FREIGHT DEMAND MODELING AND LOGISTICS PLANNING FOR ASSESSMENT OF FREIGHT SYSTEMS’ ENVIRONMENTAL IMPACTS

BY

TAE SUNG HWANG

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Civil Engineering in the Graduate College of the University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Associate Professor Yanfeng Ouyang, Chair
Professor Christopher P.L. Barkan
Associate Professor Tami C. Bond
Assistant Professor Bumsoo Lee
ABSTRACT

This dissertation research aims at examining the U.S. freight transportation systems and the relationship between freight shipment activities and the related environmental issues such as air pollution and greenhouse gas emissions in the nation. This study develops freight demand models to forecast freight movements between and within the U.S. geographical regions via two major shipment modes, truck and rail. Freight flow is categorized into two types: inter-regional and intra-regional freight flow. For the inter-regional freight flow, the well-known four-step freight demand forecasting model is adopted which consists of trip generation, trip distribution, modal split, and traffic assignment. In case of the intra-regional freight movements, various network modeling and logistics systems optimization methodologies are applied to address a large-scale freight delivery problem in the U.S. freight zones and an individual truck routing problem on stochastic congested roadway networks.

Following the four-step freight demand forecasting framework, we first propose a methodology to estimate future freight demand for all commodity types that begin and end in each geographical region in the U.S., and the amount of freight that moves between all origin-destination pairs. This procedure corresponds to trip generation and trip distribution for inter-regional freight demand. Using future economic growth factors, the amounts of freight production and attraction in each geographical region are forecasted and taken as given. Then, an efficient matrix balancing method, an RAS algorithm, is applied to distribute the estimated freight shipment demand for all origin-destination pairs.

Various freight shipment modes have significantly different impacts on air quality and environmental sustainability, and this highlights the need for a better understanding of inter-regional freight shipment mode choices. This dissertation work develops a binomial logit market
share model to predict the U.S. inter-regional freight modal share between truck and rail, as a function of freight and shipment characteristics. This step corresponds to modal split procedure in the four-step freight demand forecasting framework. A set of multi-year freight shipment and geographical information databases as well as crude oil price information were integrated to construct regression models for typical freight commodities. The atmospheric impact levels incurred by different freight mode choice decisions are analyzed to provide insights on the relationship among freight modal split, oil price change, and air quality.

In addition to ‘mode choices,’ ‘route choices’ in freight deliveries can significantly affect national and regional air quality. Therefore, as the last step of the inter-regional freight flow modeling framework, truck and rail freight shipment assignment is conducted while network congestion effect is taken into consideration. Carriers’ route choices are assumed to follow a user equilibrium principle. A traditional convex combinations algorithm is used to solve for traffic routing equilibria for truck flow in the U.S. highway network. Link cost function is modified to consider traffic volume that already exists on the highway network. A customized network assignment model is proposed for rail freight shipment demand, where single- and double-track lines are represented by an equivalent directed graph with railroad-specific link traffic delay functions. An adapted convex combinations algorithm is developed to find shipment routing equilibrium. Our models are applied to an empirical case study for the U.S. highway and rail networks and solutions are found within a short computation time.

For the intra-regional freight demand, we first focus on developing a methodology for freight distribution and collection within the U.S. geographical regions where a large number of spatially distributed freight demand and supply points need to be served. This problem is formulated as a large-scale vehicle routing problem and solved by an modified ring-sweep
algorithm. A set of closed-form formulae is constructed to estimate the asymptotic total travel distance of a fleet of trucks. A case study is conducted to forecast regional freight delivery cost for the U.S. geographical regions that include major metropolitan areas. Numerical results under three urban development scenarios show that the proposed methodology can effectively estimate the total cost and the related emissions.

Lastly, a microscopic urban freight truck routing problem on a stochastic network is addressed. Freight trucks are known as a major source of air pollutant and greenhouse gas emissions in the U.S. metropolitan areas. Therefore, emissions from freight trucks during their deliveries need to be considered by the trucking service sector when they make routing decisions. This study proposes a model that incorporates total delivery time, various emissions from freight truck activities, and a penalty for late or early arrival into the total cost objective of a stochastic shortest path problem. We focus on urban networks in which random congestion state on each link follows an independent probability distribution. Our model finds the best truck routing on a given network so as to minimize the expected total travel cost. This problem is formulated into a mathematical model and two solution methods including a dynamic programming approach and a deterministic shortest path heuristic are proposed. Numerical examples show that the proposed algorithms perform very well even for the large-size U.S. urban networks.

This dissertation study will be useful for transportation planners and decision makers in public and private sectors to assess how freight mode and route choices on the national scale will affect air quality and eventually human health in a variety of future global economic growth and environmental policy scenarios. Also, the estimated freight shipment activities in the regional level can be used to infer the human exposures to emissions from freight delivery in large urban areas.
ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my advisor, Professor Yanfeng Ouyang, for his thoughtful guidance and patient support in completion of the dissertation work. It was he who provided me with insights and helped me with a lot of excellent ideas and advice throughout my Ph.D. program. I am truly fortunate to study under his guidance.

I greatly appreciate the valuable guidance and advice over the years from my committee members, Professor Christopher P.L. Barkan, Professor Tami C. Bond, and Professor Bumsoo Lee. I am grateful for the opportunity to work in an interdisciplinary research group. I would also like to express my appreciation to all UIUC professors from whom I have taken courses and learned knowledge in various fields.

I would like to thank Liang Liu, Steven J. Smith, and Fang Yan who collaborated with me on the project. In particular, I would like to thank Sungwon Lee who worked very closely with me over the past four years; I greatly appreciate his help and suggestions.

I am grateful to all former and current colleagues in our research group for their support and friendship all the time. Special thanks go to Yun Bai, Seungmo Kang, and Fan Peng who collaborated with me on other research topics. My special thanks are extended to Leila Hajibabai, Xiaopeng Li, Seyed Mohammad Nourbakhsh, Ryan Smith, Xin Wang, and Weijun Xie who provided me with helpful feedback and excellent suggestions throughout my graduate study. In a similar vein, I am also grateful to all my friends at UIUC and around the U.S. Special thanks to Inyong Kim (and Jungyoo Kim) for the encouraging words and friendship. I am further thankful to Koohong Chung for his steadfast support and Dosoo Moon for his great help in preparing my Ph.D. Defense. I would also like to extend my heartfelt thanks to my great friends in Korea for their everlasting support in all aspects of my life.
Last but not least, I warmly thank my family, especially my parents, brother, sister-in-law, and niece. They are my most important people and it is their never-ending support, love, encouragement, dedication, and prayers that supported me to finish my doctoral study. I would like to thank God for providing me with this opportunity and accompanying me at important times of my journey. Thank you for everything Dear Lord.

My work was supported by the U.S. Environmental Protection Agency through graduate research assistantship. The U.S. National Science Foundation also partially supported the first two years of my graduate study at the University of Illinois, Urbana-Champaign.
# TABLE OF CONTENTS

CHAPTER 1 – INTRODUCTION .................................................................................................................. 1

CHAPTER 2 – LITERATURE REVIEW .................................................................................................... 10

CHAPTER 3 – FORECASTING FREIGHT DEMAND CONSIDERING ECONOMIC GROWTH FACTORS .................................................. 30

CHAPTER 4 – FREIGHT TRANSPORTATION MODE CHOICE AND ITS ENVIRONMENTAL IMPACTS ........................................................................... 41

CHAPTER 5 – ASSIGNMENT OF FREIGHT SHIPMENT DEMAND IN CONGESTED TRANSPORTATION NETWORKS .............................................................................. 63

CHAPTER 6 – LOGISTICS SYSTEMS PLANNING FOR REGIONAL FREIGHT DELIVERY .............................................................................................. 103

CHAPTER 7 – URBAN FREIGHT TRUCK ROUTING UNDER STOCHASTIC CONGESTION AND EMISSION CONSIDERATIONS .............................................. 121

CHAPTER 8 – CONCLUSIONS AND FUTURE RESEARCH PLAN ................................................................... 139

REFERENCES ............................................................................................................................................. 145

APPENDIX A – SOFTWARE MANUAL FOR THE FOUR-STEP INTER-REGIONAL FREIGHT DEMAND MODEL .............................................................................. 164
CHAPTER 1
INTRODUCTION

1.1 Background

With rapid global industrialization and ever-increasing demand for freight movements, freight transportation has become a major source of air pollution. According to ICF Consulting (2005), most freight transportation modes such as trucks, locomotives, and ships are powered by diesel engines which are significant emission sources of national nitrogen oxides (NO\textsubscript{X}), particulate matter (PM) as well as greenhouse gases such as carbon dioxide (CO\textsubscript{2}). For example, NO\textsubscript{X} and PM\textsubscript{2.5} emissions from freight trucks and trains respectively constitute about 32% and 21% of the total NO\textsubscript{X} and PM\textsubscript{2.5} emissions from the U.S. transportation sector (Bickford, 2012). In the case of greenhouse gas emissions, freight transportation contributes about 25% of the total emissions from all transportation sources in the U.S. (ICF Consulting, 2005). The emissions from freight transportation activities affect climate change on the global scale and deteriorate air quality and human health in regional and urban areas. Thus, the freight delivery system needs to be thoroughly investigated to understand its impact on the environment.

This dissertation work is a part of a collaborative research project that aims to develop an integrated modeling framework to estimate future emissions from freight transportation systems at global, regional, and urban levels based on future economic growth and climate policy scenarios, projections of urban spatial structure, and vehicle emission characteristics. The research project has been conducted by four research subgroups, each of which is in charge of one of the four main tasks including global economic forecast models, urban spatial structure and
The global economic forecast models can be used to generate projections of various kinds of economic factors (e.g., Walmsley, 1998). For example, our collaborative project forecasts real
GDP, population, and capital and labor inputs for 24 influential countries (or geographical regions), and estimates exports and imports of 26 major U.S. industries from 2005 to 2100 in five-year increments (Edmonds et al., 1995; Edmonds et al., 2004; Vanek and Morlock, 2007). Outputs from the global economic forecast models will serve as inputs to various analyses such as those in the urban spatial structure and I/O models, freight transportation system models, and air quality and climate impact models. This is illustrated in Figure 1.1.

The urban spatial structure and I/O models in general produce two different projections based on the results of the global economic forecast models. First, they predict I/O commodity value forecasts for each geographic zone in the U.S. (Isard, 1951, 1960; Leontief and Strout, 1963; Wilson, 1970a, 1970b). Second, assuming each employee is an endpoint as well as starting point of the freight delivery systems, employment distributions within existing geographic zones (which include 22 major metropolitan areas in which the number of total population are greater than or equal to 2,000,000 in Year 2000) in the U.S. are forecasted (Song, 1994; Anas et al., 1998; Lee, 2007; Lee and Gordon, 2011). In these analyses, locations of activity centers are forecasted using GIS data. The two different outcomes from the urban spatial structure and I/O models serve as inputs to the analyses in the freight transportation system models.

The freight transportation system models are composed of two main parts: (i) inter-regional freight flow (e.g., from Los Angeles to Chicago) which deals with the freight transportation problem in a national-wide point of view, and (ii) intra-regional freight flow (e.g., within Chicago urban area) which narrows down the scope of the problem within each geographic zone. The concept of Freight Analysis Zone (FAZ), originally defined in the Freight Analysis Framework (FHWA U.S. DOT, 2011) database, is adopted to represent the U.S. geographical regions with regard to freight activities. The set of FAZs is composed of a total of
123 domestic regions: (i) 74 metropolitan areas, (ii) 33 regions representing the remaining parts of the states that these 74 metropolitan areas belong to, and (iii) 16 remaining regions, each of which represents an entire state. Figure 1.2 shows a map of the recent definition of FAZs, adapted from FHWA U.S. DOT (2011).

Figure 1.2 Domestic Freight Analysis Zones (FHWA U.S. DOT, 2011)

For inter-regional freight demand modeling, the four-step freight commodity transportation demand forecasting model (Cohen et al., 2008) is adopted, which includes trip generation, trip distribution, modal split, and traffic assignment procedures. Detailed explanation of the four-step freight demand forecasting framework is provided in the next section. Among those four steps, freight trip generation analysis uses the I/O commodity value growth forecasts generated from the I/O model to estimate future freight production and attraction at each FAZ. Also, output price index predictions generated from the global economic forecast models are applied to modal split procedure to forecast various economic factors (e.g., freight value and crude oil price) that affect shippers’ mode choice decisions. Efforts for intra-regional freight
transportation modeling are composed of two parts including a large scale freight delivery model in FAZs and a microscopic point-to-point truck routing problem. Both employment distribution data from the urban spatial structure model and output from the modal split procedure in the four-step inter-regional freight flow model are applied to the large scale freight delivery problem to find near-optimal freight distribution (and collection) in the FAZs.

Finally, required freight transportation activities at inter-regional level will be applied to estimate emissions from both long-haul trucks and rails, and those at intra-regional level will serve as inputs to assess emission problems caused by short-haul trucks in the U.S. The emission estimations generated from both inter-regional and intra-regional levels will be used as basis in obtaining emission inventory in the U.S. regions. In case of global air quality and climate impact analysis, predictions from the global economic forecast models and results from the U.S. regional emission model will be used (Bond et al., 2004; Streets et al., 2004; Marshall et al., 2005; Collins et al., 2006; Bond, 2007; Yan and Bond, 2008).

1.2 Objectives and Contributions

The freight demand modeling and logistics planning in this study forecast freight movements between and within the U.S. geographical regions (i.e., FAZ) and by modes (e.g., truck or rail) and load them onto the respective transportation networks. The rationale behind this work is that the mode and routing choices of freight shipments between and within major regions significantly affect regional and urban air quality.

As mentioned previously, the freight flow in the U.S. is divided into two levels, i.e., inter-regional and intra-regional. For the inter-regional freight flow, the four-step freight commodity transportation demand forecasting model (Cohen et al., 2008) is adopted, which is very similar to
the urban passenger travel demand forecasting model and Figure 1.1 describes its procedure in detail.

Given a forecast of economic growth factors for each FAZ, freight trip generation step predicts daily or annual production and attraction of freight shipments by different commodity types for all FAZs. Output from freight trip generation includes entering and exiting freight demand of each zone. Such information will be used as an input to trip distribution analysis, which generates production and attraction of commodity-specific freight shipment flow between all FAZ origin-destination (O/D) pairs. Output of this component, in the form of zonal O/D freight demand for each commodity type, serves as an input to either the modal split step if more than two modes are considered or the traffic assignment step if only one mode is involved. The modal split procedure forecasts shares of zonal freight delivery among the freight modes based on the relative utility (i.e., benefit or preference) of each mode between all FAZ O/D pairs. This step generates zonal O/D freight demand by shipment mode which will be an input to the traffic assignment step. Traffic assignment estimates the modal routes each unit of the freight O/D demand will use to traverse the transportation networks. It yields freight flow on each link of the network (for each mode separately) that experience the minimum path cost between each FAZ O/D pair. In this process, congestion delay caused by limited roadway capacity is typically considered to estimate congestion pattern and assigned traffic demand on each network link. The final results from the traffic assignment step will be used to estimate emissions from long-haul trucks and rails in future years.

In case of intra-regional freight movements, various network optimization models (e.g., vehicle routing problems) and solution approaches can be applied to find near-optimal freight distribution (and collection) within each FAZ considering future urban forms. Finally, we
consider the microscopic level of the freight transportation systems where individual truck finds the best route from one origin to one destination. These origin and destination are generic; they could be the truck terminal and/or consecutive delivery points.

As such, the work presented in this dissertation is part of the efforts on demonstrating local to global air quality and climate impacts of the freight transportation systems for any given economic growth and environmental policy scenarios. This study focuses on developing freight demand and logistics models for the U.S. The modeling efforts in this study will be useful for transportation planners and decision makers in evaluating freight handling decisions that contribute to reducing adverse impacts on air quality and climate change, and eventually, enhancing human health and social welfare. In addition to the detailed technical contributions described above, the proposed dissertation research also has the following contributions:

1) Development of a comprehensive modeling framework for freight shipment systems. Our study addresses both inter-regional and intra-regional freight delivery, with activities ranging from initial freight collecting systems in production areas, to freight movements and routing at the national scale, and then to final freight distributing systems in attraction areas.

2) Bridging the gaps among multiple traditionally separated research fields, including global economic forecast models, urban spatial structure and I/O models, freight transportation system models, and air quality and climate impact models. Our study provides deeper understanding of the interdependencies and connections among economic growth scenario, urban spatial change, vehicle emission distribution, and air quality and climate impacts.

3) Development and integration of a decision-supporting tool. We integrated the proposed four-step inter-regional freight demand forecasting models and algorithms
into one software package to assess impacts of various truck and rail freight shipment activities on network efficiencies. Detailed manual is included in an APPENDIX A. It will have the capability of exploring both the environmental impacts of freight mode and routing choices and the local delivery decisions in a number of future economic growth and climate policy scenarios. This tool-kit can be used to facilitate decision-making processes in the freight industries and the government agencies.

4) Extension and application of the methodologies to other transportation studies. Final results from our freight demand models include a lot of useful information such as predicted traffic flow distribution and congestion pattern in freight transportation networks in specific future year. Such information could be used to address many other related problems such as transportation network capacity expansion and maintenance (e.g., highway capacity expansion, railroad wayside sensor location design, and railroad track maintenance) as well as traffic safety prediction.

1.3 Outline

This dissertation is organized as follows. Chapter 2 reviews the related literature on both inter-regional and intra-regional freight demand modeling. In case of inter-regional freight flow, previous studies related to the four-step freight demand modeling are thoroughly reviewed. For intra-regional freight flow, previous research on logistics systems planning and various stochastic network optimization models are provided with solution algorithms. From Chapter 3 to Chapter 5, the four-step inter-regional freight demand forecasting framework is presented step by step. Chapter 3 describes trip generation and trip distribution procedures with application of an RAS algorithm. Then, freight modal split model is presented in Chapter 4. Chapter 5 proposes a user equilibrium traffic assignment procedure based on a convex combinations algorithm for truck and rail freight shipment demand. Chapters 6 and 7 are related to the intra-regional freight demand modeling. Chapter 6 presents a logistics systems model to serve spatially distributed
freight demand within an FAZ. Chapter 7 shows a mathematical model for an urban freight truck routing problem under stochastic congestion and emission considerations. Finally, concluding remarks and discussions on future work are provided in Chapter 8.
CHAPTER 2
LITERATURE REVIEW

This chapter reviews existing studies on both inter-regional and intra-regional freight demand modeling. In case of inter-regional freight flow, related literature on the four-step freight demand forecasting framework is thoroughly reviewed. Then, previous studies on the intra-regional freight demand modeling will be provided; our focus will mainly be on large-scale freight demand delivery problems in FAZs and a shortest path problem in a stochastic network setting.

2.1 Inter-regional Freight Demand Modeling

This section first reviews the state-of-the-art literature on each step of the four-step inter-regional freight demand forecasting framework. The traffic assignment problem, the last step of the four-step analysis, is also reviewed to provide a methodological background of the literature.

2.1.1 Four-step freight commodity transportation demand forecasting model

Early attention has been given to the inter-regional freight transportation and commodity flows. Isard (1951, 1960) and Moses (1955) considered inter-regional commodity flow analysis in terms of an I/O method. Leontief and Strout (1963) suggested a gravity-type model formulation in which the shipment distance between two regions is addressed in the form of a friction factor. Wilson (1970a, 1970b) presented several methods for analyzing a system of inter-regional commodity flows, including a Newtonian gravity model, an I/O framework, an entropy maximizing method, and a hybrid gravity and I/O modeling approach. Kim et al. (1983) proposed an inter-regional commodity flow I/O model and provided some empirical applications.
Rho et al. (1989) used small and large scale networks to compare solution techniques for the inter-regional commodity flow model in Wilson (1970a, 1970b). Despite these early efforts, development and application of the inter-regional freight shipment models have been far less advanced compared with their counterparts on the urban passenger side, probably due to the lack of freight flow data (Jiang et al., 1999; Ham et al., 2005). Freight demand information is usually expensive to obtain and often kept confidential.

