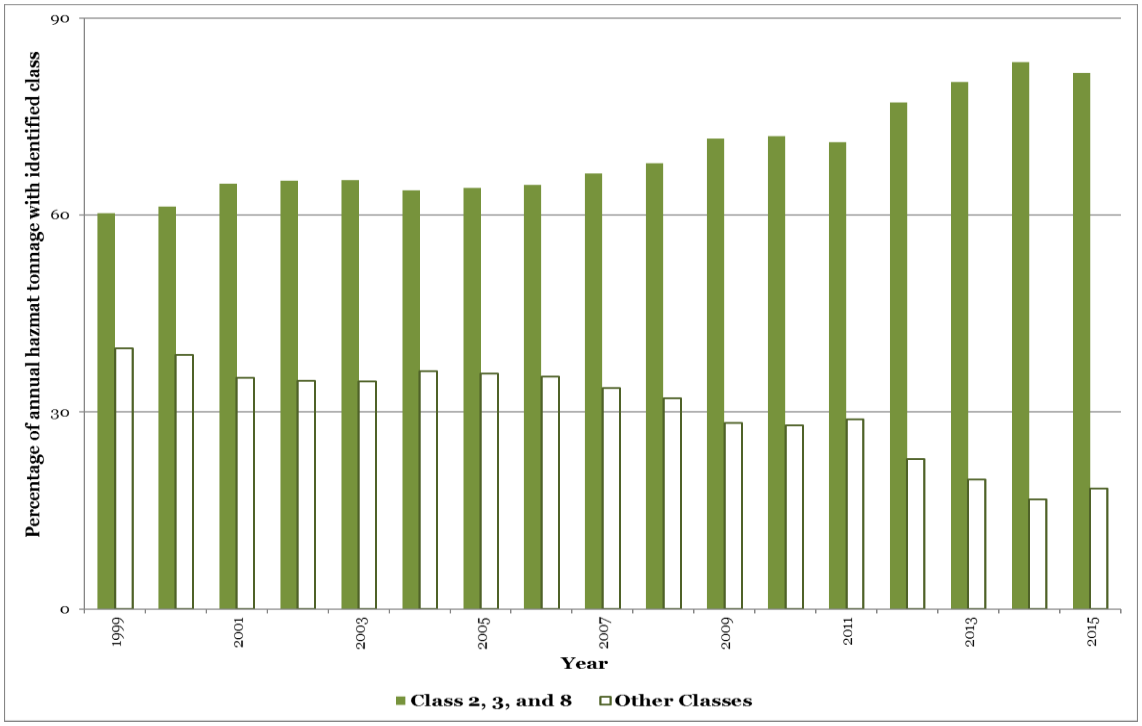


(a) hazmat traffic

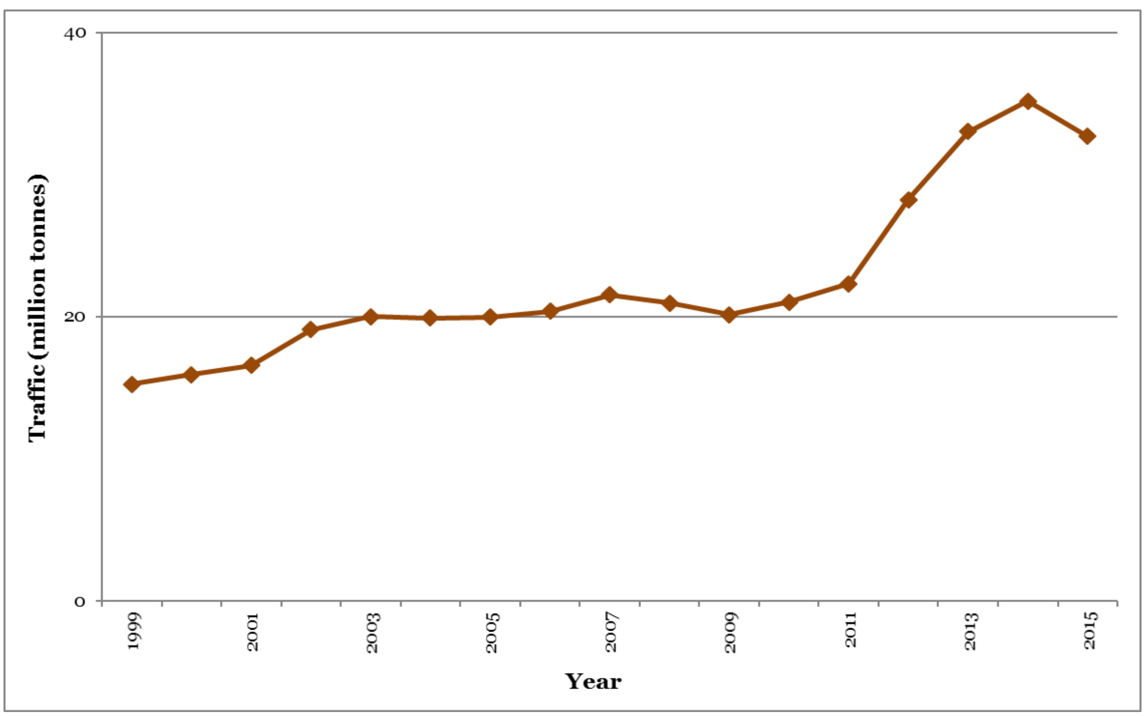


(b) percent of hazmat by class

Figure 3.3. Estimate on railroad network in Canada from 1999 to 2015



(a) Comparison



(b) Growth

Figure 3.4. Estimate of classes 2, 3 and 8 on railroad network in Canada from 1999 to 2015

3.3.3 Generate a Time-series Model & Develop Forecast

The first two steps of the proposed methodology entailed analysis of historical data, which provided sufficient insights to develop forecasts of possible volumes in the future periods. This step is important because it will throw light on the possible changes in traffic in the future, and the need to update hazmat risk and emergency response preparedness. Hence, we estimate rail hazmat traffic from 2016 to 2030 for the three dominant classes of hazmat discussed in the previous section. It should be evident from Figure 3.3 (b) and Figure 3.4 (b) that hazmat volume for all the three dominant classes increased from 1999 to 2015. As indicated earlier, the increase was the highest for Class 3, and it is expected that the need to move crude oil to refineries will continue to drive volume until about 2018, after which some of the proposed pipeline projects are expected to go online. Thus, any forecast for this hazmat class should incorporate the impact of new pipelines (See Section 3.5). On the other hand, for the other two classes, we assume that past trends will not change drastically over the next fifteen years. However, for all three classes, we make use of scenarios to account for uncertainty in the predictive estimates.

Given the temporal data at our disposal, a linear or nonlinear forecasting model that uses time-series would be appropriate. It is pertinent that the empirical research conducted by Makridakis et al. (1998) demonstrated that nonlinear models do not provide more accurate forecasts than linear models. In fact, our own experiments reveal that linear trend equation model yielded more reasonable forecasts than a variety of nonlinear models[4]. For

[4] We experimented with 5 types of nonlinear forecasting models: quadratic; cubic; quartic; logarithmic; and, exponential. Though some of them fit the historical dataset well, but failed to provide reasonable forecasts.

instance, the quadratic forecasting model estimated declining volume for Class 8 only because of the equation of a downward parabola, i.e., the first part of the parabola comprised of the historical data until 2015, and the forecasts for the fifteen periods was on the second part that was pointing downward.

		Class 2			**Class 3**			**Class 8**		
Year	$\boldsymbol{x}$	$y_t = 6125332 + 89705x$			$y_t = 3912589 + 922706x$			$y_t = 2894721 + 47555x$		
		Low	*Ref*	*High*	*Low*	*Ref*	*High*	*Low*	*Ref*	*High*
2016	18	6.29	7.74	9.19	14.67	20.52	26.38	3.27	3.75	4.23
2017	19	6.36	7.83	9.30	15.48	21.44	27.41	3.31	3.80	4.28
2018	20	6.42	7.92	9.42	16.29	22.37	28.44	3.35	3.85	4.34
2019	21	6.48	8.01	9.54	17.09	23.29	29.49	3.39	3.89	4.40
2020	22	6.53	8.10	9.66	17.88	24.21	30.55	3.42	3.94	4.46
2021	23	6.59	8.19	9.79	18.66	25.13	31.61	3.46	3.99	4.52
2022	24	6.64	8.28	9.92	19.43	26.06	32.68	3.50	4.04	4.58
2023	25	6.69	8.37	10.04	20.20	26.98	33.76	3.53	4.08	4.64
2024	26	6.74	8.46	10.17	20.96	27.90	34.84	3.57	4.13	4.70
2025	27	6.79	8.55	10.30	21.72	28.83	35.93	3.60	4.18	4.76
2026	28	6.84	8.64	10.44	22.47	29.75	37.03	3.63	4.23	4.82
2027	29	6.88	8.73	10.57	23.22	30.67	38.13	3.67	4.27	4.88
2028	30	6.93	8.82	10.70	23.96	31.59	39.23	3.70	4.32	4.94
2029	31	6.97	8.91	10.84	24.69	32.52	40.34	3.73	4.37	5.01
2030	32	7.02	9.00	10.98	25.43	33.44	41.45	3.76	4.42	5.07

Table 3.2. Forecast of traffic (million tonnes) for various scenarios for classes 2, 3 and 8

Table 3.2 depicts the linear trend equation and the forecasts for the three hazmat classes from 2016 to 2030. The *'Ref'* column lists the estimate for the specific year, whereas the *'Low'* and *'High'* indicates the lower and upper limit, respectively, of the 95% prediction interval created around *'Ref'*. For Class 2, R^2=0.37 and SE (standard error) =605,270; The corresponding values for Class 3 are R^2=0.79 and SE=2,449,418 and for Class 8 are R^2 = 0.61 and SE=199,531. As alluded, we generate a bound around each

While quartic provided negative values of forecast, the logarithmic model failed to account for recent spurt in traffic thereby underestimating volume.

individual estimate to account for uncertainties stemming from the model itself, politics, and the environment. For instance, for Class 2, since $x = 1$ represents year 1999, substituting $x = 18$ in the linear trend equation as the number of period for year 2016 will yield 7,740,030 as the *'Ref'* estimate. At 95%, the t-value for the given dataset is 2.13145 while the standard deviation of the forecast for year 2016 is 678,700, which together results in a margin of error of 1,446,614. The margin of error is subtracted from *'Ref'* to yield *'Low'* value of 6,293,416, and added to *'Ref'* to yield *'High'* value of 9,186,644. In mathematical terms, $PI: \hat{y} \pm ME = \hat{y} \pm t^*_{n-2} \times SE \times \sqrt{1 + \frac{1}{n} + \frac{(x* - \bar{x})^2}{(n-1)\, s_x^2}}$. The first term is the predicted value, i.e., *'Ref'*, and the second term is the margin of error (ME), which is calculated based on a t-score with $n - 2$ degrees of freedom (n: sample size), a standard error (SE), the desired value of x to predict ($x *$), mean of explanatory variables ($\bar{x}$), and standard deviation of the sample (s_x). Note that other values in the table and prediction intervals at other levels can be similarly generated.

Though we expect the prediction intervals to contain the impact of uncertainties from the indicated sources for classes 2 and 8, the same cannot be assumed for Class 3. This is because some new pipelines are expected to become available sometime after 2018, and we reckon that the crude oil shipments on railroad will be impacted in subsequent years and thus appropriate adjustments should be made to the forecasts. The evaluation of such impact on crude oil rail shipments is illustrated in Section 3.5

3.3.4 Dis-aggregate Hazmat Traffic

The aggregate hazmat traffic information, for a given hazmat class, will be dis-aggregated to rail-link and rail-yard volumes using a three-stage process under this step of the methodology: *first*, determine the origins for shipments; *second*, determine destinations for shipments; and *third*, solve a transhipment model. For expositional reasons, we will delineate the details only for Class 2, and note that information for the other two classes could be extracted similarly. Furthermore, we will demonstrate the process using the *'Ref'* estimate for year 2016 Class 2 from Table 3.2, and then provide a comparative assessment with the remaining 44 scenarios (i.e., three each from 2017 to 2030, plus the two scenarios for 2016).

Class 2 hazmat is primarily composed of the three commodity groups: gaseous hydrocarbons; other refined petroleum and coal products; and, fertilizers (Table 3.1). However, on closer inspection of the historical dataset, it was noticed that gaseous hydrocarbons accounted for around 80% of the traffic for this class. Thus, to facilitate discussion and demonstrate the dis-aggregation process, we assume that the entire Class 2 is composed of only gaseous hydrocarbons, but note that the process could be replicated for the other two commodity groups. In addition, based on the existing literature, we assume that gaseous hydrocarbons are moved on the railroad network in the form of propane and butane, and collectively grouped under Liquefied Petroleum Gas, i.e., LPG (WLPGA, 2016). It is estimated that, worldwide, approximately 60% of LPG is recovered during the extraction of natural gas and oil from the earth, and the remaining 40% is produced during the refining of crude oil. Consequently, we assume that the natural gas fields (or gas pools) are the origins for *direct* production of LPG, and oil refineries for *indirect* production. In the absence of precise information, we make a reasonable assumption that Class 2 hazmat are consumed at the major population centers and are exported to the United States (i.e., destinations).

Origins: Table 3.3 provides a snapshot of the origins of Class 2 hazmat, and is organized by the direct and indirect sources of production. According to Natural Resources Canada, there is a total of seventeen direct sources of production in Canada (NRC, 2015a), but we consider only the relevant fourteen gas fields/ gas pools for which gas production data is available through National Energy Board (NEB, 2016a). To facilitate estimation, we make the following assumptions: *first*, if a gas field is spread over a large geographical area, designate the center of such a region as the point of origin for Class 2 hazmat; *second*, for the offshore gas field in NS, treat the closest onshore location (i.e., Halifax) as the point of origin; and *third*, in the absence of field-level information for provinces/ territories with multiple fields, divide the volume equally amongst the fields. A two-stage process was used to develop the 2016 forecast: *first*, the marketable natural gas production database from 2000 to 2015 was analyzed to identify a trend, and develop a linear forecasting equation; and *second*, the forecast equation was suitably adapted to incorporate scenarios wherein the production amount could be non-positive. The adapted linear model was used to estimate gas production at the fourteen fields, which in turn was converted into a percent and volume in tonnes. For instance, the five gas fields in AB is projected to produce 257,639 cubic meters of gas that would account for 64.61% of the total production, and is equivalent to a total of 3,000,655 tonnes equally divided amongst the five gas fields. Other numbers can be interpreted similarly. Note that the 2016 *‘Ref’* estimate is 7,740,030 from Table 3.2, and 60% of which results from direct production, i.e., $0.6 \times 7{,}740{,}030 = 4{,}644{,}018$ tonnes.

	Province/ Territory	Gas Field/ Refinery Identity	Gas Production / Refinery Capacity		Percent of sub-total	Volume (tonnes)
Gas Fields (or gas pools)	AB	High Level	cubic meters per day	257,639	0.6461	3,000,655
		Beaverlodge				
		Westlock				
		Stettler				
		Brooks				
	BC	Prespatou		125,709	0.3153	1,464,098
	NB	Sussex		63	0.0001	729
	NS	Sable Island (Halifax)		4,646	0.0117	54,111
	ON	Hamilton		251	0.0006	2,918
	SK	Eston		10,240	0.0257	119,259
	Territories	Fort Liard		193	0.0005	2,805
		Colville Lake				
		Tuktoyatuk				
		Eagle Plains				
Total for Gas Fields					***1.0000***	***4,644,018***
Oil refineries	AB	Imperial Oil – Edmonton	barrels per day	187,000	0.1011	313,117
		Suncor Energy –Edmonton		142,000	0.0768	237,768
		Shell Canada –Fort S		100,000	0.0541	167,443
	BC	Husky Energy –Prince G		12,000	0.0065	20,093
		Chevron Canada -Burnaby		55,000	0.0297	92,093
	NB	Irving Oil –Saint John		313,000	0.1693	524,095
	NF	North Atlantic –Come by C		115,000	0.0622	192,559
	ON	Imperial Oil –Sarnia		121,000	0.0654	202,605
		Imperial Oil –Nanticoke		112,000	0.0606	187,536
		Suncor Energy –Sarnia		85,000	0.0460	142,326
		Shell Canada –Corunna		75,000	0.0406	125,582
	QC	Valero –Levis		265,000	0.1433	443,723
		Suncor Energy –Montreal		137,000	0.0741	229,396
	SK	Federated Co-op –Regina		130,000	0.0703	217,675
Total for oil refineries					***1.000***	***3,096,012***

Table 3.3. Origins and Class 2 hazmat traffic for ‘Ref’ scenario for 2016

Finally, the list of active refineries and their respective production capacities made available via Natural Resources Canada (NRC, 2015b), were used to estimate the remaining 40% of Class 2 hazmat (i.e., 3,096,012 tonnes). For instance, the largest refinery in the country, i.e., Irving Oil in NB, has a production capacity of 313K barrels per day, which amounts of 16.93% of the capacity in Canada, and would thus produce 524,095 tonnes of Class 2 hazmat. Other values could be interpreted similarly, and all of them when summed would yield the sub-total for Class 2 hazmat traffic originating at the refineries.

Destinations: Table 3.4 depicts a high level overview of the destinations, and the proportion and the volume of overall traffic to be received at each location. An analysis of the propane and butane shipments statistics released by National Energy Board (NEB, 2016b) revealed two major consumption points for the Class 2 hazmat: major population centers; and, exports only to the United States. For brevity, we assume that Class 2 hazmat are only demanded at the top twenty population centers in Canada, and note that the approach can easily be extended to other population centers. In addition, finer analysis of the NEB statistics from 1999 to 2015 threw light on the volume of LPG (i.e., propane plus butane) being exported to the United States, which was then compared to the Class 2 hazmat volume on railroad network in Canada over the same time period (i.e., Table 3.1). The export percentage ranged from 38% to 54% over the period in question for an average of 44.56%, which is the forecast proportion for 2016. Hence, 55.44% of Class 2 hazmat (i.e., $0.5544 \times 7{,}740{,}030 = 4{,}291{,}073$ tonnes) would be consumed at the top twenty centers in proportion relative to their 2015 population census estimate. For instance, Toronto would be the destination for 26.52% of the non-export Class 2 hazmat shipments, and together with Montreal and Vancouver consume more than half of these shipments.

	Provinces	**Population Centers / Export Terminals**	**Population Estimate/ Export Zones**	**Percent of sub-total**	**Volume (tonnes)**
Population Centers	AB	Calgary	1,439,756	0.0623	267,329
		Edmonton	1,363,277	0.0589	253,128
	BC	Vancouver	2,504,340	0.1084	464,997
		Victoria	365,291	0.0158	67,826
	MB	Winnipeg	793,428	0.0343	147,321
	NS	Halifax	417,847	0.0181	77,584
	NF	St. John’s	214,285	0.0093	39,788
	ON	Hamilton	771,703	0.0339	143,287
		K-C-W	511,319	0.0213	94,940
		London	506,418	0.0219	94,030
		Ottawa –Gatineu	1,332,001	0.0576	247,321
		Oshawa	388,956	0.0168	72,220
		St. C. -Niagara	408,222	0.0177	75,797
		Toronto	6,129,934	0.2652	1,138,184
	QC	Montreal	4,060,692	0.1757	753,975
		Quebec	806,359	0.0349	149,722
		Sherbrooke	214,485	0.0093	39,788
	SK	Regina	241,422	0.0105	44,826
		Saskatoon	304,975	0.0132	56,627
Total destined for Population Centers				***1.0000***	***4,291,073***
Export Terminals	AB	Coutts	PADD 4	0.0457	78,766
	BC	Kingsgate			78,766
		Matsqui	PADD 5	0.1304	224,950
		Mission			224,950
	MB	Emerson (2 terminals)	PADD 2+ PADD 3	0.5315+ 0.0048= 0.5363	2×205,511=411,022
		Middlebro			205,511
	ON	Niagara Falls	PADD 1	0.2876	330,642
		Fort Erie			330,642
		Fort Frances	PADD 2 PADD 3	**	
		Sault Ste. Marie			
		Sarnia			
		Windsor			
	QC	Cantic	PADD 1	*	
	SK	Estevan	PADD 2 PADD 3	**	
		Bienfait			
Total destined for Export Terminals				***1.0000***	***3,448,957***
where, * implies same entry as for PADD 1, and ** that for PADD 2 and 3.					

Table 3.4. Destinations and Class 2 hazmat traffic for ‘Ref’ scenario for 2016

The export data for Class 2 hazmat, on the other hand, is organized by regions in the United States called PADD, i.e., petroleum administration for defense districts: PADD 1 refers to East Coast; PADD 2 to Midwest; PADD to Gulf Coast; PADD 4 to Rocky Mountain; and PADD 5 to West Coast, Alaska, and Hawaii (US EIA, 2012). As indicated earlier, the total volume of Class 2 hazmat going to the five PADD in 2016 is forecast at $0.4456 \times 7{,}740{,}030 = 3{,}448{,}957$ tonnes. Exports to the United States could happen only via the export terminals of the two railroad companies in Canada (i.e., CN and CP export terminals), and they are listed in Table 3.4. To arrive at the reasonable allocation for the five PADD regions, we analyzed the NEB database to determine the average of the relative proportion of export volume being shipped to each region from 1999 to 2015, which was then assumed as the forecast for 2016.

It is relevant that more than four-fifths of the shipments are destined for PADD 1 and PADD 2 regions of the United States, and entails the allocation of most of the available terminal capacity. In addition, 28.76% or 991,926 tonnes of Class 2 hazmat are destined to PADD 1 via the three export terminals in ON and QC, wherein each is handling one-third, i.e., 330,642 tonnes. Other values can be interpreted similarly. Note that PADD 3 does not share a border with Canada, and thus relevant traffic is assigned to the export terminal handling shipments for PADD 2 because the former region is located below the latter.

Transhipment Problem: Having ascertained the production/ consumption locations and the respective volumes, the next step is to determine the best path to send shipments from the origins to the destinations from a carrier's perspective. Note that since a path is comprised of rail-links and rail-yards, such path determination would also enable us to estimate the hazmat volume at each traversed link and yard. We assume that carriers, i.e., railroad companies, would like to focus on cost, and hence solve a transhipment problem that seeks to minimize system wide transportation cost (Hillier and Lieberman, 2014). We assume that each railroad company can use tracks owned by the other company, which particularly makes sense in Canada since the two major railroad companies in the country have numerous agreements to share tracks. The problem will make use of the processed data developed so far. For instance, for the year 2016, the requisite inputs for hazmat Class 2 are: origins and respective supply amounts from Table 3.3; destinations and respective demand numbers from Table 3.4; and, length of rail-links would be a proxy for transportation cost. Thus, for a directed and connected railroad network with *n* nodes, the decision variable is:

x_{ij} = volume of traffic on directed arc $(i, j), i \to j$

and the parameters are:

l_{ij} = length of directed arc $(i, j), i \to j$

b_i = net flow generated at node i

(P1)

$$Minimize\ \sum_{i=1}^{n}\sum_{j=1}^{n} l_{ij}x_{ij} \tag{3.1}$$

subject to

$$\sum_{j=1}^{n} x_{ij} - \sum_{j=1}^{n} x_{ji} = \begin{cases} +b_i & if\ i\ is\ a\ supply\ node \\ -b_i & if\ i\ is\ a\ demand\ node \\ 0 & if\ i\ is\ a\ transhipment\ node \end{cases} \qquad \forall i \tag{3.2}$$

$$x_{ij} \geq 0 \qquad \forall (i, j) \tag{3.3}$$

(P1) seeks to find the minimum cost solution such that demands at the consumption points are satisfied using the available supply at the production locations. It is an uncapacitated version of the minimum cost flow problem with transhipment attribute, which could allow us to determine the path between different origins and destinations for each hazmat individually, and then aggregate the resulting values to get a picture of the flow through each link and each yard. However, we solve the problem for hazmat Class 2 and for 2016 *'Ref'* problem instance in the next section and note that it could be solved similarly for other hazmat classes.

3.3.5 Depict Solution

(P1) requires 3 pieces of input parameters: *first*, origin locations and available supply; *second*, demand volume and locations; and *third*, the railroad network in Canada. Recall that we already have information about the first two parameters through Table 3.3 and Table 3.4, and thus elaborate on only the last parameter. Over 95% of rail freight is moved by two railroad operators in Canada, i.e., Canadian National (CN) and Canadian Pacific (CP). Hence, we chose to focus on them and employed a three-step procedure to recreate their networks: *first*, rail-links and rail-yards were identified, named, and recorded in Google Earth Pro, version 7.1.5.1557 (Google Inc., 2016); *second*, the resulting map was imported into ArcGIS, version 10.3.1 (ESRI, 2016); and *third*, individual layers of CN and CP were combined to obtain the Canadian railroad network, which could now be consulted to determine the length of any link or coordinates for any yard. It is important that the resulting network is a directed graph with 464 nodes and 940 arcs, and that origins/ destinations that are not on the railroad network are projected to the closest point on the network such that a virtual node is created.

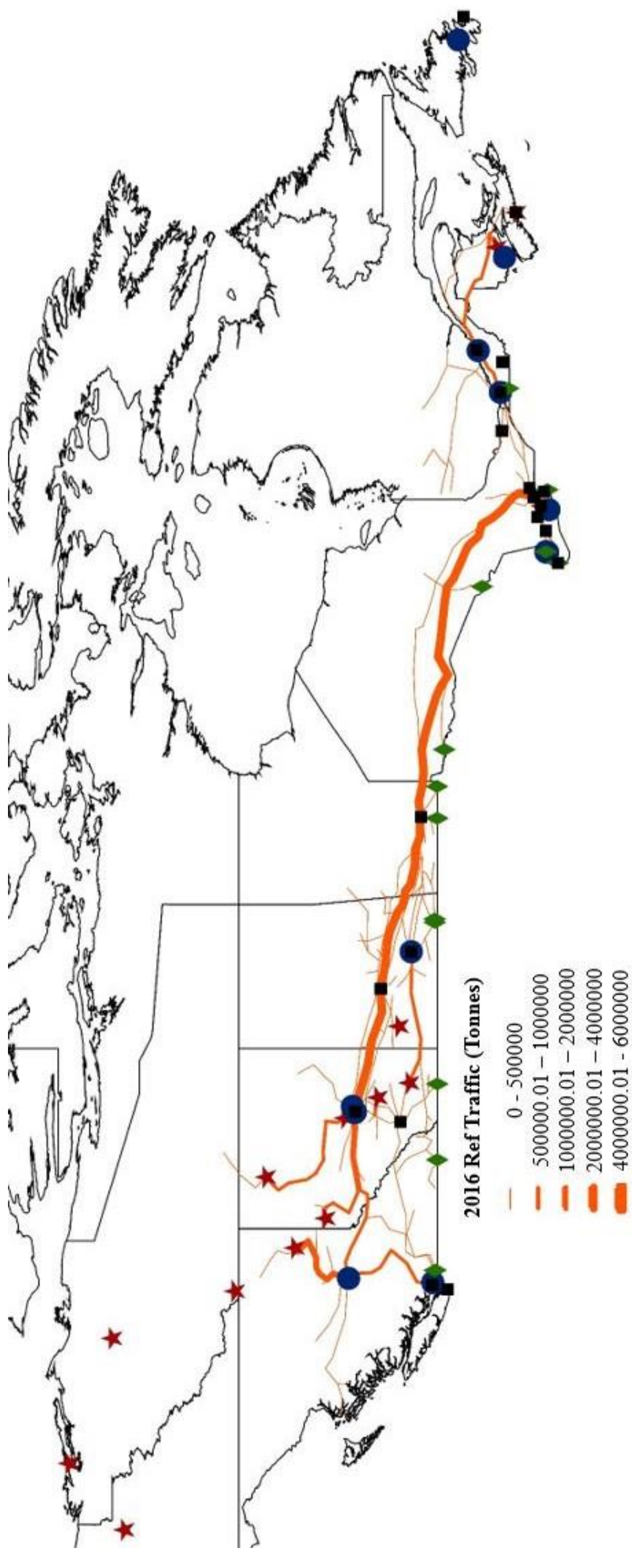


Figure 3.5. Annual Class 2 hazmat traffic for 2016 'Ref' problem instance

(P1) was solved using CPLEX 12.6.2.0 (IBM, 2016), and the resulting solution is depicted in Figure 3.5. Note that origins are represented using stars for gas fields and circles for refineries, whereas destinations are indicated using squares for population centers and diamonds for export terminals. For visual ease, we map the hazmat volume traversing a link (or yard) by varying the thickness of the line, which is also highlighted through the use of legends. It should be evident that, given the reluctance of railroad operators, such dis-aggregated information could be extremely useful to both emergency response providers and to the researchers in the insurance industry and in the academia.

3.4 Solution Discussion & Analysis

In this section, we will further analyze the solution for the 2016 *'Ref'* problem instance to identify the high-traffic segments of the railroad network in Canada, and then comment on any changes vis-à-vis forecast for other scenarios. Finally, we will comment on any distinct attributes noticed while analysing the solutions for the other two hazmat classes.

3.4.1 Identifying High-traffic Points in the Network

The link and yard level values used to generate Figure 3.5 were sorted for Class 2 hazmat volume, and the top twenty under each are reported in Table 3.5. It is not a co-incidence that the top twenty links and yards are located in the three western provinces, i.e., AB, MB, and SK. This is because these provinces produce more LPG (a surrogate for Class 2 hazmat) than the intra-provincial demand, and thus the surplus is shipped to the major population centers in the eastern provinces. Clearly, the indicated nodes and links have high-traffic, and appropriate emergency preparedness and adequate insurance tools should be in place. However, of the top hundred links and nodes, approximately one-quarter of each is located in ON that has the highest population. Though we provide a snapshot of province level statistics in the next sub-section, it should be evident that a volume-based ranking of nodes and links could be developed for each province.

3.4.2 Comparison with Other Scenarios

Table 3.6 depicts the objective function values of the forty-five scenarios, the 2016 *'Ref'* problem instance and the remaining forty-four introduced in Table 3.2.

<table>
<tr><th colspan="3">Rail-Yards/ Nodes</th><th colspan="3">Rail-Links/ Arcs</th></tr>
<tr><th>Name</th><th>Province</th><th>Volume (tonnes)</th><th>Name</th><th>Province</th><th>Volume (tonnes)</th></tr>
<tr><td>Winnipeg</td><td rowspan="2">MB</td><td rowspan="2">3,461,115</td><td>Portage la Prairie -Winnipeg</td><td>MB</td><td>3,461,115</td></tr>
<tr><td>Portage la Prairie</td><td>Saskatoon-Colonsay</td><td rowspan="7">SK</td><td rowspan="12">3,099,204</td></tr>
<tr><td>Saskatoon</td><td rowspan="6">SK</td><td>3,155,831</td><td>Colonsay -Lanigan</td></tr>
<tr><td>Colonsay</td><td rowspan="11">3,099,204</td><td>Lanigan -Wynward</td></tr>
<tr><td>Lanigan</td><td>Wynward –Foam Lake</td></tr>
<tr><td>Wynyard</td><td>Foam Lake -Yorkton</td></tr>
<tr><td>Foam Lake</td><td>Yorkton –Bredenbury</td></tr>
<tr><td>Bredenbury</td><td>Bredenbury -Marchwell</td></tr>
<tr><td>Harrowby</td><td rowspan="3">MB</td><td>Marchwell -Harrowby</td><td rowspan="5">MB</td></tr>
<tr><td>Solsgirth</td><td>Harrowby -Solsgirth</td></tr>
<tr><td>Minnedosa</td><td>Solsgirth -Minnedosa</td></tr>
<tr><td>Marchwell</td><td rowspan="2">SK</td><td>Minnedosa -Gladstone</td></tr>
<tr><td>Yorkton</td><td>Gladstone –Portage la Prairie</td></tr>
<tr><td>Gladstone</td><td>MB</td><td>Edmonton -Ryley</td><td rowspan="3">AB</td><td rowspan="7">2,615,089</td></tr>
<tr><td>Edmonton</td><td rowspan="2">AB</td><td rowspan="6">2,615,089</td><td>Ryley -Viking</td></tr>
<tr><td>Wainwright</td><td>Viking -Wainwright</td></tr>
<tr><td>Biggar</td><td rowspan="3">SK</td><td>Wainwright -Artland</td><td rowspan="4">SK</td></tr>
<tr><td>Winter</td><td>Biggar -Saskatoon</td></tr>
<tr><td>Artland</td><td>Scott -Biggar</td></tr>
<tr><td>Ryley</td><td>AB</td><td>Winter -Scott</td></tr>
</table>

Table 3.5. Top 20 nodes and links for hazmat Class 2 for 2016 'Ref' problem instance

For each year, the *'Ref'* problem instance solution is containing between the intervals defined by the *'Low'* and *'High'*, wherein the latter incorporate possible uncertainties. It should be clear that the prediction intervals exhibit non-linearity (Table 3.2), and this phenomenon is accentuated because the relative share of production locations are changing over time. More specifically, the production levels at the gas fields are expected to change over time, which in turn could be expected to impact the relevant nodes and links (i.e., high-traffic segments of the network). Note that, for a given year, the volume-based ranking of nodes and links will be preserved for all the three problem settings, and hence we attempt to investigate possible changes by analysing only hazmat traffic related to the *'Ref'* problem instances highlighted in Table 3.6.

Year	*Low*	*Ref*	*High*
2016	10,134,381,762	12,463,894,244	14,793,391,628
2017	10,322,916,950	12,715,295,864	15,107,674,340
2018	10,509,507,290	12,969,149,552	15,428,819,305
2019	10,696,106,583	13,227,763,712	15,759,424,691
2020	10,883,232,838	13,491,626,238	16,100,017,169
2021	11,071,111,219	13,760,852,728	16,450,610,343
2022	11,259,868,802	14,035,530,956	16,811,202,335
2023	11,447,604,850	14,313,112,982	17,178,641,698
2024	11,614,279,626	14,568,351,408	17,522,438,882
2025	11,781,009,983	14,827,424,511	17,873,829,942
2026	11,948,258,527	15,090,765,235	18,233,253,525
2027	12,116,128,216	15,358,390,273	18,600,629,077
2028	12,284,827,259	15,630,440,229	18,976,065,855
2029	12,454,578,563	15,907,139,189	19,359,692,981
2030	12,621,304,563	16,183,144,197	19,744,978,419

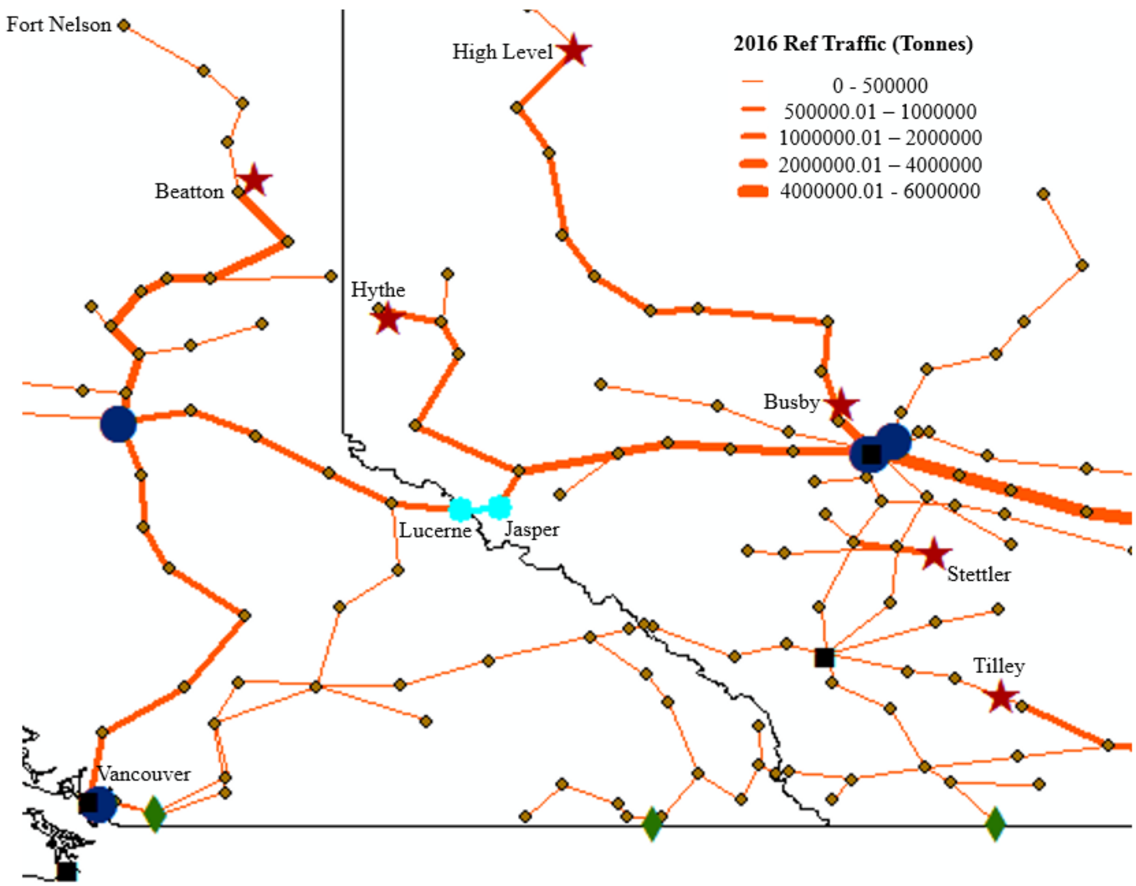
Table 3.6. Objective function value (tonne-km) for the 45 problem instances of Class 2 hazmat

Table 3.7 provides a snapshot of the top hundred nodes (*N*) and links (*L*), with the highest traffic, for the *'Ref'* problem instances highlighted in Table 3.6. It is clear that the changes in production level at the gas fields are going to impact the relative riskiness of several nodes and links in the AB, and to a certain extent in MB and SK. In addition, owing to the expected shifts, it is expected that demand in BC will increase substantially thereby bringing several additional nodes and links into the high-traffic category. On the other hand, the number of high-traffic nodes and links in the eastern provinces would remain largely unchanged, except marginal increment in ON.

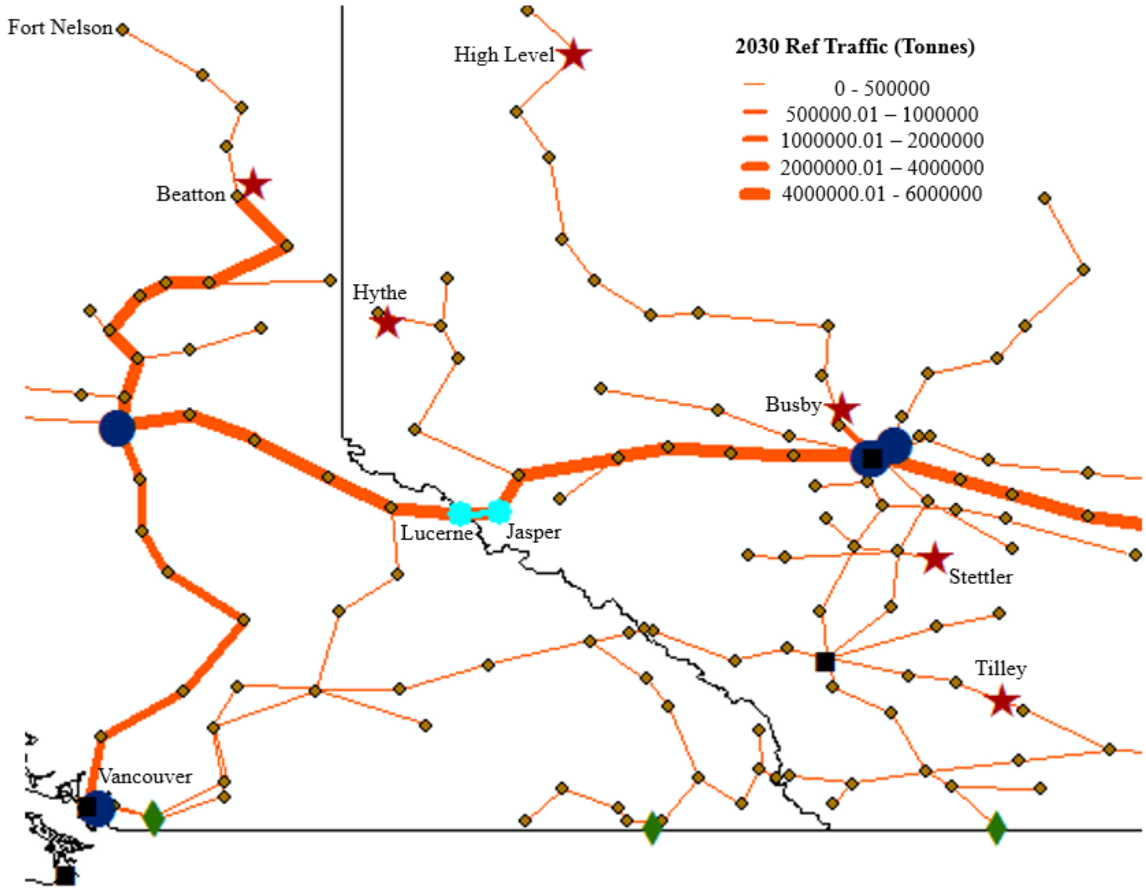
Year	**AB**		**BC**		**MB**		**SK**		**ON**		**QC**		**NB**	
	N	***L***	***N***	***L***	***N***	***L***	***N***	***L***	***N***	***L***	***N***	***L***	***N***	***L***
2016	23	22	16	15	8	9	12	15	27	26	8	8	6	5
2018	18	20	21	20	8	9	12	12	27	26	8	8	8	5
2021	15	17	21	20	8	9	12	12	27	26	8	8	9	8
2024	15	17	21	20	8	9	12	12	27	26	8	8	9	8
2027	12	14	21	20	10	11	12	12	28	27	8	8	9	8
2030	11	14	21	20	10	11	12	12	29	27	8	8	9	8

Table 3.7. Top hundred nodes and links for specific 'Ref' problem instances

The decoded solutions of the forty-five problem instances depicted in Table 3.6 reveal the following. *First*, the top twenty-three nodes and top eighteen links will not change in ranking over the indicated period. *Second*, other nodes and links would move up or down on the volume-based ranking list depending on the changing share of the 14 gas fields. For instance, the nodes at Jasper and Lucerne will become more active and thus move up, whereas the one at Busby will become less risky. As a result, the links connecting Jasper and Lucerne would experience increased hazmat traffic, while that linking Busby to Edmonton would register a decrease.



(a) 2016 'Ref' problem instance



(b) 2030 'Ref' problem instance

Figure 3.6. Class 2 hazmat traffic around *Jasper* node and *Jasper-Lucerne* link

For expositional reasons, we will focus on a specific node and link, and note that other locations in the network could be similarly analyzed. The decoded solution for *'Ref'* problem instances for 2016 and 2030 was analyzed, and the traffic around Jasper was coded into Arc GIS to generate Figure 3.6. The decoded solution told us that the supply at Beatton would increase from 1.4 million tons in 2016 to 3.4 million tons in 2030, whereas supply will decrease at the following seven locations, i.e., High Level, Hythe, Busby, Stettler, Tilley, Fort Nelson, and Hay River. On the other hand, demand in BC is expected to increase over the same period. However, the increased supply at Beatton would be used to meet part of the higher demand in BC thereby reducing the inflow from AB, and continuing through AB to the eastern part of the country. Figure 3.6 (a) depicts the current situation, whereas the thicker edges between Jasper and Lucerne in Figure 3.6 (b) represent higher flow from Beatton transiting through AB for eastern provinces. In closing, we note that changes in hazmat volume should be expected in other parts of the network, which in turn would impact the relative volume-based ranking of the 464 nodes and 940 links.

3.4.3 Comments About Hazmat Classes 3 & 8

Class 3 hazmat, in general, corresponds to four specified group categories as indicated in Table 3.1. However, based on the available data, it seems reasonable to approximate the entire class using a single category, i.e., *fuel oils and petroleum products*, or more specifically *crude oil*. Thus, the major production location would be AB followed by SK, while the refineries and the export terminals would be the destinations. *Excluding* the impact of pipeline projects, our analysis of the 2016 *'Ref'* problem instance, and the remaining scenarios revealed two things. *First*, just like Class 2 hazmat, traffic will flow from west to east. *Second*, the top twenty high-traffic nodes and links are located in MB, SK, and ON. Note that not even a single node or link resides in AB, which is interesting, but in part could be explained by the location of refineries and export terminals in eastern Canada. An extensive analysis specific to crude oil is conducted in Section 3.5, which also includes the impact of proposed pipeline projects.

Class 8 hazmat is primarily denoted by *sulphuric acid* and *potash*, wherein the former accounts for more than 90% of the volume and hence could represent the class. Note that sulphuric acid is manufactured in acid plants, which are mostly located in ON and QC. On the other hand, the destination locations are fertilizer plants and the export terminals that are spread across the country. Hence, unlike the other two hazmat classes, no clear traffic direction could be identified. Each of the top twenty nodes and links is located either in ON or QC. In addition, given the assumption of proportionately linear changes in both the production and consumption patterns, the volume-based ranking of nodes and links did not change from 2016 to 2030.

3.5 Crude Oil Rail Shipments in Canada[5]

In this section, we conduct a focused analysis on the railroad transportation of an important hazmat commodity, i.e., crude oil, in Canada.

Crude oil, the lifeblood of the modern economy, accounts for the largest share of the total world primary energy demand and is forecasted to remain so until 2040 (IEA, 2016). Canada has the third largest reserves, and is the sixth largest global producer of crude oil. In 2016, crude oil production totalled 3.85 million barrels per day in Canada, and is expected to increase to 5.12 million barrels per day by 2030 (CAPP, 2017). Approximately 94% of the crude oil production, both from conventional technique and from oil sands, takes place in the two western provinces of Alberta and Saskatchewan, which are then transported to refineries to produce gasoline and fuel oil (NEB, 2017a). It is important that both these provinces are landlocked with no direct access to water

[5] Sections 3.5 has appeared in a journal article with following citation details:
Vaezi, A., & Verma, M. (2018). Railroad transportation of crude oil in Canada: Developing long-term forecasts, and evaluating the impact of proposed pipeline projects. *Journal of Transport Geography*, *69*, 98-111.

transportation infrastructure, and primarily relies on pipelines to access refineries located in different regions of Canada and the United States. In addition, technological advances over the past decade have facilitated vast and economically viable production of crude oil (and gases) from tight and shale formation in Alberta (Canada) and in the Dakotas (United States), thereby yielding unprecedented excess supply. However, the near capacity utilization of the existing pipelines and the lack of access to water transportation infrastructure necessitated using alternative modes of transportation to move the excess supply to refineries. The economies of scale and the ability to move to different markets in response to demand rendered railroad as a viable alternative for moving bulk crude oil shipments. As a result, in Canada, rail transport of crude oil has been steadily increasing since 2009, except a minor blip stemming from the wildfire episode in Alberta in 2016 (CAPP, 2017). Note that the preceding statistic underlines the long-term viability of this alternate transport mode, which in part was shaped by the proactive efforts of the railroad industry to respond to this opportunity by introducing unit trains that are formed at the loading terminals and travel non-stop to the refineries (i.e., demand locations), thereby eliminating the need for any intermediate yard operations and making the overall duration of a given trip much shorter.

It should be evident that unit trains provide economies of scale because a large volume of crude oil could be shipped together. However, crude oil is a hazmat commodity, and could potentially be dangerous to human life and the environment. The inherent risk associated with such trains carrying crude oil (and any other hazmat) cannot be underestimated, even given the good safety statistics of railroads (Verma and Verter, 2013), or the low probability –high consequence nature of multi railcar incidents (Sherali et al., 1997). For instance, the most catastrophic incident in recent history involved the derailment of a unit train carrying 72 railcars of crude oil from the bakken shale region of North Dakota (United States) to the largest refinery in Canada, i.e., Irving refinery in Saint John on the east coast of Canada. 63 of the derailed railcars leaked about six million litres of crude oil, and the resulting fire and explosion left 47 people dead in a small town of Lac-Mégantic in the eastern province of Quebec (TSB, 2014). In this section, we answer the following research questions. *Is it possible to mitigate network risk by developing better shipment plans? What is the likely impact of the proposed pipeline projects?* To that end, we (1) analyze historical data on crude oil shipments; (2) select appropriate models for short-term and long-term forecasting; (3) develop forecast; (4) employ different schemes to route traffic; (5) estimate crude oil traffic at various rail links and yards; and (6) evaluate the impact of proposed pipeline projects in Canada.

3.5.1 Analysing Historical Data

Figure 3.7 provides a snapshot of the monthly historical data from January 1999 to March 2017 for the non-intermodal traffic class, which is responsible for around 80% of the railroad traffic in Canada (Statistics Canada, 2017a). In general, the volume of crude oil and petroleum products has increased consistently, except for a downturn in 2016 resulting from a wildfire in Fort McMurray, i.e., home to most extensive and complex crude oil production infrastructure. However, the crude volume on railroads has since rebounded as evidenced by higher month-to-month traffic in both February and March of 2017. In fact, the crude oil shipments in March 2017 were 40% higher than that in March

2016. It is pertinent that crude oil shipments account for 78% of the shipments in the broader-category depicted in Figure 3.7 (Transport Canada, 2017), however, we make use of the aggregate number in each month in our analysis.

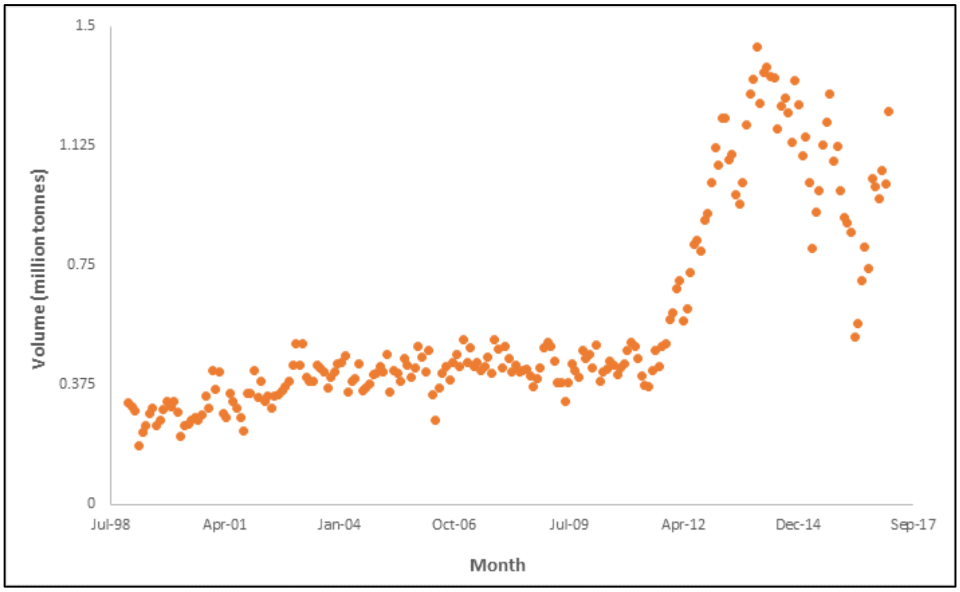


Figure 3.7. Fuel oil and crude petroleum (non-intermodal) on Canadian railroads

3.5.2 Selecting a Forecasting Model

As alluded earlier, we intend to prepare forecast from 2018 to 2030 by first selecting an appropriate forecasting model that closely mimics the pattern observed over the nineteen-year period depicted in Figure 3.7. It is important that the end date is consistent with the timeline over which forecasts are prepared by the petroleum industry (CAPP, 2017). Note that Figure 3.7 provides information until March 2017, and hence we adopt a 2-stage process to estimate the traffic for the remaining months: *first*, select a suitable short-term forecasting model to estimate the crude oil traffic for the month of April; and *second*, incorporate other pertinent information to estimate annual traffic for 2017. The resulting value will enter as a new entry thereby augmenting the dataset depicted in Figure 3.7, which will then be used to determine the best time-series model to forecast annual traffic from 2018 to 2030.

1st stage: Given the objective of preparing a forecast for the next period, we selected three popular (short-term) predictive analytics techniques, viz., weighted moving average (WMA); double exponential smoothing (DES); and, linear trend equation (LTE). While WMA prepares forecast by taking an average of a certain number of most recent actual demand (or production) values, DES is more versatile since it can also incorporate trends. Finally, LTE is based on fitting a trend line to a series of historical data points and then projecting the line into the future. Consistent with the approach in the literature, we ascertained model fit by examining various measures of error, i.e., mean absolute deviation (MAD); mean squared error (MSE); and, mean absolute percent error (MAPE) (Heizer et al., 2017). Table 3.8 provides a snapshot of the relevant computations.

Given the clear upward trend in Figure 3.7, for *WMA*, we placed more weights on recent periods. For example, under *2-WMA*, the most recent period received twice the weight that of the previous period. Similarly, for *3-WMA*, the most recent period received three times the weight of the last period, whereas the middle period received twice the weight of the last period (i.e., 3, 2 and 1, respectively, for the three periods starting from the most recent). Clearly, *2-WMA* outperforms *3-WMA* for this given dataset. For *DES*, the data points are attached weights by an exponential function to compute a smoothing average of the data and then a trend adjustment component (Heizer et al., 2017). Since the smoothing constant for the average (i.e., α) and that for the trend (i.e., β) are between 0 and 1, and in the absence of any pertinent information, we resorted to simulation to randomly generate 100 combinations. For expositional reasons, and for brevity, we depict the error measures of only ten of those combinations. It is clear from Table 3.8 that *2-WMA* is outperforming any combination of *DES*, and thus should be chosen over the latter. Finally, the error measures returned by *LTE* are much worse for short-term forecasting, and thus should be discarded. It is important to note that the forecasting accuracy of any time-series model depends on a variety of factors including the type of data (i.e., yearly, monthly, etc.) and the time horizon (Makridakis et al., 1982). Thus, it would be incorrect to claim that any specific model will outperform other models across all situations. For example, for preparing short-term forecast using the dataset depicted in Figure 3.7, *2-WMA* outperforms other approaches, and results in an estimated volume of 1.16 million tonnes for April 2017 (=[2× $\text{Traffic}_{\text{March 2017}}$ +1× $\text{Traffic}_{\text{February 2017}}$] / [2+1] = [2×1.234+1×1.003]/3).

Forecasting Method			**Mean Absolute Deviation**	**Mean Squared Error (billion)**	**Mean Absolute Percent Error**
Weighted Moving Average	2-period		55,061	5.68	10.24%
	3-period		57,061	6.25	10.52%
Double Exponential Smoothing	α=0.1	β=0.1	91,330	18.0	16.54%
		β=0.4	75,181	12.2	13.75%
		β=0.9	86,613	17.5	15.11%
	α=0.4	β=0.1	61,481	7.47	11.22%
		β=0.4	66,716	8.77	12.24%
		β=0.9	68,173	9.24	12.71%
	α=0.8	β=0.1	55,095	5.84	10.38%
		β=0.4	60,084	6.68	11.51%
		β=0.9	69,878	8.40	13.36%
	α=1.0	β=1.0	83,679	12.6	15.97%
Linear Trend Equation			148,088	36.7	27.52%

Table 3.8. Forecasting techniques and errors

2nd stage: The annual traffic for 2017 is estimated by comparing the first four months of traffic for years 2016 and 2017. More specifically, the annual traffic for 2016 is divided by the traffic for the first four months, which yields a multiplier of 2.7455 (=9,939,465/3,620,271). Assuming the same ratio for the first four months in 2017, would yield a volume of 12,207,435 (=2.7455 × Traffic $_{\text{first four months 2017}}$ = 2.7455 × 4,446,338).

3.5.3 Preparing Forecast till 2030

The (estimated) annual traffic for 2017 was added to the crude oil traffic dataset, and the underlying dis-aggregated monthly data was tested for seasonality. Time series regression model that makes use of dummy variables to capture seasonality was used in the investigation (Bowerman et al., 2009). It is important to note that unlike explanatory regression models, the time series version treats the system as a black-box, i.e., the model only relies on historical data to prepare forecasts, and does not seek to establish any causal relationship (Makridakis et al., 1998). Hence, it just shows that crude oil traffic in any month has not been significantly different than that from other months, but does not eliminate the influence of one or more external factors in the shipment volume. Consequently, working with annual traffic data seems reasonable. Given the temporal data at our disposal, either a linear or a nonlinear forecasting model that makes use of time-series may be appropriate, however, empirical research supports using the former over the latter (Makridakis et al., 1998). Also, our own experiments reveal that linear trend equation model yielded more reasonable forecasts than a variety of nonlinear models. Thus, we use linear trend projection, which is suitable for medium to long-range forecasts (Heizer et al., 2017). In addition, given the time horizon of interest, it should be evident that *both WMA* and *DES* are not appropriate since they are myopic, i.e., can prepare forecasts only for the next period. Finally, we also report on the forecast scheme proposed by the Canadian Association of Petroleum Producers (CAPP).

Year	$y_t = 1397041 + 569559x$			CAPP
	Low	*Ref*	*High*	
2018	7.70	12.79	17.88	12.82
2019	8.19	13.36	18.53	13.46
2020	8.68	13.93	19.18	14.13
2021	9.16	14.49	19.84	14.41
2022	9.64	15.07	20.49	14.70
2023	10.11	15.64	21.17	14.99
2024	10.57	16.21	21.84	15.29
2025	11.04	16.78	22.52	15.60
2026	11.49	17.34	23.19	15.91
2027	11.95	17.91	23.88	16.23
2028	12.39	18.48	24.57	16.56
2029	12.84	19.05	25.26	16.89
2030	13.28	19.62	25.96	17.23

Table 3.9. Forecast of traffic (million tonnes) using proposed LTE equation and CAPP scheme

Table 3.9 (and Figure 3.8) depicts the projected crude oil traffic from 2018 and 2030 using the *LT* equation generated from the augmented historical dataset, and the forecast using the scheme outlined by the petroleum industry (CAPP, 2017). Since uncertain events are expected to influence the actual realized values in the future, we also develop a 95% prediction interval around each forecast value, which is in line with the existing literature (Makridakis et al., 1998). Note that the *low* and *high* values are not only based on the extent of variation of the data, but also on the forecast horizon. They could be used to build the most optimistic and most pessimistic scenarios, and the interval becomes larger as the period of interest gets farther away from the given time series. Finally, according to the CAPP scheme, crude oil supply in Canada is forecast to grow by 5% every year until 2020, and then slow down to 2% a year until 2030. It is evident from both Table

3.9 and Figure 3.8 that the *'Ref'* values suggested by *LTE* are lower than the equivalent CAPP numbers until 2020, and higher thereafter. Two observations are pertinent in this context. *First*, given the uncertainty associated with pipeline approval process and subsequent construction, we reckon that all of the scenarios outlined in Table 3.9 (and Figure 3.8) are plausible. *Second*, forecasts developed using the CAPP scheme is contained within the prediction interval for the specific year. Thus, our analysis is based on the dataset procured from Statistics Canada (2017a, b) and Transport Canada (2017), and not the CAPP scheme.

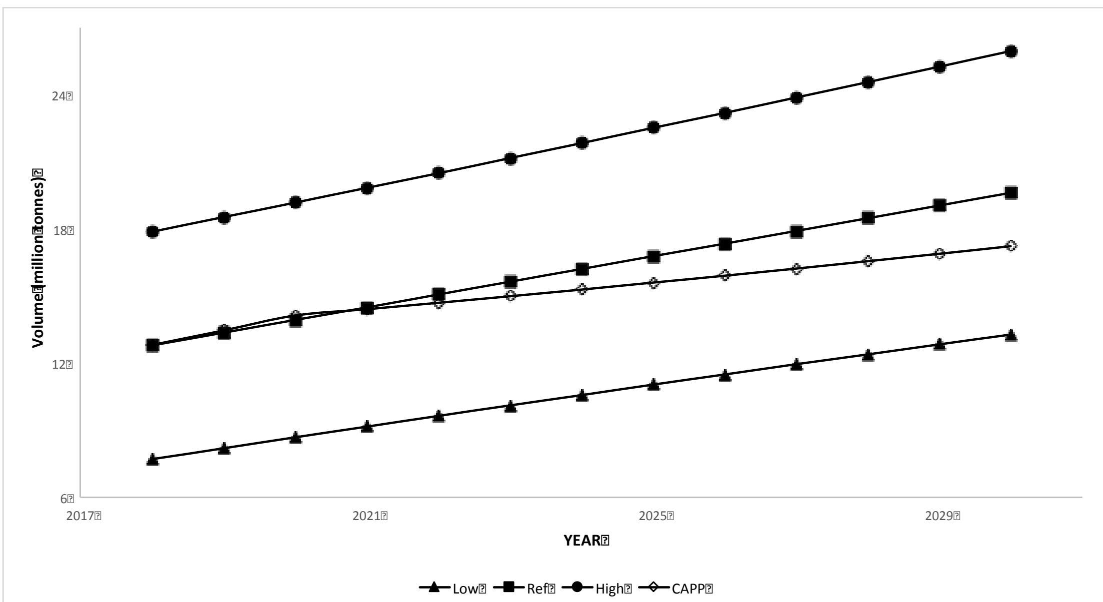


Figure 3.8. Forecast of crude oil traffic from 2018 to 2030 using LT and CAPP scheme

3.5.4 Preparing Shipment Routes

In this sub-section, we will propose different ways to route the crude oil shipments over the available railroad network in Canada. Given the nature of hazmat shipments, we will make use of bi-objective routing model that will capture the two extreme perspectives, i.e., railroad companies that intend to minimize cost, and regulators who want to minimize risk, and a set of intermediate solutions (Verma et al., 2011). Any routing model requires a set of inputs: origin; destination; demand; supply; cost to move shipment; risk from each shipment; and, paths from each origin to each destination. Recall that we only have annual crude oil traffic, and hence the first task is to intelligently locate the origins and destinations, and respective volume available or demanded at those locations.

ORIGINS & SUPPLY: In Canada, crude oil is primarily extracted from either oil fields or from oil sands. Table 3.10 depicts the CAPP forecast for crude oil by provinces/ territories expressed in thousand barrels per day. It is clear that Alberta is responsible for the majority of crude oil production viz. 82.7%, followed by Saskatchewan at 9.6%, in 2018. Hence, AB is the origin for 82.7% of crude oil shipments in 2018, i.e., 0.827 * 12,788,233 = 10,585,077 tonnes of crude oil in the 'Ref' scenario depicted in Table 3.9. Available supply for each of the other regions of the country can be determined similarly.

<table>
<tr><th rowspan="2">Year</th><th colspan="2">Eastern Canada</th><th colspan="5">Western Canada</th><th rowspan="2">Total</th></tr>
<tr><th>ON</th><th>Rest</th><th>AB</th><th>BC</th><th>SK</th><th>MB</th><th>NWT</th></tr>
<tr><td>2018</td><td rowspan="13">1</td><td>231</td><td>3671</td><td>70</td><td>424</td><td>33</td><td rowspan="2">6</td><td>4435</td></tr>
<tr><td>2019</td><td>264</td><td>3751</td><td>71</td><td>420</td><td>31</td><td>4543</td></tr>
<tr><td>2020</td><td>278</td><td>3820</td><td>72</td><td>416</td><td>30</td><td rowspan="4">5</td><td>4622</td></tr>
<tr><td>2021</td><td>267</td><td>3870</td><td rowspan="3">71</td><td rowspan="2">414</td><td>29</td><td>4657</td></tr>
<tr><td>2022</td><td>279</td><td>3912</td><td rowspan="2">28</td><td>4710</td></tr>
<tr><td>2023</td><td>297</td><td>3970</td><td>416</td><td>4786</td></tr>
<tr><td>2024</td><td>306</td><td>4009</td><td>70</td><td>421</td><td rowspan="2">27</td><td rowspan="7">4</td><td>4838</td></tr>
<tr><td>2025</td><td>286</td><td>4064</td><td rowspan="2">69</td><td>427</td><td>4878</td></tr>
<tr><td>2026</td><td>252</td><td>4131</td><td>432</td><td>26</td><td>4915</td></tr>
<tr><td>2027</td><td>227</td><td>4141</td><td>68</td><td>439</td><td rowspan="2">25</td><td>4905</td></tr>
<tr><td>2028</td><td>211</td><td>4176</td><td>67</td><td>448</td><td>4931</td></tr>
<tr><td>2029</td><td>197</td><td>4283</td><td>66</td><td>456</td><td rowspan="2">24</td><td>5032</td></tr>
<tr><td>2030</td><td>185</td><td>4377</td><td>65</td><td>461</td><td>5118</td></tr>
<tr><td colspan="9">AB: Alberta; BC: British Columbia; MB: Manitoba;
ON: Ontario; SK: Saskatchewan: NWT: Northwest Territories.</td></tr>
</table>

Table 3.10. CAPP crude oil production forecast

More detailed information is required about the locations where crude rail tank cars are loaded and shipped to the market. Based on the approximate locations of crude oil and natural gas resources, made available through NRC (2015a), we identified the main oil fields and oil sands regions in the country, each of which is then assigned to a rail node, i.e., origin for rail crude oil shipment. This is summarized in Table 3.11. Four points should be noted in this regard: *First*, crude oil production in the Atlantic Provinces is assumed to be entirely sourced from three main fields located offshore of Newfoundland and Labrador, i.e., Hibernia, Terranova, and White Rose. This is because production in other Atlantic provinces is considered negligible compared to that of Newfoundland and Labrador (CAPP, 2017; NEB, 2017a). *Second*, since no major oil field was previously identified in Manitoba (MB), the relative share of this province was given to the neighbour province, i.e., Saskatchewan (SK). *Third*, each oil field/oil sands region is assigned to the closest rail node in the network. Particularly, the offshore oil fields in NL are assigned to Halifax, which is considered as the origin for rail shipments of crude oil coming from these fields. *Fourth*, in the absence of additional information, we assume that the total volume coming from Alberta (AB) is equally divided between the identified crude oil sources; the same logic would be also extended to the Northwest Territories (NWT), where multiple oil fields have been identified.

For illustration purposes, the relative share of each province, as well as the corresponding originating traffic in each rail node in 2018, has been shown in the second and fifth columns, respectively. Furthermore, the 2018 crude oil supply associated with each rail node is expressed in terms of the number of railcars in the last column. To convert the tonnes to railcars, we have used the estimated capacity of the safer TC-117 tank cars, i.e., 28,371 gallons (Treichel and Barkan, 2016), and a gallons-to-tonnes conversion factor of 0.00325 (BP, 2016). This would be equivalent to assume that a railcar would approximately contain 92.21 tonnes of crude oil.

Provinces/ Territories	**2018 Relative Share (%)**	**Oil Field/Oil Sands Region**	**Assigned Node**	**2018 Traffic (tonnes)**	**2018 Traffic (railcars)**
AB	*82.77201*	*High Level*	High Level	2,117,015	22,959
		Beaverlodge	Hythe		
		Westlock	Busby		
		Stettler	Stettler		
		Brooks	Tilley		
BC	*1.57248*	*Prespatou*	Beatton	201,092	2,181
NL	*5.20827*	*Hibernia; Terranova; White Rose*	Halifax	666,045	7,223
ON	*0.02464*	*Hamilton*	Hamilton	3,151	34
SK	*10.28833*	*Eston*	Brock	1,315,695	14,268
NWT	*0.13428*	*Fort Liard*	Fort Nelson	4,293	47
		Colville Lake	Hay River		
		Tuktoyaktuk			
		Fort McPherson	Fort Nelson		
Total	***100***	***All fields***	***All nodes***	***12,788,233***	***138,686***

AB: Alberta; BC: British Columbia; NL: Newfoundland and Labrador; ON: Ontario; SK: Saskatchewan: NWT: Northwest Territories.

Table 3.11. Origins & Supply of rail crude oil shipments

DESTINATIONS & DEMAND: Next, we identify destinations for rail crude oil shipments. Refineries are considered the main points of crude oil consumption in Canada and in the United States. However, given that the scope of our study is Canada only, we consider that crude oil shipments either travel to *Canadian refineries* or toward the borders for *export*. Combining crude oil exports by rail data (NEB, 2016c) with railway commodity flow data (Statistics Canada, 2017a), we estimated that between 2012 and 2015, on average, 34.57% of the crude oil volume moved by rail was exported. Hence, we assume that 34.57% of the Canadian oil production is exported and the rest, i.e., 65.43%, is transported to refineries across the country for processing. It is important to note that in the last few years, crude oil exports have been mainly made to the United States (CAPP, 2017), and hence we neglect exports to other countries.

Table 3.12 lists the destinations for rail crude oil shipments under two categories: Oil refineries and export terminals. The former consists of 14 refineries across Canada (NRC, 2016a). Asphalt plants, petrochemical plants, and upgraders, as well as proposed refineries, are excluded from our analysis. The latter category includes 15 identified rail nodes near the US border, through which crude oil could be transported to the US refineries. This category is also organized by regions in the United States called PADD, i.e., petroleum administration for defense districts (US EIA, 2012).

Three points should be considered with regard to destinations. *First*, as indicated earlier, approximately 65.43% of the rail crude oil shipments are destined for Canadian oil refineries (14 refineries listed in Table 3.12) and the rest is intended for export to the US (through 15 rail nodes listed in Table 3.12). The corresponding volumes in 2018 are 8,367,341 (90,742 railcars) and 4,420,892 tonnes (47,944 railcars), respectively. *Second*, refinery capacity information has been used to estimate the volume of crude oil destined for each refinery. For instance, the largest refinery in the country, i.e., Irving Oil in NB, has a capacity of 313K barrels per day, which amounts of 16.93% of the capacity in Canada. This is equivalent to 1,416,591 tonnes (15,363 railcars) of crude oil coming to this refinery by rail in 2018. *Third*, LPG historical export data between 1999 and 2015, accessed through NEB (2016b), was analyzed as a proxy to obtain an average share for each PADD in terms of crude oil export. For instance, we found that about 28.76% of the exported portion travels toward the US refineries located in PADD 1; furthermore, in the absence of additional information, we simply assume that the three export terminals assigned to PADD 1, i.e., Niagara Falls, Fort Erie, and Cantic, have equal shares in terms of receiving the crude oil coming from oil fields or oil sands regions. As a result, each aforementioned export terminal is forecast to receive 423,816 tonnes (4,596 railcars) of crude oil by rail in 2018.

Routing Shipments: Having ascertained the origin/ destination locations and the respective volumes, the next step is to determine the best path to send shipments from the origins to the destinations. Note that since a path is comprised of rail-links and rail-yards, such path determination would also enable us to estimate the hazmat volume at each traversed link and yard. As indicated earlier, railroad companies focus on cost and hence solving a routing problem that minimizes transportation cost would be enough. However, in an effort to demonstrate cognizance of catastrophic events such as the Lac Mégantic disaster, we would also adopt a risk-based approach to route shipments. It should be evident that the routing program will make use of the data developed so far, i.e., Table 3.11 for origins and supply amounts, and Table 3.12 for destinations and demand. In addition, length of rail-links would be a proxy for transportation cost; and, average population density within a specified distance from the rail-links a measure for unit transport risk.

	Prov.	Refinery Identity / Export Terminals	Refinery Capacity (bpd)/ Export Zones	Percent of sub-total	2018 Volume (tonnes)	2018 Volume (railcars)
Oil Refineries	AB	Imperial Oil – Edmonton	187,000	0.1011	845,938	9,174
		Suncor Energy – Edmonton	142,000	0.0768	642,612	6,969
		Shell Canada –Fort S	100,000	0.0541	452,673	4,909
	BC	Husky Energy –Prince G	12,000	0.0065	54,388	590
		Chevron Canada - Burnaby	55,000	0.0297	248,510	2,695
	NB	Irving Oil –Saint John	313,000	0.1693	1,416,591	15,363
	NL	North Atlantic –Come by C	115,000	0.0622	520,449	5,644
	ON	Imperial Oil –Sarnia	121,000	0.0654	547,224	5,935
		Imperial Oil – Nanticoke	112,000	0.0606	507,061	5,499
		Suncor Energy –Sarnia	85,000	0.0460	384,898	4,174
		Shell Canada –Corunna	75,000	0.0406	339,714	3,684
	QC	Valero –Levis	265,000	0.1433	1,199,040	13,003
		Suncor Energy – Montreal	137,000	0.0741	620,020	6,724
	SK	Federated Co-op – Regina	130,000	0.0703	588,224	6,379
Total destined for Oil Refineries				***1.0000***	***8,367,341***	***90,742***
Export Terminals	AB	Coutts	PADD 4	0.0457	101,017	1,096
	BC	Kingsgate			101,017	1,096
		Matsqui	PADD 5	0.1304	288,242	3,126
		Mission			288,242	3,126
	MB	Emerson (two terminals)	PADD 2 PADD 3	0.5363	526,872	5,714
		Middlebro			263,436	2,857
	ON	Niagara Falls	PADD 1	0.2876	423,816	4,596
		Fort Erie			423,816	4,596
		Fort Frances	PADD 2 PADD 3	**		
		Sault Ste. Marie				
		Sarnia				
		Windsor				
	QC	Cantic	PADD 1	*		
	SK	Estevan	PADD 2 PADD 3	**		
		Bienfait				
Total destined for Export Terminals				***1.0000***	***4,420,892***	***47,944***
where, * implies same entry as for PADD 1, and ** that for PADD 2 and 3.						