However, over the past decade freight demand modeling and analyses have received a great amount of attention. A number of models have been developed by various groups including transportation engineers, public policy makers, and planning agencies. Implementation and application of the analytical freight demand models have also emerged since the comprehensive commodity flow survey data from the U.S. Bureau of the Census became available in the 1990s (Ham et al., 2002, 2005). Cohen et al. (2008) summarizes a standardized toolkit for forecasting freight movements at the state level. This report presents five freight demand forecasting models (i.e., the direct facility flow factoring method, the O/D factoring method, the truck model, the four-step commodity model, and the economic activity model), each of which contains a subset of six basic model components (i.e., direct factoring, trip generation, trip distribution, modal split, traffic assignment, and economic/land use modeling). This report also provides a number of case studies of model implementations by state agencies. For instance, the State of Florida adopted the four-step freight demand forecasting framework to develop a statewide freight model using TRANSEARCH database (Cambridge Systematics, Inc., 2002), and Washington State Department of Transportation adopted economic class spatial I/O model to conduct Cross-Cascade Corridor analysis in which passenger and freight transportation demand are forecasted (Cohen et al., 2008). In this dissertation work, we develop a comprehensive nationwide freight
shipment demand forecasting model ranging from initial collecting systems, to freight movements and routing at the national scale, and then to final distributing systems. The four-step freight demand modeling framework is adopted, which is composed of trip generation, trip distribution, modal split, and traffic assignment. In this section, previous studies related to each step are thoroughly provided and reviewed.

2.1.1.1 Trip generation and trip distribution in inter-regional freight demand modeling

There have been a large number of previous studies on trip generation and trip distribution steps of the passenger travel demand modeling. In general, there are two modeling methods for passenger trip generation (Ortuzar and Willumsen, 1995): (i) a regression modeling approach defines the trip generation forecast as an independent variable and socio-economic characteristics as explanatory variables. It is simple and easy to implement; however its inherent linear functional structure might fail to represent reality, and (ii) a cross-classification model groups households into different types and estimates trip generation rate using specific coefficients assigned to each type of household. It allows nonlinear relationship at the expense of being time consuming. ITE (1997) provided a guideline for passenger trip generation with an extensive introduction of available resources. Anderson and Olander (2002) developed a single purpose trip generation model applicable for traffic analysis zones in small urban areas. The authors showed that the suggested model can simplify procedures and reduce complexity and data requirement although results are similar with those from a multiple purpose conventional trip generation model. Gamas et al. (2006) adopted a spatial regression model instead of a traditional ordinary least square estimation method to generate work, shopping, and school trips in Mexico City, Mexico. Roorda et al. (2010) analyzed trip generation rates of three vulnerable
population groups including single-parent families, low income households, and the elderly in three major Canadian metropolitan areas. The authors adopted an ordered probit model with spatially expanded coefficients. Previous trip generation analyses in the passenger travel demand modeling side provided trip generation analyses in the freight demand modeling side with basic methodological background.

For freight demand modeling, ITE (1998) presents a number of initial freight truck trip generation studies. Thornton et al. (1998) provides a commercial vehicle travel model in Atlanta, Georgia to estimate truck emissions. The authors developed a truck trip generation model using data collected from commercial vehicles surveys. Trucks were divided into two classes including light-duty and heavy-duty vehicles, and each of them is associated with different land use/employment categories to estimate truck trip generation rates. Fischer et al. (2001) identifies a set of truck trip generation data available for transportation engineers and travel demand modelers, and reviews the current state of the practice of data applications in various research fields. This report classifies trip purposes and trip generating activities associated with appropriate categories of land uses, and shows how those factors affect the truck trip generation data and truck trip rates. Transportation Engineering and Planning, Inc. (2003) studied truck trip generation in the City of Fontana, California. This research investigates a total of nine land use categories that generate heavy truck traffic volume and suggests a set of equations which can be used for predicting truck trip generation rates considering the land use categories. However, results might not be appropriate for national level studies since the analysis is conducted mainly based on locally collected data. Also, this study is confined to the freight trip production side and does not consider freight trip attraction. Tolliver et al. (2006) developed trip generation equations for grain elevators in North Dakota (which have been known as major sources of truck traffic).
using various databases including land use and highway traffic data. In this study, a trip attraction equation is applied to analyze how elevator storage capacity and side track capacity affect elevator throughout. As such, trip generation analyses in freight demand modeling side are usually very similar to those in passenger travel demand modeling side; the former produces annual or daily freight trip generation rate by commodity type, which is a function of employment number by industry type or total population in a geographic zone.

For the trip distribution step in passenger travel demand modeling, a Fratar method was first developed and used for urban transportation planning (OTDMUG, 2012). It assumes the base-year travel pattern between all O/D zones will remain the same in the future, and future production and attraction of each analysis zone are expected to be scaled up or down according to simple growth factors. The base-year trip distribution matrix is iteratively modified until future trip distribution converges to future production and attraction of all analysis zones. A gravity model was developed later based on Newton’s law of gravitation. This method is still widely used for passenger travel demand forecasting. The basic expression can be formulated as follows (Voorhees, 1955; Easa, 1993):

\[
Q_{ij} = P_i \left( \frac{A_j F_{ij} K_{ij}}{\sum_{k=1}^{n} A_k F_{ik} K_{ik}} \right). \tag{2.1}
\]

It assumes travel demand from origin zone \( i \) to destination zone \( j \) \((Q_{ij})\) is directly proportional to the total trip production at the origin zone \( i \) \((P_i)\) and the total trip attraction at the destination zone \( j \) \((A_j)\), while the friction factor \((F_{ij})\) denotes the reluctance or impedance (e.g., travel time) for making the trip from origin zone \( i \) to destination zone \( j \). Parameter \( K_{ij} \) represents a
socioeconomic adjustment factor for the given O/D travel and \( n \) denotes the total number of analysis zones. Goncalves and Ulyssea-Neto (1993) developed a gravity-opportunity model a mixture of both gravity and intervening-opportunity models. The authors showed that the conventional gravity model is a special case of the suggested model, and an application in public transit passenger flow in southern Brazil is provided. Levinson and Kumar (1995) developed multimodal trip distribution impedance functions for a metropolitan area in Washington. They showed the model reflects changes in transportation supply better than the conventional gravity model which uses the impedance of automobile only.

In case of freight demand modeling, Rawling and DuBoe (1991) used employment distribution to estimate truck trip distribution in Chicago area. The City of Portland’s Office of Transportation (1994) directly created regional distribution of truck freight demand based on Port of Portland commercial vehicle survey data. The gravity model that has been commonly used in the passenger travel demand analysis is adopted and widely used to distribute freight shipment demand in many studies (Cohen et al., 2008). A simple gravity model is adopted in Transmode Consultants, Inc. (1995) in which distance, travel time, and travel cost are included as transportation performance measures. A fully constrained gravity model (or an entropy model) was investigated in Wilson (1970a) and further applied in Black (1999). Mao and Demetsky (2002) studied a commodity based gravity model for freight flow distribution of truck mode. The authors defined four freight flow scenarios to consider freight movements within Virginia and between Virginia and other regions. Also, friction factors which represent difficulties associated with moving freight among zones were calculated and calibrated accordingly. As an application of the suggested gravity model, future freight flow distribution was forecasted based on production and attraction estimations. The Fratar method was also adopted from the conventional
passenger travel planning side for the freight demand distribution process. For example, Cambridge Systematics, Inc. et al. (1996) introduce the Fratar method and its applications to estimate future freight flow at the statewide level. Sorratini and Smith (2000) studied a freight demand forecasting model only for the truck mode in the state of Wisconsin. They first generated trip production and attraction rates using various public and private databases. Resulting travel demand was categorized into four trip types and distributed to each traffic analysis zone using both the gravity model and the Fratar method.

In this dissertation work, freight trip generation and trip distribution are investigated in Chapter 3. In general, previous studies on trip generation estimated freight production and attraction in study regions using a simple linear regression on the total employment number (by industry type) or population number. However, in our study, I/O commodity value growth forecasts for all FAZs and all commodity types are given exogenously; the global economic forecast models provide initial projections of various economic factors, which serve as inputs to the urban spatial structure and I/O models. Then, the amount of freight movements that begin and end in each FAZ for all commodity types can be directly estimated by scaling base-year freight production and attraction. For freight trip distribution analysis, Cohen et al. (2008) report that the gravity model has been commonly used in various statewide trip distribution analyses. However, it may not be suitable for our nationwide freight shipment analysis for the following reasons. First, in national scale freight movements, shipping cost per unit weight can be a more important friction factor than vehicle travel time or distance. For instance, bulk of heavy and time-insensitive goods can be transported through a far distance from its origin to its destination using a transportation mode that provides the lowest shipment cost per unit weight. Second, the socioeconomic adjustment factors in the gravity model are hard to be predicted for future years.
under multiple global economic growth and environmental policy scenarios. Lastly, a base-year travel time matrix between all shipment O/D pairs is not available in our study, but that is required for calculation of the friction factors in the gravity model. Thus, a two-dimensional matrix balancing RAS method is used to distribute the estimated future freight production and attraction among all pairs of freight shipment O/D zones. This approach utilizes a base-year freight demand distribution matrix which already implies impedance associated with freight movements such as unit shipment cost, shipment distance or time among all shipment O/D pairs for various commodity types. Also, this approach requires neither socioeconomic adjustment factors nor travel time matrix between all shipment O/D pairs. Result obtained from this step will be used as input to the next module of the four-step framework, modal split.

2.1.1.2 Freight transportation mode choice and its environmental impacts

There have been some studies related to the freight demand mode choice modeling. Allen (1977) and Daughety (1979) proposed microeconomic freight transport demand models to show that optimal total flow and mode choice can be obtained via shippers’ profit maximization. Winston (1981) developed a freight transportation mode choice model based on utility maximization of individual decision makers. Winston (1983) compared aggregate and disaggregate freight transportation demand models, and Gray (1982) reviewed three types of freight mode choice models, including economic positivist, technological positivist, and perceptual approaches. Abdelwahab and Sargious (1991) and Holguín-Veras (2002) proposed joint discrete-continuous decision processes on shipment size and freight transportation mode choices. In their models, the decisions on shipment size take continuous values while those on mode choice are discrete. Windisch et al. (2010) presented a transportation chain and shipment size choice model in which
the shipment size is categorized into 18 discrete levels. Multinomial logit models (and different variants) have been widely applied to freight shipment mode choice (Golias and Yannis, 1998; Catalani, 2001; Arunotayanun and Polak, 2011). Nam (1997) developed a set of logit mode choice models for shipments of six freight commodity types in South Korea. Shinghal and Fowkes (2002) examined determinants of freight shipment mode choice in India. This study used stated preference empirical survey data in 1998 and analyzed mode choice attributes including service frequency, reliability, service time, and cost. Jiang et al. (1999) developed a nested multinomial logit mode choice model using a large-scale, national freight demand survey database in France in 1988. The study concluded that French shippers tend to show the highest likelihood of selecting public road transportation if the shipping distance is approximately 700 km, while that of choosing rail transportation peaks around 1,300 km. A similar trend is also shown in the U.S., as Bryan et al. (2007) summarize that trucks have been appealing for local or regional freight shipments in urban areas, while rail and intermodal are competitive for inter-regional traffic shipments spanning several hundred miles or more.

These previous studies provided very useful insights on how freight transportation mode choice decisions are influenced by various factors. However, they have not explicitly considered the effect of oil price change, which has taken a large share in freight transportation operation cost across all modes. In addition, the high energy efficiency has become the key factor for choosing the transportation mode recently (TEMS, Inc., 2008). Therefore, this dissertation work aims to incorporate oil price as an independent variable in a mode choice model so that this model will be useful for decision-makers to evaluate future oil price effect on freight delivery mode choice decisions. Furthermore, it will play an important role in estimating the impacts of consequent mode choice decisions on air quality and climate change.
2.1.1.3 Freight demand network assignment under congestion

Freight shipment routing across highway and rail networks fall into the category of traffic assignment problems. Multiple O/D freight demand share the same infrastructure (i.e., network) and the freight flow is loaded onto the each modal network to satisfy certain traveler/shipper objectives, for instance, minimizing travel time of each vehicle. Such problems are also sometimes called the network equilibrium problems. Although the general principles of network equilibrium and traffic flow route choice can be applied to both highway and rail networks, a critical operational difference exists in the U.S. system. In a highway network, two nodes are often connected by two separate directed links in opposite directions. On the other hand, two nodes in a rail network should be connected by one undirected link since the same infrastructure is often shared by traffic flow in both directions.

Methodologies for solving the traffic assignment problems have been studied extensively, so a brief overview is provided as follows. Following the seminal work by Frank and Wolfe (1956), which formulated the problem into a quadratic program that can be solved by the convex combinations algorithm, Von Hohenbalken (1971, 1975) modified the convex combinations algorithm into a simplicial decomposition algorithm, which was later modified further into a restricted version (Hearn et al., 1987) and a disaggregate version (Larsson and Patriksson, 1992). Column generation algorithms were also used to solve the traffic assignment problem in Gibert (1968) and Dafermos and Sparrow (1969). Recently, Jayakrishnan et al. (1994) suggested a gradient projection method and Bar-Gera (2002) developed an origin-based assignment method. Previous studies on the traffic assignment problems and solution algorithms are mainly for automobile traffic on the highway networks, e.g., Sheffi (1985) (i.e., directed graph network); they have not considered the unique features of the rail network (i.e., undirected graph network).
Only more recently, traffic assignment problem formulations and related solution approaches have been developed for freight transportation systems. Winebrake et al. (2008) proposed a geospatial intermodal freight transportation model developed in a GIS platform. They combined road, rail, and waterway networks into one intermodal network with modal transfer points, in which each intermodal link is associated with travel time, cost, and emission. This model finds the path with the least delivery time, least cost, and least emission in the intermodal freight transportation network for a given O/D pair so that tradeoffs between these different criteria can be evaluated. The model in Winebrake et al. (2008) is applied in Comer et al. (2010) to investigate the use of marine vessels instead of heavy-duty trucks in the U.S. Great Lakes regions. In these studies, operational difference between highway network and rail network has not been investigated; rail network is also represented by a directed graph. Mahmassani et al. (2007) proposed a dynamic freight network simulation-assignment platform to analyze the multiproduct intermodal freight transportation systems in Europe, which was later applied and validated by Zhang et al. (2008). Assuming known time-dependent freight demand for each O/D pair, the least-cost paths through a sequence of shipment modes (e.g., truck, train, and ferry) are constructed iteratively while at the same time the overall network flow pattern is adjusted until convergence. Their generalized transportation cost (as perceived by shippers) included not only network link travel time but also transfer delay at intermodal transfer terminals, classification yards, and ports. The rail network is represented by a directed graph, and travel time on each link (including delay caused by trains’ meets and overtakes) is assumed to be captured by given time tables. This approach may be reasonable for these studies because train service in Europe is well-known to operate on schedule. However, it will be difficult to be applied to the U.S. freight rail
systems because substantial deviations from even daily schedule exist, and thus terminal managers frequently modify original train operation plan on an ongoing basis (Sussman, 2000).

Freight traffic assignment analyses can be used to help decision making in various ways. For example, Nair et al. (2008) used the network simulation platform in Mahmassani et al. (2007) to investigate the potential competitiveness of rail-based intermodal service along the RERIENT network in Europe under several operational and policy scenarios. In this study, new service design options are proposed from market-based research and expert opinion; it turned out that operations under high level of service combined with favorable rail policies attract significant demand. Kuo et al. (2008) evaluated three collaborative decision-making strategies for international rail-based intermodal freight service operations by multiple carriers. This study used a discrete-time carrier collaboration simulation-assignment platform, an extension of the earlier framework in Mahmassani et al. (2007). It was found that the proposed strategies not only attract more demand but also save cost and reduce shipment delay.

The existing studies suggested very efficient algorithms to solve the network assignment problems and provided examples of how the models and the solution algorithms can be applied to various fields. However, they have not addressed carefully the operational difference between highway network (i.e., directional flow on a separate link) and rail network (i.e., bi-directional flow on one shared link). In this dissertation work, we aim to conduct freight traffic demand network assignment analysis considering congestion effect on truck and rail networks. The fundamental premise is that freight shipments cause congestion in networks and in turn increase transportation cost. In the case of truck freight demand assignment problem, the basic traffic assignment model is developed using the simplified representation of the entire U.S. highway network and 2007 national freight demand data. Link cost function is modified to capture the
effect of background traffic volume that already exist in the given network. The conventional
convex combinations algorithm is applied to obtain optimal solution based on the user
equilibrium principle. In the case of rail freight shipments, a customized network assignment
model based on the user equilibrium principle is proposed; our model addresses a practical issue
of the rail network, where traffic flow in two opposite directions generally need to be loaded on
one shared link. A railroad-specific link cost function adjusted for single and double tracks is
constructed and incorporated into our model. The modified convex combinations algorithm is
applied to an empirical case study with the full-scale U.S. rail network and national freight
shipment data in Year 2007. The results will be one of the important measures to estimate the
environmental impacts from the freight transportation systems from a macroscopic view. Such
effort will also contribute to public benefit by reducing adverse social impacts imposed by
network traffic congestion.

2.1.2 User equilibrium in transportation networks
Detailed explanation about the traffic assignment problem is reviewed in this section, primarily
following the notations in Sheffi (1985). The last part of the four-step inter-regional freight
demand forecasting framework is freight traffic assignment. Given freight shipment demand and
a graph representation of the network, the traffic assignment problem determines the optimal
freight flow pattern between all O/D travel demand in a network. The link performance function
defines relationship between the link travel time and the assigned traffic flow on the link
assuming the link cost (i.e., link travel time) increases as the traffic flow on the link increases
due to the congestion caused by limited link capacity. The traffic flow on each link is sum of the
flow on several routes connecting many possible O/D pairs.
The traffic assignment problem requires a rule for individual travelers to select their routes. There can be generally two different rules including user equilibrium and system optimum. Following the user equilibrium principle, each motorist is assumed to know all network information (i.e., travel time for all possible routes) and select the shortest travel time route (given the congestion pattern collectively caused by all travelers) between their O/D. Eventually, when equilibrium is reached, all used routes connecting each O/D pair have the same cost less than or equal to the costs of unused routes. Thus, no drivers can reduce their travel time by unilaterally choosing another route. Following the system optimum principle, the total system-wide travel time spent by all motorists in the network is minimized. In the system optimum state, motorists might be able to reduce their travel time by unilaterally changing their routes; e.g., some travelers have to sacrifice their own interests for the benefit of the entire system. This implies that system optimum state often requires centrally controlled decisions, and it might not be a proper route choice rule for modeling independent travelers’ behavior. In this study, traffic assignment problem subject to the user equilibrium principle is formulated for both highway and rail networks. To solve the problem, the convex combinations algorithm is widely used. Following this algorithm, traffic demand for all O/D pairs can be assigned on the network such that the travel times on all used routes connecting each O/D pair equal each other and it is less than or equal to the travel times on unused routes between the O/D.

2.2 Intra-regional Freight Demand Modeling

This section reviews previous studies related to the intra-regional freight shipment problems. Intra-regional freight delivery modeling addresses two areas: (i) investigating a logistics systems model to connect a large number of demand as well as supply points distributed over the FAZs,
and (ii) developing a decision making model for a freight truck driver on an urban road network in a stochastic congestion state. The second problem addresses the microscopic level of the freight transportation systems where individual truck routing from one origin to one destination is the focus. These origin and destination are generic; they could be the truck terminals and/or consecutive delivery points in a freight delivery tour within an FAZ.

2.2.1 Logistics systems planning for a large-scale freight delivery problem

After completing the four-step inter-regional freight demand forecast, the bulk of freight arrives at a number of truck and railroad terminals in each FAZ, which needs to be disaggregated and delivered to widely distributed individual customers within delivery region (i.e., the final destinations of freight demand). Also, a fleet of short-haul trucks collects freight from a large number of supply points in each origin FAZ to the truck and railroad terminals. As such, the vehicle routing problem (VRP) is closely related to our intra-regional freight distribution and collection problem since a logistics systems model to connect the freight demand and supply points by a fleet of vehicles needs to be investigated. The VRP is one of the combinatorial optimization problems, in which vehicles that start and end their delivery service at a central depot (or terminal) need to serve a number of spatially distributed customers. The objective is to minimize the total cost for freight delivery service (Toth and Vigo, 2002). Since Dantzig and Ramser (1959) introduced the VRP, numerous studies have been presented to solve the problem. For example, Solomon (1987) and Potvin and Rousseau (1993) proposed constructive heuristics, and Or (1976), Savelsbergh (1991), Thompson and Psaraftis (1993), Potvin and Rousseau (1995), and Taillard et al. (1997) studied local search algorithms to solve the VRP. The VRP with Time Windows (VRPTW) is an extension of the traditional VRP in which each customer needs to be
visited within a certain time interval which is called as a time window constraint (Solomon, 1987; Savelsbergh, 1991; Potvin and Rousseau, 1993, 1995; Tan et al., 2006; Kallehauge, 2008). Another variation of the VRP is a VRP with Pickup and Delivery (VRPPD) in which each customer has two types of demand, pickup and delivery service (Toth and Vigo, 2002; Xu et al., 2003; Lin, 2011). Although extensive studies have been conducted on the VRP and its variations, and numerous solution algorithms have been proposed by many researchers, they are practically hard to be implemented in our problem since this study is related to a large-scale freight distribution and collection problem.