Table 3.12. Destinations & Demand for crude oil shipments

Consequently, for a directed and connected railroad network with n nodes, the decision variable is:

x_{ij} = volume of traffic on directed arc $(i,j), i \rightarrow j$

and the parameters are:

l_{ij} = length of directed arc $(i,j), i \rightarrow j$
p_{ij} = average population density within 1km of the directed arc $(i,j), i \rightarrow j$
b_i = net flow generated at node i

(P2)

$$\begin{aligned} &Minimize \\ &Cost\text{: } \sum_{i=1}^{n} \sum_{j=1}^{n} l_{ij} x_{ij} \\ &Risk\text{: } \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} x_{ij} \end{aligned} \tag{3.4}$$

subject to

$$\sum_{j=1}^{n} x_{ij} - \sum_{j=1}^{n} x_{ji} = \begin{cases} +b_i & if\ i\ is\ a\ supply\ node \\ -b_i & if\ i\ is\ a\ demand\ node \\ 0 & if\ i\ is\ a\ transhipment\ node \end{cases} \qquad \forall i \tag{3.5}$$

$$x_{ij} \geq 0 \qquad \forall (i,j) \tag{3.6}$$

(P2) is a bi-criteria optimization model with cost and risk objectives as presented in (3.4). The cost objective represents the total railcar-kilometers travelled, whereas hazmat risk is represented as the product of population density in the vicinity of the rail-links and the volume of crude oil traversing them. Clearly, the current expression for hazmat risk accepts the impacted population within a pre-specified distance and the volume of crude oil as the two inputs, but could be expanded to include other elements such as damage to the environment, property, loss of flora/ fauna. However, in an effort to undertake focused analysis, we consider only the two indicated inputs. In addition, we make use of population census data provided by Statistics Canada to determine the impacted population (Statistics

Canada, 2017c). It should be evident that if zero weight is placed on the risk coefficient, one ends up solving an equivalent of the traditional *minimum cost* flow problem with transhipment constraints, i.e., **(P1)**. On the other hand, assigning zero weight to the cost coefficient would reduce **(P2)** to an equivalent of a *minimum risk* problem. Finally, simultaneous assignment of positive weights to both objectives would yield other intermediate solutions.

Four of seven inputs required to solve **(P2)** were indicated in Table 3.11 and Table 3.12, and the remaining three, i.e., cost to move shipment, risk from each shipment, and, path from each origin to each destination, will be estimated using the railroad network in Canada. We focus only on the networks of Canadian National (CN) and Canadian Pacific (CP) since the two together account for 95% of the annual rail tonne-kilometers in Canada (Transport Canada, 2011). As indicated earlier, their networks were recreated, which resulted in a graph with 464 nodes and 940 arcs. Figure 3.10 shows the resulting network in Canada, and will provide information on the three remaining inputs. For expositional ease, we use stars to denote origins of crude oil shipments (i.e., Table 3.11), and hollow circles for destinations (i.e., Table 3.12). Additionally, the company to which a specific oil refinery or export terminal belongs is shown in the parenthesis besides the corresponding label.

Now, we have all the seven inputs required to solve **(P2)**. Two of the most common techniques for solving multi-objective models are pre-emptive optimization and weighted sums (Rardin, 1998). The former calls for sequential solution process, while the latter associates weights to different objective values. We solve **(P2)** using the latter approach, and vary weights over a large interval so that the extreme stance of the railroad company (i.e., minimum cost) and the regulator (i.e., minimum risk) is represented, along with several intermediate solutions.

3.5.5 Discussion and Analysis

In this section, we first solve **(P2)** from the perspective of the railroad company and then compare the resulting solution to those obtained via alternative routing schemes, and conclude with an assessment of the impact of the major pipeline projects.

3.5.5.1 *Current Practice*

(P2) was solved using CPLEX 12.6.2.0 (IBM, 2016), and using the 2018 *'Ref'* scenario as the input. Figure 4 depicts the resulting solution. As discussed in the previous section, the railroad companies intend to minimize system wide transportation cost, which is depicted as *Min Cost* in Figure 3.9.

Figure 3.9. Solutions using different routing schemes on 2018 'Ref' scenario

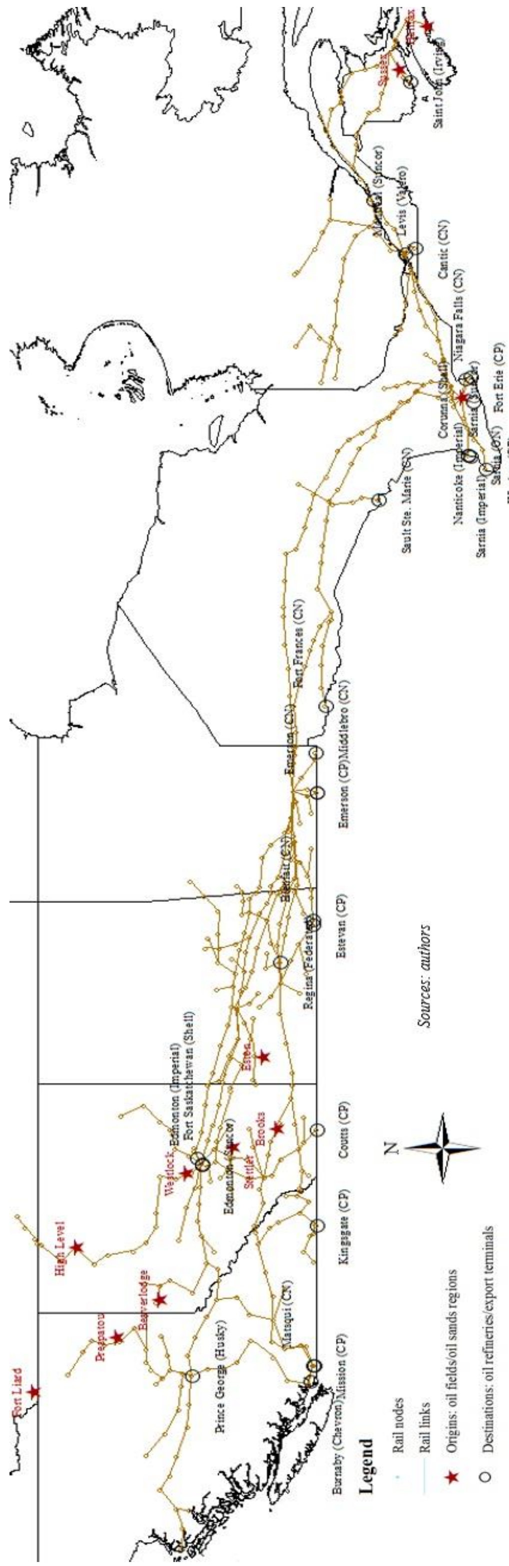


Figure 3.10. Railroad network in Canada and crude oil origins & destinations

The current practice will result in a cost of $323 million to route crude oil shipments from origins to their respective destinations, and expose 536 million individuals across the railroad network.

In an effort to better understand the implications of the current practice, and for expositional reasons, we delineate the top twenty rail-links based on hazmat volume and when both volume and population density is considered (Table 3.13). It should be evident that we have information about all the 940 links depicted in Figure 3.10, but have chosen to focus on the ones with high traffic volume. If crude oil volume is used as the only indicator, then sixteen of the top twenty links are found in the provinces of Manitoba (MB) and Saskatchewan (SK), which is not unreasonable because crude oil flows from the western provinces to the demand locations in eastern Canada.

The crude oil volume on each of the twenty links are indicated in terms of the number of railcars, and reported in column 4. However, if both crude oil volume and population density are considered, then sixteen of the top twenty links would belong to Ontario (ON) and Quebec (QC), which are the two most populous provinces in Canada. Note that Manitoba (MB) still finds three links in this list, including two that are common to the list with only volume (and denoted by asterisks). It is pertinent that population census in 2011 (Statistics Canada, 2017c) is used, and the analyses assume no significant change in the population density of the considered census divisions over the next decade. The last column in the table depicts the risk, expressed as volume*population density, for the top twenty links.

<table>
<tr><th colspan="4">Volume</th><th colspan="4">Volume*Population Density</th></tr>
<tr><th colspan="2">LINK</th><th rowspan="2">Pro</th><th rowspan="2">No. of Railcars</th><th colspan="2">LINK</th><th rowspan="2">Pro</th><th rowspan="2">Risk</th></tr>
<tr><th>#</th><th>Name</th><th>#</th><th>Name</th></tr>
<tr><td>385</td><td>Winnipeg-Portage la Prarie</td><td>MB</td><td>86,554</td><td>465</td><td>Montreal-Saint Luc</td><td>QC</td><td>127,925,607</td></tr>
<tr><td>369</td><td>Saskatoon-Colonsay</td><td rowspan="7">SK</td><td rowspan="13">75,689</td><td>215</td><td>Toronto-Beaverton</td><td rowspan="2">ON</td><td>77,172,316</td></tr>
<tr><td>370</td><td>Colonsay-Lanigan</td><td>226</td><td>Toronto-Oshawa</td><td>76,706,507</td></tr>
<tr><td>375</td><td>Lanigan-Wynyard</td><td>466</td><td>Les Coteaux-Saint Luc</td><td>QC</td><td>54,357,943</td></tr>
<tr><td>376</td><td>Wynyard-Foam Lake</td><td>217</td><td>Toronto-Mississauga</td><td rowspan="2">ON</td><td>49,153,396</td></tr>
<tr><td>377</td><td>Foam Lake-Yorkton</td><td>449</td><td>Toronto-Guelph Junction</td><td>26,150,611</td></tr>
<tr><td>378</td><td>Yorkton-Bredenbury</td><td>423*</td><td>Winnipeg-Molson</td><td>MB</td><td>12,233,504</td></tr>
<tr><td>379</td><td>Bredenbury-Marchwell</td><td>467</td><td>Saint Luc-Delson</td><td rowspan="2">QC</td><td>10,931,334</td></tr>
<tr><td>380</td><td>Marchwell-Harrowby</td><td rowspan="8">MB</td><td>255</td><td>Montreal-Drummondville</td><td>9,873,798</td></tr>
<tr><td>381</td><td>Harrowby-Solsgirth</td><td>385*</td><td>Winnipeg-Portage la Prairie</td><td>MB</td><td>9,017,206</td></tr>
<tr><td>382</td><td>Solsgirth-Minnedosa</td><td>220</td><td>Mississauga-Hamilton</td><td>ON</td><td>8,819,071</td></tr>
<tr><td>383</td><td>Minnedosa-Gladstone</td><td>258</td><td>Val Alain-Quebec City</td><td>QC</td><td>7,847,016</td></tr>
<tr><td>384</td><td>Gladstone-Portage le Prairie</td><td>214</td><td>Washago-Beaverton</td><td rowspan="4">ON</td><td>7,790,584</td></tr>
<tr><td>423</td><td>Winnipeg-Molson</td><td rowspan="7">75,126</td><td>227</td><td>Oshawa-Coubourg</td><td>6,096,934</td></tr>
<tr><td>424</td><td>Molson-Rennie</td><td>452</td><td>Guelph Junction-London</td><td>3,511,520</td></tr>
<tr><td>425</td><td>Rennie-Ingolf</td><td>443</td><td>Sudbury-Cartier</td><td>3,030,957</td></tr>
<tr><td>426</td><td>Ingolf-Keewatin</td><td rowspan="4">ON</td><td>66</td><td>Edmonton-Busby</td><td>AB</td><td>2,780,535</td></tr>
<tr><td>427</td><td>Keewatin-Vermilion Bay</td><td>210</td><td>Torrance-Washago</td><td>ON</td><td>2,554,705</td></tr>
<tr><td>428</td><td>Vermilion Bay-Dyment</td><td>421</td><td>Winnipeg-Dufrost</td><td>MB</td><td>2,051,701</td></tr>
<tr><td>429</td><td>Dyment-Ignace</td><td>464</td><td>Chesterville-Les Coteaux</td><td>ON</td><td>2,010,971</td></tr>
<tr><td colspan="8">AB: Alberta; MB: Manitoba; ON: Ontario; QC: Quebec; SK: Saskatchewan; Pro: Province.</td></tr>
</table>

Table 3.13. Top twenty links as per current practice (i.e., *Min Cost*)

Note that the *Min Cost* approach (i.e., current practice) does not consider the impact on population or society, and hence a hazmat incident could result in catastrophic consequences, such as in Lac Mégantic. To better illustrate the above observation we make use of Figure 3.11 to depict traffic in the provinces of Alberta (AB) and British Columbia (BC). Note that the thickness of a specific rail-link is an indicator of the volume of crude oil through it. Additionally, origins are depicted as stars and destinations as solid circles.

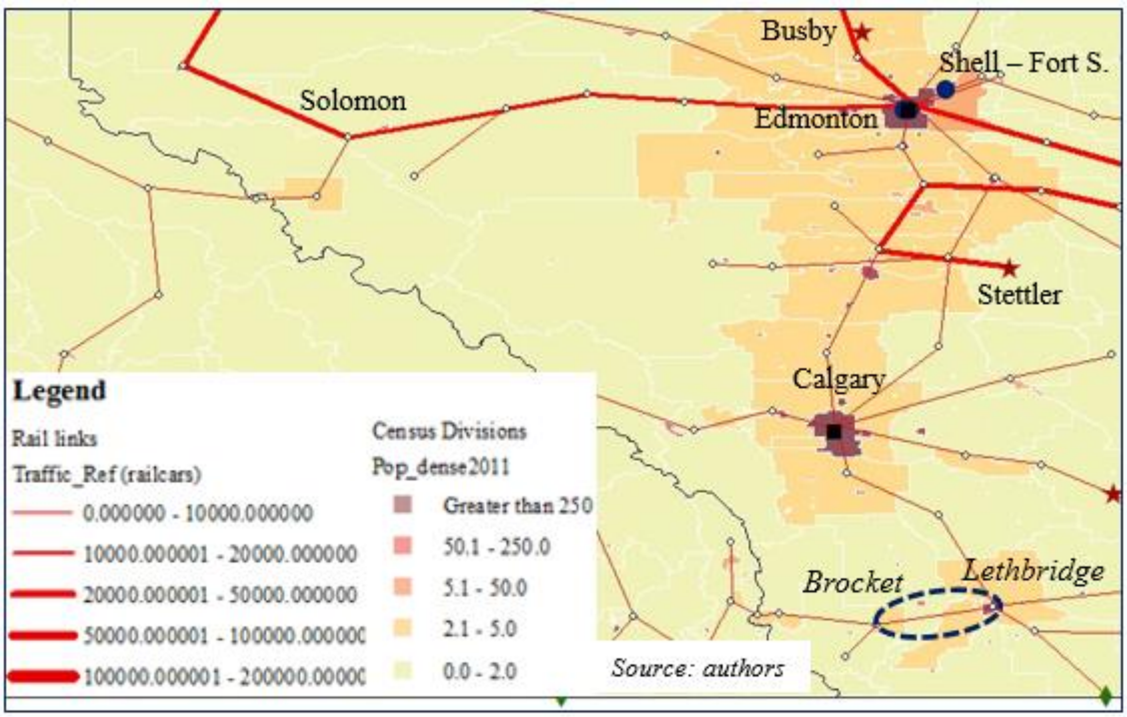


Figure 3.11. Impact of current practice

In general, we noted that a major proportion of the eastbound crude oil shipments were going through the city of Edmonton, despite the existence of alternative routes with much lower risk. For instance, the (highlighted) link between Brocket and Lethbridge would carry much lower population exposure risk, however, it is not being used to move any crude oil shipments. It should be clear that the current practice of the railroad industry (i.e., minimization of cost) could be troublesome to both the regulators and the society in general. Hence, we make an attempt to propose alternative routing schemes that could potentially achieve a reduction in risk at little or no incremental cost.

3.5.5.2 *Alternate Routing Schemes*

Alternate routing schemes were implemented by attaching complementary weights to the cost and risk objective in **(P2)** for the 2018 *'Ref'* scenario. A total of ten problem instances were solved, which yielded six distinct solutions that are also depicted in Figure 3.9. For example, *A* depicts the solution when cost coefficient carries 90% and 80% of the weights, while risk coefficient carries 10% and 20%, respectively. Figure 3.9 can be used to deduce the following points. *First*, *Min Risk* solution is equivalent to *F*, i.e., cost coefficient has only 10% weight. *Second*, *B* and *C* have the same risk, but the former has lower cost and thus would dominate the latter. Similarly, *D* dominates *E*. *Third*, if the railroad company can be persuaded to spend $18 million then the hazmat risk in the network will decrease by 44 million, i.e., 2.45 fewer individuals at risk for every additional dollar spent. *Fourth*, perhaps the railroad company is willing to spend an additional million dollars and decrease the risk by thirty-five million (i.e., moving from *Min Cost* to *A*). Such a trade-off between cost and risk is present in many business problems, examples of which can be also found in hazmat transport literature (Verma, 2009). *Finally*, it appears that solutions are influenced by cost more than that by risk, e.g., *C* is closer to the *Min Cost* than to *Min Risk*, or even equidistant between the two. The latter implies that the perspectives of both the stakeholders (i.e., the railroad company and the regulator) cannot be captured by attaching equal weights to their objectives. Hence, we will use the *Min Risk* routing approach to denote the alternate technique and contrast the results with those discussed in Section 3.5.5.1. It is important that such routing technique is consistent with the prevalent practice around the world such as the instance of including the number of individuals exposed to hazmat in risk assessment, and consequent routing decisions (Fabiano et al., 2002).

It is noteworthy that, with crude oil volume as the only indicator, fifteen of the top twenty links are located in the province of Ontario under the alternate routing scheme. But link 385 that connects Winnipeg to Portage la Prairie, once again, sends the largest number of railcars in the entire network. This, we reckon, is because of the absence of alternate links connecting western and eastern regions of Canada. Furthermore, higher volume of crude oil shipments is being forced onto links with lower population density. For instance, using volume*population density as the indicator, link 465 connecting Montreal to Saint Luc carried the highest risk (Table 3.13), which was replaced by link 467 connecting Saint Luc to Delson under the alternate routing scheme (Table 3.14). There were several other instances of links displacing others to move up the ranks, but in general and if possible, crude oil shipments were routed over links with lower population density but higher transportation cost. Finally, we focus on the region delineated earlier. Recall that most of the traffic was routed through the densely populated region of Edmonton under the current practice of the railroad industry even though alternate links were available (i.e., Figure 3.11). However, under the *Min Risk* approach, shipments do start using the link between Brocket and Lethbridge. In fact, around 14.5K carloads of crude oil are using this link, thereby reducing an equivalent number going through Edmonton.

Volume				Volume*Population Density			
LINK		Pro	No. of Railcars	LINK		Pro	Risk
#	Name			#	Name		
385	Winnipeg-Portage la Prairie	MB	86,544	467	Saint Luc-Delson	QC	90,630,107
174	Winnipeg-Elma		75,126	215	Toronto-Beaverton	ON	77,172,316
175	Elma-Decimel			226	Toronto-Oshawa		76,706,507
176	Decimel-Winnitoba			466	Les Coteaux-Saint Luc	QC	54,357,943
177	Wonnitoba-Minaki			449	Toronto-Guelph Junction	ON	45,799,190
178	Minaki-Jones	ON		469	Montreal-Saint Jean sur Richelieu	QC	38,599,899
179	Jones-Amesdale			255	Montreal-Drummondville		9,873,798
180	Amesdale-Alcona			_385*_	Winnipeg-Portage la Prairie	MB	9,017,206
181	Alcona-Savant Lake			468	Delson-St Jean	QC	8,952,383
182	Savant Lake-Collins			_174*_	Winnipeg-Elma	MB	8,644,739
183	Collins-Armstrong Station			258	Van Alain-Quebec City	QC	7,847,016
184	Armstrong Station-Ombabika			_214*_	Washago-Beaverton	ON	7,790,584
185	Ombabika-Nakina			227	Oshawa-Cobourg		6,096,934
186	Nakina-Longlac			450	Guelph Junction-Hamilton		4,587,599
187	Longlac-Hillsport			452	Guelph Junction-London		3,511,520
188	Hillsport-Hornepayne			66	Edmonton-Busby	AB	2,780,535
189	Hornepayne-Oba			_210*_	Torrance-Washago	ON	2,554,705
209	Parry Sound-Torrance		72,269	421	Winnipeg-Dufrost	MB	2,051,701
210	Torrance-Washago			464	Chesterville-Les Coteaux	ON	2,010,971
214	Washago-Beaverton			205	Felix-Sudbury		1,510,419

Table 3.14. Top twenty links as per alternate routing scheme (i.e., *Min Risk*)

Doing so, clearly, increases risk locally in the region indicated by the hollow ellipse but results in a decrease in network-wide risk (Figure 3.12).

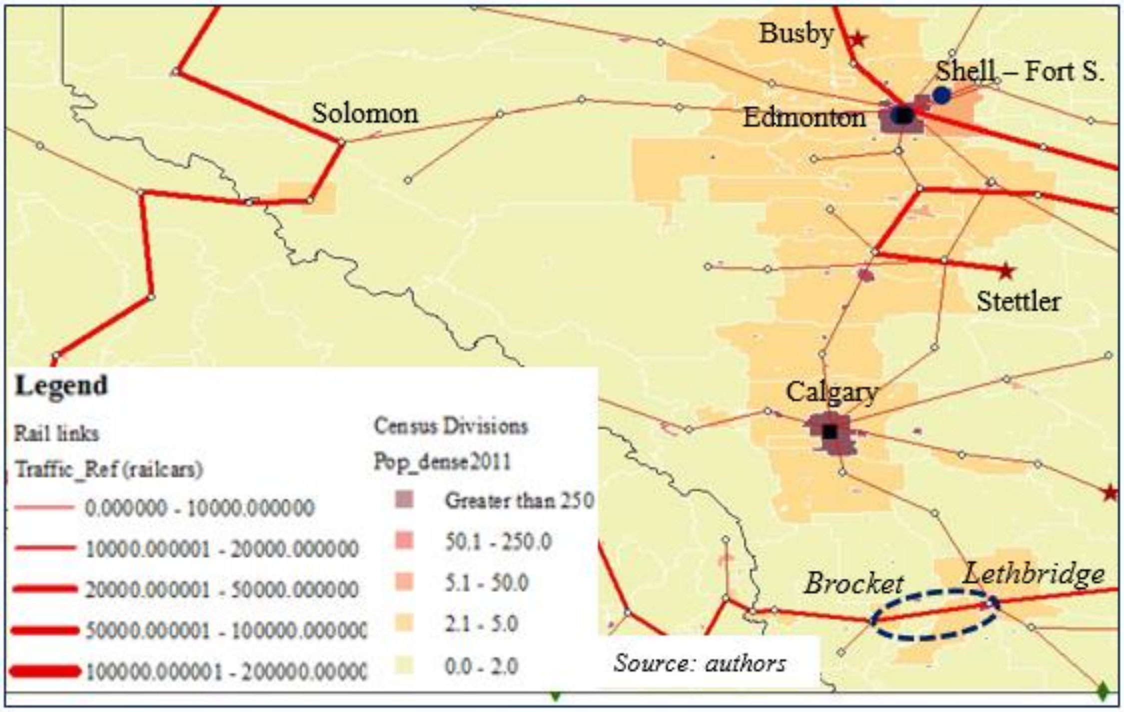


Figure 3.12. Impact of alternate routing scheme

3.5.5.3 *Projected Time Period: Current v/s Alternate*

In an effort to estimate the cost and risk values over the chosen time horizon, an additional 44 problem instances were simulated and analyzed using the *'Ref'* values for four years: 2021, 2024, 2027, and 2030, and a snapshot of the results are presented in Table 3.15. Finally, for brevity, we are not reporting link information, such as Table 3.13 and Table 3.14, over the forecasted time period.

Current Practice: It was noticed that, with crude oil volume as the only indicator, unlike in 2018 most of the top twenty links shifts from the current domicile provinces of Manitoba (MB) and Saskatchewan (SK) to Ontario (ON) towards the end of the time horizon. This can be explained by the projected change in the configuration of supply and demand locations. However, when both volume and population density are considered, then most of the top links continue to remain in the provinces of Ontario (ON) and Quebec (QC), which is not surprising since we have made use of census data which projects these two provinces to be most populous in the country until 2030.

<table>
<tr><th colspan="2" rowspan="2">Weights</th><th colspan="2">2021</th><th colspan="2">2024</th><th colspan="2">2027</th><th colspan="2">2030</th></tr>
<tr><th>Cost</th><th>Risk</th><th>Cost</th><th>Risk</th><th>Cost</th><th>Risk</th><th>Cost</th><th>Risk</th></tr>
<tr><td colspan="2">Min Cost</td><td>363</td><td>599</td><td>402</td><td>660</td><td>458</td><td>763</td><td>512</td><td>857</td></tr>
<tr><td>α=0.9</td><td>β=0.1</td><td rowspan="2">364</td><td rowspan="2">560</td><td rowspan="2">403</td><td>616</td><td rowspan="2">459</td><td rowspan="2">714</td><td rowspan="2">513</td><td rowspan="2">804</td></tr>
<tr><td>α=0.8</td><td>β=0.2</td><td>615</td></tr>
<tr><td>α=0.7</td><td>β=0.3</td><td rowspan="3">366</td><td rowspan="3">554</td><td rowspan="3">405</td><td rowspan="3">609</td><td>461</td><td>707</td><td>515</td><td>796</td></tr>
<tr><td>α=0.6</td><td>β=0.4</td><td rowspan="2">462</td><td rowspan="2">706</td><td rowspan="2">516</td><td rowspan="2">795</td></tr>
<tr><td>α=0.5</td><td>β=0.5</td></tr>
<tr><td>α=0.4</td><td>β=0.6</td><td rowspan="2">369</td><td rowspan="2">552</td><td rowspan="2">408</td><td rowspan="3">606</td><td>465</td><td rowspan="3">703</td><td rowspan="2">520</td><td rowspan="2">792</td></tr>
<tr><td>α=0.3</td><td>β=0.7</td><td rowspan="2">466</td></tr>
<tr><td>α=0.2</td><td>β=0.8</td><td>370</td><td>551</td><td>409</td><td>521</td><td>791</td></tr>
<tr><td>α=0.1</td><td>β=0.9</td><td rowspan="2">383</td><td rowspan="2">549</td><td rowspan="2">424</td><td rowspan="2">604</td><td rowspan="2">483</td><td rowspan="2">700</td><td rowspan="2">539</td><td rowspan="2">789</td></tr>
<tr><td colspan="2">Min Risk</td></tr>
</table>

Table 3.15. Cost and Risk for the 44 problem instances

Alternate Scheme: With crude oil volume as the only indicator, the top twenty links continue to reside in the provinces of Manitoba (MB) and Ontario (ON), i.e., just like Table 3.14. However, with crude oil volume*population density as the measure, most of the high risk links continue to be in the provinces of Ontario (ON) and Quebec (QC) over the time-horizon in question, albeit there were some changes in rankings. It is pertinent that conceivably, a more comprehensive risk measure that includes the impact on the environment, marine life, etc., could be developed, which in turn may yield a dis-similar shipment network. However, in the absence of such a composite measure, we have attempted to propose a simple technique that could provide useful information to policy makers whose interest is in establishing risk-reduction regulations.

3.5.5.4 *Impact of the Proposed Pipeline Projects*

It is anticipated that the combined capacity of the existing and the proposed pipelines would exceed the crude oil production of western provinces by 2030 (CAPP, 2017). However, the aforementioned is unlikely to end rail crude oil shipments because of the following reasons. *First*, as indicated earlier, railroads offer greater flexibility in reaching wider markets, which would keep them as an attractive transport mode. *Second*, proposed pipelines would likely impact the supply and demand at those locations served by railroads that are directly on the pipeline route, and easily accessible. This would still leave plenty of other locations un-impacted. *Third*, pipeline infrastructure and not-too-low prices are likely to maintain an upward pressure in crude oil production until 2030 (NEB, 2017b). Nevertheless, we assume that pipelines would have an impact on rail crude oil shipments, which we elaborate later. Note that our analyses so far have neglected the pipelines' impact on crude oil rail shipments.

Table 3.16 lists the four major crude oil pipelines proposed in Canada along with their receipt points, delivery points, capacities, and the expected date in service. Primary delivery points for "Enbridge Line 3 Restored" and "TransCanada Keystone XL" are in the US. Since the scope of our study is Canada, we consider the closest border location as their destination within Canada, i.e., Gretna (MB) and Grassy Creek (SK), respectively. For brevity, and for relevance, we analyze the impact of pipelines on rail crude oil shipments in selected years in our forecast horizon. Given the target in service dates, and in the absence of additional information, we simply assume that Enbridge, Kinder Morgan, Keystone XL and Energy East pipelines will be in service in 2019, 2020, 2021, and 2022, respectively.

Pipeline	Primary Receipt Point	Primary Delivery Point	Capacity (thousand bpd)	Target in Service
Enbridge Line 3 Restored	Hardisty, AB	Superior, WI [Gretna, MB]	370	2019
Kinder Morgan Trans Mountain Expansion	Strathcona County, AB	Burnaby, BC	590	End 2019
TransCanada Keystone XL	Hardisty, AB	Steele City, NE [Grassy Creek, SK]	830	2020+
TransCanada Energy East	Hardisty, AB	Quebec City, QC/ St. John, NB	1,100	2021+
Total Capacity = 2.89 million barrels per day				
AB: Alberta; WI: Wisconsin; MB: Manitoba; BC: British Columbia; NE: Nebraska; SK: Saskatchewan; QC: Quebec; NB: New Brunswick; bpd: barrels per day.				

Table 3.16. Proposed crude oil pipelines existing western Canada

We next outline the two steps required to incorporate the proposed pipelines, and then discuss the consequent impact on rail crude oil shipments.

Step 1: Assign railcar supply/ demand locations to pipeline receipt/delivery points.

For example, "*Enbridge Line 3 Restored*" pipeline project is expected to be in service in 2019, and the possible impact would be felt thereafter. Let *S* denote the total number of railcars with crude oil coming from the five[6] supply points that are closest to the pipeline receipt point of Hardisty. These supply points were identified in Table 3.11 and are: Busby; Stettler; Tilley; Hythe; and, Brock. Similarly, let *D* denote the total number of

[6] We have chosen five locations for two reasons. *First*, given that the receipt points for all proposed pipelines are geographically close to each other in Alberta, and in the absence of additional information, applying the reduction to only one supply point could result in unrealistic scenarios. *Second*, given the number of supply/demand points identified for rail crude oil shipments and their spatial distribution relative to pipeline routes, equally assigning the total reduction to five origin/destination seems to be reasonable based on our judgement.

railcars with crude oil going to destinations closest to the pipeline delivery point of Gretna. These demand locations were identified in Table 3.12 and are: Emerson; Middlebro; Fort Frances; Estevan; and, Bienfait. Finally, let C be the capacity of the pipeline which is 370K barrels per day, or 18.43 million tonnes per year. The resulting pipeline with the relevant information is depicted in Figure 3.13.

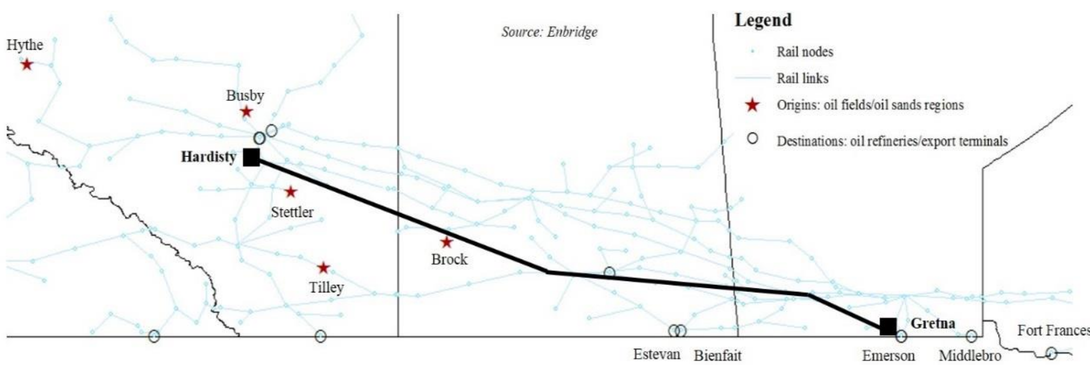


Figure 3.13. Enbridge Line 3 Restored pipeline

<u>Step 2: Determine the reduction in the number of railcars from the pipeline project.</u>

For Enbridge Line 3 project, if R is the reduction in the number of railcars with crude oil, then $R = Min\{D, S, C\}$. Our data suggests that demand is the most relevant parameter in every case, which in turn implies that supply at each of the five origins is reduced by $D/5$, and that the new demand at the five destinations is zero.

We have solved **(P2)** with updated supply/demand configuration to analyze the possible impact of proposed pipelines on rail crude oil shipments and on the corresponding spatial risk distribution. Figure 3.14 depicts the country-wide impact of pipelines on crude oil shipments in 2019 and 2022 for the two routing schemes.

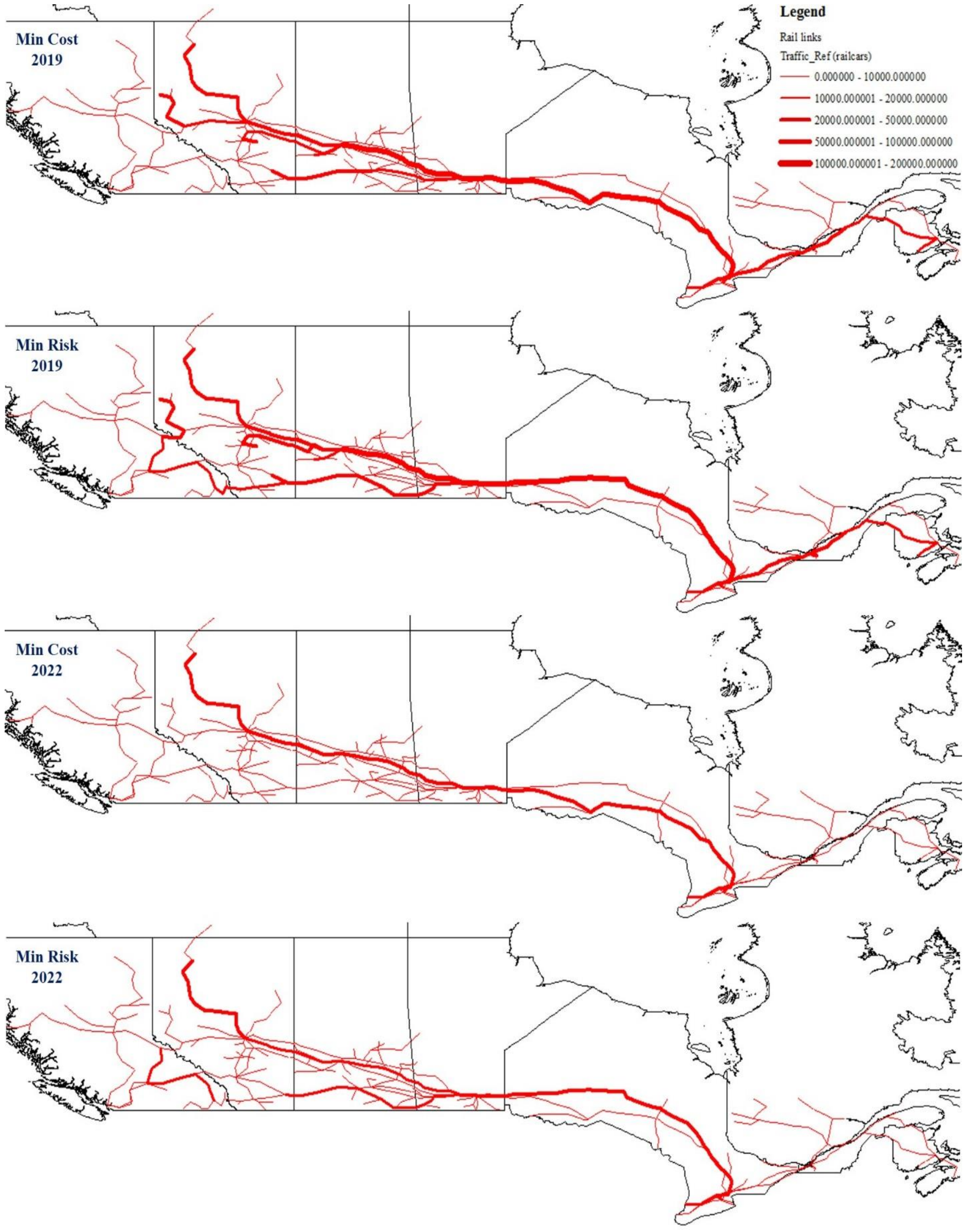


Figure 3.14. Snapshot of the impact of proposed pipelines for 'Ref' scenarios in 2019 and 2022

We found that the country-wide incremental impacts of Enbridge, Kinder Morgan, and Keystone XL pipelines on rail crude oil shipments are hardly visible due to relatively low demands in MB, BC, and SK. Therefore, we only represent the country-wide impact of Energy East pipeline in 2022 compared to 2019, when only the Enbridge pipeline is in service. We observe quite a significant reduction in crude oil movement by rail as shown by reduced rail-link thickness in Figure 3.14. This is mainly because of the eastbound direction of Energy East pipeline and the fact that crude oil supply points, i.e., oil fields and oil sands regions, are mainly located in Western provinces, whereas demand points such as refineries are mainly located in Eastern and Atlantic provinces. Since one of the delivery points of the Energy East pipeline is St. John in New Brunswick, the crude oil movement on the rail-links leading to Atlantic provinces will be substantially reduced in 2022. In both 2019 and 2022, the *Min Risk* scheme (i.e., alternate routing) routes a part of the eastbound traffic through less risky parts of the network, e.g., through Golden, Brocket, and Lethbridge to reduce the volume passing through the densely populated area in Edmonton, as illustrated before.

We next provide finer details to help decision makers get a better sense of the impact of the proposed pipelines on rail links or regions. To that end, we will comment on the riskiest links in the network from the perspective of the railroad company (i.e., *Min Cost*) and that of the regulators (i.e., *Min Risk*).

For the four indicated years, and using crude oil volume as the sole indicator of rail-risk, link 385 connecting Winnipeg to Portage la Prairie is the busiest under all the routing scheme, except for 2022 for the *Min Risk* setting. In this solitary instance, link 385 is replaced by the link connecting Toronto to Guelph (i.e., link 449) in the province of Ontario (ON). It is important that this shift has resulted from the completion of Energy East pipeline in 2022, and that most of the top twenty links were located in the provinces of ON and MB. In general, the relative share of rail crude oil shipments for ON, vis-à-vis other provinces, are expected to increase although the absolute volume will be lower once pipeline projects are completed.

The impact of the proposed pipelines is more noticeable when crude oil volume*population density is the measure of risk. Unsurprisingly, most of the links are in the two most populous provinces of ON and QC. For brevity, we are not reproducing similar to those presented in earlier sections. However, we did notice the impact of the completion of the proposed pipeline projects and the resulting changes in the supply/demand configuration. For instance, the completion of the Energy East pipeline project in 2022, rendered the link between city of Hamilton and Niagara Falls in the province of Ontario (ON) much riskier. In addition, given the current practice of the railroad industry, the link between Montreal to Saint Luc in the province of Quebec (QC) is the riskiest in 2019, but will be replaced by the link connecting Toronto and Mississauga in 2022. Similar observations could be made when using the alternate routing scheme. For instance, link 467 running connecting Saint Luc to Delson in Quebec (QC) will be replaced by link 449 connecting Toronto to Guelph Junction in the province of Ontario (ON). Such observations provide a better understanding of the likely evolution of the system, particularly from the risk perspective, which would be useful to railroad companies, regulatory agencies, and researchers in the field. It is important that alternate routing scheme will yield a risk reduction of 8% in 2019, and of 19% in 2022 over the risk values resulting from the current practice.

3.6 Conclusion

Though the last two decades have witnessed the development of numerous initiatives to tackle the assessment and management of rail hazmat shipments, the true potential benefits of these efforts could not be aptly demonstrated because of the absence of rail-link and rail-yard level information about hazmat volume. We make a first attempt to develop a methodology that makes use of popular analytics techniques to country-level commodity flow data, and entails a number of steps to extract yard and link level data.

The proposed methodology was applied to estimate the volume of the top three classes of hazmat that collectively account for over eighty percent of the rail hazmat shipments in Canada. For each hazmat class, appropriate sources of uncertainties were accounted for through the generation of prediction intervals around every base-case forecast. Furthermore, for each hazmat class, solutions and analyses of the different problem instances enabled us to identify the high-traffic yards and links in the network, which in turn could help with emergency response planning. It is pertinent to mention that such information is also useful for undertaking a more precise risk assessment and risk management exercises by academic and industry researchers. Also, it was possible to discern the impact of geographical shifts in hazmat supply, and how it is likely to impact the relative volume-based ranking of various rail yards and rail-links in the near future.

Finally, a focused analysis was conducted on crude oil shipments in Canada to develop long-term forecasts, compare alternate routing schemes in terms of network-wide risk, and evaluate the impact of new pipeline projects. In particular, the availability of the proposed pipeline infrastructure would change the supply and demand location configurations over the forecast horizon, with the maximum changes to the current crude oil traffic flow pattern stemming from the completion of the Energy East pipeline project.

In closing, though the insights are specific to Canada, the proposed methodology could be applied to other jurisdictions once relevant modifications to capture the region-specific topographical attributes are made. The information on rail hazmat traffic could be invaluable for designing risk management and mitigation measures such as a data-driven emergency response network.

Chapter 4 Emergency Response Planning for Rail Hazmat Incidents in Canada

4.1 Introduction

Timely response is crucial in keeping the hazmat incident consequences relatively low. While first responders are typically able to reduce damages to some extent, the presence of specialized response teams and equipment is necessary to effectively respond to hazmat emergencies. For example, the Railway Association of Canada (2015) has special response teams drawn from railroad companies, who work with the chemical industry and local public security agencies to mitigate damages in case of rail hazmat emergencies. According to Transport Canada (2012b), in hazmat incidents, first, the material name or the associated identity number should be identified. Then, first responders are advised to call the emergency response telephone number, which is usually listed on the shipping papers. A potential hazmat incident site could be deemed covered in terms of emergency response if it could be reached from an emergency response station within a specified time, which would depend on the incident scenario. Due to resource constraints, it would be impractical to plan for responding to every hazmat incident regardless of location, hazmat class, and release volume; hence, some sort of prioritization is required, which is usually applied by means of a risk measure. For instance, *population exposure* has been considered as the measure of highway transport risk in the literature (Berman et al., 2007); as indicated earlier in Chapter 2, this measure is defined as the total number of individuals that are exposed to hazmat shipments (Batta and Chiu, 1988). Our approach is, however, to plan for emergency response to incident situations based on a *composite risk measure* that accounts for both incident probabilities and total estimated costs associated with incidents that are not responded to in a timely manner.

This study deals with facility location and resource allocation at the strategic and the tactical levels, respectively. We propose a two-stage stochastic integer programming problem to account for uncertainties in some model parameters. We will solve this model over part of the Canadian railroad network, and specifically provide answers to the following questions: *where to locate emergency response facilities*; *what types and how many equipment packages to stockpile at each facility*; and, *how they are going to be utilized and dispatched in response to hazmat incidents*. To the best of our knowledge, there is no study in the literature that integrates the differentiating aspects of the rail hazmat transport into planning for hazmat incident response.

The rest of this chapter is organized as follows. Section 4.2 provides the details on our mathematical modeling approach, followed by a detailed description of the proposed solution methodology in Section 4.3. We present the case study in Section 4.4, which is followed by conclusion in Section 4.5.

4.2 Mathematical Model

As mentioned earlier, we are interested in optimally locating emergency response facilities, acquiring and stockpiling specialized equipment packages, and allocating such resources to rail hazmat incidents. In this problem, some factors or parameters are deterministic and some are stochastic. For example, fixed costs related to the construction of emergency response facilities as well as estimates of scenario-specific release volumes and associated damages are the deterministic factors. On the other hand, exact location, volume, and class of released hazmat commodity are classified as the stochastic factors. To model the abovementioned emergency response planning problem, we make several assumptions:

1. To define a hazmat incident situation, we consider the hazmat class involved, the rail-link where it occurs, and the profile (type) of release incident. The last attribute is general enough to make our model flexible in terms of characterizing various hazmat incident situations. However, when solving our model on a real network, we will assume that a hazmat incident profile is simply expressed in terms of the volume released in the incident. Hence, we ignore the influence of other attributes that could define an incident such as weather conditions. Note that the incident profile can represent a more comprehensive factor once more accurate data becomes available.

2. We assume that equipment packages that are suitable to respond to a given hazmat class are different only in size; therefore, whether conditions, for instance, would not affect the type of equipment packages that are going to be used.

3. In contrast to the model by Berman et al. (2007), in which fractional coverage of arcs is possible, we assume that in our case, an arc (link) is either completely covered, or otherwise counted as uncovered. So, if only parts of a given link are covered, the link is not counted as covered. Regarding the coverage area -reachable within a critical time-, in

the preprocessing stage, we identify those arcs over the network that can be completely covered by any of the potential response facilities.

4. It is assumed that in case an incident is responded to in a timely manner, the only cost incurred is transportation and operating costs related to specialized response packages. Otherwise, potential harmful consequences on the affected population, property, and the environment would have to be considered.

5. Making decisions on whether to evacuate the impact area depends on a lot of factors, and hence excluded from our analysis. In other words, we assume that evacuation decisions are made independently and therefore the associated costs are not taken into account. However, we consider that the initial evacuation and isolation distance for every hazmat, recommended by CANUTEC (Transport Canada, 2012a) would be enforced by first responders at the incident site.

6. Class 2, 3, and 8 (i.e., gases, flammable liquids, and corrosives) accounted for almost 80% of the rail hazmat tonnage in Canada (Provencher, 2008). Our data extraction and analysis endeavor validates this statement and shows that the abovementioned three classes constitute the majority of hazmat transported by rail during the last decade (See Figure 3.4). Consequently, we focus on only these three hazmat classes.

Because of the uncertain aspects of location, class, and profile of rail hazmat incidents, we use a two-stage stochastic programming with recourse, which is cited as a general-purpose technique to deal with uncertainty in model parameters. In our model, the first stage is related to strategic decisions i.e., *facility location* and *equipment acquisition*, and the second stage is related to tactical/operational ones i.e., *dispatching* decisions. We next introduce the sets/indices, variables, and parameters, followed by the model for each of the two stages:

First-stage problem (P3)

Sets and indices

I: set of candidate facility sites, indexed by i;
E: set of equipment packages, indexed by e;
M: set of hazmat classes (Class 2, 3, and 8), indexed by m;

Decision variables

Y_i= 1 if facility at site i is open, and 0 otherwise;
U_i^{em}: number of equipment packages type e acquired and stockpiled at i, suitable to respond to incidents with Class m;

Parameters

F_i: fixed cost to open an emergency response facility at site i;
B^{em}: acquisition (buying) cost of an equipment package type e, suitable to respond to incidents with Class m;
K: capacity of a given facility, expressed in terms of the number of equipment packages.

Minimize

$$\sum_{i\in I} F_i Y_i + \sum_{i\in I}\sum_{e\in E}\sum_{m\in M} B^{em} U_i^{em} \tag{4.1}$$

subject to

$$KY_i - U_i^{em} \geq 0 \qquad \forall i \in I, \forall e \in E, \forall m \in M \tag{4.2}$$

$$Y_i \in \{0,1\} \qquad \forall i \in I \tag{4.3}$$

$$U_i^{em} \geq 0 \;\; integer \qquad \forall i \in I, \forall e \in E, \forall m \in M \tag{4.4}$$

In this model, (4.1) is the objective function requiring facility location and equipment acquisition decisions. Constraint set (4.2) ensures that equipment packages are acquired only in open facilities, and also their number for each type in any open facility does not exceed a standard capacity K. The remaining constraints, i.e., (4.3) and (4.4), define the ranges of the variables.

The second-stage problem is characterized by random variables that correspond to an outcome $\omega: \xi(\omega) = (a(\omega), m(\omega), t(\omega))$. In other words, a given realized value represents a specific hazmat incident situation/scenario depending on its location in the network, the hazmat class involved, and the incident profile.

Second-stage problem (P4)

Sets and indices

A: set of arcs or links in the railroad network, indexed by a;
T: set of incident profiles (types), indexed by t;
$a(\omega) \in A$: set of possible hazmat incident locations;
$m(\omega) \in M$: set of possible hazmat classes involved in an incident;
$t(\omega) \in T$: set of possible hazmat incident profiles;
I_a: set of facilities that are able to cover arc a in response to hazmat incidents;
A_i: set of arcs that can be covered by the facility at location i in response to hazmat incidents;

Decision variables

N_{iat}^{em}: number of equipment packages type e dispatched from site i to arc a, to respond to an incident with Class m and profile t;
Z_{at}^{m}= 1 if the incident with Class m and profile t along arc a is covered, and 0 otherwise.

Parameters

EC_{at}^{m}: estimated costs (sum of population exposure, property loss, and environmental cleanup costs) resulting from not responding to an incident with Class m and profile t along arc a;

v_t: volume of hazmat involved in an incident with profile t;

C^e: containment capacity of an equipment package type e;

OC_t^m: cost to operate one unit of equipment package type e for an incident with profile t;

TC_{ia}^{e}: transport cost to move one unit of equipment package type e from site i to arc a;

l_a: length of arc a;

α^m: minimum portion of required network coverage for incidents involving Class m;

τ: a relatively large number such as 1000.

For any given incident scenario (a, m, t) and first-stage solution (Y, U), the second-stage recourse problem is $Q(Y, U, \xi) = Q(Y, U, a, m, t)$:

Minimize

$$EC_{at}^{m}\,(1 - Z_{at}^{m}) + \sum_{i \in I} \sum_{e \in E} (TC_{ia}^{e} + OC_t^m)\, N_{iat}^{em} \tag{4.5}$$

subject to

$$\max_{a \in A_i} N_{iat}^{em} \leq U_i^{em} \qquad \forall i \in I, \forall e \in E \tag{4.6}$$

$$N_{iat}^{em} \leq \tau\, Z_{at}^{m} \qquad \forall i \in I, \forall e \in E \tag{4.7}$$

$$\sum_{i \in I_a} \sum_{e \in E} C^e\, N_{iat}^{em} \geq v_t\, Z_{at}^{m} \tag{4.8}$$

$$l_a\, v_t\, (Z_{at}^{m} - \alpha^m) \geq 0 \tag{4.9}$$

$$N_{iat}^{em} \geq 0 \quad integer \qquad \forall i \in I, \forall e \in E \tag{4.10}$$

$$Z_{at}^{m} \in \{0,1\} \tag{4.11}$$

In this model, objective function (4.5) contains the equipment dispatching decisions, which have to be made in response to hazmat incidents. In particular, the total cost incurred in this stage is either resulting from not covering the specific incident situation, i.e., *non-response*, or the combination of transport costs and operating costs associated with dispatching specialized equipment packages to the incident site.

Constraint set (4.6) ensures that the maximum number of equipment packages dispatched in response to a specific incident scenario, does not exceed the acquired equipment packages of the same type and class in the facility.

Constraint set (4.7) is based on the logic that no equipment packages should be dispatched in response to a specific incident situation, unless the model decides to respond to the incident[7]. Constraint set (4.8) ensures that the total capacity of the equipment packages dispatched in response to a specific incident scenario is at least equal to the hazmat volume involved in the incident.

Constraint (4.9) is a coverage policy constraint. If an expectation is taken over all arcs and profiles, it ensures at least a predetermined level of volume-weighted response coverage on the entire network, for each hazmat class. The remaining constraints, i.e., (4.10) and (4.11) are related to the decision variables.

[7] It is important to note that this constraint set is necessary since incident probability P_{at}^{m}, based on historical data, is zero for some release profiles, which mistakenly leads to non-zero solutions for the corresponding N_{iat}^{em} values, even if the associated Z_{at}^{m} is zero.

Considering the probability associated with each possible incident scenario, i.e., P_{at}^{m}, the two-stage stochastic programming problem can be equivalently stated as a single optimization problem:

(P5)

Minimize

$$\sum_{i\in I} F_i Y_i + \sum_{i\in I}\sum_{e\in E}\sum_{m\in M} B^{em} U_i^{em} + \sum_{a\in A}\sum_{m\in M}\sum_{t\in T} P_{at}^{m} \left[EC_{at}^{m}\,(1 - Z_{at}^{m}) + \sum_{i\in I}\sum_{e\in E} (TC_{ia}^{e} + OC_{t}^{m})\, N_{iat}^{em} \right] \quad (4.12)$$

subject to

$KY_i - U_i^{em} \geq 0$	$\forall i \in I, \forall e \in E, \forall m \in M$	(4.13)
$\max_{a\in A_i} N_{iat}^{em} \leq U_i^{em}$	$\forall i \in I, \forall e \in E, \forall m \in M, \forall t \in T$	(4.14)
$N_{iat}^{em} \leq \tau Z_{at}^{m}$	$\forall i \in I, \forall a \in A, \forall e \in E, \forall m \in M, \forall t \in T$	(4.15)
$\sum_{i\in I_a}\sum_{e\in E} C^{e} N_{iat}^{em} \geq v_t Z_{at}^{m}$	$\forall a \in A, \forall m \in M, \forall t \in T$	(4.16)
$\sum_{t\in T}\sum_{a\in A} l_a v_t (Z_{at}^{m} - \alpha^{m}) \geq 0$	$\forall m \in M$	(4.17)
$Y_i \in \{0,1\}$	$\forall i \in I$	(4.18)
$U_i^{em} \geq 0$ $integer$	$\forall i \in I, \forall e \in E, \forall m \in M$	(4.19)
$N_{iat}^{em} \geq 0$ $integer$	$\forall i \in I, \forall e \in E, \forall a \in A, \forall m \in M, \forall t \in T$	(4.20)
$Z_{at}^{m} \in \{0,1\}$	$\forall a \in A, \forall m \in M, \forall t \in T$	(4.21)

(P5) is a *nonlinear integer programming* problem and can be solved by commercial solvers for not very large networks. However, if we take a closer look, we will see that the constraint set (4.14) can be linearized. We define an auxiliary variable W_{it}^{em} such that $W_{it}^{em} = \max_{a \in A_i} \mathrm{N}_{\mathrm{iat}}^{\mathrm{em}}$. This is equivalent to the following two constraints:

$$W_{it}^{em} \geq N_{iat}^{em} \qquad \forall i \in I, \forall a \in A, \forall e \in E, \forall m \in M, \forall t \in T \qquad (4.14')$$

$$W_{it}^{em} \leq U_{i}^{em} \qquad \forall i \in I, \forall e \in E, \forall m \in M, \forall t \in T \qquad (4.14'')$$

The constraint set (4.14) in **(P5)** can be replaced by (4.14') and (4.14'') to create a *linear intege*r problem. The validity of such an equivalence was verified for sample instances using CPLEX solvers (IBM, 2016). We refer to the resulting problem as Emergency Response Planning Problem **(ERPP)**.

4.3 Solution Methodology

(ERPP) is an *integer programming* (IP) problem, which is *NP-hard*. Due to the *combinatorial* nature of the problem, the computing time needed to find the optimal solution to realistic size problems could be extremely long. Consequently, we propose a *heuristic* solution methodology, which combines a *Decomposition Method* with *Lagrangian Relaxation* to efficiently solve the problem.

First, using the special structure of the model, we decompose **(ERPP)** into three sub-problems, each of which corresponds to the problem instance specific to a given hazmat class. The resulting sub-problem is a simplified version of **(ERPP)**, where hazmat class indices are dropped:

(P6)

Minimize

$$\sum_{i\in I} F_i Y_i + \sum_{i\in I}\sum_{e\in E} B^e U_i^e + \sum_{a\in A}\sum_{t\in T} P_{at}\left[EC_{at}\,(1-Z_{at}\,) + \sum_{i\in I}\sum_{e\in E}(TC_{ia}^e + OC_t)\,N_{iat}^e\right] \tag{4.22}$$

subject to

$$KY_i - U_i^e \geq 0 \qquad \forall i\in I, \forall e\in E \tag{4.23}$$

$$W_{it}^e \geq N_{iat}^e \qquad \forall i\in I, \forall a\in A, \forall e\in E, \forall t\in T \tag{4.24}$$

$$W_{it}^e \leq U_i^e \qquad \forall i\in I, \forall e\in E, \forall t\in T \tag{4.25}$$

$$N_{iat}^e \leq \tau\, Z_{at} \qquad \forall i\in I, \forall a\in A, \forall e\in E, \forall t\in T \tag{4.26}$$

$$\sum_{i\in I_a}\sum_{e\in E} C^e\ N_{iat}^e \geq v_t\, Z_{at} \qquad \forall a\in A, \forall t\in T \tag{4.27}$$

$$\sum_{t\in T}\sum_{a\in A} l_a\, v_t\,(Z_{at} - \alpha) \geq 0 \tag{4.28}$$

$$Y_i \in \{0,1\} \qquad \forall i\in I \tag{4.29}$$

$$U_i^e \geq 0 \ \ integer \qquad \forall i\in I, \forall e\in E \tag{4.30}$$

$$N_{iat}^e \geq 0 \ \ integer \qquad \forall i\in I, \forall e\in E, \forall a\in A, \forall t\in T \tag{4.31}$$

$$Z_{at} \in \{0,1\} \qquad \forall a\in A, \forall t\in T \tag{4.32}$$

Second, after some computational experiments, constraint (4.28) is identified as the "complicating" constraint. We remove this constraint from **(P6)**, and bring it to the objective function (4.22) along with a *Lagrangian multiplier* λ, to form a *Lagrangian Lower Bound Program* as follows:

(P7)

Minimize

$$\sum_{i\in I} F_i Y_i + \sum_{i\in I}\sum_{e\in E} B^e U_i^e + \sum_{a\in A}\sum_{t\in T}\left\{P_{at}\left[EC_{at}\,(1-Z_{at}\,) + \sum_{i\in I}\sum_{e\in E}(TC_{ia}^e + OC_t)\,N_{iat}^e\right] + \lambda\, l_a\, v_t\,(\alpha - Z_{at})\right\} \tag{4.33}$$

subject to

$$KY_i - U_i^e \geq 0 \qquad \forall i \in I, \forall e \in E \tag{4.34}$$

$$W_{it}^e \geq N_{iat}^e \qquad \forall i \in I, \forall a \in A, \forall e \in E, \forall t \in T \tag{4.35}$$

$$W_{it}^e \leq U_i^e \qquad \forall i \in I, \forall e \in E, \forall t \in T \tag{4.36}$$

$$N_{iat}^e \leq \tau\, Z_{at} \qquad \forall i \in I, \forall a \in A, \forall e \in E, \forall t \in T \tag{4.37}$$

$$\sum_{i\in I_a}\sum_{e\in E} C^e\, N_{iat}^e \geq v_t\, Z_{at} \qquad \forall a \in A, \forall t \in T \tag{4.38}$$

$$Y_i \in \{0,1\} \qquad \forall i \in I \tag{4.39}$$

$$U_i^e \geq 0 \;\; integer \qquad \forall i \in I, \forall e \in E \tag{4.40}$$

$$N_{iat}^e \geq 0 \;\; integer \qquad \forall i \in I, \forall e \in E, \forall a \in A, \forall t \in T \tag{4.41}$$

$$Z_{at} \in \{0,1\} \qquad \forall a \in A, \forall t \in T \tag{4.42}$$

Third, **(P7)** is solved for each hazmat class sequentially, in a chosen sequence. More specifically, in first sub-problem of the sequence, an optimal value of λ is chosen such that the best lower bound is achieved. Next, the facility locations associated with this first solution are fixed, and given as parameters to the second sub-problem of the sequence, which is similarly solved to find the best lower bound. Next, the facility locations suggested by this second solution are fixed, and given as parameters to the third sub-problem in the sequence. *Finally*, the solutions to the three problems can be integrated to obtain an approximate best lower bound for the original problem, i.e., **(ERPP)**. A detailed description of the aforementioned solution methodology is captured in Figure 4.2.

We will also propose an alternate way for solving **(ERPP)**. This technique involves dropping the policy constraint (4.17), and applying scaling factors to incident probabilities such that the 2nd stage costs are magnified and hence able to compete with the 1st stage costs. Further details about this technique are provided in Section 4.4.2.2

4.4 Case Study

As mentioned in the previous chapter, we have recreated the railroad network using the available information from the two Class I railroad operators, i.e., Canadian National (CN) and Canadian Pacific (CP).

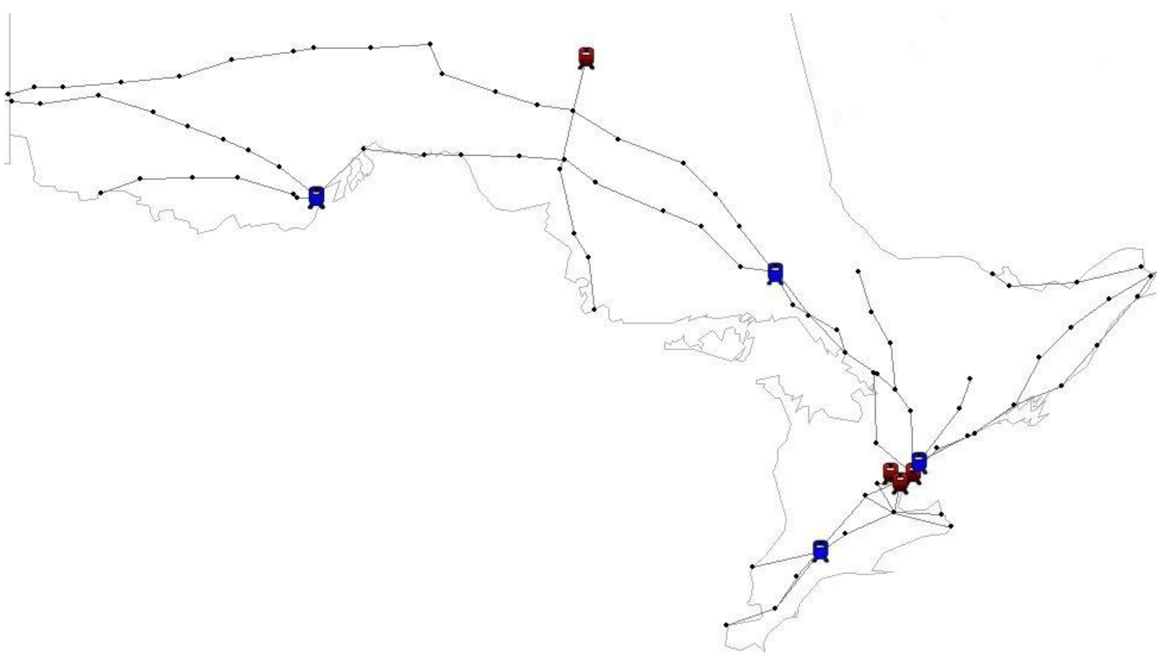


Figure 4.1. Ontario railroad network based on CN and CP networks

Step 1. Decompose the problem into three sub-problems based on hazmat classes.

Step 2. Relax the complicating constraint (4.28) in **(P6)**, and bring it to the objective function (4.22), to form a Lagrangian lower bound program **(P7)**.

Step 3. Select an appropriate sequence for solving the three sub-problems, and solve them sequentially. Start with the 1st sub-problem in the sequence.

3.1. Choose two positive distinct initial Lagrangian multipliers λ_0 and λ_1 (e.g., 0.1 and 1). Also, choose an initial appropriate step size ρ (e.g. 0.2). Set i=1.

3.2. Let ε be a desired tolerance (e.g., 0.01). While $|\lambda_i-\lambda_{i-1}| > \varepsilon$, solve **(P7)** with both multipliers:

3.2.1. If the constraint (4.28) in **(P6)** is violated in both cases, set $\lambda_{i+1}= \max\{\lambda_i,\lambda_{i-1}\}+ \rho$; update $\lambda_{i-1} \leftarrow \max\{\lambda_i,\lambda_{i-1}\}$ & $\lambda_i \leftarrow \lambda_{i+1}$

- After *m* subsequent iteration (e.g., 5) where this condition is held true, double the step size, i.e. update $\rho \leftarrow 2\rho$

3.2.2. else if the constraint (4.28) in **(P6)** is satisfied with a positive slack with respect to Z_{at} in both cases, set $\lambda_{i+1}= \min\{\lambda_i,\lambda_{i-1}\}- \rho$; update $\lambda_{i-1} \leftarrow \min\{\lambda_i,\lambda_{i-1}\}$ & $\lambda_i \leftarrow \lambda_{i+1}$

- After *m* subsequent iteration (e.g., 5) where this condition is held true, double the step size, i.e. update $\rho \leftarrow 2\rho$

3.2.3. else if the constraint (4.28) in **(P6)** is satisfied without any slack in at least one of the cases, record the corresponding optimal solution for the sub-problem. Go to *Step 4*.

3.2.4. else if the constraint (4.28) in **(P6)** is violated in one case and is satisfied with a positive slack in another case, set $\lambda_{i+1}= (\lambda_i+\lambda_{i-1})/2$;

- While $|\lambda_i-\lambda_{i-1}| > \varepsilon$, solve **(P7)** in both cases. Update λ_{i-1} & λ_i values such that the condition in *3.2.4* is held true (i.e., the constraint (4.28) is violated for one value and is satisfied with a positive slack for the other).

Step 4. Fix the located facilities, and continue with the next sub-problem in the sequence, if there is any, and go to *3.1*. Otherwise, stop.

Figure 4.2. The proposed solution methodology for (ERPP)

The railroad network in Ontario is captured in Figure 4.1, which has 100 nodes and 103 arcs.

4.4.1 Parameter Estimation

The model input parameters are estimated by analyzing publicly available sources and by referring to the scientific literature. In the following, we only provide the values used in our experiments. Complete estimation details will appear in the appendices.

a) *Facility Cost (F_i)*

The construction cost estimates for fire stations, suggested by RSMeans (2016a), are used as a proxy to emergency response facility costs. Table 4.1 list such estimates for 20 cities across Ontario. Further details are provided in Appendix A.

Location	Cost-per-square-foot (C $)	Total Construction Cost (C $)
Barrie	164.06	984,360
Brantford	165.27	991,620
Cornwall	162.84	977,040
Hamilton	167.09	1,002,540
Kingston	163.60	981,600
Kitchener	158.14	948,840
London	165.12	990,720
North Bay	161.48	968,880
Oshawa	162.99	977,940
Ottawa	165.72	994,320
Owen Sound	162.99	977,940
Peterborough	162.69	976,140
Sarnia	165.57	993,420
Sault Ste. Marie	157.53	945,180
St. Catharines	157.23	943,380
Sudbury	156.62	939,720
Thunder Bay	157.38	944,280
Timmins	161.48	968,880
Toronto	170.43	1,022,580
Windsor	156.77	940,620

Table 4.1. Construction cost estimates for fire stations across Ontario

b) Equipment Acquisition Cost (B^{em})

We use a base number ($450K) suggested by Verma et al. (2013), to estimate the acquisition cost of an equipment package suitable to respond to a 50-ton release of flammable liquids, i.e., Class 3. Nonlinear extrapolation is used to estimate larger-size equipment packages, as shown in Table 4.2.

Equipment package size	*Size 1*	*Size 2*	*Size 3*	*Size 4*
Release volume of flammable liquids (tons)	*50*	*100*	*200*	*400*
Acquisition Cost ($ thousands)	**450**	675	900	1125

Table 4.2. Acquisition cost of equipment packages suitable to respond to flammable liquid incidents

Additionally, we use initial evacuation distances recommended by Transport Canada (2012b), to estimate equipment acquisition cost for the other two classes (Table 4.3). Complete details are included in Appendix B.

Hazmat class	**Significant commodity**	**Evacuation distance (meters)**	**Assigned factor**
Class 2	*Hydrocarbon gas*	1600	1.5
Class 3	*Crude oil*	800	**1**
Class 8	*Sulphuric acid*	800	1

Table 4.3. Factors assigned to hazmat classes to estimate the cost of their specialized equipment package

c) Incident Probability (P_{at}^{m})

We propose the following equation for estimating the probability associated with each hazmat incident scenario: (See Appendix C for details.)

$$P_{at}^{m} = P(R_{network}) \times \frac{N_a^m l_a}{\sum_{o=1}^{O}\sum_{j=1}^{J} N_j^o l_j} \times P(Q_{at}^{m} | R_a^m, H_a^m, D_a, G_a) \qquad (4.43)$$

Year	P(G)	P(D|G)	P(H|D,G)	P(R|H,D,G)	P($R_{network}$)
2000	0.00000413	0.70801	0.03106	0.53354	4.85E-08
2001	0.00000425	0.73900	0.02778	0.51557	4.50E-08
2002	0.00000376	0.72644	0.03285	0.47449	4.26E-08
2003	0.00000406	0.70653	0.03351	0.41978	4.03E-08
2004	0.00000440	0.71935	0.02732	0.46593	4.03E-08
2005	0.00000414	0.70576	0.02646	0.63934	4.94E-08
2006	0.00000368	0.73282	0.03177	0.42980	3.68E-08
2007	0.00000339	0.71816	0.03640	0.65341	5.79E-08
2008	0.00000321	0.72108	0.02795	0.42000	2.72E-08
2009	0.00000286	0.71653	0.03645	0.44059	3.29E-08
2010	0.00000270	0.70084	0.03611	0.43629	2.98E-08
2011	0.00000283	0.72337	0.03027	0.47297	2.93E-08
2012	0.00000241	0.73314	0.03462	0.58036	3.55E-08
2013	0.00000247	0.70849	0.04234	0.32452	2.40E-08
2014	0.00000244	0.70680	0.04032	0.28160	1.96E-08
2015	0.00000259	*0.70567*	*0.03197*	*0.39535*	2.31E-08
Average	**0.00000333**	**0.71700**	**0.03295**	**0.46772**	**3.64E-08**

Table 4.4. Network-wide components of hazmat release probability from 2000 to 2015

Number of railcars	Observed frequency	Probability
1	1781	0.9989
2	0	0.0000
3	0	0.0000
4	0	0.0000
5	1	0.0006
6	0	0.0000
7	0	0.0000
8	0	0.0000
9	0	0.0000
10	0	0.0000
11	0	0.0000
12	0	0.0000
13	0	0.0000
14	1	0.0006
15	0	0.0000

Table 4.5. A discrete probability distribution for Class 2 incident profiles

Equation (4.43) estimates incident probability for each link, class, and profile based on three main components: first, the probability of hazmat release on the entire network, which is shown in the last column of Table 4.4; second, the ratio of the train-kilometers traveled on the link carrying the specific hazmat class to the total train-kilometers travelled on the entire network carrying hazmat, which is calculated based on the outcome of Chapter 3; and, third, the conditional probability of the specific incident profile, which is captured in the last column of Table 4.5 for Class 2 incidents. Extensive details, including the mathematical developments, are provided in Appendix C.

Hazmat class	**Class 2**	**Class 3**	**Class 8**
Representative chemical	Propane	Hexane	Nitric Acid
Wind speed (m/s)	10	10	10
Wind 102measurement height (m)	5	5	5
Ground roughness	Urban or forest	Urban or forest	Urban or forest
Cloud cover	Partly cloudy	Partly cloudy	Partly cloudy
Air temperature (°C)	20	20	20
Stability class	C	C	C
Inversion height option	No inversion	No inversion	No inversion
Humidity (%)	70	70	70
Tank	Horizontal cylinder	Horizontal cylinder	Horizontal cylinder
Volume (gallon)	33700	28400	15700
Length (ft)	70	60	50
Diameter (ft)	9.05	8.98	7.31
Tank contains	Liquid	Liquid	Liquid
Liquid level (%)	100	100	100
Type of tank failure	BLEVE	BLEVE	None
Other specifications	-	100% of mass in fireball	Leak through a 10” circular opening
Threat model	Thermal radiation from fireball	Thermal radiation from fireball	Gaussian
Red threat zone	(10.0 kW/(sq m) = potentially lethal within 60 sec)	(10.0 kW/(sq m) = potentially lethal within 60 sec)	(92 ppm = AEGL-3 [60 min])
Orange threat zone	(5.0 kW/(sq m) = 2nd degree burns within 60 sec)	(5.0 kW/(sq m) = 2nd degree burns within 60 sec)	(24 ppm = AEGL-2 [60 min])
Yellow threat zone	(2.0 kW/(sq m) = pain within 60 sec)	(2.0 kW/(sq m) = pain within 60 sec)	(0.16 ppm = AEGL-1 [60 min])

Table 4.6. Assumptions and input parameters in ALOHA

d) *Incident Cost (EC_{at}^{m})*

Total incident cost in the absence of effective response, is divided into three components: Population Exposure Costs, Property Loss Costs, and Environmental Cleanup Costs. Impact areas are estimated using ALOHA (US EPA, 2016a) with the assumptions and parameters shown in Table 4.6.

For illustration, Figure 4.3 identifies three threat zones for Class 3 incidents, which depends on the number of railcars that release hazmat contents.