The large-scale VRP is typically solved by a “cluster-first, route-second” approach in which the total delivery region is partitioned into many vehicle routing zones (VRZs) such that each zone contains a certain number of delivery demand and the VRP is conducted within each zone. Daganzo (1984a, b) presented an easy manual recipe to construct the tour zones and a near optimal travel cost was obtained from simple formulae provided in the literature. Newell and Daganzo (1986a, b) developed guidelines for constructing the VRZs in a large-scale network, in which a Continuum Approximation (CA) optimization scheme is applied assuming stochastic delivery points can be represented by a continuous customer demand density function that may vary slowly in a region. They also provided simple formulae to calculate the near optimal total delivery cost. Since it is an asymptotic approximation method for large-scale problems, better result can be obtained as more delivery points are included in the delivery region. Newell (1986) improved the methodologies in Newell and Daganzo (1986a, b) by considering an inventory cost for delivering valuable commodities. Recently, Ouyang (2007) suggested algorithms to automatically design the VRZs and obtain a near optimal solution for the large-scale problem.
based on the CA approach. A set of zoning techniques including a disk model from Ouyang and Daganzo (2006) was used.

In this dissertation work, a ring-sweep algorithm (Newell and Daganzo, 1986a) is adopted to estimate the total freight delivery cost within the FAZ, which is sum of the total line-haul distance and the total local travel distance. Since the ring-sweep algorithm assumes that each demand point is composed of identical customers, the same amount of identical freight is required to be delivered to each demand point from the single source. However, each demand point in real world might consist of a number of customers depending on the industries involved. Also, there can be many terminals or depots where delivery starts and ends. Thus, in this study, the ring-sweep algorithm is modified to address these issues. To obtain the total cost for collecting the freight, we can assume a large number of supply points at origin FAZ, instead of demand points at destination FAZ, need to be served and the same approach can be applied to the collection of goods in the freight delivery region using freight production data. The output will be useful to estimate human exposure to emissions from freight delivery in large urban areas in the U.S.

2.2.2 Freight truck routing problem in a stochastic urban network considering emission effect

After the efficient logistics plan for freight distribution (and collection) within each FAZ has been determined, a truck driver needs to make routing decisions whenever he/she travels from one to another freight demand (or supply) points. This is a microscopic level problem as well as the last step analysis in the entire freight transportation system models. Since roadway congestion in large urban areas is stochastic due to many uncertain factors such as unexpected car accidents, this study needs to be investigated within a stochastic network framework.
Route/path optimization problems in a stochastic network setting have been studied extensively in computer science and operations research fields. For example, Nielsen and Zenios (1993), Glockner and Nemhauser (2000), and Boyles and Waller (2010) considered stochastic network flow problems, whereas Leipala (1978), Krauth and Mezard (1989), Percus and Martin (1999), and Tang and Miller-Hooks (2005) discussed stochastic traveling salesman problems. Among them, the closest literature studies are stochastic shortest path problems.

A number of algorithms for the shortest path problems in a network with fixed, deterministic link travel cost have been proposed and the standard shortest path algorithms such as dynamic programming and Dijkstra’s algorithm (Papadimitriou and Steiglitz, 1998) are very well known. Frank (1969) introduced a problem of estimating probability distribution of length of the shortest path, in which the nonnegative, real number cost or time associated with each link is a random variable. Many follow-up studies have been conducted to propose the shortest path algorithms under stochastic link cost. Sigal et al. (1980) presented the stochastic shortest path problem in a directed, acyclic network with independent, random arc length, in which a path optimality index is proposed to provide a measure for finding an optimal path. Loui (1983) defined a utility function which represents preference of each candidate path in a stochastic optimal path problem where weight on each edge is nonnegative, real-valued, and independent random variable with known probability distribution.

The stochastic shortest path problems have been extended in various ways. Hall (1986) introduced a time-dependent stochastic shortest path problem where link travel time is a stochastic process and depends on arrival time at link starting node. The author suggested a dynamic programming based time-adaptive route choice rule and a small size transit network example was provided. Fu and Rilett (1998) studied a dynamic and stochastic shortest path
problem and suggested $k$-shortest path algorithm based heuristic to find the expected shortest path. Another variant is a stochastic shortest path problem with recourse where network information is revealed throughout whole period of problem, and a decision maker needs to recalculate the expected cost for remaining routes at each decision point based on the information disclosed (Croucher, 1978; Andreatta and Romeo, 1988). Miller-Hooks and Mahmassani (2000) considered a recourse problem in stochastic, time-dependent transportation networks and suggested two algorithms to find the least expected time path and the lower bound on the least expected cost. More recently, Waller and Ziliaskopoulos (2002) developed a similar problem in a static network with limited spatial and temporal inter-arc dependencies. One-step spatial dependence assumes that once predecessor arc information is provided, the information on further arcs has nothing to do with the expected cost of the current arc. Limited temporal dependence implies that link travel time is realized once a traveler reaches the link starting node and every visit to the link is an independent stochastic trial.

These previous research efforts proposed efficient solution techniques to the various stochastic shortest path problems. However, they seemed to have considered only travel time as a total cost component. In this study, we present a methodology to obtain the minimum expected total cost of a freight truck delivery in a stochastic congestion network. We incorporated into the total travel cost not only the stochastic truck travel time, but also cost of various emissions and a penalty for late or early truck arrival at delivery destination. Although significant efforts have been made recently in combining environmental effects from vehicle emissions with various transportation studies, they are not stochastic but deterministic network modeling and optimization problems (Nagurney, 2000; Yin and Lawphongpanich, 2006; Zhang et al., 2010; Bektaş and Laporte, 2011; Aziz and Ukkusuri, 2012) or traffic control problems such as signal
timing optimization at urban intersections (Liao and Machemehl, 1998; Stevanovic et al., 2009). In this study, various vehicle emissions are assumed to follow U-shaped functions with respect to vehicle speed (TRL, 1999). Also, penalty cost will be assigned if freight is delivered earlier or later than the scheduled arrival time at the destination. Thus, solution satisfying the least expected travel time does not necessarily guarantee the least expected “total” cost in our problem. This study can provide a scientific basis for policy makers to evaluate the impacts of freight truck operations in urban areas on air quality and to develop strategies to mitigate the negative environmental consequences especially under roadway congestions and various delivery constraints.
CHAPTER 3
FORECASTING FREIGHT DEMAND CONSIDERING ECONOMIC GROWTH FACTORS

3.1 Introduction

According to Cohen et al. (2008), the amount of freight movement within the U.S. has grown dramatically in recent three decades. For example, the total U.S. freight shipment in ton-miles increased more than 70% between 1970 and 2000 (BTS, 2012). Globalization and the corresponding changes in transportation and logistics systems affected the growth in freight shipment (Hesse and Rodrigue, 2006). On the other hand, the sharp increase in freight activities has generated great amount of emissions from various freight shipment modes (ICF Consulting, 2005), which degraded air quality and affected human health and welfare significantly. Although there has been a lot of effort to improve energy efficiency to reduce total energy consumption in freight transportation and the following emissions, emissions from freight shipment modes are still expected to increase further (Schipper et al., 1997).

In this regard, freight transportation modes have been taking a large share in generating various air pollutants and greenhouse gases, and this problem motivated our study to develop comprehensive freight transportation system models. According to Jiang et al. (1999), a lot of research has been done on forecasting movement of passenger travel, but considerably less focus has been put on modeling freight demand. The freight demand forecasting models in this study are divided into two parts including inter-regional and intra-regional freight flows. In inter-regional freight demand modeling, future commodity specific freight demand within each FAZ will be estimated and distributed to each O/D zone pair, and they will be split into each freight
mode and assigned on each modal network considering congestion effect. In intra-regional freight demand modeling, freight movement within each FAZ and its total travel cost will be estimated using suitable logistics systems models.

The primary purpose of this chapter is to forecast future freight demand for all commodity types that begin and end in each FAZ and the amount of freight that moves between all pairs of zones under various combinations of potential global economic growth and environmental policy scenarios. These procedures correspond to the trip generation and trip distribution steps in the four-step inter-regional freight demand forecasting framework as denoted by the grey boxes in Figure 3.1.

![Trip Generation → Trip Distribution → Modal Split → Traffic Assignment](image)

**Figure 3.1** Trip generation and trip distribution in the four-step freight demand forecasting model

Previous studies related to the trip generation analysis typically estimated freight production and attraction in a region using simple linear regression equations in which total employment number by industry type or population number are often used as dependent variables (Fischer et al., 2001; Cohen et al., 2008). However, since future I/O commodity value growth at all FAZs for all commodity types under various global economic development and environmental policy scenarios are exogenously given from the urban spatial structure and I/O models (Isard, 1951, 1960; Leontief and Strout, 1963; Wilson, 1970a, 1970b), the amount of freight movements that begin and end in each FAZ can be directly forecasted by scaling the current year freight production and attraction. Then, the estimated freight production and attraction for each FAZ can be distributed on all shipment O/D pairs using an RAS algorithm, a
two-dimensional matrix balancing approach. The result obtained from this chapter will serve as inputs to the next step, the modal split procedure.

3.2 Freight Demand Forecasting Model

This section presents procedure for forecasting future freight demand and distributing them on all shipment O/D pairs. The RAS algorithm is adopted to iteratively allocate future freight production and attraction on all freight demand O/D zone pairs proportionally to the current freight demand distribution. Since its introduction in Deming and Stephan (1940) and further development by Csiszár (1975) and Bishop et al. (1975), the RAS algorithm has been widely used in various research fields including economics and statistics.

To apply the RAS algorithm, several assumptions are made in this section: (i) forecast of economic growth factors are given for all commodity types and all FAZs, (ii) current FAZ structure does not change (i.e., neither new zones will appear nor currently existing zones will disappear), and (iii) distribution of future freight demand is proportional to the that of current freight demand between all FAZ O/D pairs.

Figure 3.2 Base-year freight demand distribution data structure
Let $Q$ denote a set of commodity types, which is assumed to be composed of $N$ different types, i.e., $Q=\{1,2,\ldots,N\}$. Figure 3.2 describes the structure of the base-year freight demand distribution data for a certain commodity type $i \in Q$. All squares on the left show origin zones denoted by set $O$, and those on the right represent destination zones denoted by set $D$. Since each FAZ can be both origin and destination of a freight shipment, the origin zone set $O$ and the destination zone set $D$ are generally composed of the same elements indexed from 1 to $Z$, i.e., $O=D=\{1,2,\ldots,Z\}$. The set of arrows connecting all FAZ O/D pairs describe freight movements. Also, we define the variables such that $P^i_o$ represents base-year total production of commodity $i \in Q$ in an origin zone $o \in O$, $A^i_d$ describes base-year total attraction of commodity $i \in Q$ in a destination zone $d \in D$, $D^i_{od}$ represents freight volume of commodity $i \in Q$ moving from origin zone $o \in O$ to destination zone $d \in D$. We use $\alpha^{i,y}_o$ to denote growth rate of commodity $i \in Q$ production in an origin zone $o \in O$ for future year $y$. Similarly, $\beta^{i,y}_d$ is used to represent growth rate of commodity $i \in Q$ attraction in a destination zone $d \in D$ for future year $y$. Then, detailed procedures can be described as follows:

Step 0: Generate base-year freight demand O/D matrix for a commodity type $i$, $\forall i \in Q$, as shown in Figure 3.3. Each row describes each origin and each column represents each destination spanning from 1 to $Z$. Let $D^i_{od}$ be base-year commodity $i$ freight movement from origin $o$ to destination $d$. The last two columns describe given production and future production for each origin zone, and currently information related to only given production is available from base-year data such that the sum of $D^i_{od}$ in a row direction is $P^i_o$. Similarly, the last two rows represent given attraction and future attraction for each destination zone, and
the cells with given attraction can be filled with the base-year data such that the sum of $D_{od}$ in a column direction becomes $A_d^i$.

![Base-year freight demand distribution matrix](image)

**Figure 3.3 Base-year freight demand distribution matrix**

Step 1: Estimate future production ($V_o^i$) and attraction ($W_d^i$) represented by grey color column and row in Figure 3.4 for all FAZs such that each $P_o^i$ and $A_d^i$ is multiplied by $\alpha_o^{i,o}$ and $\beta_d^{i,d}$ respectively, i.e., $V_o^i = \alpha_o^{i,o} \cdot P_o^i$, $W_d^i = \beta_d^{i,d} \cdot A_d^i$, $\forall o \in O$, $d \in D$.

![Freight demand distribution matrix with future production and attraction](image)

**Figure 3.4 Freight demand distribution matrix with future production and attraction**

Step 2: Since estimations of future input and output commodity value growth are modeled separately, the total future production summed across all origin zones
(i.e., $\sum_{o \in O} V^i_o$) could be different from the total future attraction summed across all destination zones (i.e., $\sum_{d \in D} W^i_d$) although theoretically the total sums of production and attraction should be the same for the whole U.S. Therefore, assuming freight commodity production is derived by attraction, we multiply future production of all origin zones by the same factor such that the total sum of modified future production is balanced with the total sum of future attraction:

$$ \text{Update } V^i_o \leftarrow V^i_o \left( \frac{\sum_{d \in D} W^i_d}{\sum_{o \in O} V^i_o} \right), \forall o \in O.$$ 

Step 3: The RAS algorithm is applied. In this algorithm, first we define the growth factor $R^i_o$, $\forall o \in O$, to adjust each entry $D^i_{od}$ in a row direction to match with the future production $V^i_o$, $\forall o \in O$. Also, we define the growth factor $C^i_d$, $\forall d \in D$, to adjust each entry $D^i_{od}$ in a column direction to match with the future attraction $W^i_d$, $\forall d \in D$. Then, we modify the matrix in a row direction first and then in a column direction, and we do so iteratively until the sums of each row and each column converge to both future production and future attraction respectively. Distribution of future freight demand obtained from the suggested RAS algorithm will be proportional to that of the base-year freight demand. The detailed algorithm is described as follows:

Define tolerance $\varepsilon \ll 1$, and let $L = $ large positive integer and $n = 1$.

Define $R^i_o = \frac{V^i_o}{\sum_{d \in D} D^i_{od}}$, $\forall o \in O$, and $C^i_d = \frac{W^i_d}{\sum_{o \in O} D^i_{od}}$, $\forall d \in D$.

While $\{(n \leq L) \text{ and } (|R^i_o - 1| > \varepsilon \text{ for some } o \in O \text{ or } |C^i_d - 1| > \varepsilon \text{ for some } d \in D)\}$

{ 
  Set $D^i_{od} \leftarrow R^i_o D^i_{od}$, $\forall o \in O$, $d \in D$.
}
Update $C'_d \leftarrow \frac{W^i_d}{\sum_{o \in O} D^i_{od}}$, $\forall d \in D$,

Set $D^i_{od} \leftarrow C^i_d D^i_{od}$, $\forall o \in O$, $d \in D$,

Update $R^i_o \leftarrow \frac{V^i_o}{\sum_{d \in D} D^i_{od}}$, $\forall o \in O$,

Update $n \leftarrow n+1,$

} 

In this algorithm, whenever $R^i_o$ (and $C^i_d$) is calculated or updated, following sub-algorithm is also conducted to avoid the case that the denominator of $R^i_o$ (and $C^i_d$) is zero. Note that the algorithm described in parenthesis is for the growth factor $C^i_d$.

If $\sum_{d \in D} D^i_{od} = 0$ for some $o \in O$ (If $\sum_{o \in O} D^i_{od} = 0$ for some $d \in D$)

If $V^i_o = 0$, update $R^i_o \leftarrow 1$.

(If $W^i_d = 0$, update $C^i_d \leftarrow 1$.)

Else, update $D^i_{od} \leftarrow \frac{V^i_o}{|D|}$, $\forall d \in D$ and $R^i_o \leftarrow 1$.

(Else, update $D^i_{od} \leftarrow \frac{W^i_d}{|O|}$, $\forall o \in O$ and $C^i_d \leftarrow 1$.)

Else, update $R^i_o \leftarrow \frac{V^i_o}{\sum_{d \in D} D^i_{od}}$. (Else, update $C^i_d \leftarrow \frac{W^i_d}{\sum_{o \in O} D^i_{od}}$.)

In this way, future freight demand generation for all commodity types and for all FAZs can be forecasted, and the results can be distributed on all shipment O/D pairs proportionally to the current freight demand distribution.
3.3 Future Freight Demand Forecast

3.3.1 Data sources and processing procedures

The proposed algorithm is applied to forecasting future freight demand distribution within the U.S. from 2010 to 2050 in five-year increments. We investigate four scenarios (proposed by the global economic forecast models) to consider variations in both global economic growth and environmental regulation. To capture uncertainty in the global economy, “low GDP growth” and “high GDP growth” scenarios are considered. For each hypothetical economic development scenario, two different environmental regulations are addressed including “business as usual” and “climate policy” scenarios. The former assumes global emission projections will follow historical trends, while the latter assumes cumulative constraints in energy-related CO₂ emissions (such as carbon tax) will be implemented. The base-year freight demand distribution matrix and the future I/O commodity value growth estimates for all scenarios are inputs to this analysis.

The base-year freight demand distribution is collected from the Freight Analysis Framework data version 3 (FAF³) obtained from the U.S. Department of Transportation’s Federal Highway Administration (FHWA U.S. DOT, 2011). The FAF³ database contains information on the freight shipment activities in Year 2007. It includes amount of freight flow in terms of tonnage and value between all shipment O/D pairs in the U.S. by 7 modes of transportation and 43 types of commodities which are defined by the Standard Classification for Transported Goods (SCTG) code. The FAF³ database consists of 131 geographical regions for both origins and destinations (i.e., 123 domestic FAZs and 8 international regions). From the total FAF³ records, data related to two major freight shipment modes in the U.S., truck and rail, are extracted. Since we are focusing on the freight shipment within the U.S. continent, data for the international freight movements and those related to Hawaii and Alaska are excluded. Then,
43 different kinds of commodities in the original dataset are grouped into 10 types based on physical and economical similarity. Finally, a total of 128,562 data records are obtained to construct the base-year freight demand distribution matrix, which include origin, destination, commodity type, and freight demand in tons.

The commodity production and attraction growth estimates for all FAZs and future years under different scenarios are obtained exogenously from the urban spatial structure and I/O models (Isard, 1951, 1960; Leontief and Strout, 1963; Wilson, 1970a, 1970b). Data include commodity production and attraction in terms of dollars for all FAZs and for 10 commodity types from 2005 to 2050 in five-year increments for four different scenarios. The base-year is set to be 2007, and the data for this year are obtained using linear interpolation between 2005 and 2010 data.

3.3.2 Forecasting results

The proposed algorithm is coded and tested on a personal computer with a 3.4 GHz CPU and 8 GB memory. The algorithm converged in a short time (i.e., within a few minutes) and the future freight demand from 2010 to 2050 in five-year increments is generated for each scenario. The results are presented in three hundred and sixty 120-by-120 matrices, each of which estimates the total freight demand among all FAZ O/D pairs for a specific future year, commodity type, and scenario. Table 3.1 summarizes the total freight demand forecast for four given scenarios, summed across all commodity types and all FAZ O/D pairs.

Row (a) of Table 3.1 shows four global economic growth and environmental policy scenarios. Column (b) represents future years from 2010 as well as the base-year 2007. Columns (c)-(f) describe the total freight demand in thousand tons, and columns (g)-(j) calculate
percentage changes of the total freight demand from the base-year data. The freight demand in Year 2007 is the observed benchmark used in all scenarios.