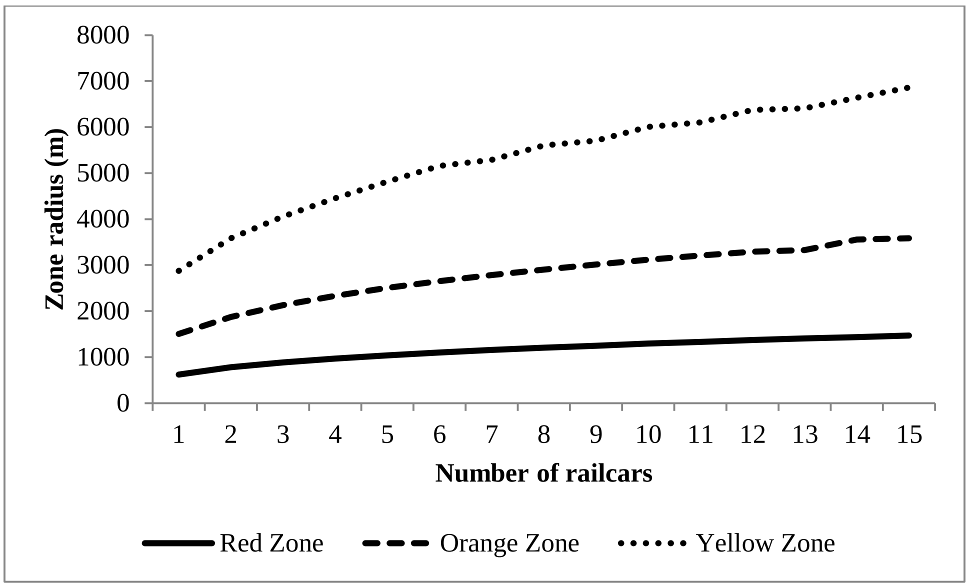


Figure 4.3. Radii of threat zones for different Class 3 release volumes

d. 1) Population Exposure Costs

To estimate the value of risk-reducing intervention actions, we use the Value of Statistical Life (VSL) estimates provided by the U.S. Department of Transportation (DOT's Office of the Secretary, 2016), which is listed in the third columns of Table 4.7.

Year	Low Value ($million)	VSL ($million)	High Value ($million)
2013	5.20	9.10	12.90
2014	5.20	9.20	13.00
2015	5.20	9.40	13.00
2016	5.40	9.60	13.40

Table 4.7. VSL estimates and alternative values by the DOT's Office of the Secretary

Similarly, values associated with preventing injuries as fractions of VSL (DOT's Office of the Secretary, 2016) are used to estimate population exposure costs in each of the three threat zones identified in Figure 4.3.

AIS Level	Severity	Fraction of VSL	AEGL Level	Modified Fraction
1	Minor	0.003	*1*	0.025
2	Moderate	0.047		
3	Serious	0.105	*2*	0.186
4	Severe	0.266		
5	Critical	0.593	*3*	0.797
6	Unsurvivable	1.000		

Table 4.8. Value of preventing injuries

d. 2) Property Loss Costs

This component is classified into three elements as follows: Material loss, carrier damage, and property damage. A base number is extracted from PHMSA (2016) for each hazmat class based on historical data. More specifically, for a given hazmat class, this number represents the average cost of property loss corresponding to one railcar incidents, which is then extrapolated to estimate values corresponding to greater number of railcars. Also, the impact of location is applied by a factor that magnifies the property damages corresponding to links located in densely populated areas.

d. 3) Environmental Cleanup Costs

Similarly, a base number is extracted from PHMSA (2016), and extrapolated to obtain the other values.

Further details on incident cost estimation are provided in Appendix D. Finally, we note that because incident probabilities (*Part c*) and incident consequences (*Part d*) are already estimated, a network-wide *Composite Risk Measure* can be calculated by multiplying probabilities and consequences for each incident situation and taking an expectation over all.

e) Equipment Transport Cost (TC_{ia}^{e})

Fuel economy information for two types of fire engines (Oak Ridge National Laboratory, 2016) is used to estimate the cost of moving equipment packages with 50-ton and 100-ton capacities (Table 4.9). A nonlinear extrapolation scheme is then applied to these two numbers to estimate transport cost for larger equipment packages. For further details, refer to Appendix E.

Gross Weight Range (lbs.)	Typical Fuel Economy Range in 2007		Average Fuel Economy (L/100km)	Average Fuel Cost (C$/100km)
	Miles per gallon	Litre per 100km		
26001-33,000	4.0-8.0	29.4-58.8	44.1	48.1
33,001-80,000	2.5-6.0	39.2-94.1	66.6	72.6

Table 4.9. Weight and fuel economy for two types of fire engines

f) Equipment Operating Cost (OC_{t}^{m})

PHMSA (2016) provides estimates of Response Cost by hazmat class and release volume, but it is not clear to us what cost components it exactly represents. In the absence of additional information, we consider such response cost as a proxy for operating cost of equipment packages in response to hazmat incidents. Further, we assume that the response personnel wages are included in this cost.

g) Facility Capacity (K)

For simplicity, we assume that each emergency response facility can accommodate up to 12 equipment packages of each type, i.e., 48 equipment packages in total.

h) Required Network Coverage (α^m)

Response to incidents of specific hazmat classes may be prioritized over the others. For example, the decision maker may be interested in reaching at least 75% volume-weighted response coverage of Class 3 hazmat incidents and 50% coverage for the other two classes under investigation, i.e., α^m = (0.5,0.75,0.5). We will experiment with several values to see how this parameter would affect the location and allocation decisions. It should be noted that the coverage distance is assumed to be 250 km. This means that a potential response facility is able to cover a hazmat emergency on a given rail link if the *entire* link is within a circle with a radius of 250 km around the facility.

4.4.2 Solution and Discussion

(ERPP) is an integer programming problem with 1,877,935 (=100+100×4×3+100×103×4×3×15+103×3×15+100×4×3×15) variables for the case under consideration. Despite the notable size and complexity of the problem, it can be solved for *integer* combinations of α^m, e.g., (0,1,1), in order of minutes. Tables 4.10, Table 4.11, 4.12 provide a summary of solutions related to all possible integer combinations. CPLEX 12.4 (IBM, 2016) has been used to solve all the problem instances. Further, high-performance computers (SHARCNET, 2018) have been utilized to provide up to 16 parallel threads and up to 200 GB of memory.

Seven important observations are made based on our computational experiments: *First*, the (0,0,0) combination, which basically removes the constraint (4.17), leads to a *trivial* zero solution, i.e., *no open facility*. We will apply a scaling technique in Section 4.4.2.2 to tackle this issue for this specific case. *Second*, the (1,1,1) combination, which requires a 100% volume-weighted coverage of the network for every hazmat class, suggests locating 5 facilities at Guelph Junction (#16), Cornwall (#38), Biscotasing (#59), Red Rock (#79), and Keewatin (#88). Each of the five facilities should accommodate a total of <u>nine</u> equipment packages of *size 4*[8], i.e., <u>three</u> equipment packages for each hazmat class. *Third*, even if the policy constraint (4.17) applies to a single hazmat class only (Table 4.11), the number of facilities would be still five. In other words, in order to cover every part of the railroad network in Ontario, regardless of class, a constant number of facilities needs to be located. *Fourth*, optimal facility locations for complete coverage of "only Class 2" and "only Class 3" are exactly the same, however, they are different from (although geographically close to) those of "only Class 8". This is because in the case of a *single* hazmat class coverage, in addition to 1st stage cost, 2nd stage cost, although much smaller, has an impact on the optimal solution too, and the 2nd stage cost depends on the traffic distribution for that specific hazmat class. For example, the rail-links over which Class 2 and Class 3 freights are routed, are similar (partly due to similar origins and destinations and west-to-east traffic flow), but both of them are quite different from those of Class 8. Review sections 3.3.4, 3.4.3, and 3.5.4 for further details. *Fifth*, if one has to choose

[8] All the equipment packages are *size 4* because constrains (4.14) and (4.16) together ensure that a complete coverage decision is associated with response to incidents that involve up to 15 railcars, i.e., 1200 tons of hazmat contents. This is exactly the combined capacity of <u>three</u> equipment packages of *size 4*, which is preferred to other sizes in this case due to economies of scale.

between the complete coverage of only one of the three hazmat classes, Class 3 should be chosen since it results in the least 1st stage and 2nd stage costs. *Sixth*, although the number of facilities and the total number of equipment packages are the same for each case shown in Table 4.11, the complete coverage of Class 2 (i.e., second column from the left) requires a substantially higher budget, which is mainly because of the relatively higher equipment acquisition cost for this hazmat class. *Finally*, all possible combinations where at least two or three hazmat classes are required to be completely covered, would result in exactly the same facility locations. This is partly because the total volume of Class 2 and Class 3 carried by rail, each is considerably greater than that of Class 8 (See Table 3.2); as a result, the complete coverage requirement for at least one of these two major classes is enough to enforce the same facility locations every time.

Problem instance	$\boldsymbol{\alpha^m = (0,0,0)}$	$\boldsymbol{\alpha^m = (1,1,1)}$
Number of facilities	0	5
Facility ID #	-	16, 38, 59, 79, 88
Number of equipment packages	-	45
Network coverage	(0,0,0)	(1,1,1)
Stage 1 cost (C$)	0	63,811,000
Stage 2 cost (C$)	1,603	0.0010261
Objective value (C$)	1,603	63,811,000
CPLEX Solution time (h:mm:ss)	0:09:20	0:06:53

Table 4.10. Computational results for integer combinations of α^m

Problem instance	$\boldsymbol{\alpha^m = (1,0,0)}$	$\boldsymbol{\alpha^m = (0,1,0)}$	$\boldsymbol{\alpha^m = (0,0,1)}$
Number of facilities	5	5	5
Facility ID #	16, 38, 59, 79, 88	16, 38, 59, 79, 88	13, 38, 56, 77, 91
Number of equipment packages	15	15	15
Network coverage	(1,0,0)	(0,1,0)	(0,0,1)
Stage 1 cost (C$)	30,061,000	21,624,000	21,624,000
Stage 2 cost (C$)	1,366	394	1,447
Objective value (C$)	30,062,366	21,624,394	21,625,447
CPLEX Solution time (h:mm:ss)	0:08:16	0:08:39	0:08:51

Table 4.11. Computational results for integer combinations of α^m

Problem instance	$\boldsymbol{\alpha^m = (1,1,0)}$	$\boldsymbol{\alpha^m = (1,0,1)}$	$\boldsymbol{\alpha^m = (0,1,1)}$
Number of facilities	5	5	5
Facility ID #	16, 38, 59, 79, 88	16, 38, 59, 79, 88	16, 38, 59, 79, 88
Number of equipment packages	30	30	30
Network coverage	(1,1,0)	(1,0,1)	(0,1,1)
Stage 1 cost (C$)	46,936,000	46,936,000	38,499,000
Stage 2 cost (C$)	156	1210	238
Objective value (C$)	46,936,156	46,937,210	38,499,238
CPLEX Solution time (h:mm:ss)	0:07:54	0:07:31	0:07:14

Table 4.12. Computational results for integer combinations of α^m

Although the problem can be solved for integer combinations of α^m, in a reasonable time, this is not the case for *fractional* values of α^m. Therefore, we either use the solution methodology outlined in Section 4.3, i.e., *decomposition & Lagrangian relaxation heuristic*, or apply a *probability scaling* technique to conduct further computational experiments.

4.4.2.1 *Decomposition & Lagrangian Relaxation Heuristic*

This solution methodology, as captured in Figure 4.2, applies Lagrangian relaxation inside each decomposed problem, which is solved sequentially for each of the three hazmat classes. It is pertinent to mention that based on our experiments, neither decomposition by class nor Lagrangian relaxation can individually solve the problem in a reasonable time.

We decided to choose the sequence *Class 3- Class 2- Class 8* to solve the decomposed problem for *fractional* values of α^m; this decision is based on the aggregate volume of hazmat traffic in Canada, which is the largest for Class 3 and the smallest for Class 8. For illustration purposes, we show the solution for the case where the decision/policy maker is interested in covering 25%, 50%, and 75% of the volume-weighted length of the network in terms of emergency response to Class 2, 3, and 8, respectively (i.e., $\alpha^m = (0.25{,}0.5{,}0.75)$).

Problem instance	$\boldsymbol{\alpha^2 = 0.5}$	$\boldsymbol{\alpha^1 = 0.25}$	$\boldsymbol{\alpha^3 = 0.75}$
Number of facilities	2	2 (*same*)	3 (*one new*)
Facility ID #	34, 74	34, 74	34, 74, 95
Number of equipment packages	6	3	9
Network coverage	0.70274	0.44113	0.87131
Stage 1 cost (C$)	8,678,700	5,062,500	11,069,000
Stage 2 cost (C$)	25	26	40
Objective value (C$)	8,678,725	5,062,526	11,069,040
CPLEX Solution time (h:mm:ss)	0:51:36	0:07:14	0:16:39
Lagrangian multiplier (λ)	0.27	0.20	0.42

Table 4.13. Computational results for the case with 25%, 50%, and 75% network coverage for classes 2, 3, and 8

Table 4.13 shows solutions to the decomposed problems, each of which is solved in the previously mentioned sequence, i.e., Class 3, 2, and 8. The integrated solution corresponds to three facilities, nine equipment packages, and a total cost of $24,819,291 (=8,678,725+5,062,526+11,069,040). As expected, with relatively negligible 2nd stage cost, the total cost associated with this case is much smaller compared to the case of full coverage of all hazmat classes (See Table 4.10). It is also interesting to note that the solution time for the first decomposed problem is considerably longer than that of the following problems, which makes sense because the first problem is less constrained and hence, the associated feasible search area is relatively larger. The Lagrangian multipliers shown in the last row are computed based on the solution methodology outlined in Figure 4.2.

The 2nd column is specific to Class 3, and suggests two locations for emergency response facilities: Washago (#34) and Marathon (#74). Each of the two facilities should accommodate three equipment packages of *size 4* suitable to respond to *Class 3* incidents. The resulting volume-weighted network coverage is around 70% as shown. We fix the located facilities and continue with the next decomposed problem.

The 3rd column is specific to Class 2, which is not associated with any new location, but three equipment packages of *size 4* suitable to respond to *Class 2* incidents should be stockpiled at Washago (#34). Also, the resulting volume-weighted network coverage is around 44% as shown.

Finally, the 4th column is specific to Class 8, and suggests a new location: Alcona (#95). Figure 4.4 shows the three facility locations and their coverage area in terms of emergency response to hazmat incidents. Each of the three open facilities should then accommodate three equipment packages of *size 4* suitable to respond to *Class 8* incidents. Also, the resulting volume-weighted network coverage is around 87% as shown.

To illustrate how the decision variables related to incident coverage and equipment dispatching can be interpreted, consider the rail-link between Marathon and White River (link #76), as shown in Figure 4.4. The integrated solution suggests $Z^1_{76,t} = 0$ and $Z^2_{76,t} = Z^3_{76,t} = 1$ for any profile t, which means that none of Class 2 incidents on this link will be covered, but all Class 3 and 8 incidents on this link will be covered, regardless of the incident profile. Additionally, the solution suggests $N^{e,1}_{74,76,t} = 0$ for any equipment package size e and for any profile t, which is obvious since the decision not to cover Class 2 incidents on this link implies that no equipment package should be sent to this link in case of an incident. Also, the solution suggests the following for

Class 3 incidents: $N^{4,2}_{74,76,t} = 1$ for t values between 1 and 5, i.e., release from 1 to 5 railcar; $N^{4,2}_{74,76,t} = 2$ for t values between 6 and 10; and, $N^{4,2}_{74,76,t} = 3$ for t values between 11 and 15. For example, if the incident on link #76 involves release from 1 up to 5 railcars, a complete coverage of the incident would require one equipment package of *size 4* to be sent from the facility located at node #74, i.e. Marathon, to the incident site (Figure 4.4). Similar number of equipment packages should be dispatched in response to Class 8 incidents[9].

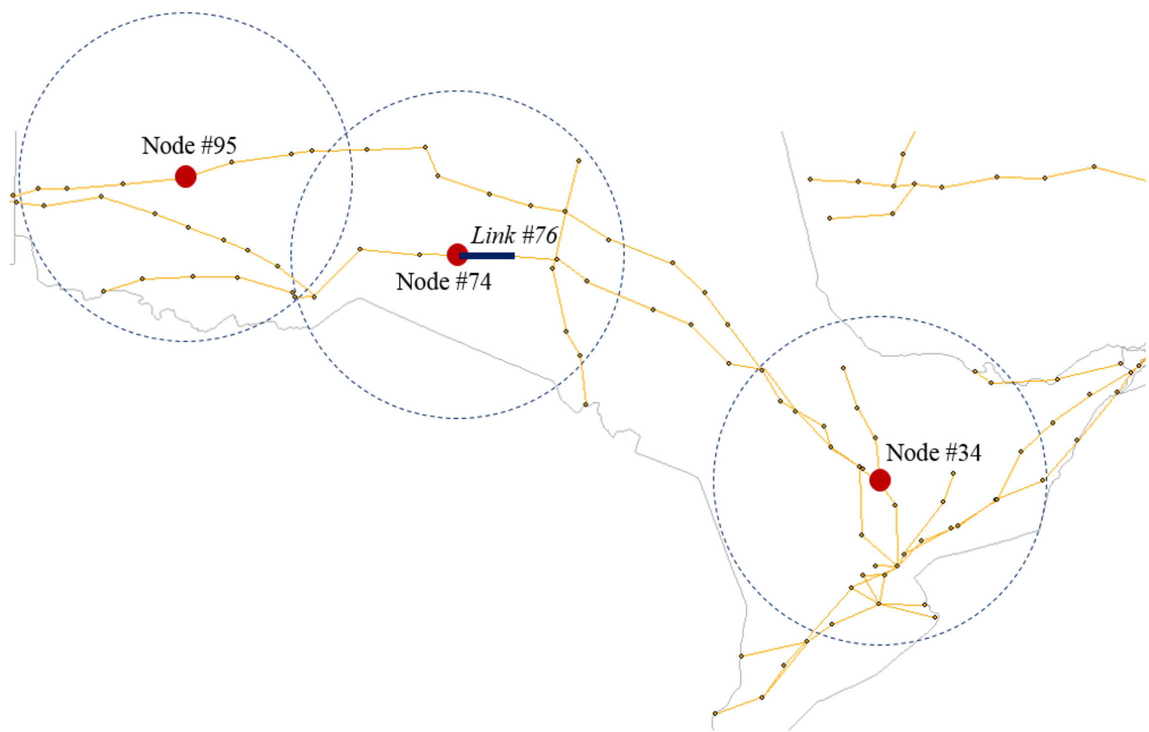


Figure 4.4. Emergency response facilities for the case with 25%, 50%, and 75% network coverage for classes 2, 3, and 8

[9] Each hazmat railcar is assumed to contain 80 tons. As a result, release from 5 railcars implies that the incident involves 400 tons of hazmat contents, which can be completely contained by an equipment package of size 4. However, the number of equipment packages dispatched to the incident site should increase proportionately if the incident involves more than 5 railcars.

Three additional points should be noted regarding the solution: (1) the aforementioned link (#76) is within the coverage area of this facility (#74), and therefore it could be reached in a timely manner; (2) for each of the hazmat classes 3 and 8, three specialized equipment packages of *size 4* are available at this facility, which can be utilized depending on the incident profile; and, (3) with a containment capacity of 400 tons, each equipment package of *size 4*, can effectively respond to a hazmat incident that involves up to five railcars.

Before introducing the alternative technique for solving **(ERPP)**, it would be interesting to see what would happen if the three decomposed problems are solved *independently* rather than sequentially.

Problem instance	$\boldsymbol{\alpha^2 = 0.5}$	$\boldsymbol{\alpha^1 = 0.25}$	$\boldsymbol{\alpha^3 = 0.75}$
Number of facilities	2	1	3
Facility ID #	34, 74	34	34, 80, 84
Number of equipment packages	6	3	9
Network coverage	0.70274	0.44113	0.90649
Stage 1 cost (C$)	8,678,700	6,046,900	12,999,000
Stage 2 cost (C$)	25	26	23
Objective value (C$)	8,678,725	6,046,926	12,999,023
CPLEX Solution time (h:mm:ss)	0:51:36	0:13:40	0:45:13
Lagrangian multiplier (λ)	0.27	0.24	0.35

Table 4.14. Computational results for the case α^m=(0.25,0.5,0.75) when solved independently

Table 4.14 shows the corresponding solutions. The 2nd column is exactly the same as that of Table 4.13, since Class 3 is the first in the sequence. Three observations are made by comparing the sequentially solved results with the independently solved results. *First*, in spite of similar network coverage requirements, the independent case suggests locating six facilities, i.e., twice the number of facilities in the sequential case. *Second*, the total cost of the independent case is $27,724,674, which is about $3 million more than that of the sequential case. This difference is mainly due to the additional cost of individual facilities for each hazmat class. *Third*, the average solution time in the sequential case is shorter than that in the independent case, which is reasonable since the search area of the former is relatively smaller compared to that of the latter.

4.4.2.2 *Probability Scaling*

In this section, we propose an alternative technique to deal with the complexity of the problem. As alluded to earlier, in case the policy constraint (4.28) is excluded from **(ERPP)**, using estimated incident probabilities would lead to a trivial zero solution, i.e., not opening any response facility. Although we drop the policy constraint, we propose a scaling technique to tackle the issue of zero solution and therefore obtain meaningful solutions that provide some emergency response coverage in the network.

Problem instance	**Scaling by 10^6**	**Scaling by 10^{12}**	**Scaling by 10^{15}**
Number of facilities	4	5	44
Facility ID #	21, 29, 60, 83	21, 33, 59, 74, 86	*[Not listed]*
Number of equipment packages	12	35	122
Network coverage	(0.00358, 0.04307, 0.00528)	(0.11248, 0.12970, 0.01493)	(0.11248, 0.12970, 0.01493)
Stage 1 cost (C$)	14,356,000	51,520,000	163,640,000
Stage 2 cost (C$)	2,236,500	1,025,800,000	1,023,200,000,000
Objective value (C$)	16,592,500	1,077,320,000	1,023,363,640,000
CPLEX Solution time (h:mm:ss)	0:52:29	0:15:56	0:08:03

Table 4.15. Computational results for three sample instances using a probability scaling technique

This technique involves applying a magnifying factor to the current incident probabilities in order to truly distinguish between rail-links in terms of incident chance, i.e., the longer the link and the more hazmat volume it carries would directly impact the corresponding incident probability, and hence 2nd stage cost. As a result, unlike before, the

2nd stage cost will be able to compete with 1st stage cost, even if the coverage policy is not enforced.

Table 4.15 summarizes the computational results for three sample instances. For example, solution to the case where incident probabilities are magnified by a factor of 10^6 (2nd column of Table 4.15) suggests locating four facilities at Brampton (#21), Kingston (#29), Eton (#60), and English River (#83).

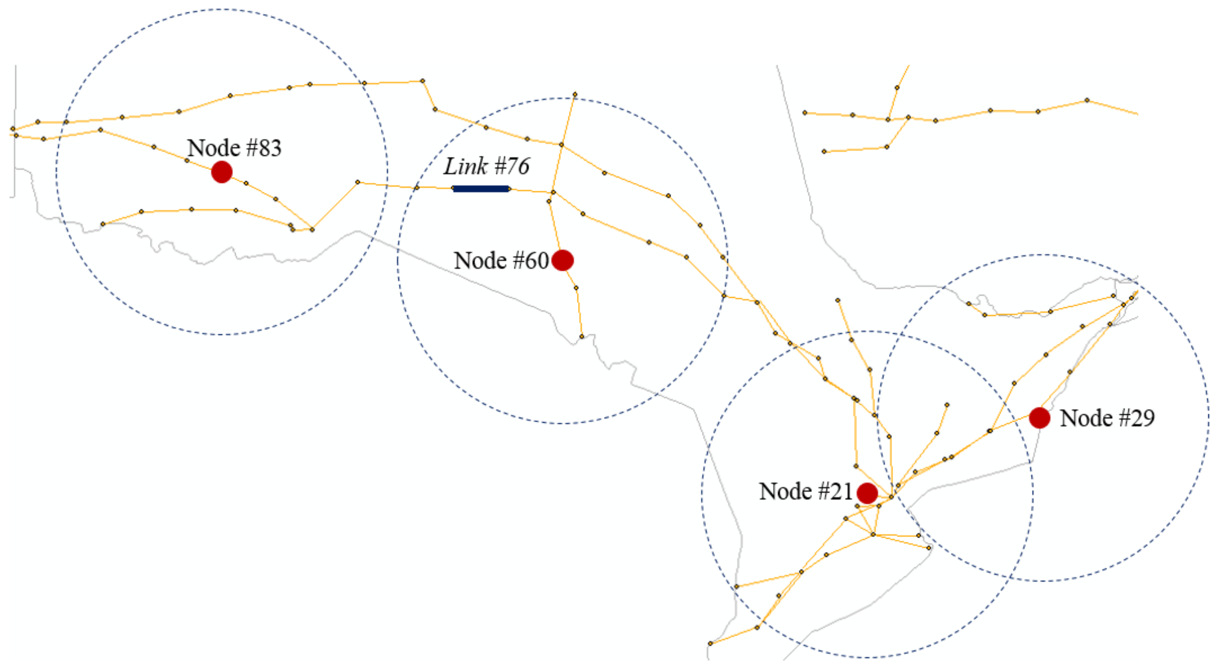


Figure 4.5. Emergency response facilities when incident probabilities are magnified by a factor of 10^6

Figure 4.5 shows these locations and the largest area of the network that each can cover in terms of emergency response to hazmat incidents. Also, a total of 12 equipment packages need to be purchased and stockpiled at these facilities. More specifically, the facility at Brampton should accommodate the following equipment packages: <u>one</u> of *size*

2 suitable to respond to *Class 2* hazmat incidents; one of *size 2* suitable to respond to *Class 8* incidents; and, three of *size 4* suitable to respond to *Class 3* incidents. Additionally, the facilities located at Kingston, Eton, and English River should accommodate three, three, and one equipment packages, respectively. For brevity, we do not provide further details on equipment packages.

Considering the same problem instance, the resulting network-wide coverage suggests only 0.36%, 4.31%, and 0.53% coverage for Class 2, Class 3, and Class 8 hazmat incidents, respectively. It is important to note that the low percentage of network coverage is mainly because of neglecting those rail-links with zero hazmat volumes (and therefore zero incident probabilities) (See Section 4.4.1). Again, consider the rail-link between Marathon and White River (link #76), which is also shown in Figure 4.5. The solution suggests $Z_{76,1}^{1} = Z_{76,1}^{2} = Z_{76,1}^{3} = 1$, which means those hazmat incidents that involve release from only one railcar will be responded to, irrespective of hazmat class. Additionally, the solution suggests $N_{60,76,1}^{2,1} = N_{60,76,1}^{2,2} = N_{60,76,1}^{2,3} = 1$, which means in such incidents, one equipment package of *size 2* needs to be dispatched from the facility at Eton (#60) to the incident site. It is important to note that (1) the aforementioned link is within the coverage area of this facility, and therefore it could be reached in a timely manner; (2) the equipment package of *size 2* is available at this facility; and, (3) with a containment capacity of 100 tons, it can effectively respond to a hazmat incident that involves release from one railcar (about 80 tons).

Comparing the 1st stage and 2nd stage costs implies that the strategic level decisions, i.e., facility location and equipment acquisition, still cost one order of magnitude more than the tactical level decisions. Finally, the solution time, which is less than one hour, suggests that the problem can be solved in a reasonable time.

As it is expected, the number of facilities increases when the incident probabilities are further magnified. In particular, for the case where incident probabilities are magnified by a factor of 10^{15} (4th column of Table 4.15), the 1st stage costs become too small relative to the 2nd stage costs such that 44 response facilities are located. The maximum coverage of the network is 11.25%, 12.97%, and 1.49% for Class 2, Class 3, and Class 8 hazmat incidents, respectively, which is, in fact, equivalent to 100% coverage of the rail-links with *non-zero* hazmat traffic.

Finally, we simulate incident probabilities to provide a more general perspective on how the solution changes depending on input probabilities. More specifically, we generate incident probabilities P_{at}^{m} between 0 and 1, and run several instances to assess the sensitivity of facility location decisions to this parameter. Table 4.16 provides a summary of computational results for three sample instances. In all instances, five facilities should be opened, three of which is common among them: Woodstock (#13), White River (#71), and Dyment (#86). In addition, the solution to all instances suggest acquiring a total of 45 equipment packages, which includes 9 equipment packages in each open facility as follows: <u>three</u> of *size 4* suitable to respond to *Class 2* hazmat incidents; <u>three</u> of *size 4* suitable to respond to *Class 3* hazmat incidents; and, three of *size 4* suitable to respond to *Class 8* hazmat incidents.

Problem instance	Simulation #1	Simulation #2	Simulation #3
Number of facilities	5	5	5
Facility ID #	13, 33, 42, 71, 86	13, 35, 42, 71, 86	13, 33, 50, 71, 86
Number of equipment packages	45	45	45
Network coverage	(1,1,1)	(1,1,1)	(1,1,1)
Stage 1 cost (C$)	63,867,000	63,880,000	63,822,000
Stage 2 cost (C$)	760,850,000	773,990,000	758,410,000
Objective value (C$)	824,717,000	837,870,000	822,232,000
CPLEX Solution time (h:mm:ss)	0:56:22	1:11:52	0:59:22

Table 4.16. Computational results for three instances using simulated probabilities

As observed, the resulting network coverage in every instance is 100% for Class 2, Class 3, and Class 8 hazmat incidents. This is happening mainly for two reasons. *First*, randomly generated probabilities include *all* rail-links in the network, and is not necessarily limited to the rail-links with estimated non-zero hazmat volume. *Second*, given that the generated numbers are between 0 and 1, they make the 2nd stage costs one order of magnitude larger than the 1st stage costs, which drive the solution toward responding to each and every incident situation, i.e. *full coverage*.

4.5 Conclusion

The timely arrival of specialized equipment packages and trained personnel at the incident site is crucial in effective and efficient response to rail hazmat incidents. We propose an optimization model that deals with facility location and resource allocation at the strategic and the tactical levels, respectively. A two-stage stochastic programming approach is taken to integrate the strategic and tactical levels, as well as to account for the uncertainties associated with model parameters. For example, exact location, volume, and class of released hazmat commodity are the stochastic input parameters of the model. To the best of our knowledge, there is no study in the literature that integrates the differentiating aspects of the rail hazmat transportation into planning for hazmat incident response.

The proposed mathematical model is a combinatorial optimization problem, which implies that the computing time needed to find the optimal solution to realistic size problems could be extremely long. We proposed a heuristic solution methodology, which uses the special structure of the model to decompose the problem into smaller sub-problems, and to apply an appropriate relaxation technique. More specifically, this solution methodology applies Lagrangian relaxation inside each of the decomposed sub-problems, which is solved sequentially for each of the three hazmat classes, and they are integrated at the end.

After parameter estimation, which is partly based on the link-level estimation efforts in Chapter 3, we provided computational experiments for several problem instances. We showed how the solution can be interpreted in each case. Although the problem is considerably large with about two million decision variables, it can be solved for *integer* combinations of α^m in order of minutes. However, for fractional values, we applied the proposed solution methodology to tackle the complicating constraint and solve the problem efficiently. Finally, we presented an alternate technique, where the coverage policy constraint is removed, and scaling factors are applied to incident probabilities.

Chapter 5 Infrastructure Investment as a Risk Mitigation Strategy in Railroad Transportation of Hazardous Materials

5.1 Introduction

As indicated in Chapter 2, academic researchers have proposed a number of risk mitigation strategies for rail hazmat shipments. For example, the problem of routing hazmat shipments through less populated areas has been studied as a risk management technique in the literature. It is important to note that in general, the government (regulator) has no authority to impose specific routes on railroad companies (hazmat carriers), but it can mitigate the associated risk through regulatory tools such as *network design* and *toll setting* policies (Bianco et al., 2013). In this chapter, we propose a risk mitigation strategy based on infrastructure investment, i.e., building new railway tracks to take hazmat traffic away from the riskiest locations across the network. In particular, this strategy consists of designing a cost-effective network and setting penalties for using high-risk links. Adopting such a strategy could result in effectively reducing the risk associated with rail hazmat shipments. For instance, the railroad network in Canada is fairly sparse and linear, which implies that in many cases, there may be only one possible route between a given origin/destination (OD) pair for hazmat shipments. As a result, it is possible that some rail-links, particularly those connected to population centers, will be overloaded with hazmat traffic and therefore become too risky. The presence of alternative links in the network would lead to a further connected network where hazmat shipments could use alternative links and hence reach their destinations through less risky routes. For instance, the unit train carrying 72 railcars of crude oil from North Dakota to the east cost of Canada, which derailed in Lac Mégantic and led to a disaster in 2013, could have been routed through less risky routes and could have involved less catastrophic consequences, only if alternative less risky links had been available for this shipment.

Such an alternative network for hazmat transport is expected to benefit both the railroad companies and the regulators; it would not only facilitate mitigation of public risk, but also translate into better insurance rates and cleaner public image for the railroad companies, and fewer catastrophic episodes involving casualties for the regulators. Additionally, it would provide growth opportunities that are of interest to both corporate and regulatory players.

To identify an effective infrastructure investment strategy, it is crucial to be able to measure the extent to which an investment decision is beneficial to a system. Furthermore, since infrastructure investment projects have a finite budget, a framework to assess the associated trade-off between costs and benefits is desirable. As a result, the methodology presented in this chapter follows a general *Cost Benefit Analysis (CBA)* model, but is tailored to the characteristics of the specific problem under study. *First*, decisions made by both the carrier and the regulator influence the outcome of the analysis performed to find an optimal investment strategy. The perspective of the former is captured through routing decisions that would minimize system-wide transportation cost, whereas that of the latter is captured via optimal investment decisions that would minimize the network-wide risk.

While there are other benefits associated with these investments, risk reduction is taken as the single surrogate measure for the societal benefits resulting from such an augmented network. *Second*, the viewpoint of the regulator is adopted in making the investment decision, assuming its willingness to invest in risk mitigation as well as its legal authority in setting regulations which would have to be followed by the carrier. *Third*, a bilevel optimization program is developed to conduct the analysis, with benefit, i.e., reduced risk, as the objective function and cost as one of the constraints. In other words, in this case, CBA is used to identify investment strategies that provide an optimal trade-off between the cost of the implemented measures and the corresponding risk reduction (Špačková and Straub, 2015). This way, we (1) avoid monetizing the expected societal benefits, and (2) provide flexibility to the decision makers to experiment with different values for the allocated budget and find the optimal strategy in every case.

The rest of this chapter is organized as follows. Section 5.2 briefly describes our methodology for identification of optimal infrastructure investment strategies. We conduct a risk assessment of the Canadian railroad network in Section 5.3, which enables us to select candidate areas and potential rail-links for investment. Section 5.4 provides the details on our bilevel programming model, followed by solution methodology in Section 5.6. We describe the problem instance and provide computational results and managerial insights in sections 5.6 and 5.7, respectively. Finally, conclusion and directions of future research are outlined in Section 5.8.

5.2 Methodology

In this section, we briefly outline our methodology to identify optimal infrastructure investment strategies that would mitigate the network risk associated with rail hazmat shipments. As indicated earlier, our methodology is based on a general CBA model, but takes into account the specificities of our study, such as the societal benefits associated with risk reduction, and the perspectives of railroad companies and regulators.

It is assumed that the regulator is willing to invest in risk mitigation and possess the legal authority to set regulations that have to be followed by the carriers, i.e., railroad companies. Consequently, we adopt the viewpoint of the regulator to make investment decisions that would maximize the benefits to society. As indicated earlier, several benefits may result from these investments, but we consider network risk reduction as the single surrogate measure for such benefits. In the context of rail track investment, this would be equivalent to an optimization problem for the regulator, wherein the total risk in the augmented network is minimized given an allocated budget. On the other hand, railroad companies still make routing decisions that would minimize their system-wide transportation cost, but need to use the available network and follow the regulations set by the regulator, e.g., tolls/penalties for using high-risk links. Therefore, they indirectly influence the investment strategy adopted by the regulator. In this chapter, we assume that the presence of new alternative links is combined with a policy tool, i.e., toll setting to dissuade the carrier from using the existing high-risk links. In such a setting, it is evident that the decisions made by both the carrier and the regulator would affect the outcome of the optimal investment strategy. As a result, a bilevel setting can be considered where the regulator is the leader, i.e., makes decision on risk-reducing infrastructure investment (and other regulations), and the carrier is the follower, i.e., makes routing decisions that minimize system-wide transportation cost in the augmented and regulated network. Although investment and routing decisions are made using different time frames, the regulator can still anticipate the carrier's routing behavior in the presence of alternative links. Figure 5.1 depicts the bilevel framework used in this chapter. It is important to note that in our framework, the regulator sets tolls/penalties for using high-risk links in the network; however, instead of being treated as decision variables in the regulator's problem (which would represent a more general case), these tolls/penalties are just considered as input parameters to the carrier's problem.

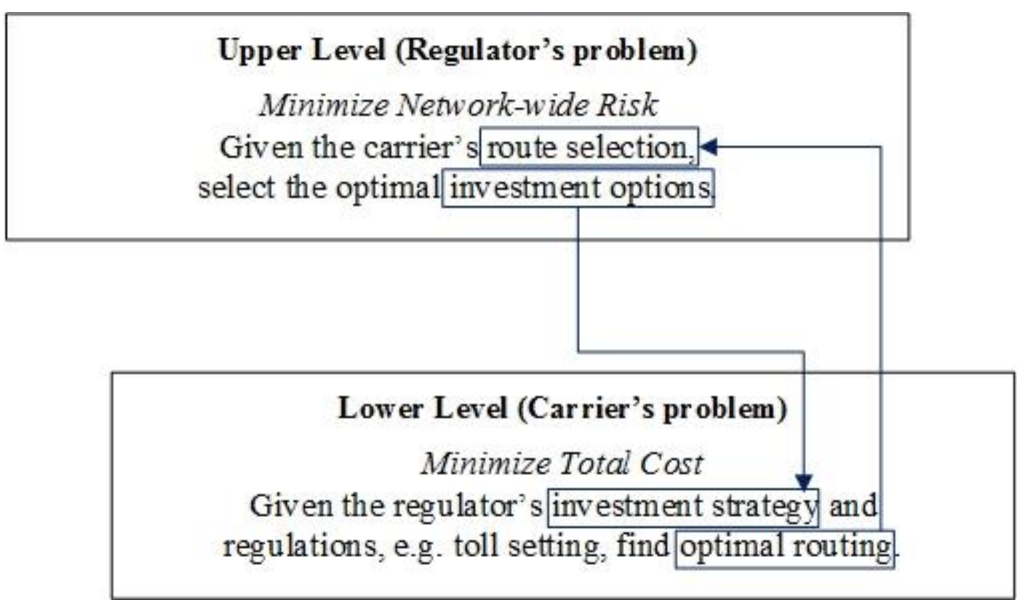


Figure 5.1. A bilevel framework for infrastructure investment

We make a first attempt to present a methodology to mitigate the risk associated with rail hazmat shipments using optimal rail-link infrastructure investment strategies. This methodology is applied to the Canadian railroad network, which is based on a 3-step procedure: *first*, we conduct a country-wide risk assessment of the railroads in Canada to ascertain hot spots and identify top ten high-risk links across the network, which will represent a number of candidate priority areas for investment; *second*, alternative rail-links are introduced in each candidate area such that a considerable volume of hazmat shipments is not anymore moved over the high-risk links; *third*, given an allocated budget to build rail-links, a bilevel optimization program is introduced with the regulator as leader and the carrier as follower. Details of the optimization program will be presented in Section 5.4.

As indicated earlier, since risk reduction represents benefit in our CBA framework, identification of candidate areas for investment would require a risk assessment of the network, which is the subject of the following section.

5.3 Ascertaining the Hot Spots

To decide the best investment strategies, we first need to identify potential areas for investment. We conduct a risk analysis of the Canadian railroad network for this purpose. In other words, hot spots in the network would serve as a basis for identification of the short-listed priority areas for infrastructure investment. Population exposure (PE) is the chosen risk measure in this chapter; as mentioned earlier in Chapter 2, this measure corresponds to a total number of people exposed to the possibility of harmful consequences due to a hazmat shipment (Verma and Verter, 2007). To estimate PE in the vicinity of each rail-link, two main inputs are required: (1) link-level hazmat traffic, and (2) population density information across the country.

Since the major railroad companies in Canada, i.e., Canadian National (CN) and Canadian Pacific (CP), usually seek to minimize their system-wide transportation cost, which seems to be the current practice, a minimum cost routing approach is taken to estimate link-level hazmat traffic all over the network. Here, we list the notations used, followed by the mathematical formulation of the routing problem **(P8)**:

Sets and indices

M: set of hazmat classes (i.e., classes 2, 3, and 8), indexed by m;

I: set of origin yards, indexed by i;

J: set of destination yards, indexed by j;

L: set of train services in the network, indexed by l;

P_{ij}: set of paths connecting yards i and j, indexed by p;

P_l: set of paths using train service l;

$(r,s) \in A$: set of rail links (arcs) in the network;

Decision variables

X_{ij}^{pm}: number of railcars with hazmat Class m using path p to travel from origin yard i and destination yard j;

N_l: number of trains of type l needed in the network;

Q_{rs}^{m}: number of railcars with hazmat Class m moving over rail-link (r,s);

Parameters

C_{ij}^{pm}: cost of moving one railcar with hazmat Class m to travel from yard i to yard j, using path p;

C_l: fixed cost to operate train service of type l;

U_l: capacity of train service of type l;

D_{ij}^{m}: number of railcars with hazmat Class m demanded on a biweekly basis at yard j from yard i;

δ_{rs}^{p}: incidence indicator that equals 1 if the link (r, s) is on path p, 0 otherwise.

(P8)

Minimize

$$\sum_{m\in M}\sum_{i\in I}\sum_{j\in J}\sum_{p\in P_{ij}} C_{ij}^{pm} X_{ij}^{pm} + \sum_{l\in L} C_l\, N_l \tag{5.1}$$

subject to

$$\sum_{p\in P_{ij}} X_{ij}^{pm} = D_{ij}^{m} \qquad \forall i \in I, \forall j \in J, \forall m \in M \tag{5.2}$$

$$\sum_{m\in M}\sum_{i\in I}\sum_{j\in J}\sum_{p\in P_{ij}\cap P_l} X_{ij}^{pm} \le U_l\, N_l \qquad \forall l \in L \tag{5.3}$$

$$Q_{rs}^{m} = \sum_{i\in I}\sum_{j\in J}\sum_{p\in P_{ij}} \delta_{rs}^{p}\, X_{ij}^{pm} \qquad \forall\, (r,s) \in A, \forall m \in M \tag{5.4}$$

$$X_{ij}^{pm} \ge 0 \;\; integer \qquad \forall i \in I, \forall j \in J, \forall m \in M, \forall p \in P \tag{5.5}$$

$$N_l \ge 0 \;\; integer \qquad \forall i \in I, \forall e \in E, \forall m \in M \tag{5.6}$$

$$Q_{rs}^{m} \ge 0 \;\; integer \qquad \forall\, (r,s) \in A, \forall m \in M \tag{5.7}$$

The objective in (5.1) contains not only the cost to transport railcars but also the fixed cost to operate each type of train service. Constraint (5.2) ensures that demand for hazmat cargo is met at the tactical level, e.g., on a biweekly basis. Constraint (5.3) states that the frequency of each train type is determined by the number of railcars, moving on different paths between origin and destination yards, using that particular train service. Constraint (5.4) converts the path flow into link flow, which is mainly used for the risk assessment purpose. Finally, constraints (5.5) to (5.7) indicate integrality and sign restrictions for the variables.

5.3.1 Parameters

We use \$0.50 as the cost to move a railcar one mile (Ahuja et al., 2007). Therefore, depending on the length of the path between a given origin i and a given destination j, it is possible to calculate C_{ij}^{pm}. In the absence of relevant information, we do not differentiate between hazmat classes in calculating this cost. It is pertinent to note that 0.50 USD/mile is approximately equal to 0.39 C\$/km. Based on the work of Verma (2009), we assume an hourly rate of \$500 for the fixed cost of operating a train service. This consists of the hourly payment rate for a driver, and engineer, a brakeman, and an engine at \$100, \$100, \$100, and \$200 respectively. To calculate C_l, the average speed of freight train is assumed to be around 22 miles per hour (Railroad Performance Measures, 2018). The hourly rate of 500 USD is approximately equal to 623 C\$. Furthermore, a capacity of 100 railcars is considered for each train service ($U_l = 100$).

To estimate the demand for each hazmat class and each origin/destination pair, we use the aggregate country-wide class-specific supply/demand information in Vaezi and Verma (2017) (See Table 3.3 and Table 3.4 for Class 2 supply/demand information), to propose an allocation scheme such that a traffic demand value D_{ij}^{m} can be generated for hazmat Class m and OD pair ij. Such a scheme is simply based on the relevant share of each origin/destination location in terms of the hazmat volume supplied or demanded at it. For instance, it was estimated by Vaezi and Verma (2017) that among all the rail shipments of Class 2 hazmat, i.e., gases, in 2016, 600,128 tonnes (9.23% of total tonnage carried) came from High Level, AB (See Table 3.3). Additionally, it was estimated that 56,627 tonnes (0.87% of total tonnage carried) were shipped to Saskatoon, SK (See Table 3.4). The abovementioned percentages, i.e., 9.23% and 0.87%, represent the shares of net supply and net demand in 2016 for High Level and Saskatoon, respectively. Consequently, we estimate that about 0.0803% (=9.23% * 0.87%) of the total volume of Class 2 moved by rail in 2016 is related to those shipments that are coming from High Level and traveling to Saskatoon, which would result in the annual traffic demand for this specific OD pair.

Considering that railroad companies make routing decisions at the tactical level, it would be useful to investigate whether the planning horizon is weekly or biweekly, given the identified origins and destinations. To that end, we choose the OD pair associated with the longest distance, and calculate the trip duration for a freight cargo travelling between the pair; in case it is less than 5 days, we will use the weekly horizon, and otherwise, we will consider a biweekly planning horizon. The recreated Canadian railroad network is a directed graph with 464 nodes and 940 arcs (See Section 3.3.5). The farthest OD pair seems to be Fort Nelson-Halifax (Node #28 to Node #17), which is related to Class 3 hazmat, i.e., flammable liquids. Solving a simple shortest path problem, we find the corresponding distance to be 5,856 km. Considering the average speed of freight train (22 miles per hour, or 35.41 km per hour), it would take at least 6.89 days for a freight train to travel between these two locations. Therefore, we use a biweekly planning horizon and hence, biweekly demands. For example, the biweekly Class 2 hazmat demand for the OD pair High Level-Saskatoon is approximately two railcars.

Finally, it should be noted that the incidence indicator δ_{rs}^{p} corresponds to a matrix in which each row corresponds to a specific link in the network, and each column corresponds to a specific path for hazmat shipments. The corresponding element in the matrix is set to 1 if the link is part of the path, and set to 0 otherwise.

5.3.2 Results

(P8) was solved using CPLEX 12.6.2.0 (IBM, 2016) based on the biweekly hazmat traffic demand by class in 2016. Here, we present some visual results to illustrate how such information could be useful in making infrastructure investment decisions. In Figure 5.2, we see traffic over the railroad network (in terms of railcars) for the three major hazmat classes. Being consistent with our observations in Chapter 3, we see that hazmat Class 3 has a higher share in railroad transportation compared to the other two classes.

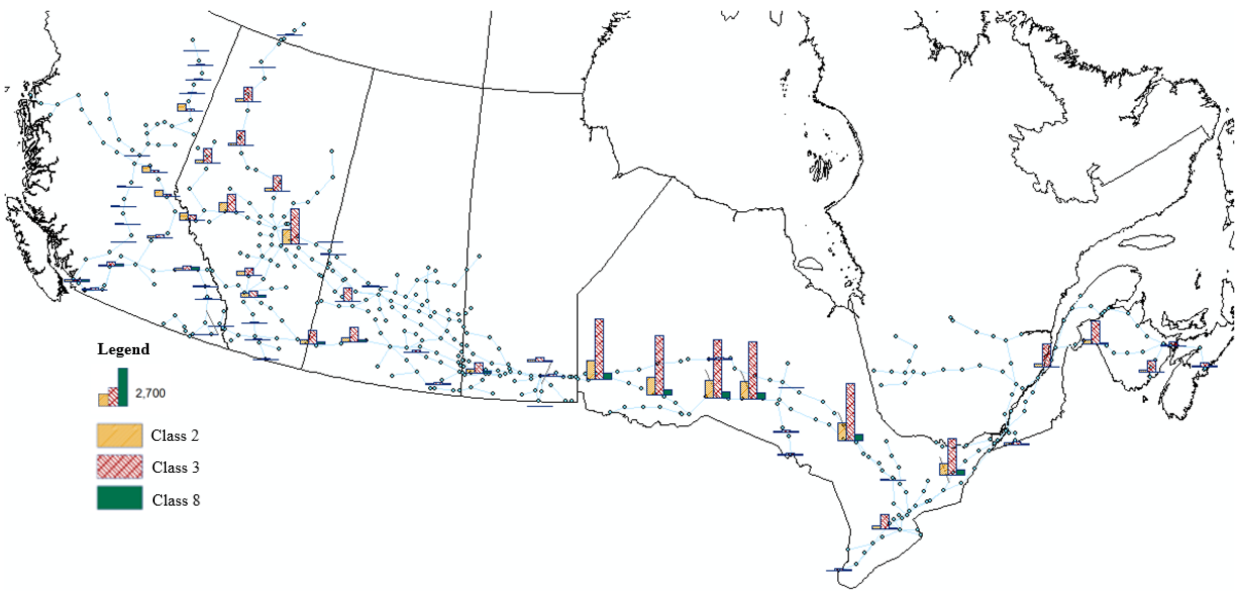


Figure 5.2. Distribution of hazmat traffic over the Canadian railroad network in 2016

To illustrate the results in finer details, Figure 5.3 shows the relative share of each hazmat class moved over the railroad network in Ontario in 2016. The size of each pie chart corresponds to the total number of railcars moved over a given part of the network.

For the specified link, i.e., the rail link between Batchewana and Sault Ste. Marie, the number of railcars for each hazmat class is shown as well. In particular, it indicates that 87, 175, and 60 railcars containing hazmat classes 2, 3, and 8, respectively, are exported to the US through the export terminal at Sault Ste. Marie, on a biweekly basis. As indicated earlier, each train is assumed to have a capacity of 100 railcars. Except for unit trains, the railcars are either mixed with regular freight (non-hazmat) or other hazmat classes in a given train. For example, 10 trains carry Class 2 along with regular cargo, and each train consists of 4 Class 2 railcars on average (which means the remaining 96 railcars contain regular commodities); additionally, 5 trains carry Class 2 along with Class 3 and regular cargo, and each train consists of 9 class 2 railcars and 25 Class 3 railcars, on average.

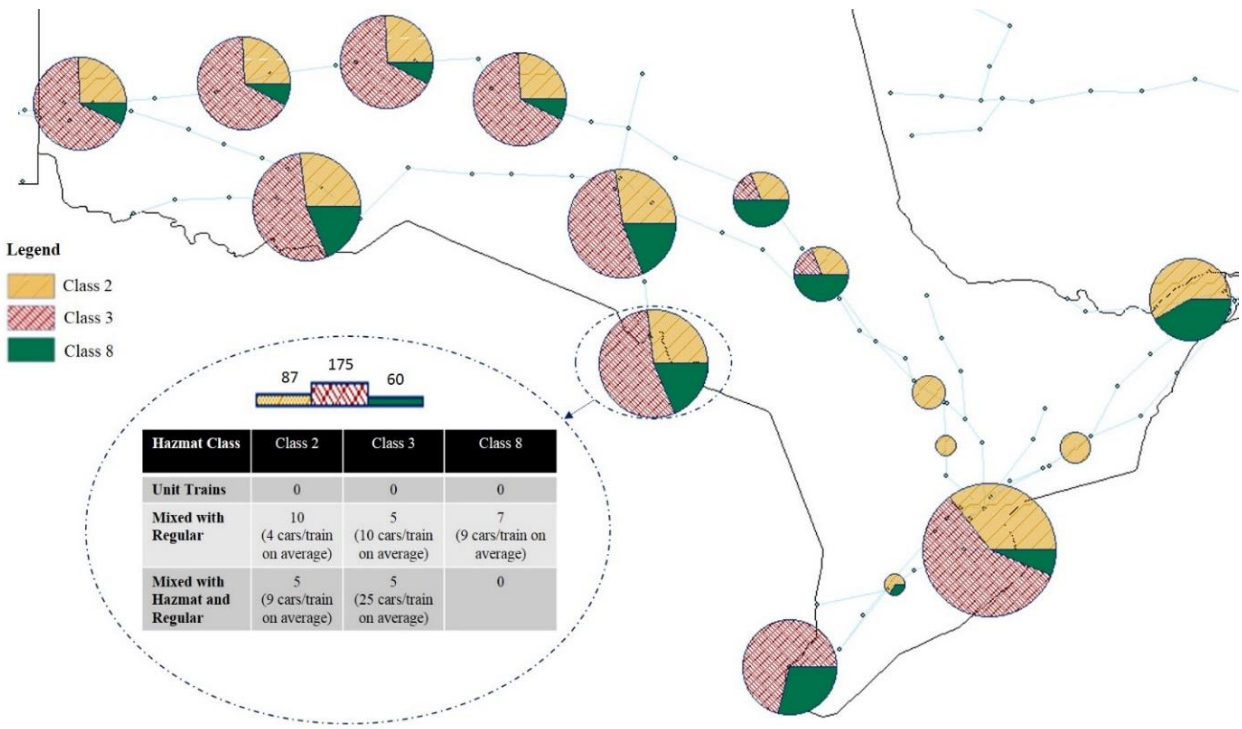


Figure 5.3. Railroad network in Ontario and link-level traffic by hazmat class in 2016

Based on a conservative approach adopted in this chapter, the most relevant information for risk analysis seems to be the maximum number of hazmat railcars in a given train, assuming it could meet with an accident and release contents on any given link with non-zero hazmat traffic. For instance, the maximum number of hazmat railcars in a given train moving over the specified link is 34 (= max {4, 10, 9, 9+25}). This corresponds to a danger area (i.e., potentially lethal within 60 seconds) with a radius of 1678 meters around the link, according to the analysis done using ALOHA software (US EPA, 2016a). This number was calculated based on Class 3, which has the largest impact area of the three hazmat classes, and a nonlinear extrapolation of the graph shown in Figure 4.3.

Population density information for different census divisions is provided by Statistics Canada (2016b). We use ArcGIS, version 10.3.1 (ESRI, 2016) to (1) create a buffer around each rail link given the corresponding danger area (which clearly depends on the maximum number of hazmat railcars that are moved over the link in a train), (2) intersect the resulting layer with the population density layer, and (3) calculate the risk, i.e., PE measure or the number of people seriously exposed to potential hazmat incidents, around each link. Consequently, a risk map could be created as shown in Figure 5.4. The top five high-risk areas on the network are those around Toronto, Montreal, Vancouver, Winnipeg, and Quebec City. In fact, the population exposure around links leading to each of these areas is over 70,000 people.

Link-level risk values would provide more detailed information on specific high-risk links which could be more helpful in proposing new alternative links for hazmat shipments. For expositional reasons, we focus on the top ten rail-links based on the selected risk measure. Table 5.1 lists the top ten links with highest risk levels in the network. As observed, these links are located in either Ontario or Quebec, near the two populous cities of Toronto and Montreal. The last two columns of the table account for all the three hazmat classes and show the total number of hazmat railcars and the maximum number of hazmat railcars moved over each of the top-10 high-risk links.

Figure 5.4. Risk distribution (PE) across the Canadian railroad network

As indicated earlier, the output of a country-wide risk assessment would be helpful to select candidate areas for investment. For example, new railway tracks can be built around one or more of the identified high-risk links/areas to mitigate the risk associated with rail hazmat shipments. The extent to which such a strategy is effective in risk mitigation depends on how the hazmat traffic is routed over the augmented network. It is conceivable that infrastructure investment projects have a finite budget, and therefore it may not be economically feasible to build an alternative track around every identified high-risk link/area. As a result, recognizing the behavioral differences between regulator and carrier in decision making, we present a bilevel programming approach which follows the general structure of a CBA, to find optimal risk-mitigating investment strategies, and to gain a better insight on how much network-wide risk reduction is feasible given an allocated budget.

LINK		Province	Risk (Exposed Population)	Hazmat Traffic (Railcars)	Maximum Hazmat Railcars
#	Name				
449	Toronto-Guelph Junction	ON	407,245	1,758	100
226	Toronto-Oshawa	ON	394,716	4,491	100
215	Toronto-Beaverton	ON	393,504	6,757	100
236	Montreal-Glen Robertson	ON	387,298	182	23
456	Toronto-Scarborough	ON	321,261	84	25
217	Toronto-Mississauga	ON	315,639	1,185	57
466	Les Coteaux-Saint Luc	QC	301,941	4,353	100
465	Montreal-Saint Luc	QC	189,487	3,643	100
448	Toronto-Alliston	ON	177,579	12	4
255	Montreal-Drummondville	QC	162,087	2,941	100

Table 5.1. Top 10 high-risk links of the Canadian railroad network

5.4 Infrastructure Investment Problem

An optimal investment problem can identify the best options for investment on alternative rail-tracks, considering the limited budget, in order to reduce the risk as much as possible (or equally minimize the total risk in the augmented network). For illustration purposes, consider a rail hazmat shipment moving from a given yard i to another yard j, as shown in Figure 5.5. Also, consider another hazmat shipment moving from yard i to yard z. Given the existing network, both shipments have to pass through a densely populated high-risk area in order to reach their destinations. Consequently, each of the three alternative rail tracks shown in Figure 5.5 is expected to provide less risky routes for these shipments. Of course, it is assumed that once an alternative new track is available, regulators can set regulations or provide incentives to encourage railroad companies to use the alternative links and hence avoid high-risk links or areas.

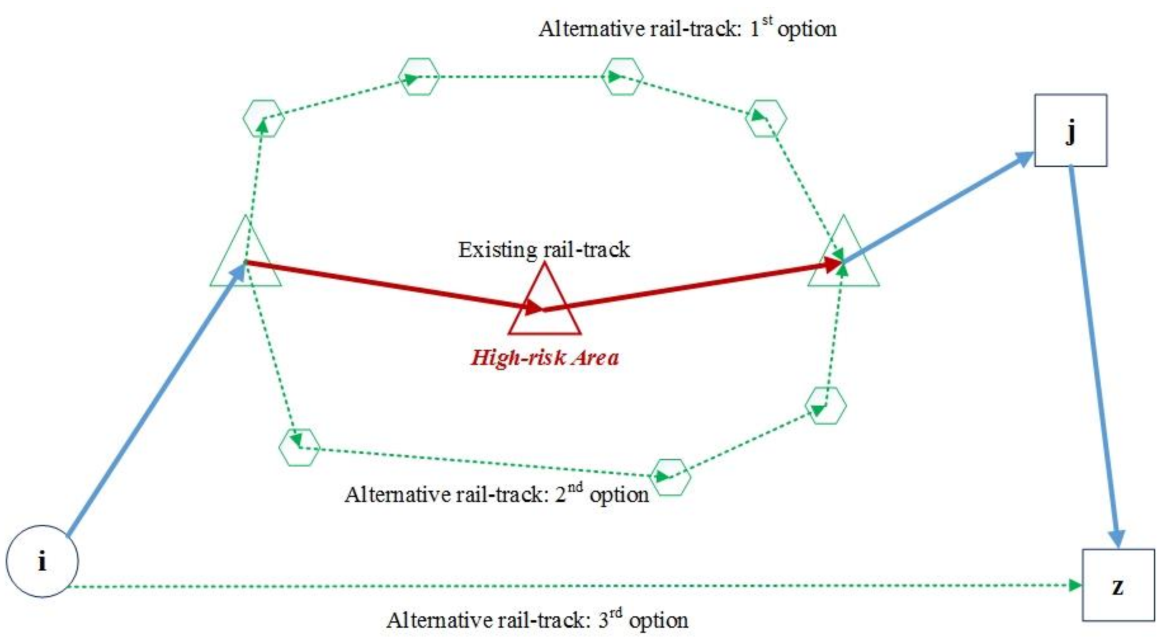


Figure 5.5. An illustration of the rail-track infrastructure investment problem

As mentioned earlier, a bilevel programming approach is appropriate in cases where each player knows the objective and feasible strategies available to the other player, i.e., *perfect information*. This seems to be the case in the railroad transportation of hazmat, where the varying perspectives of the railroad companies and the regulators are usually captured in terms of cost and risk minimization, respectively. If the regulator is willing to invest in risk mitigation and has the authority to enforce regulations, the presence of a bilevel setting is conceivable, which motivates us to take a bilevel programming approach.

The upper level (outer) problem would represent the regulator's risk minimizing behavior in infrastructure investment whereas the lower level (inner) problem would represent the carrier's cost minimizing behavior in routing hazmat shipments. We use the following notations followed by our mathematical formulation for the Bilevel Infrastructure Investment Problem **(BIIP)**:

Sets and indices

W: set of potential high-risk areas for rail-track investment, indexed by w;

K_w: set of alternative rail-track investment options around high-risk area w;

$(r,s) \in A$: set of rail links (arcs) in the network;

Decision variables

Y_w^k: decision variable that equals 1 if the investment option k is chosen around the high-risk area w, and 0 otherwise;

Q_{rs}^m: number of railcars with hazmat Class m moving over rail-link (r,s);

Parameters

ρ_{rs}^m: population density in the vicinity of rail-link (r,s) within the impact areas corresponding to potential Class m hazmat incidents;

Γ_{ij}^{pm}: cost of moving one railcar with hazmat Class m to travel from yard i to yard j, using path p, modified according to the toll/penalty policy on high-risk links;

(BIIP)

Minimize

$$Risk = f\ (Y_w^k, Q_{rs}^m, \rho_{rs}^m) \tag{5.8}$$

subject to

$$\sum_{w \in W} \sum_{k \in K_w} Y_w^k \leq B \tag{5.9}$$

$$Y_w^k \in \{0,1\} \qquad \forall w \in W, \forall k \in K_w \qquad (5.10)$$

Where Q_{rs}^m solves **(P8)** over the augmented network with modified transport cost matrix:

Minimize

$$\sum_{m\in M}\sum_{i\in I}\sum_{j\in J}\sum_{p\in P_{ij}} \Gamma_{ij}^{pm} X_{ij}^{pm} + \sum_{l\in L} C_l\, N_l \qquad (5.11)$$

subject to

$$\sum_{p\in P_{ij}} X_{ij}^{pm} = D_{ij}^m \qquad \forall i \in I, \forall j \in J, \forall m \in M \qquad (5.12)$$

$$\sum_{m\in M}\sum_{i\in I}\sum_{j\in J}\sum_{p\in P_{ij}\cap P_l} X_{ij}^{pm} \le U_l\, N_l \qquad \forall l \in L \qquad (5.13)$$

$$Q_{rs}^m = \sum_{i\in I}\sum_{j\in J}\sum_{p\in P_{ij}} \delta_{rs}^p\, X_{ij}^{pm} \qquad \forall\, (r,s) \in A, \forall m \in M \qquad (5.14)$$

$$X_{ij}^{pm} \ge 0\ \ integer \qquad \forall i \in I, \forall j \in J, \forall m \in M, \forall p \in P \qquad (5.15)$$

$$N_l \ge 0\ \ integer \qquad \forall i \in I, \forall e \in E, \forall m \in M \qquad (5.16)$$

$$Q_{rs}^m \ge 0\ \ integer \qquad \forall\, (r,s) \in A, \forall m \in M \qquad (5.17)$$

The objective of the outer problem in (5.8) indicates the network-wide calculated risk in the presence of new rail tracks. This risk measure cannot be expressed in the closed form and is a function of the following: the investment option chosen by the regulator, the link-level hazmat traffic as the output of the carrier's routing problem, and population density around rail-links within the impact area of the related hazmat class. Constraint (5.9) ensures that the total cost required for the selected investment options does not exceed the available budget. Constraint (5.10) contains the binary decision variables for investment options. Link-level hazmat traffic, i.e., Q_{rs}^m, is obtained by solving **(P8)** over the augmented network, where the cost of using the top ten high-risk links (as listed in

Table 5.1) for hazmat shipments is increased substantially for the carrier, e.g., 10 times regular cost, which is shown by Γ_{ij}^{pm} to indicate the new input parameter. This could be interpreted as a toll/penalty, set by the regulator, to discourage the carrier from using high-risk links.

5.5 Solution Methodology

As alluded to earlier, bilevel programs are intrinsically difficult to solve. The bilevel programming problem proposed in Section 5.4 involves a great deal of complexity. The inner level problem is an Integer Programming (IP) problem, and the outer level problem contains an objective function that cannot be represented in a closed form. Consequently, it is difficult to simplify the model, e.g., reformulate the existing bilevel model as a single-level model. Therefore, we propose a *heuristic* solution methodology with three main steps as follows. *First*, for a given budget, we identify the corresponding number of rail-links that could be added to the network, and then find different ways to select the investment options by generating all possible combinations. For example, if a total of ten investment options, i.e., candidate areas for investment, are identified, and the available budget is enough to build only two new rail tracks, a total of $\binom{10}{2} = 45$ possible distinct investment strategies exist. The best investment strategy is identified by solving the inner problem for every combination and selecting the one that that results in minimum risk in the entire network. It is important to note that the strategy identified will be optimal since all possible combinations are taken into account. *Second*, for every increment in the available budget, the network augmented through investment decisions in previous step will be considered as given. Among all possible new combinations, the best investment strategy is similarly identified by solving the inner problem for every combination over the augmented network and selecting the one that results in minimum network-wide risk. *Third*, the augmented network is updated in each step, and the process continues until no further risk reduction is possible in the network. The structure of this solution methodology is outlined in Figure 5.6.

Step 1. For a given allocated budget:

(a) Generate all possible combinations, i.e., investment strategies;
(b) Solve the inner problem for every single combination;
(c) Evaluate the network-wide risk in every case based on the routing choices made by the carrier in the inner problem;
(d) Select the case that corresponds to minimum risk as the best investment strategy.

Step 2. For an increment in the available budget:

(a) Update the network based on investment decisions made in previous step;
(b) Generate all new combinations;
(c) Solve the inner problem for every combination;
(d) Evaluate the network-wide risk in every case;
(e) Select the case that corresponds to minimum risk as the best investment strategy.

Step 3. If further network-wide risk reduction is achieved by the investment, repeat previous step. Otherwise, stop.

Figure 5.6. The proposed solution methodology for (BIPP)

5.6 Problem Instance

In this section, we describe a problem instance to which our model was applied. First, we need to create the potential augmented network according to the previously-identified high-risk areas and propose candidate investment options. We intend to select the potential alternative links intelligently, which is partly based on the current volume and direction of hazmat traffic around the two identified high-risk areas, i.e., Toronto and Montreal.

To illustrate, consider Figure 5.7, which shows part of the railroad network in Ontario. More specifically, it shows Class 3 volume (in railcars), flow direction over the links connected to the Toronto yard, and population density distribution (in term of people per square kilometers). Because Toronto is not an origin or destination for Class 3 hazmat shipments, the net flow at Toronto yard is zero, however it plays a key role as a transshipment point. For instance, based on the minimum cost routing approach taken by the railroad companies, i.e., *current practice*, over 4000 railcars enter the Toronto yard on a biweekly basis, which are then moved either toward Quebec or toward the four rail nodes that represent refineries/export terminals, i.e., Sarina, Windsor, Niagara Falls and Fort Erie. As a result, if alternative links in less populated areas are present in the network where needed, a significant portion of rail hazmat shipments (particularly Class 3) can bypass the Toronto yard, which would result in a reduction in total risk because of reduced population exposure. Based on such an argument, we select a total of 10 links around the two previously identified high-risk areas of the network, i.e., Toronto and Montreal, as potential options for rail track investment.

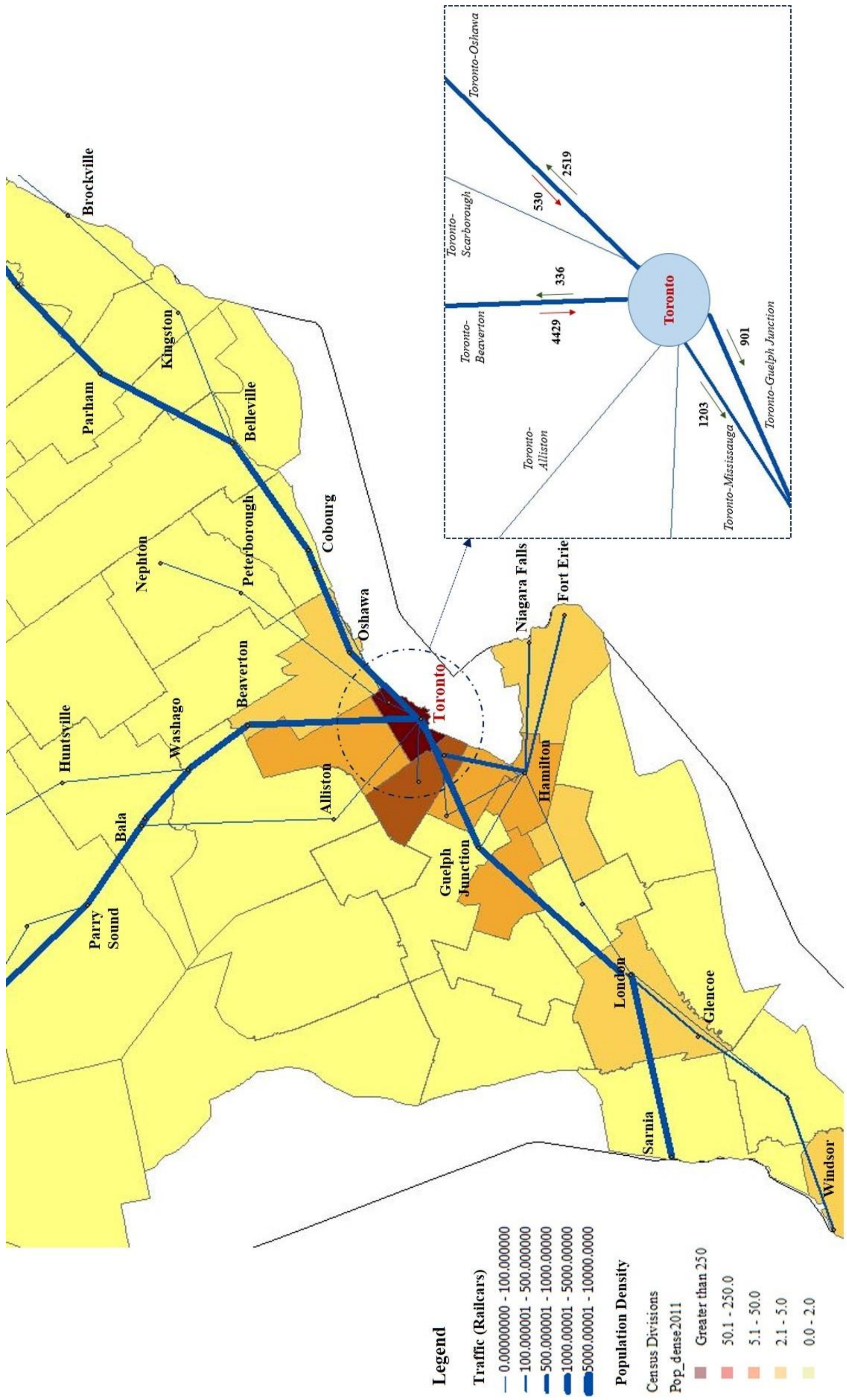


Figure 5.7. Class 3 hazmat volume, flow direction, and population density around Toronto

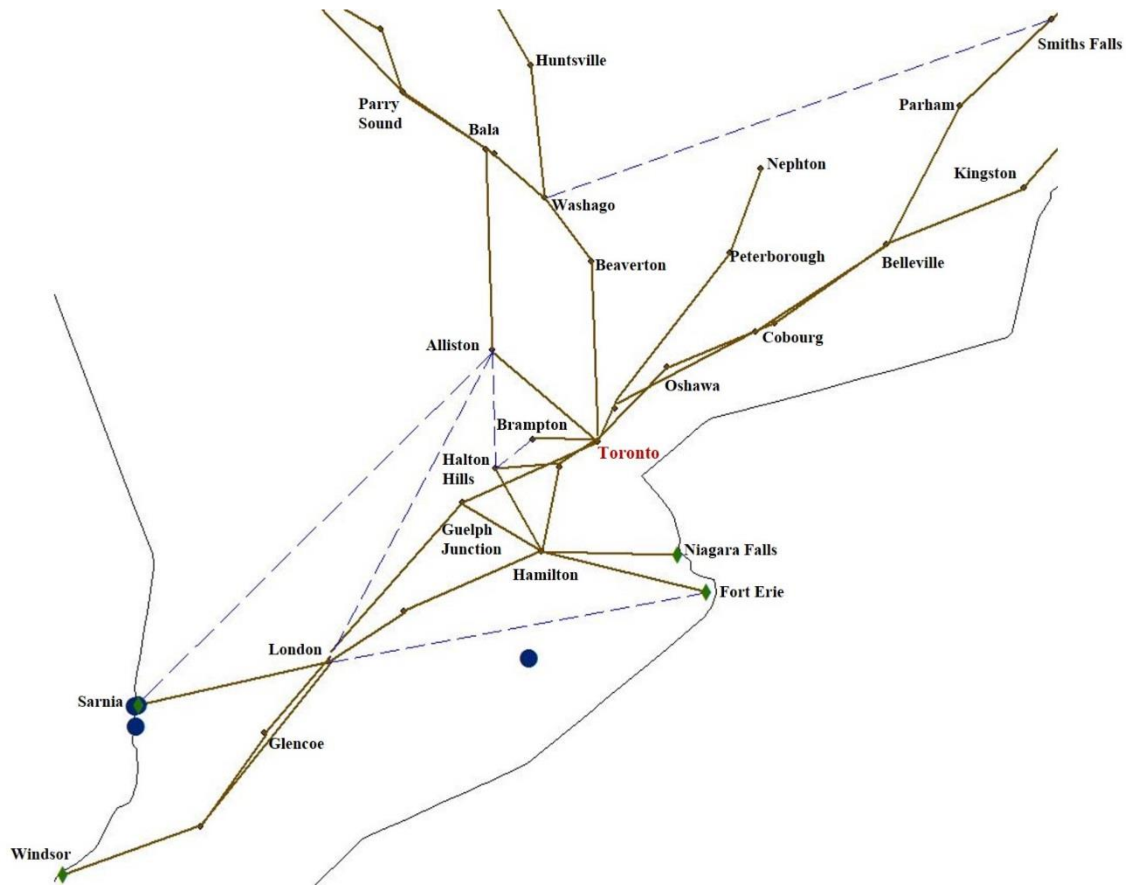


Figure 5.8. Proposed alternative links for hazmat transport around Toronto (candidate investment options)

Figure 5.8 shows the selected candidate links as investment options around Toronto. Six candidate links are shown by dashed lines and include Alliston-Halton Hills (#471), Alliston-London (#472), Alliston-Sarnia (#473), London-Fort Erie (#474), Brampton-Halton Hills (#475), and Washago-Smiths Falls (#476). The proposed rail links would provide alternative less risky routes for those hazmat shipments that use Toronto yard as a transhipment point. In particular, the proposed links provide alternative routes for Class 3 shipments that are coming from Western Canada and are traveling either toward Quebec or toward refineries (shown by blue circles) or export terminals (shown by green diamonds). It is important to note that each of the proposed links is assumed to be a straight line between the start and end nodes. Although this might not be the case in reality due to constraints such as land acquisition and geometry, we assume that the implementation of such rail track infrastructure projects is practically feasible.