Table 3.1 Computational results for forecasting future freight demand

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Year</th>
<th>Total freight demand forecasted (thousand ton)</th>
<th>% change</th>
<th>Total freight demand forecasted (thousand ton)</th>
<th>% change</th>
<th>Total freight demand forecasted (thousand ton)</th>
<th>% change</th>
<th>Total freight demand forecasted (thousand ton)</th>
<th>% change</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Scenario</td>
<td>(b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 1:</td>
<td>2007</td>
<td>15,059,745</td>
<td>0.00</td>
<td>15,059,745</td>
<td>0.00</td>
<td>15,059,745</td>
<td>0.00</td>
<td>15,059,745</td>
<td>0.00</td>
</tr>
<tr>
<td>High GDP growth with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>business as usual</td>
<td>2010</td>
<td>15,703,789</td>
<td>4.28</td>
<td>15,648,288</td>
<td>3.91</td>
<td>15,528,787</td>
<td>3.11</td>
<td>15,494,244</td>
<td>2.89</td>
</tr>
<tr>
<td>High GDP growth with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>climate policy</td>
<td>2020</td>
<td>19,431,308</td>
<td>29.03</td>
<td>18,780,540</td>
<td>24.71</td>
<td>18,355,956</td>
<td>21.89</td>
<td>17,742,894</td>
<td>17.82</td>
</tr>
<tr>
<td>Low GDP growth with</td>
<td>2025</td>
<td>21,438,103</td>
<td>42.35</td>
<td>20,650,764</td>
<td>37.13</td>
<td>19,755,145</td>
<td>31.18</td>
<td>19,023,791</td>
<td>26.32</td>
</tr>
<tr>
<td>business as usual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 3:</td>
<td>2030</td>
<td>23,693,953</td>
<td>57.33</td>
<td>22,780,286</td>
<td>51.27</td>
<td>21,271,576</td>
<td>41.25</td>
<td>20,435,507</td>
<td>35.70</td>
</tr>
<tr>
<td>Low GDP growth with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>climate policy</td>
<td>2035</td>
<td>26,034,285</td>
<td>72.87</td>
<td>24,945,108</td>
<td>65.64</td>
<td>22,725,696</td>
<td>50.90</td>
<td>21,747,683</td>
<td>44.41</td>
</tr>
<tr>
<td>Low GDP growth with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>climate policy</td>
<td>2040</td>
<td>28,697,929</td>
<td>90.56</td>
<td>27,356,813</td>
<td>81.66</td>
<td>24,523,312</td>
<td>62.84</td>
<td>23,339,737</td>
<td>54.98</td>
</tr>
<tr>
<td>Low GDP growth with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>climate policy</td>
<td>2045</td>
<td>31,574,234</td>
<td>109.66</td>
<td>29,893,810</td>
<td>98.50</td>
<td>26,377,074</td>
<td>75.15</td>
<td>24,903,553</td>
<td>65.37</td>
</tr>
<tr>
<td>Low GDP growth with</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>climate policy</td>
<td>2050</td>
<td>34,673,664</td>
<td>130.24</td>
<td>32,621,827</td>
<td>116.62</td>
<td>28,351,364</td>
<td>88.26</td>
<td>26,573,564</td>
<td>76.45</td>
</tr>
</tbody>
</table>

Table 3.1 shows that estimations of the total future freight demand consistently increase for all global economic growth and environmental regulation scenarios. Around Year 2030 in scenarios 1 and 2, Year 2035 in scenario 3, and Year 2040 in scenario 4, future freight demand increase more than 50% compared to the base-year freight demand. Note that the total freight demand for truck and rail in the U.S., in fact, decreased by 9% from 2007 to 2011 (FHWA U.S. DOT, 2011) due to the recession that began in late 2007. This discrepancy was caused by the fact that the global economic forecast models which provide initial projections of various economic factors are not able to capture unexpected short-term economic fluctuations. Outputs obtained in this section will be used as inputs to the modal split procedure in the four-step framework.
3.4 Conclusion

Due to the globalization and worldwide industrialization, demand for freight delivery has been persistently increasing in recent decades. The fast increase in the amount of freight movements induced environmental problems such as global warming and degraded air quality in urban areas. The emissions generated from various freight shipment modes have become major sources of air pollution and they have been affecting human health and social welfare. Therefore, the relationship between the freight delivery activities and the environmental problems motivated this dissertation work to construct the freight transportation system models.

In this study, the four-step freight demand forecasting framework is adopted for the inter-regional freight shipments. Specifically, this chapter focuses on trip generation and trip distribution steps, the first two modeling procedures among the four-step analysis. Assuming the estimates of future economic growth factors for all commodity types and FAZs are given, the total freight demand that will be produced and attracted can be forecasted for all zones. Then, the RAS algorithm is applied to distribute the estimated future freight demand on all shipment O/D pairs. Distribution of the future freight demand is assumed to be proportional to that of the base-year freight demand between all FAZ O/D pairs, and the base-year FAZ structure is assumed not to be changed in the future years. The proposed methodology is applied to generate future freight demand distribution from 2010 to 2050 in five-year increments considering four economic growth and environmental regulation scenarios. The FAF³ database is selected to construct the base-year freight demand matrices and the commodity production and attraction growth estimates are generated from the urban spatial structure and I/O models. As a result, the algorithm converged in a short time and the future freight demand distribution is obtained.
CHAPTE R 4
FREIGHT TRANSPORTATION MODE CHOICE AND ITS ENVIRONMENTAL IMPACTS\(^1\)

4.1 Introduction

Emissions from freight transportation operations usually vary significantly across different modes. For example, the U.S. EPA (2008) reports that for each ton-mile of freight shipment, truck, rail, waterborne craft, and aircraft respectively produce 0.297, 0.0252, 0.048, and 1.527 kg of CO\(_2\) emission; 0.0035, 0.002, 0.0041, and 0.0417 g of CH\(_4\) emission; and 0.0027, 0.0006, 0.0014, and 0.0479 g of N\(_2\)O emission. Also, NRDC (2012) shows that truck, rail, water, and air transportation modes generate 92, 13, 25, and 119 mg of PM\(_{10}\) emission for each ton-mile of freight movement. Therefore, even to transport the same amount of freight for the same O/D pair, different shipment modes can result in different emissions, which in turn affect air quality and human health. As of today, the truck mode carries the largest percentage of the total national freight movements in the U.S., and this percentage becomes even more prominent in the case of freight shipments within large states (Chin et al., 1998). However, compared to the trains, trucks exhibit significantly lower fuel efficiency and higher emission levels. In order to understand the environmental impacts from the freight systems, therefore, we need to investigate mode choice for freight shipment demand.

\(^1\) This chapter has been adapted from “Hwang, T. and Ouyang, Y. (2014b). Freight shipment modal split and its environmental impacts: An exploratory study. Journal of the Air & Waste Management Association, 64(1), 2-12.” This article is reprinted by the written permission of copyright owner, the Air & Waste Management Association (http://www.awma.org).
Overall, shippers’ freight mode choice decisions are influenced by many factors, such as strength of regional economy, infrastructure capacity, and shipping distance or time. The dramatic surge in oil price during the past decade has become a critical issue in the U.S. freight transportation market, as the fuel cost comprises more than 50% of the total operating cost for the transportation industry nowadays (TEMS, Inc., 2008). Since the sensitivity of transport operating cost to the oil price change varies significantly across shipment modes, oil price has become an important factor in freight mode choices. Unfortunately, previous studies have largely ignored the effect of oil price on freight mode choice decisions, and few have tried to connect oil price to freight transportation emissions. Most of the existing work focused on theoretical model developments; e.g., inter-regional commodity flow analysis in an I/O framework, freight demand mode choice based on shipper’s profit maximization, and route selection in multimodal transportation networks. The empirical implementation of freight modal split models was also quite rare (possibly due to lack of data). Although more multi-year freight data have become available in recent years, few efforts have been made to clarify the relationship among various economic factors, freight transportation mode choice, and freight transportation emissions.

This dissertation work aims to fill these gaps by constructing an aggregated binomial logit market share model that estimates modal split between truck and rail (the two major freight shipment modes in the U.S.) for 10 groups of typical commodities, which corresponds to the step three in the four-step inter-regional freight demand forecasting framework as denoted by the grey box in Figure 4.1.

![Figure 4.1 Modal split in the four-step freight demand forecasting model](image_url)
In this model, we explicitly incorporate various economic and engineering factors as explanatory variables in the utility function in order to quantify their effects on freight transport mode choices. Such explanatory variables include crude oil price, truck and rail shipping distances, freight value per unit weight for each type of commodity. The quantitative models are obtained using empirical freight transportation demand data between origins and destinations (i.e., FAZs), and the models are validated using extra empirical data. We used four years of available data that can be found in the public domain. This study also provides discussions on freight transportation demand data statistics, interpretations of the parameter estimations, insights on their effects on freight mode choice, and environmental impact assessments.

The remainder of this chapter is organized as follows: A simple freight transportation modal split model is presented in Section 4.2. Section 4.3 describes the empirical data sources and the data cleaning procedures. Section 4.4 presents estimation results, model validations, and an illustrative example. Proposed model application for future modal split prediction is shown in Section 4.5. Finally, concluding remarks are provided in Section 4.6.

**4.2 Freight Transportation Mode Choice Modeling**

In this work, we focus on developing a freight modal split model within a four-step freight demand forecasting framework which is similar to the urban passenger travel demand forecasting model (Cohen et al., 2008). Assuming that a set of O/D freight demand data is given, we develop a macroscopic logit market share model for mode choice decisions as a function of a set of explanatory variables (e.g., crude oil price, freight value, shipment distance, etc.) because our databases only contain aggregated annual freight shipment observations at the freight zone level.
Our study focuses on two dominating freight modes, truck and rail, in a binomial logit model framework.

In this model, the annual market share of truck shipments (in terms of tonnage) between any O/D pairs is a dependent variable (whose value is between 0 and 1). Due to data availability, we consider four different explanatory variables for each commodity type: commodity value per ton ($/ton, denoted by $VALUE$), the average shipment distance for each mode (mile, denoted by $DIST_T$ for truck and $DIST_R$ for rail) and crude oil price ($/barrel, denoted by $OILPRC$). There might be additional factors affecting freight transportation mode choice. However, the independent variables used in this study are reasonably comprehensive. They include not only the majority of the most frequently-used independent variables in this context (Gray, 1982), but also those in the recent study for the State of Florida (Cambridge Systematics, Inc., 2002). In this analysis, $DIST_T$ and $DIST_R$ may differ even for the same O/D because they are measured in the U.S. highway and rail networks respectively. Crude oil price is selected as a proxy to represent a single oil price index because the crude oil price is a dominating factor in determining diesel fuel price (U.S. EIA, 2008). It shall be noted that although trucks and trains both use diesel oil as fuel, the unit diesel prices for truck and railroad are different; even for the same railroad company, diesel price varies significantly across fueling locations (Nourbakhsh and Ouyang, 2010).

As such, we assume that the utility functions for truck ($U^n_T$) and rail ($U^n_R$), for commodity type $n \in \{1, 2, \ldots, N\}$, can be defined as follows:

$$U^n_T = a_{1n} + b_{1n} \cdot VALUE + c_{1n} \cdot DIST_T + d_{1n} \cdot OILPRC,$$  \hspace{1cm} (4.1)

$$U^n_R = a_{2n} + b_{2n} \cdot VALUE + c_{2n} \cdot DIST_R + d_{2n} \cdot OILPRC.$$  \hspace{1cm} (4.2)
We define $P_T^n$ and $P_R^n$ respectively as the market share of truck and rail modes for commodity type $n$. These percentages can be constructed as follows (Gruca et al., 1991):

$$P_T^n = \frac{e^{U_T^n}}{e^{U_T^n} + e^{U_R^n}} = \frac{e^{U_T^n - U_R^n}}{e^{U_T^n - U_R^n} + 1},$$

(4.3)

$$P_R^n = \frac{e^{U_R^n}}{e^{U_T^n} + e^{U_R^n}} = \frac{1}{e^{U_T^n - U_R^n} + 1}.$$  

(4.4)

Then, the binomial logit model can be transformed into the form of equation (4.5) as follows (Pindyck et al., 1998):

$$\ln \left( \frac{P_T^n}{1 - P_T^n} \right) = U_T^n - U_R^n$$

$$= (a_{1n} - a_{2n}) + (b_{1n} - b_{2n}) \cdot VALUE + (c_{1n}) \cdot DIST_T + (-c_{2n}) \cdot DIST_R + (d_{1n} - d_{2n}) \cdot OILPRC.$$  

(4.5)

This equation takes a generalized linear form with four explanatory variables. The intercept $a_{1n} - a_{2n}$ and the coefficients $b_{1n} - b_{2n}, c_{1n}, -c_{2n}, d_{1n} - d_{2n}$ can be estimated via linear regression for each commodity type.

### 4.3 Freight Transportation Data

To construct the freight transportation demand model, we collected datasets from the Freight Analysis Framework data version 2 and 3 (FAF$^2$ and FAF$^3$) from the FHWA U.S. DOT (2011), the Commodity Flow Survey (CFS) data from RITA U.S. DOT (2011), and the crude oil price information from Economagic, LLC (1996).
4.3.1 Data sources and processing procedures

In order to develop the abovementioned model, we have to merge data from multiple sources into one useable dataset. The FAF\textsuperscript{2} and FAF\textsuperscript{3} record commodity shipment flow and related freight transportation activities between the U.S. geographical regions in Years 2002 and 2007 respectively. The CFS datasets provide freight transportation activities in Years 1993 and 1997, such as the volume and value of different types of commodities shipped between various origins and destinations by different modes. FAF\textsuperscript{2}, FAF\textsuperscript{3}, and 1997 CFS use the Standard Classification of Transported Goods (SCTG) and 1993 CFS adopted the Standard Transportation Commodity Code (STCC) to define commodity types. Average shipment distances of truck and rail modes are extracted from the CFS datasets. Since our research aims to address the impact of oil prices on freight shipment mode choices, we incorporate the West Texas Intermediate (WTI) crude oil price from Economagic, LLC (1996) as a proxy of oil price of each study year. In our study, WTI crude oil is selected among the three major crude oil benchmarks (i.e., WTI, Brent Blend, and Dubai) since WTI is used as a primary benchmark not only in the U.S. but also in the international market (Wikipedia Contributors, 2012).

When we process the data, each of the four datasets (FAF\textsuperscript{2} for Year 2002, FAF\textsuperscript{3} for Year 2007, 1993 CFS, and 1997 CFS) is combined with distance data based on origin, destination, and mode. Figure 4.2 shows the maps of the freight analysis zones in the FAF\textsuperscript{3} database and the national transportation analysis regions (NTAR) in the 1993 CFS database. Freight analysis zones in FAF\textsuperscript{2} are largely the same as those in FAF\textsuperscript{3}, while 1997 CFS define zones along state boundaries. We only use the data related to truck and rail modes, and different kinds of products (originally defined in the datasets) are grouped into 10 types of commodities based on physical and economical similarity. The yearly averaged WTI crude oil price remained rather stable from
$18.46 per barrel in 1993, $20.6 per barrel in 1997, and $26.1 per barrel in 2002, until a sharp increase to $72.36 per barrel occurred in 2007. Finally, information from these four sources is joined into one dataset which has 69,477 observations in total, while each observation corresponds to a year, a commodity type, and a shipment O/D pair. The dataset also contains information on commodity value per unit weight, truck and rail shipment distances, WTI crude oil price, and observed truck and rail shipment shares.

![Maps of freight zones](image)

(a) FAF$^3$ analysis zones  (b) National transportation analysis regions

Figure 4.2 Maps of freight zones (FHWA U.S. DOT, 2011; RITA U.S. DOT, 2011)

4.3.2 Freight transportation demand data statistics

Table 4.1 shows definition of commodity groups, number of data observations, and the total tonnage shipped by trucks and railroads for each type of commodity in the four years of interest. From column (d), we see that the truck mode seems to serve a broad spectrum of commodities, while commodity types 3 (i.e., stones, nonmetallic minerals, and metallic ores), 7 (i.e., base metal and machinery), 4 (i.e., coal and petroleum products), and 1 (i.e., agriculture products and fish) show the highest percentage share in shipment tonnage. However, it can be noticed from column (e) that the rail mode serves a very concentrated market; e.g., commodity type 4 (i.e.,
coal and petroleum products) occupies a dominant share (i.e., more than 50%) of the total rail shipment tonnage.

Table 4.1 Data statistics by mode and commodity type (four years total)

<table>
<thead>
<tr>
<th>(a) Commodity type</th>
<th>(b) Commodity description</th>
<th>(c) Number of observations</th>
<th>(d) Truck</th>
<th>(e) Rail</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Agriculture products and fish</td>
<td>5,704</td>
<td>4,585,159</td>
<td>530,930</td>
</tr>
<tr>
<td>2</td>
<td>Grain, alcohol and tobacco products</td>
<td>8,202</td>
<td>2,485,347</td>
<td>158,335</td>
</tr>
<tr>
<td>3</td>
<td>Stones, nonmetallic minerals, and metallic ores</td>
<td>5,630</td>
<td>7,830,644</td>
<td>550,560</td>
</tr>
<tr>
<td>4</td>
<td>Coal and petroleum products</td>
<td>4,657</td>
<td>5,339,089</td>
<td>2,999,961</td>
</tr>
<tr>
<td>5</td>
<td>Basic chemicals, chemical and pharmaceutical</td>
<td>8,824</td>
<td>1,928,319</td>
<td>585,510</td>
</tr>
<tr>
<td>6</td>
<td>Logs, wood products, and textile and leather</td>
<td>9,102</td>
<td>3,197,305</td>
<td>261,422</td>
</tr>
<tr>
<td>7</td>
<td>Base metal and machinery</td>
<td>9,053</td>
<td>6,377,410</td>
<td>291,342</td>
</tr>
<tr>
<td>8</td>
<td>Electronic, motorized vehicles, and precision</td>
<td>7,651</td>
<td>597,028</td>
<td>51,440</td>
</tr>
<tr>
<td></td>
<td>instruments</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Furniture, mixed freight, and miscellaneous</td>
<td>7,561</td>
<td>3,326,713</td>
<td>145,849</td>
</tr>
<tr>
<td></td>
<td>manufactured products</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Commodity unknown</td>
<td>3,093</td>
<td>649,901</td>
<td>4,428</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>69,477</td>
<td>36,316,916</td>
<td>5,579,778</td>
</tr>
</tbody>
</table>

Figure 4.3(a) shows that while the total annual freight tonnage increases steadily for both modes from 1993 to 2007, the increase is larger for railroad. Figure 4.3(b) compares the freight value per ton shipped by trucks and railroads. It can be seen that the truck mode consistently carries much more valuable goods on average.
(a) Total freight tonnage increasing rate

Figure 4.3 Total tonnage and freight value per ton in four years

(b) Freight value per ton

Figure 4.4 shows the cumulative percentage of the four-year total tons and ton-miles shipped by truck and rail (for all commodities). Truck and rail shipments display very different distributions across distances; e.g., Figure 4.4(a) shows that over 90% of truck tonnages are shipped for less than 300 miles, but more than 40% of railroad tonnages are shipped to destinations more than 700 miles away. A similar trend is also shown in Figure 4.4(b) such that over 50% of truck ton-miles are transported within 550 miles, however, more than 50% of railroad ton-miles involve trips of more than 1,200 miles. This suggests a strong relationship between shipment distance and mode choice.
4.4 Estimation, Validation, and Application

4.4.1 Estimation results

In this analysis, statistical software package (R version 2.12.1) was used to estimate the intercept and coefficients in equation (4.5). We divided the database into two sets for each commodity type: two thirds of the observations are included in a “training dataset” to estimate the model, and the remaining data are assigned to a “test dataset” to validate the suggested model. The estimation results for each commodity type, based on the “training dataset”, are included in Table 4.2.
Table 4.2 Estimation results and the goodness of fit