Based on a similar logic, four alternative links are proposed around Montreal (not shown), which consist of Cornwall-Cantic (#477), Les Coteaux-St Jean (#478), Cantic-Drummondville (#479), and Curry Hill-Joliette (#480).

5.7 Discussion and Analysis

We apply the solution methodology described in Section 5.5, to the case of rail hazmat transport in Canada. *First*, we simply assume that the available budget is enough to build only two new rail tracks. As a result, our problem instance depends on the choice of two out of ten investment options (each of which corresponds to a proposed alternative rail link). This would lead to a total of $\binom{10}{2} = 45$ distinct combinations, i.e., 45 possible investment strategies. We consider this set of problem instances as our *base-case instance*. *Then*, the inner problem is solved for each of the 45 cases, which would provide the estimated link-level hazmat traffic by class in each case. The network-wide risk is evaluated in each case similar to the post-processing analyses described in Section 4. The case that results in the minimum network-wide risk is selected as the optimal investment strategy given the specified budget, i.e., budget for two rail-links. *Finally*, given the identified optimal investment option, an incremental analysis is conducted to identify the relationship between the investment budget and maximum achievable network-wide risk reduction. For example, when the allocated budget is enough to build one more new rail track, a total of $\binom{8}{1} = 8$ cases should be solved in the inner problem and then evaluated for risk to decide which link would be the next optimal investment option.

5.7.1 Solution to Base-case Instance

We start with the base-case instance where the allocated budget corresponds to a maximum of two rail-links. **(BIIP)** was solved using CPLEX 12.6.2.0 (IBM, 2016), a summary of which is shown in Table 5.2. As observed, the optimal investment option corresponds to the rail-links #471 and #478, which represent Alliston-Halton Hills (near Toronto area) and Les Coteaux-St Jean (near Montreal area), respectively. This investment option is expected to result in 25.73% risk reduction in the network. Not surprisingly, this investment option also results in maximum risk reduction in top-10 links. It is important to note that local risk may be increased in some parts of the network, but as we will see later, a better overall risk distribution is achieved compared to the current state (See Figure 5.9). Finally, it is clear in Table 5.2 that the minimum risk solution does not necessarily coincide with the maximum routing cost. In fact, the 25.73% risk reduction can be achieved at the expense of only 4.15% increased total routing cost for the carrier.

Table 5.3 provides further details on the previously identified optimal investment option, i.e., investment on links #471 and #478. Five of the top-10 high-risk links, shown by asterisks, are the same as those in the current state (See Table 5.1). It is interesting to note that one of the top-10 links in the new state is located in BC around the city of Vancouver. The local risk in some of the links has increased, e.g., Toronto-Scarborough

(#456). However, as indicated earlier, the network-wide risk has been reduced by 25.73% in the augmented network.

Link ID	**Routing Cost ($)**	**Total Risk (population)**	**Total Risk Reduction (%)**	**Top-10 Links Risk (population)**	**Top-10 Links Risk Reduction (%)**
None (status quo)	12,211,674	5,541,025	0.00	3,050,757	0.00
471-472	13,432,750	4,617,761	16.66	2,548,394	16.47
471-473	13,430,950	4,417,548	20.28	2,548,909	16.45
471-474	13,460,690	4,422,887	20.18	2,548,910	16.45
471-475	13,460,690	4,422,887	20.18	2,548,910	16.45
471-476	12,886,240	4,279,133	22.77	2,360,962	22.61
471-477	12,736,840	4,315,222	22.12	2,421,035	20.64
471-478	*12,718,260*	*4,115,167*	*25.73*	*2,319,929*	*23.96*
471-479	13,459,990	4,452,631	19.64	2,548,919	16.45
471-480	12,960,540	4,653,656	16.01	2,501,514	18.00
472-473	13,511,490	4,683,050	15.48	2,704,101	11.36
472-474	13,519,990	4,719,074	14.83	2,704,102	11.36
472-475	13,520,650	4,659,468	15.91	2,704,102	11.36
472-476	13,055,590	4,515,717	18.50	2,516,157	17.52
472-477	12,796,800	4,551,803	17.85	2,576,227	15.55
472-478	12,778,220	4,351,748	21.46	2,479,561	18.72
472-479	13,519,950	4,689,212	15.37	2,704,111	11.36
472-480	13,020,500	4,890,237	11.74	2,651,161	13.10
473-474	13,563,090	4,641,341	16.24	2,740,029	10.19
473-475	13,563,090	4,641,341	16.24	2,740,029	10.19
473-476	12,996,010	4,484,636	19.06	2,552,089	16.35
473-477	12,839,240	4,533,676	18.18	2,612,154	14.38
473-478	12,820,660	4,333,621	21.79	2,515,488	17.55
473-479	13,562,390	4,671,085	15.70	2,740,038	10.18
473-480	13,062,940	4,872,110	12.07	2,687,088	11.92
474-475	13,706,090	4,671,339	15.70	2,772,852	9.11
474-476	13,145,480	4,611,184	16.78	2,665,058	12.64
474-477	12,982,240	4,563,674	17.64	2,644,977	13.30
474-478	12,963,660	4,363,619	21.25	2,548,311	16.47
474-479	13,705,390	4,701,083	15.16	2,772,861	9.11
474-480	13,205,840	4,740,555	14.45	2,719,911	10.84
475-476	13,145,480	4,611,184	16.78	2,665,058	12.64
475-477	12,982,240	4,563,674	17.64	2,644,977	13.30
475-478	12,963,660	4,363,619	21.25	2,548,311	16.47
475-479	13,705,390	4,701,083	15.16	2,772,861	9.11
475-480	13,205,840	4,902,108	11.53	2,719,911	10.84

476-477	12,825,380	4,505,753	18.68	2,537,183	16.83
476-478	12,402,950	4,303,464	22.33	2,440,517	20.00
476-479	13,144,780	4,640,928	16.24	2,665,067	12.64
476-480	12,905,460	4,842,659	12.60	2,615,671	14.26
477-478	12,956,410	4,390,233	20.77	2,548,311	16.47
477-479	12,939,430	4,327,594	21.90	2,580,107	15.43
477-480	12,946,730	4,794,799	13.47	2,615,566	14.27
478-479	12,962,960	4,393,362	20.71	2,548,320	16.47
478-480	12,939,950	4,611,594	16.77	2,521,668	17.34
479-480	13,205,240	4,931,851	10.99	2,719,919	10.84
Minimum	12,402,950	4,115,167	10.99	2,319,929	9.11
Maximum	13,706,090	4,931,851	25.73	2,772,861	23.96

Table 5.2. Base-case instances and the impacts on the network-wide risk reduction

It is also interesting to note that the hazmat traffic has substantially increased in some links. For instance, the number of hazmat railcars moved over the link Toronto-Scarborough (#456) has increased from 84 to 4,544. However, the total number of hazmat railcars in the top-10 links has slightly decreased from 25,406 to 24,113, i.e., 5% reduction.

LINK		**Province**	**Risk (Exposed Population)**	**Hazmat Traffic (Railcars)**	**Maximum Hazmat Railcars**
#	**Name**				
456*	Toronto-Scarborough	ON	423,430	4,544	100
448*	Toronto-Alliston	ON	388,463	4,830	100
236*	Montreal-Glen Robertson	ON	387,203	182	23
226*	Toronto-Oshawa	ON	229,149	31	7
469	Montreal-Saint Jean sur Richelieu	QC	189,809	3,706	100
459	Scarborough-Port Hope	ON	169,833	4,544	100
255*	Montreal-Drummondville	QC	162,111	2,941	100
467	Saint Luc-Delson	QC	124,799	363	21
37	Vancouver-Matsqui	BC	124,786	29	24
471	Alliston-Halton Hills	ON	120,346	2,943	100

Table 5.3. Top 10 high-risk links in the augmented network including links #471 and #478

A more relevant measure, which we have used for the evaluation of link-specific hazmat risk in this chapter, is the maximum number of hazmat railcars. Among the five links that are common in Table 5.1 and Table 5.3, this measure has increased for links #456 and #448, decreased for link #226, and remained the same for the other two. For the link Toronto-Scarborough (#456), this measure has increased from 25 to 100, which is not a significant increase compared to the increase in total traffic, i.e., from 84 to 4,544 hazmat railcars. This highlights a limitation of PE as a risk measure, which does not account for the *total traffic* and therefore ignores the *frequency* of trains over rail-links.

Another observation is the resulting risk redistribution in the network. Not only the number of links with risk values greater than 10,000 is reduced from 56 to 50, but also the risk is redistributed in the network in a way that fewer number of links are exposed to relatively large hazmat risks. In other words, the histogram corresponding to the augmented network is more right-skewed, which implies a shift toward lower risk values (Figure 5.9). Although additional policy constraints may be set by the regulator to achieve a more *equitable* risk distribution in the network, such an analysis is out of the scope of this thesis.

At the end of this section, we visualize some results to see how the spatial distribution of hazmat traffic and risk around the Toronto area would change in the augmented network. For illustration purposes, Figure 5.10 (a) and (b) shows Class 3 hazmat traffic around Toronto area before and after the investment. It is clear that a lower number of railcars would go to Toronto rail yard in the augmented network. In the presence of the new link, i.e., Alliston-Halton Hills (#471), which is circled in Figure 5.10 (b), a significant portion of the shipments coming from the Western provinces would avoid the Toronto yard and take alternative routes to reach their destinations, i.e., Sarina, Windsor, Niagara Falls and Fort Erie. As a result, the risk would be considerably reduced in this part of the network. Similar observations and arguments can be made for the other two hazmat classes.

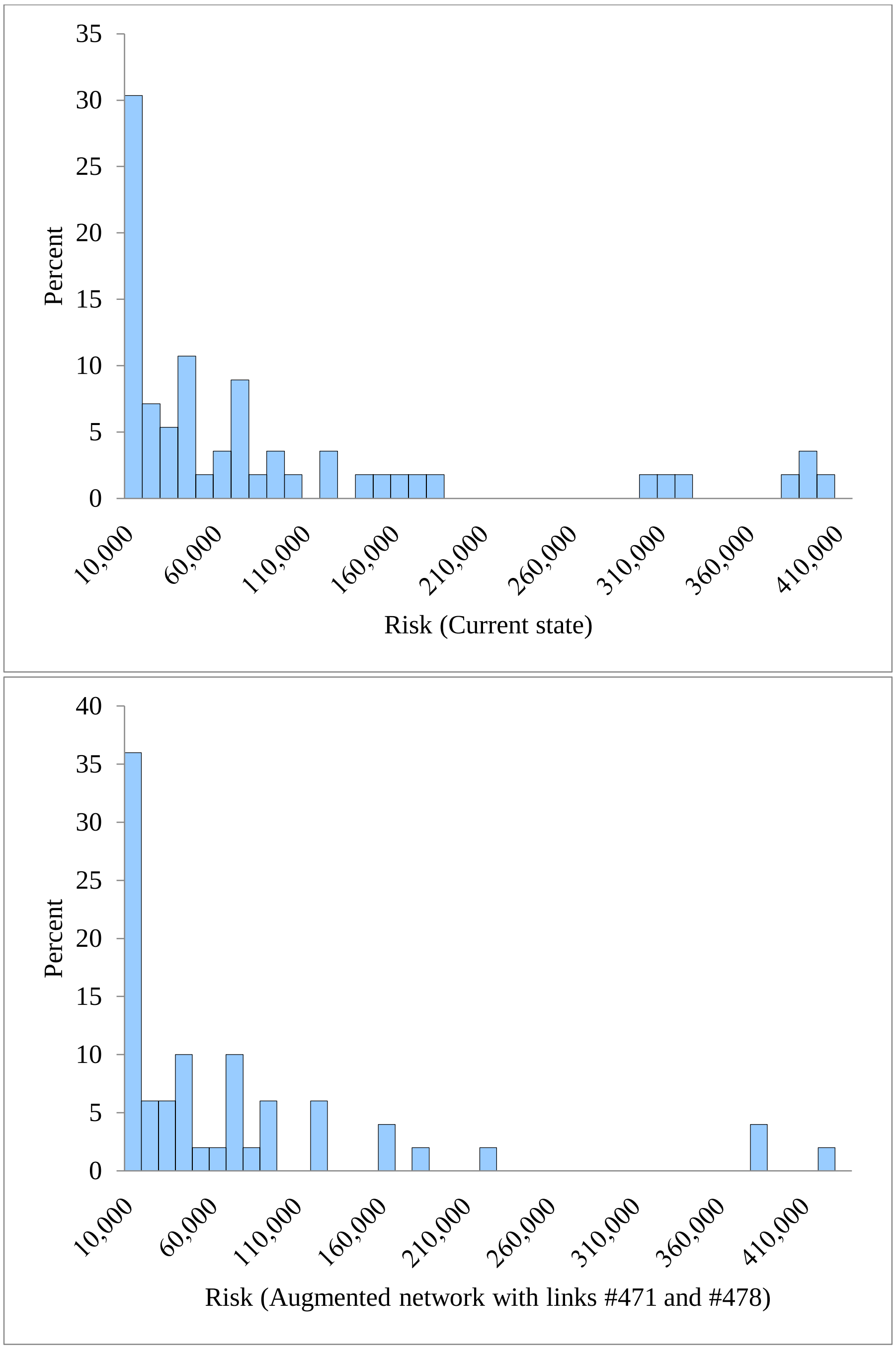


Figure 5.9. Redistribution of the risk in the network (Only links with risk values more than 10,000 are considered.)

(a) Current state

(b) Augmented network

Figure 5.10. Class 3 hazmat volume around Toronto

(a) Current state

(a) Augmented network

Figure 5.11. Hazmat risk (PE) around Toronto

Figure 5.11 shows how the population exposure risk is redistributed in this part of the network. As mentioned earlier, in PE risk calculation, the *maximum* number of hazmat railcars is directly relevant rather than the *total* number of hazmat railcars. In general, local risk has decreased in some parts of the network that are densely populated, but has increased in some other parts of the network that are less populated.

5.7.2 Incremental Analysis

In this section, we conduct an incremental analysis to estimate the relationship between investment budget and maximum achievable risk reduction in the network. As alluded to earlier, this analysis depends on the previously identified optimal investment options, and in each step provides the optimal investment solution if an additional budget is available to build one more new rail track. Consequently, a total of $\binom{8}{1}+\binom{7}{1}+\binom{6}{1}+\binom{5}{1}+\binom{4}{1}+\binom{3}{1}+\binom{2}{1}+\binom{1}{1}=36$ cases are investigated. The last term represents the situation where all the 10 candidate rail-links are selected for investment.

Figure 5.12 shows how much network-wide risk reduction is possible given the available budget. We have already seen a detailed analysis of the base-case problem instance: When the budget is available for only two rail-links, the total network risk can be reduced as much as 25.73%. Although further risk reduction is still possible, it happens at a decreasing rate with respect to the available investment budget. For instance, as illustrated in Figure 5.12, 28.33% and 28.42% risk reduction can be achieved if the budget is enough to build three links (point **b**) and four links (point **c**), respectively. However, no more network-wide risk reduction is possible even if more budget is available.

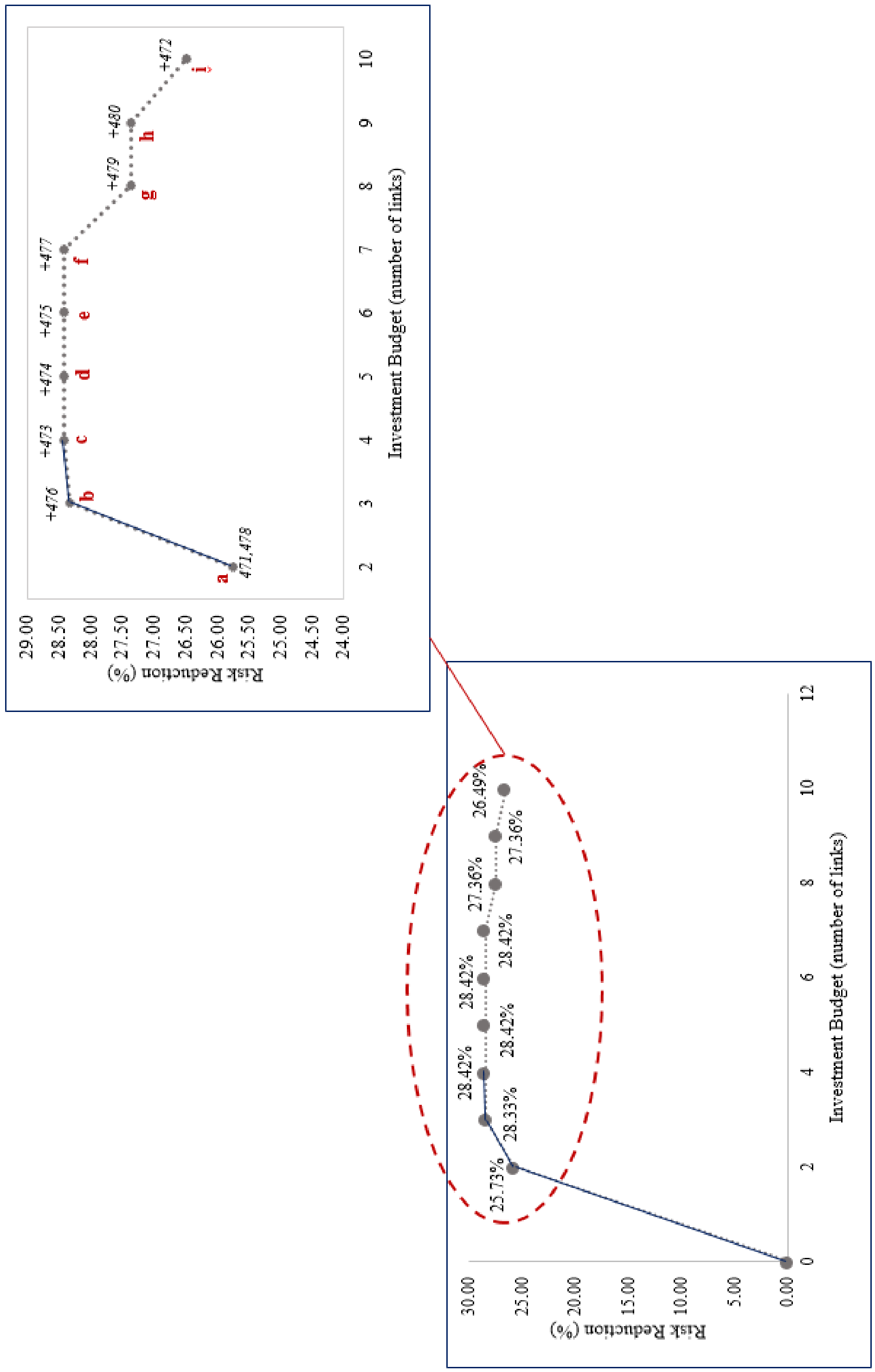


Figure 5.12. Relationship between investment budget and network-wide risk reduction

As shown in Figure 5.12, points **c**, **d**, **e**, and **f** are all associated with the same level of risk reduction, i.e., 28.42%. This is probably because after point **c**, none of the next three minimum-risk investment decisions (i.e., **d**, **e**, and **f**) will be on the shortest routes taken by the carrier. Therefore, they do not make any difference in terms of risk reduction after point **c**, which corresponds to investment on links #471, #478, #476, and #473. After point **f**, availability of more links would increase the network risk, probably because the augmented network would provide new shortest routes for some shipments, which might not be necessarily less risky compared to the alternative routes. As a result, the new routing decisions made by the carrier would involve transporting some hazmat shipments through more populated areas.

Augmented Network (471-478-476 & 471-478-476-473)		
LINK		Province
#	Name	
448	Toronto-Alliston	ON
236	Montreal-Glen Robertson	ON
456	Toronto-Scarborough	ON
469	Montreal-Saint Jean sur Richelieu	QC
255	Montreal-Drummondville	QC
459	Scarborough-Port Hope	ON
226	Toronto-Oshawa	ON
467	Saint Luc-Delson	QC
37	Vancouver-Matsqui	BC
471	N_Alliston-Halton Hills	ON

Table 5.4. Top 10 high-risk links in the augmented network with 3 and 4 new links

Finally, as listed in Table 5.4, the optimal investment strategies with three and four new links both result in the same ranking of the top 10 high-risk links. It is pertinent to note that these links are exactly the same as those shown in Table 5.3, however their relative ranking has changed. For instance, the link between Toronto and Alliston (#448) has become relatively riskier, and the link between Toronto and Oshawa (#226) has become relatively less risky.

5.8 Conclusion

We proposed a risk mitigation strategy based on infrastructure investment. This strategy involves building new railway tracks to deviate a portion of hazmat traffic away from the riskiest locations across the network.

Infrastructure investment projects typically have a finite budget, which motivated us to adopt a cost-benefit analysis framework to evaluate the societal benefits associated with different investment strategies. In hazmat transport literature, the government/regulator is typically concerned about transport risk whereas the carrier seeks to minimize routing costs while satisfying the safety requirements set by the regulator. Given such an interactive decision making environment, we considered a bilevel setting, where the regulator is in a leader position and makes decision on risk-reducing infrastructure investment as well as the corresponding regulations, and the carrier follows such regulations but still makes routing decisions that minimize system-wide transportation cost in the augmented network.

A 3-step procedure (risk assessment, alternative rail-links identification, and bilevel optimization) was used to apply the proposed methodology to the problem of rail hazmat transportation over the Canadian railroad network.

The proposed bilevel programming problem consists of (1) an outer level problem, which represents the regulator's risk minimizing behavior in infrastructure investment, and

(2) an inner level problem, which represents the carrier's cost minimizing behavior in routing hazmat shipments. The inner level problem is an Integer Programming (IP) problem, and the outer level problem contains an objective function that cannot be simply represented in a closed form expression. Consequently, it is extremely challenging to apply simplifying methods such as model reformulation. Therefore, we proposed a heuristic solution method, which is based on enumerating some combinations, solving the inner level minimum routing cost problem, evaluating the network-wide risk in the post-processing stage, and conducting an incremental analysis to assess the impact of further investment choices.

We started with the base-case instance where the allocated budget corresponds to a maximum of two rail-links. Then, we conducted an incremental analysis to estimate the relationship between investment budget and the maximum achievable risk reduction in the network. We acknowledge that the obtained result is partly influenced by our choice of potential rail-tracks for investment as described in Section 5.6.

In conclusion, the proposed methodology was useful in identifying the best rail-link investment strategies given the allocated budget, considering the interaction between the two main stakeholders in the railroad transportation of hazmat. Our computational experiments on the Canadian setting showed that significant network-wide risk reduction is possible if hazardous shipments are routed using some of the proposed alternative rail tracks. Application of the proposed methodology to other contexts or railroad networks is possible with some modifications. This methodology can be adopted to other settings to identify the best regulatory policies combining network design and toll/penalty setting.

Chapter 6 Conclusions and Future Research

The objectives of this thesis can be summarized as follows. The *first* objective is to analyze historical rail hazmat traffic flows in Canada to extract link-level information about hazmat volume. The *second* objective is to use a mathematical modeling approach to design a network for emergency response to rail hazmat incidents. The *third* objective is to identify the optimal set of infrastructure investment options to reduce the network risk associated with hazmat shipments.

In the following, we summarize the conclusions of this dissertation and present ideas for future research.

6.1 Conclusions

We identified three research gaps in rail hazmat transport literature. *First*, integrated analytical approaches on rail hazmat transport have not gained much attention in the past, and therefore a crucial input for risk assessment and management efforts, i.e., link-level hazmat volume, is missing from the related literature. *Second*, specific features of rail hazmat transport are not studied to propose optimization-based planning approaches to respond to rail hazmat incidents. *Third*, to the best of our knowledge, infrastructure investment has not been studied in hazmat transport literature as a risk mitigation measure.

Consequently, we focused on a number of works that address each identified gap. Chapter 3 addresses the first research gap using an integrated analytical approach to estimate link-level hazmat traffic over the railroad network in Canada. This chapter also includes an even more detailed analysis on crude oil rail shipments in Canada, mainly to investigate the relationship between routing schemes and network-wide risk, and to evaluate the impact of proposed pipeline projects on crude by rail at the link-level. Chapter 4 addresses the second research gaps by proposing a mathematical model to plan emergency response to rail hazmat incidents. Finally, Chapter 5 addresses the third research gap through studying infrastructure investment as a risk mitigation strategy in railroad transportation of hazmat.

In Chapter 3, we used data extraction, predictive modeling, and optimization to project the spatial distribution of Class 2, 3, and 8 traffic flow on the railroad network up to 2030. In addition, a focused analysis was conducted on crude oil shipments in Canada to develop long-term forecasts, compare alternate routing schemes in terms of network-wide risk, and evaluate the impact of new pipeline projects. The insights provided in this chapter would be also useful for undertaking more precise risk assessment and risk management exercises by academic and industry researchers.

We used a two-stage stochastic programming approach in Chapter 4 to integrate strategic and tactical decision levels in response to hazmat incidents. Relevant datasets, scientific literature, and several software packages were used to estimate the model parameters. A risk calculation scheme, called composite risk measure, was incorporated in the model to account for both incident probabilities and incident consequences.

Chapter 5 was concerned with another risk management technique in rail hazmat transport. We studied an infrastructure investment strategy to provide alternative routes to the riskiest parts of the network. We used a 3-step procedure (risk assessment, alternative rail-links identification, and bilevel optimization) to apply the proposed methodology to the problem of rail hazmat transportation over the Canadian railroad network. Given the interactions between the two main stakeholders, i.e., regulator and carrier, we proposed a bilevel programming approach to assess the effectiveness of the optimal investment strategies in terms of network-wide risk reduction.

6.2 Future Research Directions

Given the scope of the three main chapters in this thesis, we list three areas for future research. *First*, further analytics approaches could be conducted in rail hazmat transport domain to more accurately assess the associated risk. For example, by collecting relevant data, one could develop explanatory regression models to predict rail hazmat traffic based on variable selection techniques such as Stepwise Regression. *Second*, one could study an emergency response planning problem where some risk equity considerations are incorporated into the model. For example, the operational readiness in terms of emergency response can be particularly strengthened in the areas of the network with highest hazmat traffic. *Finally*, it would be interesting to study a bilevel programming problem for hazmat transportation, where both network design and toll setting policies are considered as decision variables in the outer level problem.

References

Abkowitz, M., Lepofsky, M., & Cheng, P. (1992). Selecting criteria for designating hazardous materials highway routes. *Transportation Research Record*, *1333*(2.2).

Ahuja, R. K., Jha, K. C., & Liu, J. (2007). Solving real-life railroad blocking problems. *Interfaces*, *37*(5), 404–419.

Assadipour, G., Ke, G. Y., & Verma, M. (2016). A toll-based bi-level programming approach to managing hazardous materials shipments over an intermodal transportation network. *Transportation Research Part D: Transport and Environment*, *47*, 208–221.

Bagheri, M. (2010). Risk-Based Model for Effective Marshalling of Dangerous Goods Railway Cars.

Bard, J. F. (2013). *Practical bilevel optimization: algorithms and applications* (Vol. 30). Springer Science & Business Media.

Barkan, C. P. L. (2008). Improving the design of higher-capacity railway tank cars for hazardous materials transport: Optimizing the trade-off between weight and safety. *Journal of Hazardous Materials*, *160*(1), 122–134.

Barkan, C. P. L., Ukkusuri, S. V, & Waller, S. T. (2007). Optimizing the design of railway tank cars to minimize accident-caused releases. *Computers & Operations Research*, *34*(5), 1266–1286.

Batta, R., & Chiu, S. S. (1988). Optimal obnoxious paths on a network: transportation of hazardous materials. *Operations Research*, *36*(1), 84–92.

Bayes, M., & Price, M. (1763). An essay towards solving a problem in the doctrine of chances. by the late rev. mr. bayes, frs communicated by mr. price, in a letter to john canton, amfrs. *Philosophical Transactions (1683-1775)*, 370–418.

Beale, E. M. L. (1955). On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society. Series B (Methodological)*, 173–184.

Belardo, S., Harrald, J., Wallace, W. A., & Ward, J. (1984). A partial covering approach to siting response resources for major maritime oil spills. *Management Science*, *30*(10), 1184–1196.

Berman, O., Verter, V., & Kara, B. Y. (2007). Designing emergency response networks for hazardous materials transportation. *Computers & Operations Research*, *34*(5), 1374–1388.

Bianco, L., Caramia, M., Giordani, S., & Piccialli, V. (2013). Operations research models for global route planning in hazardous material transportation. In *Handbook of OR/MS*

Models in Hazardous Materials Transportation (pp. 49–101). Springer.

Bianco, L., Caramia, M., Giordani, S., & Piccialli, V. (2015). A game-theoretic approach for regulating hazmat transportation. *Transportation Science*, *50*(2), 424–438.

Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming*. Springer Science & Business Media.

Bowerman, B. L., O'Connell, R. T., Schermer, J. A., & Adcock, J. R. (2009). *Business Statistics in Practice*. Mc Graw Hill.

BP. (2016). *Approximate Conversion Factors - Statistical Review of World Energy*. Retrieved from https://www.bp.com/content/dam/bp/pdf/energy-economics/statistical-review-2016/bp-statistical-review-of-world-energy-2016-approximate-conversion-factors.pdf

Canadian Association of Petroleum Producers (CAPP). (2017). *Crude Oil Forecast, Markets and Transportation*. Retrieved from http://www.capp.ca/publications-and-statistics/publications/303440

Chateau, B., & Lapillonne, B. (1978). Long-term energy demand forecasting a new approach. *Energy Policy*, *6*(2), 140-157.

Cheng, J., Verma, M., & Verter, V. (2017). Impact of train makeup on hazmat risk in a transport corridor. *Journal of Transportation Safety & Security*, *9*(2), 167–194.

Chiou, S.-W. (2016). A bi-objective bi-level signal control policy for transport of hazardous materials in urban road networks. *Transportation Research Part D: Transport and Environment*, *42*, 16–44.

Chow, J. Y. J., Yang, C. H., & Regan, A. C. (2010). State-of-the art of freight forecast modeling: lessons learned and the road ahead. *Transportation*, *37*(6), 1011–1030.

Chrobok, R., Kaumann, O., Wahle, J., & Schreckenberg, M. (2004). Different methods of traffic forecast based on real data. *European Journal of Operational Research*, *155*(3), 558–568.

Church, R., & Velle, C. R. (1974). The maximal covering location problem. *Papers in Regional Science*, *32*(1), 101–118.

Colson, B., Marcotte, P., & Savard, G. (2007). An overview of bilevel optimization. *Annals of Operations Research*, *153*(1), 235–256.

Crainic, T. G., Florian, M., & Leal, J.-E. (1990). A model for the strategic planning of national freight transportation by rail. *Transportation Science*, *24*(1), 1–24.

Dantzig, G. B. (1955). Linear programming under uncertainty. *Management Science*, *1*(3–4), 197–206.

Daskin, M. S., & Owen, S. H. (1999). Two New Location Covering Problems: The Partial P-Center Problem and the Partial Set Covering Problem. *Geographical Analysis*, *31*(3), 217–235.

Davenport, T. H., & Harris, J. G. (2007). *Competing on analytics: The new science of winning*. Harvard Business Press.

De Jong, G., Gunn, H., & Walker, W. (2004). National and international freight transport models: an overview and ideas for future development. *Transport Reviews*, *24*(1), 103–124.

der Vlies, A. V, & Suddle, S. I. (2008). Structural measures for a safer transport of hazardous materials by rail: The case of the basic network in The Netherlands. *Safety Science*, *46*(1), 119–131.

Erkut, E., & Gzara, F. (2008). Solving the hazmat transport network design problem. *Computers & Operations Research*, *35*(7), 2234–2247.

Erkut, E., Tjandra, S. A., & Verter, V. (2007). Hazardous materials transportation. *Handbooks in Operations Research and Management Science*, *14*, 539–621.

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

Esfandeh, T., Kwon, C., & Batta, R. (2016). Regulating hazardous materials transportation by dual toll pricing. *Transportation Research Part B: Methodological*, *83*, 20–35.

ESRI. (2016). ArcGIS Version 10.3.1. Retrieved from http://www.esri.com/arcgis/about-arcgis

Evans, J. R. (2012). Business analytics: the next frontier for decision sciences. *Decision Line*, *43*(2), 4–6.

Fabiano, B., Curro, F., Palazzi, E., & Pastorino, R. (2002). A framework for risk assessment and decision-making strategies in dangerous good transportation. *Journal of hazardous materials*, *93*(1), 1-15.

Fang, P., & Reed, H. D. (1979). Strategic positioning of railroad cars to reduce their risk of derailment. *US DOT Volpe Transportation Systems Center (DOT/TSC), Cambridge, Mass*.

Federal Railroad Administration (FRA). (2016). Office of Safety Analysis. Retrieved from http://safetydata.fra.dot.gov/OfficeofSafety/default.aspx

Flammini, F., Gaglione, A., Mazzocca, N., & Pragliola, C. (2008). Quantitative security risk assessment and management for railway transportation infrastructures. In *International Workshop on Critical Information Infrastructures Security* (pp. 180–

189).

Fontaine, P., Crainic, T., Gendreau, M., & Mendoza, J. E. (2016). *Population-based Risk Equilibration for the Multi-mode Hazmat Transport Network Design Problem*. CIRRELT, Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport= Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation.

Fontaine, P., & Minner, S. (2018). Benders Decomposition for the Hazmat Transport Network Design Problem. *European Journal of Operational Research*.

Garriguet, D. (2008a). *Beverage consumption of Canadian adults*.

Garriguet, D. (2008b). *Beverage consumption of children and teens*.

Glickman, T. S. (1983). Rerouting railroad shipments of hazardous materials to avoid populated areas. *Accident Analysis & Prevention*, *15*(5), 329–335.

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

Google Inc. 2016. Google Earth Pro, Version 7.1.5.1557. https://www.google.com/earth/desktop/

Government of Canada. (2014). *Consolidated Transportation of Dangerous Goods Regulations including Amendment SOR/2012-245*.

Heizer, J., Render, B., & Griffin, P. (2017). *Operations Management: Sustainability and Supply Chain Management* (Second Can). Pearson Higher Ed.

Hesse, M., & Rodrigue, J.-P. (2004). The transport geography of logistics and freight distribution. *Journal of Transport Geography*, *12*(3), 171–184.

Hillier, F.S., Lieberman, G.J. 2014. Introduction to Operations Research. McGraw-Hill Higher Education, New York.

Hosseini, S. D., & Verma, M. (2017). A Value-at-Risk (VAR) approach to routing rail hazmat shipments. *Transportation research part D: transport and environment*, *54*, 191-211.

Hosseini, S. D., & Verma, M. (2018). Conditional value-at-risk (CVaR) methodology to optimal train configuration and routing of rail hazmat shipments. *Transportation Research Part B: Methodological*, *110*, 79-103.