<table>
<thead>
<tr>
<th></th>
<th>Type 1</th>
<th>Type 2</th>
<th>Type 3</th>
<th>Type 4</th>
<th>Type 5</th>
<th>Type 6</th>
<th>Type 7</th>
<th>Type 8</th>
<th>Type 9</th>
<th>Type 10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>z-statistic</strong></td>
<td>+00</td>
<td>+00</td>
<td>+00</td>
<td>-01</td>
<td>+00</td>
<td>+00</td>
<td>+00</td>
<td>+00</td>
<td>+00</td>
<td>-01</td>
</tr>
<tr>
<td>**Pr(&gt;</td>
<td>z</td>
<td>)**</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Estimate</strong></td>
<td>0.00</td>
<td>0.29</td>
<td>0.80</td>
<td>0.00</td>
<td>0.20</td>
<td>0.90</td>
<td>0.90</td>
<td>0.40</td>
<td>0.80</td>
<td>0.40</td>
</tr>
<tr>
<td><strong>Value per ton</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>z-statistic</strong></td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-04</td>
<td>-04</td>
<td>-04</td>
<td>-04</td>
<td>-04</td>
<td>-03</td>
</tr>
<tr>
<td>**Pr(&gt;</td>
<td>z</td>
<td>)**</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Estimate</strong></td>
<td>8593</td>
<td>7124</td>
<td>1211</td>
<td>2538</td>
<td>7289</td>
<td>5238</td>
<td>4593</td>
<td>1948</td>
<td>3655</td>
<td>1545</td>
</tr>
<tr>
<td><strong>Avg. truck distance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intercept</strong></td>
<td>-1.532E</td>
<td>-1.766E</td>
<td>-1.930E</td>
<td>-1.663E</td>
<td>-1.531E</td>
<td>-1.904E</td>
<td>-3.142E</td>
<td>-4.025E</td>
<td>-1.901E</td>
<td>-2.042E</td>
</tr>
<tr>
<td><strong>z-statistic</strong></td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
</tr>
<tr>
<td>**Pr(&gt;</td>
<td>z</td>
<td>)**</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Estimate</strong></td>
<td>-2976</td>
<td>-1680</td>
<td>-2488</td>
<td>-3390</td>
<td>-2418</td>
<td>252</td>
<td>-3714</td>
<td>-2113</td>
<td>-1792</td>
<td>-472</td>
</tr>
<tr>
<td><strong>Avg. rail distance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>z-statistic</strong></td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
</tr>
<tr>
<td>**Pr(&gt;</td>
<td>z</td>
<td>)**</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Estimate</strong></td>
<td>-2258</td>
<td>5</td>
<td>-4958</td>
<td>-5019</td>
<td>485</td>
<td>2912</td>
<td>1613</td>
<td>1494</td>
<td>234</td>
<td>-138</td>
</tr>
<tr>
<td><strong>WTI crude oil price</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>z-statistic</strong></td>
<td>-03</td>
<td>-03</td>
<td>-02</td>
<td>-02</td>
<td>-03</td>
<td>-03</td>
<td>-03</td>
<td>-02</td>
<td>-02</td>
<td>-02</td>
</tr>
<tr>
<td>**Pr(&gt;</td>
<td>z</td>
<td>)**</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Estimate</strong></td>
<td>1634</td>
<td>-965</td>
<td>-5993</td>
<td>-14669</td>
<td>-2758</td>
<td>-818</td>
<td>-389</td>
<td>963</td>
<td>4948</td>
<td>432</td>
</tr>
<tr>
<td>**Pr(&gt;</td>
<td>z</td>
<td>)**</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Estimate</strong></td>
<td>0.00</td>
<td>0.59</td>
<td>0.00</td>
<td>0.90</td>
<td>0.30</td>
<td>0.90</td>
<td>0.90</td>
<td>0.40</td>
<td>0.80</td>
<td>0.40</td>
</tr>
<tr>
<td>**Pr(&gt;</td>
<td>z</td>
<td>)**</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Number of data used:** 3.802 5.468 3.753 3.105 5.883 6.068 6.035 5.100 5.041 2.062

**Pseudo R-squared**
- McFadden: 0.348 0.427 0.241 0.659 0.270 0.381 0.133 0.203 0.134 0.438
- Nagelkerke: 0.391 0.456 0.261 0.747 0.311 0.410 0.143 0.229 0.143 0.445

Table 4.2 shows that all types of commodities have positive intercepts, implying that everything else being equal, truck is more likely to be chosen. The coefficients of “value per ton” are also positive in all commodity types, indicating that truck tends to ship higher value goods than rail. The coefficients of “average truck distance” are mostly negative except commodity type 6 (i.e., logs, wood products, and textile and leather), which implies that as truck shipping distance increases the utility of truck decreases. In case of the coefficients of “average rail distance”, type 1, type 3, type 4, type 6, and type 10 commodities have a negative sign. Considering equations (4.2) and (4.5), this can be interpreted as positive coefficients of “average
“rail distance” variable ($c_{2n}$) in rail utility equation. Hence, as the distance between origin and destination increases, shippers prefer rail service for these types of commodities. On the other hand, commodities type 2, type 5, type 7, type 8, and type 9 have positive coefficients of “average rail distance”, which reflects decreasing of rail utility for shippers as distance between origin and destination increases. However, for those cases decreasing rate of truck utility caused by unit increase of “average truck distance” is larger than rail utility drop rate caused by unit increase of “average rail distance”. Note that for commodity type 6, rail utility increasing rate is larger than truck utility increasing rate for the unit increment of rail and truck average distances. Thus, rail service still has an advantage for long distance freight shipment. Finally, type 2, type 3, type 4, type 5, type 6, and type 7 commodities have negative coefficients of “WTI crude oil price”, which means as oil price increases shippers prefer rail than truck service. However, the other four types of commodities have positive oil price coefficients. For these types of commodities, the trend of using rail freight service might be decreasing regardless of oil price increase. For example, commodity type 1 includes time sensitive products (e.g., live animals and fish, fresh or chilled vegetable, fruit, meat, seafood, and animal origin product). Since railroads do not provide fast and flexible delivery service, it might be the reason for rail to lose the freight shipping share gradually in this market. Also, commodity type 8, type 9, and type 10 comprise only about 3.6% in the total rail shipment during recent 15 years as shown in Table 4.1. It might be the reason that the railroad service does not concentrate on the market of these commodity types and loses the freight delivery share. To test if any of the estimated intercept and coefficients is statistically different from zero, $z$-statistics and their $p$-values are also included in row (a). The absolute values of the $z$-statistics are very large for all cases, and all $p$-values are

52
less than 0.001. Hence, we can conclude that all estimates of the model coefficients are statistically significant for all types of commodities.

Table 4.2 also shows two pseudo R-square measures (i.e., indicating the goodness of fit of the logit model) in row (c). McFadden and Nagelkerke pseudo R-squares (Menard, 2001; Faraway, 2006), the most commonly used tests, generate very close measures. In case of commodities type 1, type 2, type 4, type 6, and type 10, the current models show more than 30% higher maximal likelihood over the intercept-only models.

4.4.2 Model implementation and validation

To test how accurately the estimated model predicts the reality, model validation is conducted using reserved data in the original database (i.e., “test dataset”). The data structure of the “test dataset” is exactly the same as that of the “training dataset”, thus we can obtain truck and rail shipment share predictions for each commodity type and O/D pair, and then compare them with the observed shares.

Column (a) of Table 4.3 represents size of the “test dataset” for each commodity type that is used to validate the proposed model. They include approximately one third of the total observations. Columns (b) and (c) of Table 4.3 respectively show the observed and the predicted total truck shipment shares for each commodity type, summed across all O/D pairs in the “test dataset”. It can be seen from column (d) that the estimated model generally yields very close predictions of the total modal shares, probably benefiting from the law of large numbers. In case of commodity type 5 (i.e., basic chemicals, chemical and pharmaceutical products), the prediction error is relatively larger but the estimated model is still acceptable.
Table 4.3 Comparison between observed and predicted total truck shares and paired comparison

<table>
<thead>
<tr>
<th>(a) Number of data used</th>
<th>Total truck shipment share (%)</th>
<th>Paired comparison (α=0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(b) Observed</td>
<td>(c) Predicted</td>
</tr>
<tr>
<td>Type 1</td>
<td>1,901</td>
<td>89.217%</td>
</tr>
<tr>
<td>Type 2</td>
<td>2,734</td>
<td>93.581%</td>
</tr>
<tr>
<td>Type 3</td>
<td>1,877</td>
<td>93.327%</td>
</tr>
<tr>
<td>Type 4</td>
<td>1,552</td>
<td>59.709%</td>
</tr>
<tr>
<td>Type 5</td>
<td>2,941</td>
<td>69.237%</td>
</tr>
<tr>
<td>Type 6</td>
<td>3,034</td>
<td>93.019%</td>
</tr>
<tr>
<td>Type 7</td>
<td>3,017</td>
<td>95.513%</td>
</tr>
<tr>
<td>Type 8</td>
<td>2,550</td>
<td>93.503%</td>
</tr>
<tr>
<td>Type 9</td>
<td>2,520</td>
<td>94.223%</td>
</tr>
<tr>
<td>Type 10</td>
<td>1,031</td>
<td>99.259%</td>
</tr>
</tbody>
</table>

To verify if there is any statistical difference between the groups of observed and predicted truck shipment shares, pairwise t-tests are conducted in columns (e) and (f) of Table 4.3. In this analysis, comparison of the matched O/D pairs for each commodity type is adopted to improve precision and reduce variability. The sample sizes of the tests are shown in column (a) of Table 4.3. The null hypothesis ( \( H₀ \) ) is that there is no obvious difference between truck shipment predictions and observations, and the alternative hypothesis ( \( H₁ \) ) is that there is significant difference between them. Column (f) of Table 4.3 shows that we do not reject the null hypothesis at the 0.05 level of significance for all types of commodities. Thus, we can conclude that there is no significant difference between the predicted truck shipment shares obtained from our proposed models and the observed truck shipment shares for all commodity types.

Figure 4.5 visually illustrates how the observed truck shipment shares ( \( x \)-axis) are consistent with the model predictions ( \( y \)-axis) for each data record in the “test dataset”. Each dot in Figure 4.5 corresponds to an observation in the “test dataset”.

54
Figure 4.5 Observed and predicted truck shipment shares for all commodities
Figure 4.5 (continued)
Note that Figure 4.5 plots the logarithm of (truck shipment tonnages + 1) because we have data records whose observed truck share is zero. The dotted lines located on both sides of the $45^\circ$ solid line represent one standard deviation from the mean. All figures show that the predictions from the models are generally well matched with the observed values, although those for commodities type 3 and type 4 in Figure 4.5 show more outliers.

4.4.3 Emission estimation

The estimated modal split model can be used to estimate the effects of oil price change on freight shipment mode choices and the environmental impacts. WTI crude oil price is frequently more than $100 per barrel nowadays and freight shipment market for the railroad mode has been expanding rapidly. A freight train is reported to move a ton of freight 436 miles on one gallon of fuel, which is three or more times as fuel-efficient as most trucks (AAR, 2008). Moreover,
freight trucks have been pointed out as a dominant source of freight transportation emissions in many studies (ICF Consulting, 2005), and it is well known that rail produces less air pollutants and greenhouse gases than trucks in terms of ton-mile unit (Bryan et al., 2007). Thus, improvement of air quality can be expected by shifting freight transportation demand from truck to rail which is induced by high oil price.

To illustrate this, we pick an arbitrary data record in the “test dataset”, which describes the freight movement of commodity type 5 (i.e., basic chemicals, chemical and pharmaceutical products) from Texas to Colorado. In this record, the freight value per unit weight is $1,240.85/ton, and the average shipping distances for truck and rail are 1,005 and 1,332 miles respectively. Given all the information for the explanatory variables above as well as crude oil price in concern, we can forecast the annual freight shipment share ratios for both modes. Then, by applying appropriate emission factors that relate the amount of emission production with freight transportation activity for each mode, we can estimate the total air pollutant and greenhouse gas inventory. In the following analysis, we adopt CO$_2$, CH$_4$, and N$_2$O emission rates from the U.S. EPA (2008) and PM$_{10}$ emission rate from NRDC (2012), such that per ton-mile of truck and rail shipments generate 0.2970 and 0.0252 kg of CO$_2$, 0.0035 and 0.0020 g of CH$_4$, 0.0027 and 0.0006 g of N$_2$O, and 0.092 and 0.013 g of PM$_{10}$ respectively. Since the total annual freight shipment demand is given as 328,000 ton in the data, we can obtain truck and rail demand split prediction and the following various emission estimations for a range of oil prices; see Table 4.4.
Table 4.4 Modal split and emission estimations under different WTI crude oil prices

<table>
<thead>
<tr>
<th>WTI crude oil price ($/barrel)</th>
<th>Truck share prediction (%)</th>
<th>Rail share prediction (%)</th>
<th>Truck CO₂ emission (ton)</th>
<th>Rail CO₂ emission (ton)</th>
<th>Total CO₂ emission (ton)</th>
<th>Truck CH₄ emission (kg)</th>
<th>Rail CH₄ emission (kg)</th>
<th>Total CH₄ emission (kg)</th>
<th>Truck N₂O emission (kg)</th>
<th>Rail N₂O emission (kg)</th>
<th>Total N₂O emission (kg)</th>
<th>Truck PM₁₀ emission (kg)</th>
<th>Rail PM₁₀ emission (kg)</th>
<th>Total PM₁₀ emission (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>66.8%</td>
<td>33.2%</td>
<td>65,412</td>
<td>3,654</td>
<td>69,066</td>
<td>290</td>
<td>1,061</td>
<td>595</td>
<td>87</td>
<td>682</td>
<td>20,262</td>
<td>1,885</td>
<td>22,147</td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>63.5%</td>
<td>36.5%</td>
<td>62,163</td>
<td>4,019</td>
<td>66,182</td>
<td>319</td>
<td>1,052</td>
<td>565</td>
<td>96</td>
<td>661</td>
<td>19,256</td>
<td>2,073</td>
<td>21,329</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>60.0%</td>
<td>40.0%</td>
<td>58,784</td>
<td>4,399</td>
<td>63,183</td>
<td>349</td>
<td>1,042</td>
<td>534</td>
<td>105</td>
<td>639</td>
<td>18,209</td>
<td>2,269</td>
<td>20,479</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>56.5%</td>
<td>43.5%</td>
<td>55,304</td>
<td>4,791</td>
<td>60,094</td>
<td>380</td>
<td>1,032</td>
<td>503</td>
<td>114</td>
<td>617</td>
<td>17,131</td>
<td>2,471</td>
<td>19,602</td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>52.9%</td>
<td>47.1%</td>
<td>51,758</td>
<td>5,189</td>
<td>56,947</td>
<td>412</td>
<td>1,022</td>
<td>471</td>
<td>124</td>
<td>594</td>
<td>16,033</td>
<td>2,677</td>
<td>18,710</td>
<td></td>
</tr>
<tr>
<td>140</td>
<td>49.2%</td>
<td>50.8%</td>
<td>48,181</td>
<td>5,591</td>
<td>53,773</td>
<td>444</td>
<td>1,012</td>
<td>438</td>
<td>133</td>
<td>571</td>
<td>14,925</td>
<td>2,885</td>
<td>17,809</td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>45.6%</td>
<td>54.4%</td>
<td>44,613</td>
<td>5,993</td>
<td>50,606</td>
<td>476</td>
<td>1,001</td>
<td>406</td>
<td>143</td>
<td>548</td>
<td>13,820</td>
<td>3,091</td>
<td>16,911</td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>42.0%</td>
<td>58.0%</td>
<td>41,091</td>
<td>6,389</td>
<td>47,480</td>
<td>507</td>
<td>991</td>
<td>374</td>
<td>152</td>
<td>526</td>
<td>12,729</td>
<td>3,296</td>
<td>16,024</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>38.5%</td>
<td>61.5%</td>
<td>37,651</td>
<td>6,776</td>
<td>44,426</td>
<td>538</td>
<td>981</td>
<td>342</td>
<td>161</td>
<td>504</td>
<td>11,663</td>
<td>3,495</td>
<td>15,158</td>
<td></td>
</tr>
<tr>
<td>220</td>
<td>35.1%</td>
<td>64.9%</td>
<td>34,324</td>
<td>7,150</td>
<td>41,474</td>
<td>567</td>
<td>972</td>
<td>312</td>
<td>170</td>
<td>482</td>
<td>10,632</td>
<td>3,688</td>
<td>14,321</td>
<td></td>
</tr>
<tr>
<td>240</td>
<td>31.8%</td>
<td>68.2%</td>
<td>31,140</td>
<td>7,508</td>
<td>38,648</td>
<td>596</td>
<td>963</td>
<td>283</td>
<td>179</td>
<td>462</td>
<td>9,464</td>
<td>3,873</td>
<td>13,319</td>
<td></td>
</tr>
<tr>
<td>260</td>
<td>28.7%</td>
<td>71.3%</td>
<td>28,120</td>
<td>7,848</td>
<td>35,968</td>
<td>623</td>
<td>954</td>
<td>256</td>
<td>187</td>
<td>442</td>
<td>8,711</td>
<td>4,048</td>
<td>12,759</td>
<td></td>
</tr>
<tr>
<td>280</td>
<td>25.8%</td>
<td>74.2%</td>
<td>25,282</td>
<td>8,167</td>
<td>33,449</td>
<td>648</td>
<td>946</td>
<td>230</td>
<td>194</td>
<td>424</td>
<td>7,832</td>
<td>4,213</td>
<td>12,044</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>23.1%</td>
<td>76.9%</td>
<td>22,638</td>
<td>8,464</td>
<td>31,102</td>
<td>672</td>
<td>939</td>
<td>206</td>
<td>202</td>
<td>407</td>
<td>7,012</td>
<td>4,366</td>
<td>11,379</td>
<td></td>
</tr>
</tbody>
</table>
From columns (a)-(o), we can conclude that as the WTI crude oil price varies from $40 to $300 per barrel, the total truck shipment share and the truck emissions considered in this analysis decrease. Also, the total CO$_2$, CH$_4$, N$_2$O, and PM$_{10}$ emissions from both modes decrease, despite the increase in rail freight share and the following emissions. The national emission estimation can be further estimated by aggregating such emission calculations across all O/D pairs and all commodity types.

4.5 Model Application for Future Prediction

The proposed binomial logit market share model has been applied to forecast future truck and rail freight demand from 2010 to 2050 in five-year increments using data obtained from the previous trip generation and trip distribution steps. Future commodity value and crude oil price are obtained from the global economic forecast models (Edmonds et al., 1995; Edmonds et al., 2004; Vanek and Morlock, 2007). Table 4.5 summarizes the results. Column (a) of Table 4.5 represents future study years. Each of columns (b)-(e) includes truck, rail, and the total freight demand across all commodities at the national level for each global economic growth and environmental policy scenario (proposed by the global economic forecast models in Chapter 3).

As columns (b)-(e) show, the total freight shipment demand for truck and rail as well as their sum continuously grow over time. In scenarios 1 and 3, when the climate policy is not implemented and carbon tax is not included in the final market prices of fossil fuel commodities, the percentage freight shipment shares for truck and rail are almost constant over time. On the other hand, in scenarios 2 and 4 there is a large modal shift to the railroad mode in the future years because the climate policy is implemented and the final market prices of fossil fuel commodities (e.g., crude oil price) are driven up by the carbon tax. Note that the national freight
demand for commodity type 4 (i.e., coal and petroleum products) in scenarios 2 and 4 are lower than those in scenarios 1 and 3 respectively due to implementation of the climate policy. Since commodity type 4 occupies the dominant share in the total rail freight shipment demand, this trend partly canceled out the effect of high crude oil price which has led to the modal shift to the railroad mode in scenarios 2 and 4.

### Table 4.5 Future modal split prediction

<table>
<thead>
<tr>
<th>(a) Year</th>
<th>(b) Scenario 1</th>
<th>(c) Scenario 2</th>
<th>(d) Scenario 3</th>
<th>(e) Scenario 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Truck ton (10^6)</td>
<td>Rail ton (10^6)</td>
<td>Total ton (10^6)</td>
<td>Truck share</td>
</tr>
<tr>
<td>2010</td>
<td>13,619.34</td>
<td>2,084.45</td>
<td>15,703.79</td>
<td>86.7%</td>
</tr>
<tr>
<td>2015</td>
<td>15,175.30</td>
<td>2,326.69</td>
<td>17,501.99</td>
<td>86.7%</td>
</tr>
<tr>
<td>2020</td>
<td>16,839.82</td>
<td>2,591.49</td>
<td>19,431.31</td>
<td>86.7%</td>
</tr>
<tr>
<td>2025</td>
<td>18,549.70</td>
<td>2,888.40</td>
<td>21,438.10</td>
<td>86.5%</td>
</tr>
<tr>
<td>2030</td>
<td>20,511.49</td>
<td>3,182.46</td>
<td>23,693.95</td>
<td>86.2%</td>
</tr>
<tr>
<td>2035</td>
<td>22,502.56</td>
<td>3,531.72</td>
<td>26,034.29</td>
<td>86.4%</td>
</tr>
<tr>
<td>2040</td>
<td>24,796.75</td>
<td>3,901.18</td>
<td>28,697.93</td>
<td>86.4%</td>
</tr>
<tr>
<td>2045</td>
<td>29,844.83</td>
<td>4,828.84</td>
<td>34,673.66</td>
<td>86.1%</td>
</tr>
<tr>
<td>2050</td>
<td>30,104.84</td>
<td>5,309.30</td>
<td>35,414.14</td>
<td>84.7%</td>
</tr>
</tbody>
</table>
4.6 Conclusion

Demand for freight transportation has been persistently increasing for several decades as a result of economic growth and globalization. However, at the same time, emissions from different freight transportation modes have contributed to a large share of air pollution and caused significant concerns over air quality and public health. Meanwhile, energy shortage and oil price surge during the past decade affected freight transportation systems significantly. Hence, the motivation of this research is to draw connections among freight transportation demand mode choice, various economic factors (e.g., oil price), and the air quality and climate impacts.

In this study, a macroscopic binomial logit market share model is proposed to study freight transportation modal split between the dominating truck and rail modes. In our model, the mode choice decision between truck and rail, for each of 10 commodity types, is assumed to be a function of not only freight and shipment characteristics (such as freight value and average shipping distance), but also crude oil price. Four years of data on freight transportation activities and characteristics are obtained from the FAF\textsuperscript{2} and FAF\textsuperscript{3} datasets and two years of CFS data to support the model development. Model validation results show that the developed models are effective in predicting the freight modal shares. Generally, it was shown that trucks tend to be chosen to handle higher value products, for shorter distance shipments. Also, probably due to the oil price surge, the freight tonnage increase for railroads is much larger than that for trucks during the study period from 1993 to 2007. Interpretations of the intercept and coefficient estimations are used to draw insights on the effects of oil price change on freight transportation mode choice decisions and their environmental impacts.
CHAPTER 5
ASSIGNMENT OF FREIGHT SHIPMENT DEMAND IN CONGESTED TRANSPORTATION NETWORKS

5.1 Introduction
The freight flow in the U.S. are grouped into two systems including inter-regional and intra-regional freight shipment, and the four-step freight demand forecasting framework is adopted for the inter-regional freight flow analysis. In the previous chapter, we have developed a macroscopic binomial logit market share model for the freight shipment mode choice decisions between truck and rail assuming freight demand and supply information are given for each geographic zone and commodity type. Using this procedure, we can forecast truck and rail annual freight shipment shares for each typical commodity based on utility of the mode between all FAZ O/D pairs. Based on the results from the modal split analysis, the next step is to implement traffic assignment represented by the grey box in Figure 5.1.

Traffic assignment step develops a systematic methodology which determines the most economical freight shipment routes for all FAZ O/D pairs on each truck and rail network. It is

---

2 This chapter has been adapted in part from “Hwang, T. and Ouyang, Y. (2014a). Assignment of freight shipment demand in congested rail networks. Presented at the 93nd Annual Meeting of the Transportation Research Board, January 2014, Washington, D.C., and accepted for publication in the 2014 series of the Transportation Research Record: Journal of the Transportation Research Board (forthcoming).” The material from this paper is reproduced with written permission of the Transportation Research Board.
assumed that the freight shipments cause congestion in each modal network and in turn increase transportation cost, and carriers or shippers will react to such changes in transportation cost by choosing alternative shipment routes. Multiple shipments from various O/D pairs share the same infrastructure (i.e., highways or rail tracks), while the freight flow are loaded onto the each modal network to satisfy certain carrier/shipper objective, e.g., minimizing travel time of each truck or train. At convergence, freight flow concentration and congestion pattern in the network can be found.

There can be two different rules for the route selection (Sheffi, 1985) as studied in Chapter 2, i.e., user equilibrium and system optimum. In user equilibrium, each carrier/shipper is assumed to have full information on network congestion (i.e., travel time for all possible routes) through repeated business practice, and it selects the shipment route with the shortest travel time (which is affected by the network congestion pattern caused collectively by all O/D shipments). Eventually, when equilibrium is reached, all used routes for a given O/D pair have the same travel time which is less than or equal to the travel time of any unused routes. At this point, no carriers/shippers can reduce their travel time by unilaterally choosing an alternative route. In system optimum principle, total system-wide travel time spent by all carriers/shippers in the network is minimized. Since in the system optimum state carriers/shippers might be able to reduce their total travel time by unilaterally changing their routes, the flow pattern in the system optimum is generally unstable, which implies it might not be a proper route choice rule for modeling the actual drivers’ behavior. In this study, we assume the user equilibrium principle is suitable for carrier/shipper’s choice of routes in both truck and rail modes. Shipment plans will eventually stabilize into a so-called network equilibrium condition as shipment flow from every O/D pair finds no incentive to unilaterally change route.
While the general principles of network equilibrium and traffic flow route choice hold for both highway and rail networks, we should note a critical operational difference in the U.S. system. In a highway network system, two nodes are often connected by two separate directed links in opposite directions, and hence the network is often represented by a directed graph. In a rail network, two nodes should often be connected by one undirected link (e.g., single tracks) since the same infrastructure may be shared by traffic flow in both directions. Although there are also double-track sections in the network where each track could be dedicated to one directional flow, in practice they are often operated as two main tracks each of which still serves bidirectional flow (Bisset et al., 2008). Moreover, the use of single track is dominating in the U.S. and only limited network parts with heaviest traffic volume are operated as double tracks (Cambridge Systematics, Inc., 2007). Furthermore, it will be difficult to use time tables to capture the travel time in the U.S. freight rail system because substantial deviations from even daily schedule exist, and thus terminal managers frequently modify original train operation plan on an ongoing basis (Sussman, 2000). Instead, the railroad link travel time must be based on both network conditions (e.g., daily traffic) and capacity constraints. Some well-known link travel time formulas for the U.S. freight railroads were first developed by Krueger (1999) and later updated by Lai and Barkan (2009) and Mitra et al. (2010). These empirical formulas generally apply to undirected railroad links a form that is quite different from its highway counterpart from the Bureau of Public Roads (Sheffi, 1985). Unfortunately, previous studies have not addressed carefully or considered the abovementioned features of the rail network.

To bridge these gaps, this dissertation work incorporates the unique features of rail network operations (e.g., bi-directional traffic on shared single tracks) in a modified network assignment model and an adjusted convex combinations algorithm. This algorithm defines link
travel delay based on traffic flow in both directions on an undirected rail network graph. In so
doing, we modify the train delay equation from Lai and Barkan (2009) to develop our railroad-
specific link cost function; this equation captures a basic principle that the link travel time in
both directions increases as the traffic flow in two opposite directions increases. In case of truck
freight shipment demand assignment, background traffic volume is considered on each highway
link to represent traffic flow which already exists in a given U.S. highway network. The
conventional network assignment model and solution approach, convex combinations algorithm
(Frank and Wolfe, 1956), are applied to achieve the truck shipment routing equilibrium. The
proposed models and algorithms are applied to the full-scale U.S. highway and rail networks to
predict freight flow and congestion pattern. Using empirical freight shipment O/D data in Year
2007 and a simplified representation of the entire U.S. highway and rail networks, our models
find the optimal shipment flow patterns within a short computation time. The model output turns
out to match very well with the empirical congestion pattern observed in the real world. As such,
the modeling framework presented in this dissertation work can be used to forecast future freight
flow and congestion pattern in both highway and rail networks. Since not only the mode choices
but also the route choices in freight deliveries can significantly affect regional and urban air
quality and eventually public health, this effort will be useful to estimate the environmental
impacts from the freight transportation systems precisely for various future freight shipment
demand.

The exposition of this chapter is as follows: Section 5.2 describes the proposed network
assignment models and solution algorithms. Section 5.3 presents empirical case studies for the
U.S. highway and rail networks. Section 5.4 compares total and average freight shipment cost
between truck and rail, and Section 5.5 presents model validation. Section 5.6 investigates
possible rail network capacity expansion scenario and its effect on network assignment. Future truck and rail freight demand assignment prediction and related emission estimations are shown in Section 5.7. Section 5.8 concludes the chapter.

5.2 Truck Freight Shipment Demand Network Assignment

This section describes a model and a solution algorithm to assign truck freight shipment demand for each FAZ O/D pair on the full-scale U.S. interstate highway network. Given the traffic assignment program, an input that contains network geometry and freight shipment demand is constructed. The program generates an output that includes freight flow pattern between freight shipment regions and estimated congestion in the network.

5.2.1 Model formulation

The standard network assignment problem of truck freight shipment demand under user equilibrium principle can be formulated as follows. We primarily follow the notation in Sheffi (1985). Suppose that the graph representation of roadway network, \( D(V, A) \), is given where \( V \) is a node set and \( A \) is a directed link set. Each highway link travel time (hours) is assumed to follow the Bureau of Public Roads (BPR) link cost function (Sheffi, 1985) modified to include the concept of background traffic volume to represent the traffic flow which already exists in a link (e.g., AADT) as follows:

\[
t_a(\omega_a) = t^f_a \left[ 1 + \alpha_a \left( \frac{\omega_a + b_a}{C_a} \right)^{\beta_a} \right], \forall a \in A,
\]

(5.1)
where \( t^f_a \) is link free flow travel time (hours), \( \omega_a \) and \( b_a \) respectively denote assigned and background traffic volume (# of vehicles/hour) on the link, \( C_a \) is highway link capacity (# of vehicles/hour), and \( \alpha_a \) and \( \beta_a \) are parameters each of which is 0.15 and 4 respectively.

Let \( R \subseteq V \) be a trip origin set and \( S \subseteq V \) be a trip destination set. For each highway link \( a \in A \), we define decision variable \( x_a \) to represent assigned traffic flow on the link \( a \). Let \( K^{r,s} \) denote a set of possible routes that connect the origin node \( r \in R \) and the destination node \( s \in S \), \( f^{r,s}_k \) be the traffic flow on any possible route \( k \in K^{r,s} \), and \( q_{r,s} \) be the traffic demand from the origin \( r \in R \) to the destination \( s \in S \). Then, assigned flow on any link \( a \in A \) can be represented as 

\[
  x_a = \sum_{r \in R} \sum_{s \in S} \sum_{k \in K^{r,s}} f^{r,s}_k \delta^{r,s}_{a,k} \text{ where } \delta^{r,s}_{a,k} = 1 \text{ if link } a \text{ is included in the route } k \in K^{r,s}, \text{ or } \delta^{r,s}_{a,k} = 0 \text{ otherwise.}
\]

Using the parameters and decision variables described above, the network assignment model for truck freight shipment demand subject to the user equilibrium can be formulated as follows:

Minimize \[
\sum_{a \in A} \int_0^{\omega_a} t_a(\omega) \, d\omega,
\]

subject to
\[
\sum_{k \in K^{r,s}} f^{r,s}_k = q_{r,s}, \quad \forall r \in R, s \in S,
\]

\[
x_a = \sum_{r \in R} \sum_{s \in S} \sum_{k \in K^{r,s}} f^{r,s}_k \delta^{r,s}_{a,k}, \quad \forall a \in A,
\]

\[
f^{r,s}_k \geq 0, \quad \forall k \in K^{r,s}, r \in R, s \in S.
\]

Objective function (5.2) minimizes sum of the link cost functions that are integrated over the link flow from zero to decision variable \( x_a \). Constraints (5.3) ensure a flow conservation principle such that sum of traffic flows on all possible routes that connect each O/D pair should be the
same as traffic demand between that O/D pair. Constraints (5.4) define the traffic flow on a link to be sum of the flows of all possible routes that the given link belongs to. Finally, constraints (5.5) enforce the flow on each possible route to be nonnegative.

5.2.2 Solution algorithm

The convex combinations algorithm (Frank and Wolfe, 1956) has been widely used to solve the standard traffic assignment problem (5.2)-(5.5). The detailed step by step procedure in Sheffi (1985) is reviewed as follows:

Step 0: Initialization. Assign each O/D freight shipment demand to the shortest travel time route based on “free flow” link cost $t^0_a = t_a (0), \forall a \in A$. Set the iteration counter $n = 1$ and the obtained flow pattern $x_n^o = x^1, \forall a \in A$.

Step 1: Update the link travel time using the current flow pattern $t^n_a \leftarrow t_a (x^n_a), \forall a \in A$.

Step 2: Assign all O/D freight shipment demand to the new shortest travel time routes based on $t^n_a$ obtained in Step 1. Let the obtained auxiliary flow pattern be $y^n_a, \forall a \in A$.

Step 3: Obtain a step size $\lambda_n$ by solving $\min_{0 \leq \lambda_n \leq 1} \sum_{a \in A} \int_0^{x^n_a + \lambda_n (y^n_a - x^n_a)} t_a (\omega) d\omega$.

Step 4: Update the link flow by convex combinations: $x^{n+1}_a = x^n_a + \lambda_n (y^n_a - x^n_a), \forall a \in A$.

Step 5: If a certain convergence criterion is satisfied (e.g., percentage change of the objective value is less than a tolerance $\varepsilon \ll 1$), terminate the algorithm. The current flow pattern, $x^{n+1}_a, \forall a \in A$ is the optimal solution; otherwise, let $n \leftarrow n + 1$ and go to Step 1.
Following this algorithm, freight shipment demand for all O/D pairs can be assigned onto the U.S. highway network. Eventually, the travel time on all used routes connecting each O/D pair will equal each other, and it will be less than or equal to the travel time on unused routes.

5.2.3 Case study: Data preparation

Input data need a graph representation of the freight truck road network, which includes link and node number, link capacity, background traffic volume, link distance, and free flow travel time. In this study, data for origin and destination nodes in a network are obtained from the centroids of FAF³ regions (‘FAF3-Zone’ file). We excluded two zones in Hawaii and one zone in Alaska, thus, a total of 120 domestic FAF³ zones are obtained as both origins and destinations and 14,400 O/D pairs are generated. To construct the U.S. freight truck road network, database on the FAF³ road geometry (‘FAF311_NET’ file) is used which also contains complete information on the annual average daily traffic (AADT) (i.e., background traffic flow) and the roadway capacity in Year 2007 in a separate sub-database (‘FAF output’ file). Note that a set of database related to the Freight Analysis Framework version 3 (e.g., FAF3-Zone, FAF311_NET, FAF output, and faf3data) was obtained from FHWA U.S. DOT (2011). There was a potential challenge in preparing input data due to the huge size of the full freight truck road network (e.g., containing more than 170,000 links). Since we are analyzing truck freight movements and route choices in a macroscopic point of view, only major interstate corridors are considered so as to keep the network simple and tractable. This assumption is reasonable since long-distance inter-regional freight truck deliveries will be conducted mostly using major interstate highways.
Figure 5.2(a) describes the entire FAF$^3$ freight truck road network (FHWA U.S. DOT, 2011) and the selected major interstate highways are represented by black thick lines. Centroids of 120 domestic FAF$^3$ regions (i.e., both origins and destinations of truck freight shipments) are shown as blue dots. Figure 5.2(b) shows the simplified freight truck road network. The centroids close to the interstate highways or the junctions of the different interstate highways are aggregated to the selected major highway network. For FAF$^3$ centroids located far from the major interstate highway network, some local roads in Figure 5.2(a) are also included to connect those nodes to the network. The resulting network in Figure 5.2(b) contains 178 nodes (i.e., 120 centroids of the domestic FAZs and 58 major junctions) and 588 links in total.

Figure 5.2 Truck transportation network for the Continental U.S.
To obtain background traffic volume and link capacity for each link, network attribute table from the FAF³ road geometry database and ‘FAF output’ file are merged in ArcGIS platform. The background traffic is calculated from daily traffic information obtained from AADT07 (annual average daily traffic in Year 2007, derived from HPMS 2008 database, volume/day/route) in the database divided by 48 since we consider two directions of traffic and 24-hour operation per day. CAP07 (link capacity estimation using the procedures in HCM 2000 and the arc geometry provided in 2008 HPMS database, volume/hour/route) in the database provides link capacity. Link distance is directly measured in ArcGIS.

Input data also contain freight shipment demand for all FAZ O/D pairs. In this chapter, we utilize FAF³ truck shipment database (‘faf3data’ file), a real freight shipment demand in Year 2007, which was obtained from FHWA U.S. DOT (2011). To convert the FAF³ freight shipment demand in tonnages into equivalent numbers of trucks that need to be assigned onto the network, we assume both class 7 and class 8 combination trucks are used for the U.S. inter-regional freight deliveries. Using information from FHWA U.S. DOT (2007a) and EPA and NHTSA (2011), average payload is estimated as 16 tons per truck. Passenger car equivalents are assumed to be 2.5 based on rolling terrain (HCM, 2000), and freight truck delivery system is assumed to operate 365 days per year and 24 hours per day. Free flow speed is set to be 65 mph (Bai et al., 2011).

5.2.4 Case study: Computation result

The solution algorithm is coded in Visual C++ and run on a personal computer with 3.40 GHz CPU and 8 GB memory. Total cost is defined as follows to represent the total vehicle-hours that all freight shipments spend in the network each hour:
Total Cost = \sum (\text{Assigned Link Flow} \times \text{Link Travel Time}) = \sum_{a \in A} x_a t_a (x_a) \cdot \tag{5.6}

Convergence is reached within a tolerance of 0.0001% after 12 iterations and 0.640 sec CPU time. The total cost based on user equilibrium principle is obtained as 699,827.88 (veh-hr/hour). The output contains link and node number, link distance, the total and assigned traffic volume, link travel time, and average link speed at equilibrium. For comparison, each O/D freight shipment demand is assigned only on the shortest-distance path (i.e., all or nothing assignment) ignoring congestion, and the total cost is obtained as 715,407.31 (veh-hr/hour). Thus, drivers can reduce the total cost by 2.18% if user equilibrium is implemented.

Figure 5.3 User equilibrium result of truck freight network assignment

Figure 5.3 illustrates the user equilibrium result of truck freight demand network assignment. Sum of assigned traffic volume on two links which connect the same pair of nodes in opposite directions are classified by various line thicknesses and colors as shown in the legend. We can observe a large amount of assigned traffic and possibly heavy congestion on some of the
highway links in Washington, Montana, California, Nevada, Kansas, Texas, Florida, the Midwest states near Chicago, and northeastern areas of the U.S.

5.2.5 Truck freight network assignment examples

This section presents the result from truck freight shipment network assignment in detail to show how the routing equilibrium is reached. The purple and blue boxes in Figure 5.4 include two different O/D pair examples.

![Figure 5.4 Two detailed examples from the truck freight network assignment](image)

Figure 5.5 shows the network geometry and assignment result for the first sample from the freight shipment data, from ‘Remainder of Pennsylvania’ to ‘Remainder of Maryland’ in database, which is an O/D pair with a relatively short shipment distance.
In Figure 5.5, the origin and the destination of the freight shipment demand are respectively represented by the square and the triangle. Other centroids of the FAZs are denoted by small blue dots, and the network links are shown as lines. The result from user equilibrium principle for this specific O/D pair is described by a set of red lines; the numbers in red near each red link represent the corresponding average vehicle speed and the link travel time. For comparison, we also consider an alternative scenario in which every O/D shipment is assigned to its shortest-distance path (i.e., as if the decision maker ignores congestion when determining shipment routes). The set of green lines denotes the shortest-distance path between this specific O/D pair; the numbers in black near each green link show the expected travel speed and link travel time under congestion if the shortest-distance paths are actually implemented.

According to the data, on average 218 vehicles per hour must go through the network from ‘Remainder of Pennsylvania’ to ‘Remainder of Maryland’. In the user equilibrium state, traffic will be split between two different routes and the total travel time associated with each route is almost the same; i.e., around 12.7 hours, as we would expect of the equilibrium. On the
other hand, it can be seen that if the shortest-distance path is implemented, only one route is used for this O/D pair and each vehicle will spend around 16.7 hours to reach the destination. Hence, the vehicles can each reduce an average travel time of 4.0 hours if they follow the user equilibrium routes (as the result of the proposed model).

Figure 5.6 shows the second freight shipment example, from ‘New York-Newark-Bridgeport, NY-NJ-CT-PA Combined Statistical Area (NY Part)’ to ‘New Orleans-Metairie-Bogalusa, LA Combined Statistical Area’ in data, which has a longer shipping distance than the first example.

![Figure 5.6 Detailed truck freight network assignment result from the second example](image)

We assign 142 vehicles per hour from the given origin to the given destination. Similar as before, the set of red links represents result from the user equilibrium principle, while the set of green lines describes the selected links to be loaded with vehicles when the shortest-distance path is applied. Average vehicle speed and link travel time associated with each link and route choice rule are not shown for simplicity. Note that there are long distance detours on the far north and
south of the possible route set under user equilibrium state to avoid congestion formed near the short-distance paths. All motorists need to spend 38.7 hours to reach his/her destination if the shortest-distance path is adopted, while the travel cost for each truck driver reduces to around 30.1 hours if the user equilibrium principle is implemented.

5.3 Rail Freight Shipment Demand Network Assignment

This section presents the model formulation and solution algorithm that assign freight shipment demand between given O/D pairs onto a rail network. We discuss the link cost function for the railroad tracks and the representation of the rail network. A modified convex combinations algorithm is proposed to solve the formulated problem. A case study is conducted using the macroscopic freight shipment data in the full-scale U.S. rail network, and the result is presented with visual illustrations.

5.3.1 Model formulation

As mentioned earlier, the rail network operates very differently from the typical highway network because link traffic flow in opposite directions share the same track infrastructure. We assume the railroad link travel time between two nodes follows the train delay versus traffic volume relationship proposed by Lai and Barkan (2009). This relationship is defined for undirected railroad links (e.g., single tracks). Intuitively, the rail network can be represented by an undirected graph $G(V,E)$, where $V$ is the set of nodes and $E$ is the set of undirected links. The traffic flow in the rail network is defined as the number of trains passing a certain point on a link per day regardless of directions. The link travel time (hours) on an undirected railroad link $e \in E$ generally increases as the traffic flow increases, as follows:
\[ t_e(\omega_e) = T_e + \frac{\alpha_e d_e}{100} e^{\beta_e \omega_e}, \forall e \in E, \]  

(5.7)

where \( T_e \) is link free flow travel time (hours), \( d_e \) is link length (miles), \( \omega_e \) is the total rail link flow (# of trains/day), and \( \alpha_e \) and \( \beta_e \) are parameters that are uniquely determined by railroad traffic and operating conditions, such as the mileage percentage of double tracks (Lai and Barkan, 2009). When railroad link \( e \) connects nodes \( i \) and \( j \) in the undirected graph, as shown in Figure 5.7(a), the total link flow \( \omega_e \) equals the sum of link flow in two opposite directions, \( x_{ij} + x_{ji} \). The travel time \( t_e(\omega_e) \) applies to traffic in both directions; as such, the travel time from node \( i \) to node \( j \) will be affected not only by the traffic flow in the same direction, \( x_{ij} \), but also by that in the opposite direction, \( x_{ji} \).