Iakovou, E., Ip, C. M., Douligeris, C., & Korde, A. (1997). Optimal location and capacity of emergency cleanup equipment for oil spill response. *European Journal of Operational Research*, *96*(1), 72–80.

International Energy Agency (IEA) (2016). World Energy Outlook 2016. Retrieved from https://www.iea.org. Acessed: February, 2018.

IBM. (2016). IBM ILOG CPLEX Optimization Studio. Retrieved from http://www-03.ibm.com/software/products/en/ibmilogcpleoptistud

Jeroslow, R. G. (1985). The polynomial hierarchy and a simple model for competitive analysis. *Mathematical Programming*, *32*(2), 146–164.

Kara, B. Y., & Verter, V. (2004). Designing a road network for hazardous materials transportation. *Transportation Science, 38*(2), 188–196.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, *52*(2), 21.

Lavelle, M. 2013. Oil Train Crash Probe Raises Five Key Issues on Cause. *National Geographic Daily News*, July 11.

Liberatore, M. J., & Luo, W. (2010). The analytics movement: Implications for operations research. *Interfaces, 40*(4), 313–324.

Louveaux, F. V. (1986). Discrete stochastic location models. *Annals of Operations Research*, *6*(2), 21–34.

Lowrance, W. W. (1976). Of Acceptable Risk: Science and the Determination of Safety.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., ... & Winkler, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of forecasting*, *1*(2), 111-153.

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). Forecasting: methods and applications. Wiley, New York.

Marcotte, P., Mercier, A., Savard, G., & Verter, V. (2009). Toll policies for mitigating hazardous materials transport risk. *Transportation Science*, *43*(2), 228–243.

Mazzarotta, B. (2002). Risk reduction when transporting dangerous goods: road or rail? *Risk, Decision and Policy*, *7*(1), 45–56.

McGinty, J. C. (2016, March 25). Why the Government Puts a Dollar Value on Life. *The Wall Street Journal*. Retrieved from http://www.wsj.com/articles/why-the-government-puts-a-dollar-value-on-life-1458911310

Milazzo, M. F., Lisi, R., Maschio, G., Antonioni, G., Bonvicini, S., & Spadoni, G. (2002). HazMat transport through Messina town: from risk analysis suggestions for improving territorial safety. *Journal of Loss Prevention in the Process Industries*, *15*(5), 347–356.

National Energy Board (NEB). (2016a). Marketable Natural Gas Production in Canada. Retrieved from https://www.neb-one.gc.ca/nrg/sttstc/ntrlgs/stt/mrktblntrlgsprdctn-eng.html

National Energy Board (NEB). (2016b). Natural Gas Liquids. Retrieved from https://www.neb-one.gc.ca/nrg/sttstc/ntrlgslqds/index-eng.html. Accessed: August, 2016.

National Energy Board (NEB). (2016c). Canadian Crude Oil Exports by Rail - Monthly Data. Retrieved from https://www.neb-one.gc.ca/nrg/sttstc/crdlndptrlmprdct/stt/cndncrdlxprtsrl-eng.html. Accessed: August, 2016.

National Energy Board (NEB). (2017a). Estimated Production of Canadian Crude Oil and Equivalent. Retrieved from https://www.neb-one.gc.ca/nrg/sttstc/crdlndptrlmprdct/stt/stmtdprdctn-eng.html. Accessed: July, 2017.

National Energy Board (NEB). (2017b). Canada's Energy Future 2016. Retrieved from https://www.neb-one.gc.ca/nrg/ntgrtd/ftr/2016/fscnstrndlppln-eng.html. Accessed: August, 2017.

Natural Resources Canada (NRC). (2015a). *Atlas of Canada 6th Edition (archival version) - Crude Oil and Natural Gas Resources*.

Natural Resources Canada (NRC). (2015b). Canadian Refineries. Retrieved from http://www.nrcan.gc.ca/energy/infrastructure/5895

Natural Resources Canada (NRC). (2016a). *Energy fact book 2015-2016*. Retrieved from https://www.nrcan.gc.ca/sites/www.nrcan.gc.ca/files/energy/files/pdf/EnergyFactBook2015-Eng_Web.pdf. Accessed: August, 2016.

Natural Resources Canada (NRC). (2016b). Average Retail Prices for Diesel in Canada. Retrieved from http://www2.nrcan.gc.ca/eneene/sources/pripri/prices_byyear_e.cfm?ProductID=5

Oak Ridge National Laboratory. (2016). *Vehicle Technologies Market Report 2015*.

Department of Transportation's (DOT) Office of the Secretary of Transportation. (2016). *Treatment of the Value of Preventing Fatalities and Injuries in Preparing Economic Analyses*. Retrieved from https://www.transportation.gov/sites/dot.gov/files/docs/VSL Guidance 2016.pdf

Owen, A. D., & Phillips, G. D. A. (1987). The characteristics of railway passenger demand. An econometric investigation. *Journal of Transport Economics and Policy*, 231–253.

Owen, S. H., & Daskin, M. S. (1998). Strategic facility location: A review. *European Journal of Operational Research*, *111*(3), 423–447.

PHMSA - Office of Hazardous Materials Safety. (2016). Incident Reports Database Search. Retrieved from https://hazmatonline.phmsa.dot.gov/IncidentReportsSearch/IncrSearch.aspx

Provencher, M. (2008). The movement and handling of dangerous goods in Canada for the year 2004. *Transport Dangerous Goods Directorate*.

Psaraftis, H. N., Tharakan, G. G., & Ceder, A. (1986). Optimal response to oil spills: the strategic decision case. *Operations Research*, *34*(2), 203–217.

Public Safety Canada. (2010). *Emergency Management Planning Guide 2010–2011*. Retrieved from https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/mrgnc-mngmnt-pnnng/index-eng.aspx

Purdy, G. (1993). Risk analysis of the transportation of dangerous goods by road and rail. *Journal of Hazardous Materials*, *33*(2), 229–259.

Quah, E., & Toh, R. (2011). *Cost-benefit analysis: Cases and materials*. Routledge.

Railroad Performance Measures. (2018). Performance Reports. Retrieved August 20, 2001, from http://www.railroadpm.org/home/rpm.aspx

Railway Association of Canada. (2015). *Ensuring the safe transportation of Essential Goods – RAC Dangerous Goods Team*.

Raj, P., & Pritchard, E. (2000). Hazardous materials transportation on US railroads: application of risk analysis methods to decision making in development of regulations. *Transportation Research Record: Journal of the Transportation Research Board*, (1707), 22–26.

ReVelle, C., Cohon, J., & Shobrys, D. (1991). Simultaneous siting and routing in the disposal of hazardous wastes. *Transportation Science*, *25*(2), 138–145.

RSMeans. (2016a). Square Foot Cost Estimates. Retrieved from http://www.rsmeans.com/Square-Foot-Cost-Estimates/

RSMeans. (2016b). RSMeans Frequently Asked Questions. Retrieved from http://www.cmdgroup.com/help/storefront/rsmeans-frequently-asked-questions/#UsingtheLocationFactors

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

Salling, K. B., & Banister, D. (2009). Assessment of large transport infrastructure projects: the CBA-DK model. *Transportation Research Part A: Policy and Practice*, *43*(9–10), 800–813.

SHARCNET. (2018). SHARCNET. Retrieved from https://www.sharcnet.ca/my

Smit, H. T. J. (2003). Infrastructure investment as a real options game: the case of European airport expansion. *Financial Management*, 27–57.

Špačková, O., & Straub, D. (2015). Cost-Benefit Analysis for Optimization of Risk Protection Under Budget Constraints. *Risk Analysis*, *35*(5), 941–959.

Statistics Canada, 2012. Monthly railway carloadings survey. Retrieved from: http://www5.statscan/gc/ca/cansim.

Statistics Canada. (2016a). Table 404-0002 Railway carloadings statistics, by commodity. Retrieved from http://www5.statcan.gc.ca/cansim/a26?lang=eng&retrLang=eng&id=4040002&&pattern=&stB

Statistics Canada. (2016b). Population and Dwelling Count Highlight Tables, 2011 Census. Retrieved from http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/hlt-fst/pd-pl/Table-Tableau.cfm?LANG=Eng&T=801&PR=0&RPP=9999&SR=1&S=3&O=D

Statistics Canada. (2017a). Table 404-0002 Railway carloadings statistics, by commodity. Retrieved from http://www5.statcan.gc.ca/cansim/pick-choisir?lang=eng&p2=33&id=4040002. Accessed: August, 2017.

Statistics Canada. (2017b). Monthly Railway Carloadings Survey. Retrieved from http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=2732. Accessed: August, 2017.

Statistics Canada. (2017c). 2011 Census - Boundary Files. Retrieved from http://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2011-eng.cfm. Accessed: August, 2017.

Talbi, E.-G. (2013). A taxonomy of metaheuristics for bi-level optimization. In *Metaheuristics for bi-level optimization* (pp. 1–39). Springer.

Thompson, R. E., Zamejc, E. R., & Ahlbeck, D. R. (1992). *Hazardous materials car placement in a train consist*.

Toregas, C., Swain, R., ReVelle, C., & Bergman, L. (1971). The location of emergency service facilities. *Operations Research*, *19*(6), 1363–1373.

Transportation Safety Boeard of Canada (TSB). (2014). Lac-Megantic runaway train and derailment investigation report (R13D0054). Retrieved from http://tsb.gc.ca. Accessed: February 2018.

Transport Canada. (2011). *Transportation in Canada*. Retrieved from https://www.tc.gc.ca/eng/policy/anre-menu-3020.htm. Accessed: August, 2017.

Transport Canada. (2012a). CANUTEC ERGO 2012. Retrieved from https://www.tc.gc.ca/media/documents/canutec-eng/ergo_english.msi

Transport Canada. (2012b). *2012 Emergency Response Guidebook*. Retrieved from https://www.tc.gc.ca/eng/canutec/guide-menu-227.htm

Transport Canada. (2017). *Transportation in Canada 2016 - Statistical Addendum*.

Treichel, T., & Barkan, C. P. (2016). *Rail Transportation of Crude Oil and Ethanol*. Retrieved from http://onlinepubs.trb.org/onlinepubs/petroleum/TreichelBarkan051216.pdf

US Department of Health and Human Services (DHHS). (2017). What is the “Superfund” Program? Retrieved from https://toxmap.nlm.nih.gov/toxmap/faq/2009/08/what-is-the-superfund-program.html

US Energy Information Administration (EIA). 2012. PADD regions enable regional analysis of petroleum product supply and movements. Retrieved from: http://www.eia.gov/todayenergy/detail.cfm?id=4890

US Environmental Protection Agency (EPA). (1992). *Hazard Ranking System Guidance Manual*. Retrieved from https://semspub.epa.gov/work/HQ/189159.pdf

US Environmental Protection Agency. (2016a). ALOHA Software. Retrieved from https://www.epa.gov/cameo/aloha-software

US Environmental Protection Agency. (2016b). About Acute Exposure Guideline Levels (AEGLs). Retrieved from https://www.epa.gov/aegl/about-acute-exposure-guideline-levels-aegls

Vaezi, A., & Verma, M. (2017). An analytics approach to dis-aggregate national freight data to estimate hazmat traffic on rail-links and at rail-yards in Canada. *Journal of Rail Transport Planning & Management*, *7* (4), 291-307.

Vaezi, A., & Verma, M. (2018). Railroad transportation of crude oil in Canada: Developing long-term forecasts, and evaluating the impact of proposed pipeline projects. *Journal of Transport Geography*, *69*, 98-111.

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

Verma, M. (2009). A cost and expected consequence approach to planning and managing railroad transportation of hazardous materials. *Transportation Research Part D: Transport and Environment*, *14*(5), 300–308.

Verma, M. (2011). Railroad transportation of dangerous goods: A conditional exposure approach to minimize transport risk. *Transportation Research Part C: Emerging Technologies*, *19*(5), 790–802.

Verma, M., Verter, V., & Gendreau, M. (2011). A Tactical Planning Model for Railroad Transportation of Dangerous Goods. *Transportation Science*, *45*(2), 163–174. http://doi.org/10.1287/trsc.1100.0339

Verma, M., Gendreau, M., & Laporte, G. (2013). Optimal location and capability of oil-spill response facilities for the south coast of Newfoundland. *Omega*, *41*(5), 856–867.

Verma, M., & Verter, V. (2013). Railroad Transportation of Hazardous Materials: Models for Risk Assessment and Management. In *Handbook of OR/MS Models in Hazardous Materials Transportation* (pp. 9–47). Springer.

Verter, V., & Kara, B. Y. (2008). A path-based approach for hazmat transport network design. *Management Science*, *54*(1), 29–40.

Viscusi, W. K., & Aldy, J. E. (2003). The value of a statistical life: a critical review of market estimates throughout the world. *Journal of Risk and Uncertainty*, *27*(1), 5–76.

Von Stackelberg, H. (2010). *Market structure and equilibrium*. Springer Science & Business Media.

Winston, C. (1983). The demand for freight transportation: models and applications. *Transportation Research Part A: General*, *17*(6), 419–427.

WLPGA. 2016. What is LPG? Retrieved from http://www.wlpga.org/about-lpg

Xie, Y., Lu, W., Wang, W., & Quadrifoglio, L. (2012). A multimodal location and routing model for hazardous materials transportation. *Journal of Hazardous Materials*, *227*, 135–141.

Zhang, A., Boardman, A. E., Gillen, D., Waters, I. I., & others. (2004). Towards estimating the social and environmental costs of transportation in Canada. *Report for Transport Canada*.

Appendix A. Estimating "Facility Cost"

To the best of our knowledge, no fixed cost data is available that is specifically related to hazmat emergency response facilities. However, we next outline a suitable technique to get an intelligent estimate of the cost. RSMeans (2016a) provides construction cost estimates (released in 2013) for *1-storey* fire stations (with an area of 6000 square feet). We use the RSMeans' cost-per-square-foot estimates assuming that hazmat emergency response facilities and fire stations have similar construction costs. We also assume that fire station building is constructed with *decorative concrete block/bearing walls*. Cost estimates are available for both *union labour* and *open shop labour*. To be on the conservative side, we use the former, which is associated with higher costs. However, for simplicity, we work with the costs released in 2013 and assume that the effect of not converting them to today's dollars is offset by using the union labor costs. For illustration purposes, Table 4.1 shows the RSMeans' construction cost estimates, in Canadian dollars, for a fire station with the abovementioned characteristics, in 20 cities across Ontario.

As recommended by RSMeans (2016b), construction costs in cities for which the data is not available can be estimated using the data of the nearest city. Since emergency response facilities can be opened in any of the 464 nodes in the railroad network, we estimate relevant cost data for all the nodes based on their location, i.e., their proximity to cities with available data. For this purpose, we created a layer in ArcGIS (ESRI, 2016), which includes all the cities with available cost data. Then, we use the *Near Table* command in Arc Toolbox to find the closest city with available cost data to every node of the railroad network.

Appendix B. Estimating "Equipment Acquisition Cost"

To the best of our knowledge, acquisition cost information for equipment packages that are specialized for hazmat response, does not exist. However, in the marine domain, Verma et al. (2013) estimated that an equipment package that is capable of responding to a 50-tons crude oil spill would cost around $450K. They then use linear and nonlinear extrapolations to estimate the costs of equipment packages for larger spills. Taking crude oil as the representative of Class 3 hazmat[10], we use their base number ($450K) to estimate the acquisition cost of an equipment package capable of responding to a rail hazmat incident involving the release of 50-tons of flammable liquids. A typical rail tank car has a capacity of about 80 tons (Erkut and Verter, 1995). It is possible that in a rail hazmat incident, only one railcar is involved in content release; it is also possible that only a portion of the hazmat content is released from the railcar; therefore, we assume that an equipment package that is capable of responding to a 50-tons release of flammable liquids, is suitable to respond to a partial hazmat content release from one railcar. As in existing literature, we make use of nonlinear extrapolation (Verma et al., 2013) to estimate acquisition costs of three larger-size equipment packages. (Table 4.2).

Given the various hazardous characteristics associated with hazmat commodities, it is reasonable to assume that equipment packages and therefore their acquisition costs will depend on the class of hazmat being considered. As mentioned earlier, *crude oil* is taken

[10] Based on our analysis of Statistics Canada's (2016a) railway commodity flow data, the third hazmat class is mostly made up of crude oil in any year during the last decade; for example, crude oil accounted for about 80% of Class 3 hazmat carried by rail in 2015

as the representative commodity for hazmat Class 3. Also, *hydrocarbon gas* and *sulphuric acid* are major commodities of their respective hazmat classes, and therefore are assumed to be representative commodities for classes 2 and 8, respectively. In order to distinguish between the three hazmat classes in terms of equipment acquisition costs, we use initial evacuation distances in case of fire as recommended by Transport Canada (2012b) in the Emergency Response Guidebook 2012 (Table 4.3). The rationale behind this differentiation is the fact that greater evacuation distance implies larger impact area; so, equipment packages of similar size should have higher response capability, which can be translated into higher acquisition cost. According to this scheme, Class 2 is associated with a factor 1.5, which means acquisition costs in Table 4.2 will be multiplied by 1.5 for this hazmat class. For Class 8, values in Table 4.2 can be used since the evacuation distance for its major commodity is the same as that of crude oil. In the absence of relevant information, the effect of facility location on acquisition costs is ignored.

Appendix C. Estimating "Incident Probability"

In this section, we estimate the probabilities associated with various incident scenarios, each characterized by a given hazmat class, rail link, and release volume. Consider a simple rail network as shown in Figure C. 1. This network is made up of four nodes and three arcs. The number below each arc represents its length in kilometers

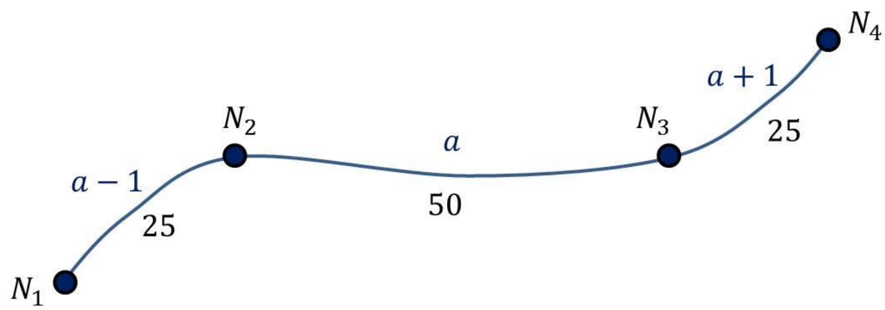


Figure C. 1. A simple rail network with three links

Based on Bayes' theorem (Bayes and Price, 1763), the probability of a hazmat incident that involves hazmat Class m and profile t on arc a is given as follows:

$$P_{at}^{m} = P(R_a^m) \times P(Q_{at}^m | R_a^m, H_a^m, D_a, G_a) \quad \text{(C.1)}$$

Where P_{at}^{m}, as defined in our mathematical program in Section 4.2, is the probability of an incident involving Class m and profile t along arc a; $P(R_a^m)$ is the probability of hazmat release incident of Class m with *any profile* along a, which is expressed by: $P(R_a^m) = P(G_a) \times P(D_a|G_a) \times P(H_a^m|D_a, G_a) \times P(R_a^m|H_a^m, D_a, G_a)$. $P(G_a)$ is the probability of accident of a train on arc a; $P(D_a|G_a)$ is the probability of derailment given that an accident occurred on arc a; $P(H_a^m|D_a, G_a)$ is the probability of derailment of at least one hazmat railcar with Class m given the derailment on arc a; $P(R_a^m|H_a^m, D_a, G_a)$ is the probability of release given the derailment of a hazmat railcar of Class m on arc a; and finally, $P(Q_{at}^m | R_a^m, H_a^m, D_a, G_a)$ in (C.1) is the probability of release with profile t given the hazmat release of Class m on arc a. Details of estimating this factor are provided in Section C.3.

Since accident rates or probabilities may not be available for every link of the network (Verma and Verter, 2013), we propose the following equation for estimating $P(R_a^m)$:

$$P(R_a^m) = P(R_{network}) \times P(R_a^m | R_{network}) \tag{C.2}$$

Where $P(R_{network})$ is the probability of hazmat release on the entire network, and expressed by: $P(R_{network}) = P(G) \times P(D|G) \times P(H|D, G) \times P(R|H, D, G)$, which is clearly the network-level version of the equation previously seen for $P(R_a^m)$. Details of estimating $P(R_{network})$ are provided in Section C.1. The other factor, i.e., $P(R_a^m | R_{network})$, is the conditional probability of release involving Class m along arc a. We will propose a methodology for estimating this probability in Section C.2.

C.1. Network-wide Release Probability: $\mathbf{P(R_{network})}$

We first use network-wide accident rates in the United States, reported by the Federal Railroad Administration (FRA). Table 4.4 shows the network-wide annual data from 2000 to 2015 (FRA, 2016), listing the four elements of $P(R_{network})$. It is important to note that FRA's hazmat derailment data are given only in terms of railcars, not trains. Therefore, to calculate the third and fourth column, we assumed that on average, 15 railcars of a given *standard* train with 100 railcars consist of hazmat railcars; this seems to be a reasonable assumption since based on our analysis of Statistics Canada's (2016a) railway commodity flow data, hazmat shipments account for around 15% of the rail traffic (Vaezi and Verma, 2017), which is also shown in Table 3.1.

The annual train accident rates (second column from the left) is the number of train accidents times 1,000,000 divided by total train miles. This is equivalent to the number of train accidents per million train miles. If we simply assume that 1 train and 1 mile of the network is involved in a given train accident, this probability measure would be a dimensionless quantity, which is desirable from the point of view of our mathematical modeling approach. The conditional probability of derailment (third column) is estimated based on the network-wide portion of train accidents that lead to derailment. For example, 1906 train accidents occurred in 2015 in the US; in 1345 instances, the train derailed; this means that the conditional probability of derailment was approximately equal to 1345/1906=0.70567 on US railroads in 2015. The conditional probability of hazmat derailment (fourth column) represents the portion of train derailments where at least one hazmat railcar of any given hazmat class is derailed. For instance, 43 out of 1345 derailments in 2015 were associated with hazmat railcar derailment; as a result, the conditional probability of hazmat derailment in 2015 is calculated as 43/1345=0.03197. The conditional probability of hazmat release (fifth column) is estimated as the portion of hazmat derailments that lead to hazmat release. For instance, 17 hazmat releases occurred out of 43 hazmat derailment events in 2015, which is equivalent to the conditional probability of hazmat release equal to 17/43=0.39535 in 2015.

Finally, the last column in Table 4.4 shows the probability of hazmat release of any profile from any given train on the US rail network. As discussed earlier, this probability is simply calculated by multiplying second, third, fourth, and fifth columns. The average value suggests that around 0.0364 of the total annual million train-miles travelled on US railroads lead to a hazmat release incident. In the absence of similar Canadian datasets, we take this value as the probability of hazmat incident on the entire rail network in Canada.

C.2. Link-specific and Class-specific Release Probability: $P(R_a^m)$

We first focus on the relationship between link-specific release probabilities and network-wide release probabilities. Several factors may differentiate rail-links in terms of hazmat release probabilities, e.g., traffic, roadbed quality, track geometry, etc. While such factors can be combined to extract more information about link-specific release probabilities, we only consider the ratio of "train-kilometers travelled along each link" to the "total train-kilometers travelled over the entire network" as the differentiating factor. It is also assumed that hazmat release incidents are independent events for different trains and for different links of the network, and the possibility of simultaneous events is ruled out. To illustrate our approach using the probability theory, let's revisit the simple rail network in Figure C. 1. Suppose a single hazmat is transported on this network in a given period of time using five trains as follows:

- 2 trains travel from N_1 to N_2;
- 1 train travels from N_1 to N_3; and
- 2 trains travel from N_3 to N_4.

Let $P(R_a)$ be the probability of hazmat incident along a, which can be calculated using the Bayes' theorem (Bayes and Price, 1763):

$$P(R_a) = P(R_{network} \cap R_a) = P(R_{network}) \times P(R_a | R_{network}) \qquad \text{(C.3)}$$

The above relationship suggests that the probability of release on arc a is the product of the network-wide probability of release and the probability of release on arc a given that a release event has occurred somewhere in the network. We focus on incidents on rail-links only. Therefore, for example, any release in the network in Figure C. 1 implies a release along exactly one of the three arcs; in other words, hazmat release incidents on different links of the network are assumed to be *mutually exclusive* and *collectively exhaustive* events. Consequently, the network-wide probability of release can be represented as follows:

$$\begin{aligned} P(R_{network}) = \ & P(R_{network} \cap R_{a-1}) + \ P(R_{network} \cap R_a) \\ & + \ P(R_{network} \cap R_{a+1}) \end{aligned} \qquad \text{(C.4)}$$

Substituting in (C.4) the probability of hazmat release on each arc according to (C.3), we get the following equation, which seems to be trivial due to our already stated assumptions.

$$P(R_{a-1} | R_{network}) + P(R_a | R_{network}) + \ P(R_{a+1} | R_{network}) = 1 \qquad \text{(C.5)}$$

Now, the question is how the abovementioned conditional probabilities should be defined such that our assumptions and the logical expressions are satisfied. To answer to this question, let's assume we know that a hazmat release has occurred on the network in Figure C. 1, but the link where the release event has occurred is unknown. We find it reasonable to assume that the number of train-miles (or train-kilometers) travelled along each link can be directly related to the conditional probability of release in that link, i.e., $P(R_{a-1}|R_{network}) \sim N_{a-1}l_{a-1}$, $P(R_a|R_{network}) \sim N_a l_a$, and $P(R_{a+1}|R_{network}) \sim N_{a+1}l_{a+1}$. In the absence of additional information, the three arcs are assumed to be similar in other aspects. Substituting in (C.5):

$$X\, N_{a-1}l_{a-1} + X\, N_a l_a + X\, N_{a+1}l_{a+1} = 1 \tag{C.6}$$

Where X is an unknown constant which is the common component among conditional release probabilities along the three links. Solving for X, we find $X = 1/(N_{a-1}l_{a-1} + N_a l_a + N_{a+1}l_{a+1})$. Then, the conditional probability of release for a given link a is given by: $P(R_a|R_{network}) = N_a l_a/(N_{a-1}l_{a-1} + N_a l_a + N_{a+1}l_{a+1})$. This relationship suggests that the conditional probability of hazmat release for a given link in the network is the *ratio* of the train-kilometers traveled on that link to the total train-kilometers travelled on the entire network. Combining the above result with (C.3), the probability of hazmat release along a is given by: $P(R_a) = P(R_{network}) \times N_a l_a/(N_{a-1}l_{a-1} + N_a l_a + N_{a+1}l_{a+1})$. In fact, we are treating this ratio as a *surrogate* measure of link risk. For instance, links with higher hazmat flow volume and length have higher incident risks. This approach enables us to capture that phenomenon.

To derive $P(R_a)$ in the simple network in Figure C. 1, we calculate the network-wide train-kilometers as well as the train-kilometers travelled on link a of the network. Given the provided information regarding the number of trains and their origins and destinations, the total train-kilometers is equal to $2 \times 25 + 1 \times (25 + 50) + 2 \times 25 = 175$, and the train-kilometres travelled along a is given by $1 \times 50 = 50$. Consequently, the probability of hazmat release along a is $P(R_a) = P(R_{network}) \times (50/175) = (2/7)\ P(R_{network})$.

For example, if the network-wide single-class probability of hazmat release for this simple network is 3.64×10^{-8} (the same as that of the Canadian rail network), then the probability of release incident along a will be around 1.04×10^{-8}.

Clearly, it is possible to extend our discussion of the simple setting in Figure C. 1 to any railroad network. However, we should account for multiple hazmat classes transported on a real network. We also note that $R_{network}$ is the release probability of all hazmat types. Consequently, the general form of the link-specific probability of hazmat release in a given network is as follows:

$$P(R_a^m) = P(R_{network}) \times \frac{N_a^m l_a}{\sum_{o=1}^{O} \sum_{j=1}^{J} N_j^o l_j} \tag{C.7}$$

Where j is any given rail-link on a network with a total of J links, and o is any given hazmat class transported on the network where a total of O hazmat classes are transported. Link-level information about the number of trains carrying each hazmat class can be extracted from the estimates provided by Vaezi and Verma (2017), details of which can be found in Chapter 3.

C.3. Link-specific, Class-specific, and Profile-specific Conditional Release Probability: $P(Q_{at}^m | R_a^m, H_a^m, D_a, G_a)$

To account for the profile (i.e., release volume) of hazmat incidents, we analyze the historical incident data by Pipeline and Hazardous Materials Safety Administration (PHMSA) (2016) from 2000 to 2015 to estimate the probability distribution of release volume for each hazmat class. We only consider those events with non-zero hazmat release volumes.

The conditional probability of a specific incident profile is estimated by the ratio of the number of incidents with that profile (expressed in terms of railcars) to the total number of incidents in the abovementioned 16-year period. For example, 1783 Class 2 release incidents occurred on the US railroads from 2000 to 2015: In 1781 incidents, only one railcar released its contents; in 1 incident, 5 railcars released their contents; and in 1 incident, 14 railcars released their contents. Consequently, a discrete probability distribution could be constructed as shown in Table 4.5. It is important to note that the probability distribution can be easily modified in to provide more realistic estimates, e.g., non-zero probabilities for incidents involving release from two railcars. However, we use the existing distribution but acknowledge these kinds of limitations associated with using historical data.

The same approach is used to construct conditional probability distributions for Class 3 and Class 8 incident profiles. Since all elements of (C.1) are estimated, we can calculate the probability of hazmat incident by link, class, and profile all over the Canadian railroad network.

Appendix D. Estimating "Incident Cost"

Total costs associated with rail hazmat incidents are divided into three components. *Population Exposure Costs* evaluate fatality and injury risks imposed on the affected population as a result of a given hazmat incident. *Property Loss Costs* quantify the damages to the adjacent public and private properties such as buildings. Finally, *Environmental Cleanup Costs* represent remediation cleanup costs associated with hazmat incidents.

Each of the three cost components is a function of the impact area associated with an incident release scenario. ALOHA (US EPA, 2016a) was used to calculate the radii of different threat zones corresponding to each scenario. For each hazmat class, we chose a representative chemical in ALOHA's database. Table 4.6 summarizes the assumptions made and input parameters given to the software to calculate the borders of threat zones in each case.

For incidents of Class 2 and 3, it is assumed that the tank fails due to a Boiling Liquid Expanding Vapor Explosion (BLEVE), in which the tank explodes and chemical burns in a fireball. On the other hand, for Class 8 incidents, the chemical is assumed to be released without any tank failure. For instance, Figure 4.3 shows the radii of the three threat zones (red, orange, yellow) for single or multi railcar Class 3 incidents. Once the impact areas are determined, the three cost components can be estimated. Details of estimating each cost category are provided in the following Sections.

D.1. Population Exposure Costs

To the best of our knowledge, no comprehensive measure exists to quantify the costs imposed on the population affected by a hazmat incident. In fact, such intangible costs are the most controversial costs to estimate (Zhang et al., 2004). However, to evaluate expensive risk-reducing regulations or investments, governments have no choice but to express in monetary terms the costs and benefits of such administrative decisions; they typically use a tool called the Value of a Statistical Life (VSL) (McGinty, 2016). VSL is defined by the U.S. Department of Transportation (DOT) as the additional cost that individuals would be willing to pay for improvements in safety so that the expected number of fatalities in a given period of time is reduced by one (DOT's Office of the Secretary, 2016). It should be noted that what is involved is the valuation of reductions in risk, not the valuation of life.

Regulatory agencies in the United States, United Kingdom, and Canada have been known to use VSL estimates in recent decades to evaluate the benefits of proposed environmental, health, and safety regulations. Also, some multinational organizations such as the European Commission have issued directions to improve cost-benefit analysis based on VSL estimates (Viscusi and Aldy, 2003). To evaluate population exposure costs, we use the values recommended by the U.S. DOT’s Office of the Secretary for valuing the reduction of fatalities and injuries. These estimates are based on nine relevant studies that provide general VSL estimates, averaged and updated to the most recent year based on changes in the price levels and incomes. Table 4.7 shows the U.S. DOT’s VSL estimates and alternative low and high values for each year between 2013 and 2016.

The U.S. DOT also provides values associated with preventing injuries as fractions of VSL. Such fractions are calculated using an interpolation method which takes into account losses in quality of life such as suffering and reduced income. Table 4.8 shows the scores corresponding to each Abbreviated Injury Scale (AIS) level (US DOT, 2016). The last column represents three average scores that we calculated based on combining injury levels to roughly represent Acute Exposure Guideline Levels (AEGLs) (US EPA, 2016b). AEGLs correspond to different levels of toxic effects on general population caused by relatively short periods of exposure to airborne chemicals. AEGLs are expressed as parts per million (or another concentration measure) of a given chemical under study associated with three thresholds above which the following could be experienced by the exposed individuals:

- *Level 1*: Notable discomfort, but transient effects;
- *Level 2*: Serious, long-lasting adverse health effects; and,

Level 3: Life-threatening health effects or death.

We use VSL estimates in Table 4.7 and fractions in Table 4.8 to calculate values of preventing injuries with different levels of exposure to hazmat incidents in 2016 (not shown). Then, given a hazmat incident scenario, we can calculate total population exposure costs using the population density of the affected zone that is specified by AEGLs. First, we use ALOHA (US EPA, 2016a) to identify such toxic threat zones; given an incident scenario, ALOHA specifies three zones, borders of which either correspond to three thermal radiation zones for Class 2 and 3, or three AEGLs for Class 8 (See Table 4.6). For simplicity, we assume that the two classifications are equivalent and show three levels of exposure with modified fractions shown in Table 4.8. Second, different buffers are created in ArcGIS (ESRI, 2016) around rail-links, based on which the number of people inside red, orange, and yellow zones is estimated, each of which is then multiplied by the corresponding VSL fraction and then added up to identify total population exposure costs due to specific incident scenarios.

D.2. Property Loss Costs

We define our Property Loss Cost component in a way so it represents three main cost elements (PHMSA, 2016):

- *Material Loss*, the value of the material lost in an incident;
- *Carrier Damage*, the cost of damages sustained by the carrier; and,
- *Property Damage*, the cost of damages to public or private property.

PHMSA (2016) provides dollar value estimates of the above cost elements for historical hazmat incidents by class and release volume on US railroads. Since relatively few instances involve hazmat release from more than one railcar, we use the PHMSA's data to estimate a base number representing the average cost of property loss corresponding to one railcar release for each hazmat class. We then use a nonlinear extrapolation scheme to estimate property loss costs for incidents that involve release from more than one railcar; it is assumed that these estimates would be also applicable to the Canadian context.

Regarding the location, we suppose that hazmat incidents on rail links located in more populated areas involve greater damage to public or private properties; this is based on the argument that more populated areas are expected to contain a greater number of properties such as buildings and vehicles around the adjacent rail segments compared to less populated areas. Consequently, to account for the effect of rail link location on property damage, we focus on the red zone previously created in ArcGIS (ESRI, 2016) and apply a proportionate factor to magnify the property damages corresponding to links located in more populated areas. In particular, in case the affected population in the red zone is more than 100 people, a magnifying factor is applied, and vice versa.

D.3. Environmental Cleanup Costs

PHMSA (2016) provides estimates of *Remediation Cleanup Costs*. Similar to the previous section, we analyze the PHMSA's data to estimate a base number representing the average cost of environmental cleanup costs corresponding to one railcar release for each hazmat class. Due to the lack of sufficient and reliable data, we do not differentiate between different rail links in terms of cleanup costs.

Appendix E. Estimating "Equipment Transport Cost"

We consider fuel cost as a proxy to estimate the cost of transporting the equipment packages. Table 4.9 shows weight and typical fuel economy ranges for two types of fire engines (Oak Ridge National Laboratory, 2016). To be on the conservative side, we assume that the characteristics of these fire engines correspond to the already defined equipment packages with smaller 50-ton and 100-ton capacities (See Table 4.2).

Average fuel costs per 100km of the two fire engines are shown in the last column of the table, which is used to estimate fuel costs associated with moving equipment packages from response facilities to incident sites. To calculate this measure, we used the average retail price of diesel in 2015, which was 109 Cents per Litre in Canada (NRC, 2016b).

To estimate the fuel cost for the equipment packages with 200-ton and 400-ton capacities, we use a nonlinear extrapolation scheme that preserves the ratio of the difference in fuel cost to the difference in equipment response capability; this would result in average fuel costs equal to 109.6 C$/100km and 165.4 C$/100km.

The transport cost of equipment packages then depends on the distance that should be travelled from emergency response facilities to incident sites. For example, the cost to move an equipment package, capable of responding to a 50-tons release, to an incident site 300km far from the facility would be calculated as $48.1 \times 3 = 144.3$ C$. We do not distinguish between different hazmat classes in terms of equipment transport cost.