![Diagram of link flow in an undirected graph](image1)

![Diagram of link flow in a directed graph](image2)

Figure 5.7 Two ways to describe the railroad link flow

For convenience of network-level modeling, we propose an equivalent directed graph representation of the undirected rail network, where each undirected link is replaced by two separate directed links in opposite directions; see Figure 5.7(b). All link properties (e.g., length, percentage of double tracks) of the two new directed links are the same as those of the previous undirected link. We now denote the set of new directed railroad links by \( A \). The railroad link travel time function (5.7) becomes the following for the directed graph:
\[ t_{ij}(x_i + x_j) = T_{ij} + \frac{\alpha_{ij} d_{ij}}{100} e^{\beta_{ij}(x_i + x_j)}, \forall (i, j) \in A. \]  

(5.8)

Obviously, formula (5.8) computes the (directional) link cost based on the traffic volume in both directions, as shown in Figure 5.7(b). Parameters \( T_{ij}, \alpha_{ij}, \beta_{ij}, \) and \( d_{ij} \) take the values of \( T_e, \alpha_e, \beta_e, \) and \( d_e \) respectively if the original undirected link \( e \in E \) connecting nodes \( i \) and \( j \) is now replaced by two directed links \( (i, j) \in A \) and \( (j, i) \in A \). The link travel times on both directed links (from node \( i \) to node \( j \) and from node \( j \) to node \( i \)) are identical, so that the directed network representation remains equivalent to the original undirected graph.

Finally, network assignment of the rail freight shipment demand under user equilibrium can be formulated as follows. Following the notation in Sheffi (1985), we let \( R \subseteq V \) be a set of shipment origins, and \( S \subseteq V \) be a set of shipment destinations. For each directed link \( (i, j) \in A \), a decision variable \( x_{ij} \) is defined to represent the assigned freight flow on link \( (i, j) \). Let \( K^{r,s} \) be a set of possible routes that connect the origin node \( r \in R \) to the destination node \( s \in S \), \( f_k^{r,s} \) be the assigned freight flow on any possible route \( k \in K^{r,s} \), and \( q_{r,s} \) be the shipment demand from the origin \( r \in R \) to the destination \( s \in S \). Then, the assigned shipment flow on any link \( (i, j) \in A \) can be represented as \( x_{ij} = \sum_{r \in R} \sum_{s \in S} \sum_{k \in K^{r,s}} f_k^{r,s} \delta_{(i,j),k} \) where \( \delta_{(i,j),k} = 1 \) if link \( (i, j) \) is part of route \( k \in K^{r,s} \), or \( \delta_{(i,j),k} = 0 \) otherwise. Using the parameters and decision variables described above, the following standard mathematical program finds the equilibrium rail freight flow pattern in the directed graph that satisfies all O/D shipment demand:
Minimize \[ \sum_{(i,j) \in A} \int_0^{t_{ij}^0} t_{ij}(\omega) d\omega, \] (5.9)

subject to \[ \sum_{k \in K^{r,s}} f_{k}^{r,s} = d_{r,s}, \quad \forall r \in R, s \in S, \] (5.10)

\[ x_{ij} = \sum_{r \in R} \sum_{s \in S} \sum_{k \in K^{r,s}} f_{k}^{r,s} \delta_{(i,j),k}, \quad \forall (i, j) \in A, \] (5.11)

\[ f_{k}^{r,s} \geq 0, \quad \forall k \in K^{r,s}, r \in R, s \in S. \] (5.12)

Objective function (5.9) minimizes sum of the railroad link travel time functions that are integrated over the link flow. Note that the upper limit of the integral includes the traffic flow in the opposite direction as well. Constraints (5.10) ensure flow conservation, such that the sum of freight flow on all possible routes between each O/D pair should be the same as the corresponding shipment demand. Constraints (5.11) postulate that the assigned link freight flow equals the sum of all route flow passing that link. Finally, constraints (5.12) enforce all flow to be nonnegative.

5.3.2 Solution algorithm

To solve the proposed model (5.9)-(5.12), we modify the convex combinations algorithm in Sheffi (1985) by addressing the dependence of link cost function on bi-directional flow. The detailed procedure is summarized as follows:

Step 0: Initialization. Assign each O/D shipment demand to the route with the shortest travel time, which is computed based on “free flow” link cost, \( t_{ij}^0 = t_{ij}(0) \), \( \forall (i, j) \in A \). Set the iteration counter \( n = 1 \) and the obtained flow pattern \( x_{ij}^n = x_{ij}^1 \), \( \forall (i, j) \in A \).
Step 1: Update the link travel time using the current flow pattern, 
\[ t^*_y \leftarrow t_y \left( x^*_y + x^*_y \right), \]
\[ \forall (i, j) \in A. \]

Step 2: Assign all O/D shipment demand to the new shortest travel time routes, now computed based on \( t^*_y, \forall (i, j) \in A \). The obtained new flow pattern is \( y^n_i, \forall (i, j) \in A \).

Step 3: Obtain a step size \( \lambda_n \) by solving

\[
\text{Minimize} \sum_{(i,j)\in A} \int_0^{(x^n_y + x^n_y + \lambda_n (y^n_y - x^n_y))} t_y(\omega) \, d\omega.
\]

Step 4: Update the link flow by convex combinations:

\[ x^n_{ij} = x^n_{ij} + \lambda_n (y^n_{ij} - x^n_{ij}), \forall (i, j) \in A. \]

Step 5: If a certain convergence criterion is satisfied (e.g., percentage change of the objective value is less than a tolerance \( \epsilon \ll 1 \)), terminate the algorithm. The current flow pattern \( x^{n+1}_{ij}, \forall (i, j) \in A \) is the optimal solution; otherwise, let \( n \leftarrow n + 1 \) and go to Step 1.

Using this algorithm, freight shipment demand for all O/D pairs can be assigned onto the rail network. At convergence, the travel time on all used routes between an O/D pair will be equal and it will be less than or equal to the travel time on any unused routes.

5.3.3 Case study: Data preparation

This section describes a procedure for preparing input to the model. The input data contain rail network topology/geometry information as well as rail freight shipment demand for each O/D pair and commodity type.

The graph representation of the rail network includes link and node number, link distance, and free flow travel time. Centroids of 120 domestic FAZs (excluding Hawaii and Alaska) which have been generated from the ‘FAF3-Zone’ file are used to represent origins and destinations of freight shipments and hence there are 14,400 shipment O/D pairs in our dataset. Data on the full-
scale U.S. rail network (‘rail_lines’ file) are obtained from ATLAS (2011). Figure 5.8(a) illustrates the entire U.S. rail network (ATLAS, 2011) and the selected major tracks are represented by black thick lines. Centroids of the domestic FAZs are represented by blue dots (i.e., origins and destinations of freight shipments). The rail freight traffic assignment also has the same, network data size issue as the truck freight traffic assignment (e.g., more than 170,000 links). Unlike the truck network data, the rail network data do not include track hierarchy information. Since we are interested in analyzing the macroscopic freight movements and route choices, we simplify the rail network by selecting only the main lines on which Class I railroads (with code AMTK, BNSF, CSXT, KCS, NS, UP, CN, and CP in the database) operate. This assumption is reasonable since Class I railroads account for more than 90% of the total freight revenue in the U.S. rail shipment industry (AAR, 2011). The resulting network for analysis is plotted in Figure 5.8(b). The FAZ centroids located near the major railroad tracks or the junctions of different railroad tracks are aggregated into the network. Tracks on which other minor railroads operate are also included in the network if they are necessary to connect some of the FAZ centroids located far from the major rail network. The resulting rail freight shipment network includes 183 nodes (i.e., 120 centroids of the domestic FAZs and 63 major junctions) and 566 links in total.

In addition, information on the U.S. multiple track mainlines is obtained from Richards and Cobb (2010) in Figure 5.8(c) as part of the network geometry input data. Such information allows us to assign more realistic values to parameters \( \alpha_j \) and \( \beta_j \) in equation (5.8) based on the percentage of double track miles on each network link. Then, the result will be able to describe traffic congestion pattern in the rail network more accurately.
Rail freight shipment demand is obtained from the FAF³ shipment database (‘faf3data’ file) in FHWA U.S. DOT (2011). We use detailed information on national freight delivery activities between the FAZs in Year 2007. We also developed a methodology to convert the
FAF\textsuperscript{3} freight shipment demand in tonnages into equivalent numbers of trainloads (i.e., the actual traffic flow that needs to be assigned onto the network) based on the types of commodities. Information on (i) average number of cars by train service type and (ii) commodity assignment to different train service is obtained from Cambridge Systematics, Inc. (2007). AAR (2007) provides gross weight and net weight of train. Table 5.1 shows different average loading weight for each commodity type. We also assume that freight train free flow speed is 60 mph (Krueger, 1999) and railroads operate 365 days a year. Finally, a total of 40,909 freight shipment data records are obtained for Year 2007.

<table>
<thead>
<tr>
<th>Commodity type</th>
<th>Commodity description</th>
<th>\textsuperscript{a}Train type</th>
<th>Average loading weight (tons/train)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Agriculture products and fish</td>
<td>B/GM/IM</td>
<td>3,357</td>
</tr>
<tr>
<td>2</td>
<td>Grain, alcohol and tobacco products</td>
<td>GM/IM</td>
<td>3,194</td>
</tr>
<tr>
<td>3</td>
<td>Stones, nonmetallic minerals, and metallic ores</td>
<td>B/GM</td>
<td>3,369</td>
</tr>
<tr>
<td>4</td>
<td>Coal and petroleum products</td>
<td>B/GM</td>
<td>3,405</td>
</tr>
<tr>
<td>5</td>
<td>Basic chemicals, chemical and pharmaceutical products</td>
<td>GM/IM</td>
<td>2,874</td>
</tr>
<tr>
<td>6</td>
<td>Logs, wood products, and textile and leather</td>
<td>GM/IM</td>
<td>2,815</td>
</tr>
<tr>
<td>7</td>
<td>Base metal and machinery</td>
<td>B/GM/IM</td>
<td>3,012</td>
</tr>
<tr>
<td>8</td>
<td>Electronic, motorized vehicles, and precision instruments</td>
<td>A/IM</td>
<td>2,436</td>
</tr>
<tr>
<td>9</td>
<td>Furniture, mixed freight, and miscellaneous manufactured products</td>
<td>GM/IM</td>
<td>2,880</td>
</tr>
<tr>
<td>10</td>
<td>Commodity unknown</td>
<td>GM</td>
<td>2,794</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Note: A = Auto; B = Bulk; GM = General Merchandise; IM = Intermodal.

5.3.4 Case study: Computation result

The modified convex combinations algorithm is coded in Visual C++ and run on a personal computer with 3.40 GHz CPU and 8 GB memory. We define the following metric (5.13) to capture the total train-hours that all freight shipments spend in the network each day.
Total Cost = \sum (\text{Assigned Link Flow} \times \text{Link Travel Time}) = \sum_{(i,j) \in A} x_{ij} t_{ij} (x_{ij} + x_{ji}). \quad (5.13)

Convergence is reached within a tolerance of 0.001% after 2,569 iterations and 25.559 seconds CPU time. The total cost is obtained as 75,425.94 (train-hr/day). The output contains link number, link origin and destination nodes, link distance, freight shipment volume (for each commodity type), link travel time, and average link speed at equilibrium. For comparison, we also consider an alternative scenario in which every O/D shipment is assigned to its shortest-distance path (i.e., as if the decision maker ignores congestion when determining shipment routes), and the total cost is obtained as 6,118,405.00 (train-hr/day). Comparing the total cost from the two different approaches, the total cost has been decreased around 98.77% when the user equilibrium is selected as a route choice rule; this is due to the form of the railroad link cost function in which the travel time of the link depends exponentially on the link traffic volume.

![Figure 5.9 User equilibrium result of rail freight network assignment](image_url)

* Unit of assigned flow: 
  * # of trains per day

85
Figure 5.9 illustrates the overall flow pattern at convergence. The total assigned traffic flow on the directed link pairs (e.g., representing the original bidirectional traffic on single tracks) are illustrated by various line thicknesses and colors as shown in the legend. We can observe large amount of freight flow and possibly heavy congestion on many links in the Western and Eastern Coastal states, plus Wyoming, Montana, North Dakota, and many of the Midwestern states near Chicago. Also, some main links connecting to Southern California, Texas, Kansas, and Georgia show high traffic flow.

5.3.5 Rail freight network assignment examples

Detailed results are provided in this section to gain insights on link or path level operations using two examples from the data. The O/D pair of the first example is included in the purple box and that of the second example is included in the blue box in Figure 5.10.

Figure 5.10 Two detailed examples from the rail freight network assignment
Figure 5.11 shows the network geometry and assignment result for the first sample from the freight shipment data, from ‘Remainder of Kentucky’ to ‘Birmingham-Hoover-Cullman, AL Combined Statistical Area’.

The square and triangle in Figure 5.11 respectively represent the origin and the destination of the specific freight shipment demand. The small blue dots denote other FAZ centroids, and all lines represent railroad links. All shipments in the network have found the shortest-time paths under congestion and user equilibrium. The set of red lines represents the shipment paths under user equilibrium for this specific O/D pair. The numbers in red near each red link represent the corresponding average train speed and the link travel time. For comparison, the result from an alternative scenario which ignores congestion when determining shipment routes and uses only the shortest-distance path for each O/D shipment assignment is also shown. The set of green lines represents the shortest-distance path between this specific O/D pair. The
numbers in black near each green link mark the expected travel speed and travel time under congestion if the shortest-distance paths are actually implemented.

In this example, the rail freight demand is 2.6 trains per day and they need to be assigned on the network from the given origin to the given destination. In the user equilibrium state, traffic will be split between two different routes, and the total travel time associated with each route is the same as around 14.1 hours. On the other hand, the total travel time is 121.9 hours and only one route is used when the shortest-distance path is used. The total cost reduction is huge in the user equilibrium state due to the form of the railroad link cost function.

Figure 5.12 shows the second example of the detailed assignment result, from ‘Remainder of Kansas’ to ‘Houston-Baytown-Huntsville, TX Combined Statistical Area’ in freight shipment data.
According to the data, 6.8 trains are assigned per day for this specific O/D pair. Again, trains in the user equilibrium state will spend around 21.7 hours with two different route choices represented by red links from the origin to the destination. On the other hand, if the trains follow the shortest-distance path denoted by green lines, the expected travel time will increase to 31.1 hours.

5.4 Comparison between Truck and Rail Freight Shipment Cost

To illustrate the difference in shipment costs along truck and rail paths for the same O/D, one data record which has both truck and rail freight shipment demand data in Year 2007 is randomly selected. This shipment record has ‘Remainder of Kentucky’ as its origin and ‘Remainder of Georgia’ as its destination. Truck freight demand is 60.06 vehicles per hour, or 384.38 tons per hour and rail freight demand is 14.67 trains per day, or 49,900.13 tons per day. The total shipment costs for truck and rail are respectively defined as equations (5.6) and (5.13), and the average shipment cost can be defined as follows:

\[
\text{Average Shipment Cost} = \frac{\text{Total Shipment Cost}}{\text{Total Freight Demand}}. \tag{5.14}
\]

In this example, the total shipment cost of truck is calculated as 583.36 (veh-hr/hour) and that of rail is 512.88 (train-hr/day). Since the units of the total cost for truck and rail are different, the average shipment cost is used as a normalized metric. The average shipment cost of truck is 9.71 (hours) or 1.52 (vehicle-hour/ton), and that of rail is 34.97 (hours) or 0.01 (train-hour/ton) since the loading capacity of a train is much larger than that of the truck. This result implies that truck is a suitable mode for delivering time sensitive or higher value commodities, while rail is preferred for transporting heavy or bulk goods which are not sensitive to the delivery time. The
results and implications are also consistent with the results derived in the modal split procedure in Chapter 4.

5.5 Model Validation

Model validation is conducted for both truck and rail freight network assignment to ensure accuracy and practicality of the proposed models and solution algorithms. In Table 5.2, columns (a) and (b) compare the total freight demand for truck and rail in Year 2007 in thousand tons, obtained from two sources: 2007 CFS (RITA U.S. DOT, 2011) and FAF³ (FHWA U.S. DOT, 2011). The FAF³ is used as input to our models to generate truck and rail freight shipment demand, i.e., the value in column (b). The total truck and rail freight demand in tonnages obtained from the FAF³ in column (b) are larger than those obtained from the 2007 CFS in column (a) and the difference is significant for trucks. The CFS data have been prepared by the Bureau of Transportation Statistics with the U.S. Census Bureau for a long time, which provide economic census data on domestic freight shipment including commodity type, value, weight, transportation mode, and shipment origin and destination (RITA U.S. DOT, 2011). The FAF data have been prepared by FHWA mainly based on the CFS data to satisfy growing demand for freight shipment data and support various government policies and legislative issues (Battelle, 2011). Generally, a variety of other databases are also included to complete the FAF³ since the 2007 CFS have a lot of missing data (Southworth et al., 2011); FAF³ is more complete than 2007 CFS. In case of rail freight shipment data, federal regulations have enforced railroad industries to record and report extensive and detailed information on most of their traffic since the early 1900’s (Sharfman, 1915). Thus, there is not much difference in rail shipment demand between different data sources. On the other hand, the freight demand data collecting system in truck
industries may not be as accurate as the one in railroad industries, which might have caused large gaps between different databases.

### Table 5.2 Freight shipment demand from different sources

<table>
<thead>
<tr>
<th>Unit</th>
<th>Source</th>
<th>Ton (10^3)</th>
<th>Ton-mile (10^6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(a) 2007 CFS</td>
<td>(b) FAF^3</td>
<td>(c) 2007 CFS</td>
</tr>
<tr>
<td>Mode</td>
<td>Truck</td>
<td>8,778,700</td>
<td>13,021,790</td>
</tr>
<tr>
<td></td>
<td>Rail</td>
<td>1,861,300</td>
<td>2,037,955</td>
</tr>
</tbody>
</table>

Columns (c) and (d) show the total truck and rail freight shipment ton-miles in Year 2007, obtained from two sources: 2007 CFS and FAF^3. Column (e) further extracts the same information based on the O/D route choice decisions from our network assignment models. In case of truck, the discrepancy between columns (c) and (d) can be mostly explained by the difference between columns (a) and (b); freight demand (in tonnages) increase in FAF^3 is huge comparing to 2007 CFS, thus it leads to huge increase in ton-miles in FAF^3 data. Additionally, we conjecture that the different methodologies for route choices might have made the gap larger. The 2007 CFS in column (c) is based on the one fixed distance (e.g., the shortest distance) between each O/D pair, but the FAF^3 for the U.S. highways in column (d) can use the least time routes (even those with longer distances) when there is congestion in the network (Battelle, 2011). Our network assignment model resulted in 2,359,810 million total truck shipment ton-miles as shown in column (e), which is almost the same as the value in column (d). In case of railroad, rail freight shipment demand (in tonnages) in column (b) is 9% larger than that in column (a), which can roughly explain the difference between columns (c) and (d). Also, additional ton-mile increase in FAF^3 might be able to be explained by difference of distance used in the network assignment procedure as explained in the case of truck. Finally, we obtained
1,693,435 million as the total ton-miles for railroad as shown in column (e). Although columns (d) and (e) present some discrepancy, the industry source (AAR, 2011) reports 1,770,545 million as freight ton-miles for railroad in Year 2007, which is larger than the values in both columns (d) and (e). Note that our result is within the range of available empirical data, and it is probably acceptable for decision making.

Figure 5.13 illustrates empirical freight flow patterns observed in the real world on the U.S. highway and rail networks. Traffic volume on each link is simply a graphical representation of survey data; it does not provide any information on how the given O/D freight demand is assigned or will be assigned in the future. Figure 5.13(a), adapted from Battelle (2011), describes the average daily truck shipment distribution on the national highway network. It is generated using the Highway Performance Monitoring System (HPMS) database (FHWA U.S. DOT, 2007b). Traffic volume on the highway network is represented by red lines with various thicknesses to show the amount of assigned daily truck flow in Year 2006. Figure 5.13(b) illustrates current rail traffic volume in major freight corridors, adapted from Cambridge Systematics, Inc. (2007). Traffic volume is represented by red lines with various thicknesses for both freight trains in Year 2005 and passenger trains in Year 2007. This map is created based on the Surface Transportation Board’s Carload Waybill Sample and published data on Amtrak and commuter passenger rail schedules (Cambridge Systematics, Inc., 2007). Considering the fact that the U.S. rail network has been dominated by freight shipment traffic, we can assume the information in Figure 5.13(b) almost represent rail freight movements alone in Year 2005. Then, our model output for truck and rail in Figures 5.3 and 5.9 can be visually compared with the traffic flow patterns in Figures 5.13(a) and 5.13(b), respectively.
In case of truck, Figure 5.13(a) describes a large amount of assigned traffic on many highway links in Washington, Oregon, California, Florida, the Midwest states near Chicago, and northeastern parts of the U.S. This trend is generally consistent with the annual freight traffic distribution in the U.S. highway network obtained from our model. Also, this figure shows high traffic flow on some main highway links that connect Southern California, Arizona, and Oklahoma, which are less emphasized in our result. In case of railroad, Figure 5.13(b) shows that high rail traffic flow concentrate around many links near California, Wyoming, Montana, many
of the Midwestern states near Chicago, northeastern regions of the U.S., and some main links that connect Southern California, Texas, and Kansas. This pattern is generally consistent with the annual freight traffic distribution obtained from our model. There are nevertheless some discrepancies on a few links around Washington, Oregon, Texas, northern Florida, and southern Pennsylvania. We suspect that such discrepancies in highway and rail networks might be partially due to two factors: (i) difference in input data on the network geometry and freight demand, and (ii) difference of the analysis year (i.e., our model uses data in Year 2007, while Figure 5.13(a) shows data in Year 2006 and Figure 5.13(b) mostly describes data in Year 2005). Note that from a high level perspective the freight flow patterns obtained from the proposed models and the empirical data generally match each other, and thus our approach could be used for predicting future congestion pattern.

5.6 Rail Network Capacity Expansion and Its Effect on Network Assignment

Freight shipment demand in the U.S. rail network is projected to increase 88% by Year 2035, and this trend is expected to result in severe congestion (Cambridge Systematics, Inc., 2007). Therefore, infrastructure investment (e.g., expansion of single tracks into double tracks) may be needed in the rail network to expand link capacity and improve network efficiency near potential chokepoints. Enhanced level of service on improved infrastructure will affect future rail freight demand assignment pattern on the national scale.

In this section, we examine the most congested railroad links in the near future. Year 2035 is chosen as a reference year since future congestion pattern in 2035 without infrastructure investment can be obtained from Cambridge Systematics, Inc. (2007) for comparison and validation. Some of the single tracks will be expanded to full double tracks so as to change the
railroad link cost functions. Lastly, the rail freight network assignment model in this dissertation will be applied again to the upgraded rail network as a subroutine to help quantify network efficiencies (i.e., the total cost and congestion pattern) in a “before and after” comparison.

(a) Congestion prediction from the proposed model in Year 2035 without infrastructure investment (Scenario: high GDP growth rate with business as usual)

(b) Congestion prediction in Year 2035 without infrastructure investment (Cambridge Systematics, Inc., 2007)

Figure 5.14 Congestion prediction in the U.S. rail network in Year 2035

Figure 5.14(a) illustrates the predicted congestion pattern in rail network in Year 2035 (without link capacity expansion). Note that average speed on each undirected link is categorized
and illustrated by line thickness and color. This result is obtained from our proposed model under the scenario that the global economic growth rate is high with business as usual (which is expected to cause the most severe congestion among the four given macro scenarios). Potential bottleneck points are the easiest to identify in this scenario. Heavy congestion is observed on many links in the Western and Eastern Coastal states, Wyoming, Montana, North Dakota, and many of the Midwestern states near Chicago. Also, some main links connecting to Southern California, Texas, Kansas, and Georgia show heavy congestion. Figure 5.14(b) shows another prediction of rail network congestion pattern in Year 2035 (without infrastructure investment), which is adapted from Cambridge Systematics, Inc. (2007); it has been created by applying economic growth rates to the traffic flow pattern in Figure 5.13(b). Although our model estimates somewhat more congestion in Wyoming, Montana, and Eastern Coastal states (possibly due to the different freight shipment demand, network geometry, and assignment methodology), the congestion pattern in this figure is generally matching that in Figure 5.14(a). Thus, the proposed methodology is effective in bottleneck identification and congestion prediction under any macro scenarios and could serve as the basis for developing strategies to relieve future congestion on the rail network. Below is a simple example.

Figure 5.14(a) contains 103 most congested undirected links whose average link speed is less than or equal to 10 mph. These links are illustrated by red color in Figure 5.14(a). Among them, 31 links are already entirely double track lines. Now we consider the scenario where the remaining 72 undirected links are upgraded to full double tracks and investigate the congestion pattern in the new network. Figure 5.15 illustrates the user equilibrium result in Year 2035 after capacity expansion. While the congestion pattern in Figure 5.15 is largely the same as that in Figure 5.14(a), except for some obvious changes in several links in Washington, Kentucky,
Tennessee, Missouri, Colorado, and some states near Chicago, the reduction in congestion and delay cost is much more significant.

Figure 5.15 Congestion prediction from the proposed model in Year 2035 after capacity expansion

Table 5.3 compares the total train-hour cost and the total ton-mile “before” and “after” the rail network capacity expansion. Link capacity expansion decreases the values of constants $\beta_c$ in equation (5.7) and $\beta_q$ in equation (5.8) for the selected heavily congested links, which leads to a significant 32.67% reduction in the total cost as shown in row (b). We can observe the total ton-miles in row (c) also decreased slightly, implying less detour toward shipment destinations. As such, this example shows how investment in railroad infrastructure can help relieve congestion on the transportation network and contribute to efficient freight movements.

<table>
<thead>
<tr>
<th></th>
<th>(a) Capacity expansion</th>
<th>Before</th>
<th>After</th>
<th>% reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>(b) Total cost ($10^3$ train-hr/day)</td>
<td></td>
<td>2,025</td>
<td>1,364</td>
<td>32.67</td>
</tr>
<tr>
<td>(c) Total ton-mile ($10^3$ ton-mile/day)</td>
<td></td>
<td>10,496,597</td>
<td>10,411,213</td>
<td>0.81</td>
</tr>
</tbody>
</table>
5.7 Model Application for Future Prediction

The proposed freight shipment assignment models have been applied to forecasting future truck and rail freight activities and estimating the related emissions from 2010 to 2050 using input data obtained from the previous modal split step. Table 5.4 summarizes the results for a number of future years and four given global economic growth and environmental policy scenarios as described in Chapter 3. Columns (b), (e), (h), and (k) each reports truck, rail, and the total freight shipment assignment results in million ton-miles summed across all commodities and O/D pairs; the total ton-mile growth and GDP growth as compared to those in Year 2010 are also included for validation purpose. Columns (c), (f), (i), and (l) each shows various emission estimations using emission factors obtained from the U.S. EPA (2008) and NRDC (2012). We assume each ton-mile of truck and rail shipments generate 0.2970 and 0.0252 kg of CO\textsubscript{2}, 0.0035 and 0.0020 g of CH\textsubscript{4}, 0.0027 and 0.0006 g of N\textsubscript{2}O, and 0.092 and 0.013 g of PM\textsubscript{10}, respectively.

As columns (b), (e), (h), and (k) show, in each scenario, the total freight shipment in ton-miles for both truck and rail, as well as their sum, continuously grow over time. We can observe the total ton-mile growth rate and the GDP growth rate are almost consistent across all study years and scenarios. This implies the proposed methodologies are effective in predicting the future trends. Since the amount of emissions produced from the freight transportation modes is proportional to the freight activities (i.e., ton-miles), we can observe in columns (c), (f), (i), and (l) that truck, rail, and the total national CO\textsubscript{2}, CH\textsubscript{4}, N\textsubscript{2}O, and PM\textsubscript{10} emissions increase as well. Also, note that scenarios 2 and 4 show a larger modal shift on the national level to the rail mode, which in turn generate lower emissions. This is due to the implementation of climate policy and carbon tax which significantly drives up the final market prices of fossil fuel commodities and makes the rail mode favorable.
Table 5.4 Future truck and rail freight demand assignment prediction and emission estimation

<table>
<thead>
<tr>
<th>Year</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Truck CO₂ emission (10^3ton)</td>
<td>Rail CO₂ emission (10^3ton)</td>
</tr>
<tr>
<td>2010</td>
<td>780,023</td>
<td>52,609</td>
</tr>
<tr>
<td>2015</td>
<td>877,916</td>
<td>59,405</td>
</tr>
<tr>
<td>2020</td>
<td>987,290</td>
<td>67,654</td>
</tr>
<tr>
<td>2025</td>
<td>1,112,674</td>
<td>76,880</td>
</tr>
<tr>
<td>2030</td>
<td>1,232,280</td>
<td>86,508</td>
</tr>
<tr>
<td>2035</td>
<td>1,387,312</td>
<td>96,548</td>
</tr>
<tr>
<td>2040</td>
<td>1,533,426</td>
<td>108,333</td>
</tr>
<tr>
<td>2045</td>
<td>1,708,324</td>
<td>119,959</td>
</tr>
<tr>
<td>2050</td>
<td>1,898,961</td>
<td>126,019</td>
</tr>
</tbody>
</table>
Table 5.4 (continued)

<table>
<thead>
<tr>
<th>Year</th>
<th>Scenario 3</th>
<th>(b) Freight demand assignment result</th>
<th>(i) Emission estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Truck</td>
<td>Rail</td>
<td>Total</td>
</tr>
<tr>
<td></td>
<td>ton-mile</td>
<td>ton-mile</td>
<td>ton-mile</td>
</tr>
<tr>
<td></td>
<td>(10^9)</td>
<td>(10^9)</td>
<td>(10^9)</td>
</tr>
<tr>
<td></td>
<td>growth</td>
<td>growth</td>
<td>growth</td>
</tr>
<tr>
<td></td>
<td>GDP</td>
<td>CO₂ emission (10^3 ton)</td>
<td>CO₂ emission (10^3 ton)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CH₄ emission (ton)</td>
<td>CH₄ emission (ton)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N₂O emission (ton)</td>
<td>N₂O emission (ton)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PM₁₀ emission (ton)</td>
<td>PM₁₀ emission (ton)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PM₂.₅ emission (ton)</td>
<td>PM₂.₅ emission (ton)</td>
</tr>
<tr>
<td>2010</td>
<td>2,595</td>
<td>2,056</td>
<td>4,651</td>
</tr>
<tr>
<td></td>
<td>2015</td>
<td>2,850</td>
<td>2,269</td>
</tr>
<tr>
<td></td>
<td>2020</td>
<td>3,111</td>
<td>2,497</td>
</tr>
<tr>
<td></td>
<td>2025</td>
<td>3,413</td>
<td>2,708</td>
</tr>
<tr>
<td></td>
<td>2030</td>
<td>3,685</td>
<td>2,989</td>
</tr>
<tr>
<td></td>
<td>2035</td>
<td>4,022</td>
<td>3,216</td>
</tr>
<tr>
<td></td>
<td>2040</td>
<td>4,338</td>
<td>3,472</td>
</tr>
<tr>
<td></td>
<td>2045</td>
<td>4,720</td>
<td>3,743</td>
</tr>
<tr>
<td></td>
<td>2050</td>
<td>5,089</td>
<td>4,017</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Scenario 4</th>
<th>(b) Freight demand assignment result</th>
<th>(i) Emission estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Truck</td>
<td>Rail</td>
<td>Total</td>
</tr>
<tr>
<td></td>
<td>ton-mile</td>
<td>ton-mile</td>
<td>ton-mile</td>
</tr>
<tr>
<td></td>
<td>(10^9)</td>
<td>(10^9)</td>
<td>(10^9)</td>
</tr>
<tr>
<td></td>
<td>growth</td>
<td>growth</td>
<td>growth</td>
</tr>
<tr>
<td></td>
<td>GDP</td>
<td>CO₂ emission (10^3 ton)</td>
<td>CO₂ emission (10^3 ton)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CH₄ emission (ton)</td>
<td>CH₄ emission (ton)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N₂O emission (ton)</td>
<td>N₂O emission (ton)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PM₁₀ emission (ton)</td>
<td>PM₁₀ emission (ton)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PM₂.₅ emission (ton)</td>
<td>PM₂.₅ emission (ton)</td>
</tr>
<tr>
<td>2010</td>
<td>2,589</td>
<td>2,059</td>
<td>4,648</td>
</tr>
<tr>
<td></td>
<td>2015</td>
<td>2,843</td>
<td>2,272</td>
</tr>
<tr>
<td></td>
<td>2020</td>
<td>2,947</td>
<td>2,614</td>
</tr>
<tr>
<td></td>
<td>2025</td>
<td>3,197</td>
<td>2,893</td>
</tr>
<tr>
<td></td>
<td>2030</td>
<td>3,397</td>
<td>3,190</td>
</tr>
<tr>
<td></td>
<td>2035</td>
<td>3,650</td>
<td>3,615</td>
</tr>
<tr>
<td></td>
<td>2040</td>
<td>3,862</td>
<td>3,965</td>
</tr>
<tr>
<td></td>
<td>2045</td>
<td>4,079</td>
<td>4,361</td>
</tr>
<tr>
<td></td>
<td>2050</td>
<td>4,304</td>
<td>4,725</td>
</tr>
</tbody>
</table>
5.8 Conclusion

In this chapter, the procedures for truck and rail freight shipment demand network assignment considering congestion effect are discussed and the results are explained in detail with visual illustrations. Generally, both ‘freight mode choices’ and ‘route choices’ can significantly affect national and regional air quality and eventually human health. Thus, it is important to investigate how to assign freight shipment demand between all FAZ O/D pairs on each modal network.

The fundamental traffic assignment model with the convex combinations algorithm is proposed to solve the truck freight shipment assignment problem under user equilibrium principle. BPR link cost function is modified to capture the effect of background traffic volume that already exists in the highway network on the link travel time. A case study is conducted using the entire U.S. highway network and the national freight shipment data in Year 2007. Convergence is reached within a short computation time and the optimal truck freight flow and congestion pattern are obtained. Outcome is the total freight ton-miles traveled on every link of the U.S. highway network. In case of the rail freight shipment assignment problem, a customized network assignment model is constructed. Our model addresses a practical issue of the rail network, where traffic flow in two opposite directions generally needs to be loaded on one shared link. A railroad-specific link cost function adjusted for single and double tracks is suggested and incorporated into our model. The proposed methodology is applied to an empirical case study with the full-scale U.S. rail network and the national freight shipment data in Year 2007. Solution can be obtained within a short computation time. Model validations are conducted and the numerical results from the proposed models are found to effectively reproduce the traffic density and congestion pattern observed in the real world for both truck and rail. Based on the proposed approach for the rail freight shipment demand assignment, we investigate
possible rail network capacity expansion scenario. We search the most congested railroad links in Year 2035 and some of single tracks on the possible bottlenecks are upgraded to full double tracks. It turned out that investment on railroad infrastructure will help reduce possible congestion and contribute to efficient rail freight deliveries. The proposed modeling frameworks will be useful for predicting future freight flow distribution and congestion pattern. Eventually, this result will form the basis for the transportation emission assessment in the national line-haul shipments.
CHAPTER 6
LOGISTICS SYSTEMS PLANNING FOR REGIONAL FREIGHT DELIVERY

6.1 Introduction

Future freight demand and its shipment between all FAZ O/D pairs in the U.S. are forecasted by the four-step inter-regional freight demand model. The final output includes not only freight flow distribution between all O/D FAZs for each commodity type and shipment mode, but also congestion estimation in each modal network. This result will be useful to predict how mode choices and route choices for various amount of future freight demand will affect air quality and eventually human health from a high level.

In this study, we assume inter-regional freight delivery is consolidated and shipped between the centroids of the FAZs. In reality, the bulk of freight arriving at the destinations (i.e., terminals) need to be broken for delivery to distributed individual customers within each FAZ. Similarly, the freight also needs to be collected from a large number of supply points to the set of origins (i.e., terminals) in each FAZ. In this regard, freight delivery activities within large urban areas are also critical because emissions from the freight shipments comprise a large share of toxic air pollutants and greenhouse gases in most metropolitan areas worldwide (OECD, 2003). Moreover, due to the rapid increase in freight demand and the significant growth in delivery activities, concerns for the air quality problems in urban areas have become more serious (Figlioizzi, 2011). The residents in metropolitan areas are more likely to be affected by the air

---

3 This chapter has been adapted from “Hwang, T., Lee, S., Lee, B. and Ouyang, Y. (2014). Regional freight delivery emission estimation under different urban development scenarios. Working Paper, University of Illinois at Urbana-Champaign, Urbana, Illinois.”
pollution problems than those in rural areas since most of them live very closely to the emission sources (e.g., commercial vehicles operated by diesel engines). This motivates us to investigate the freight shipment modeling and logistics planning at the intra-regional level.

In this chapter, we will study the freight delivery problem to/from a large number of freight demand/supply points within an FAZ. This problem can be formulated as a large-scale VRP and a ring-sweep algorithm (Newell and Daganzo, 1986a) will be adopted to estimate the network delivery efficiency. The ring-sweep algorithm assumes the identical freight is delivered from a single source (i.e., terminal or depot) to each demand point composed of identical customers. However, in reality, all demand points are mostly composed of different numbers of customers in different industries, and thus they require different amount and types of commodities. Moreover, there can be many terminals or depots since we consider freight delivery problem within FAZs. Thus, in this chapter, the algorithm in Newell and Daganzo (1986a) will be modified to address these issues and construct a logistics system model that connects the spatially distributed demand/supply points which are large scale within each FAZ. A case study is conducted to estimate future regional freight activities and the related emissions from 2010 to 2050 in 30 FAZs which cover 22 major metropolitan areas. The modeling framework presented in this chapter can be used to infer the emission distribution and estimate the human exposures to emissions from the freight delivery in the intra-regional level. Eventually, this effort will be able to contribute to enhance public benefit by decreasing the social cost incurred by vehicle emissions in large urban areas in the U.S.

The exposition of this chapter is as follows: Section 6.2 briefly reviews the ring-sweep algorithm and provides a logistics system planning for the intra-regional freight delivery. Section
6.3 presents an empirical case study for the selected FAZs, in which regional freight shipment from truck and railroad terminals are considered separately. Section 6.4 concludes the chapter.

6.2 Large Scale Freight Delivery Modeling

The ring-sweep algorithm is briefly introduced to explain the basic concept of the methodology applied in this chapter. Then, the original ring-sweep algorithm is modified to solve the intra-regional freight delivery problem.

6.2.1 Ring-sweep algorithm review

The ring-sweep algorithm proposed by Newell and Daganzo (1986a) is based on the Continuum Approximation (CA) optimization scheme in which customer demand is assumed to be a continuous density function that may vary slowly over space. This asymptotic approximation method is suitable for problems that involve a large number of demand points.

![Figure 6.1 Possible zoning and delivery plan example (Ouyang, 2007)](image)

Figure 6.1 Possible zoning and delivery plan example (Ouyang, 2007)

Figure 6.1 adapted from Ouyang (2007) illustrates the basic concept of the algorithm. The 10-by-10 square represents a freight delivery region with randomly distributed customers
and a truck terminal is located at the bottom left-hand corner of the square. The objective is to minimize the total vehicle delivery distance needed to satisfy the freight demand of the large number of customers. The ring-sweep algorithm assumes the identical freight is delivered from a single source to each demand point composed of identical customers. The algorithm partitions the given region into many delivery zones (e.g., the small rectangle elongated toward the terminal) such that one vehicle can serve freight demand in one delivery zone. As this figure shows, the vehicle needs to make a round trip between the terminal and the edge of its delivery zone (i.e., the line-haul movement) and then make a tour within a partitioned delivery zone to distribute the freight to each demand point (i.e., the local travel). Sum of the line-haul distance and the local travel distance across all the properly partitioned freight zones were shown to form a near optimal solution to the total delivery problem (Ouyang, 2007). Simple formulae to estimate the near-optimal total vehicle-distance are also provided in the literature (Newell and Daganzo, 1986a). To obtain the total cost for collecting the freight, the same approach can be applied assuming the large number of supply points, instead of demand points, need to be served in freight origin regions.

6.2.2 Regional freight delivery modeling in an FAZ

The ring-sweep algorithm assumes that the demand points in a region are identical, and thus the same amount of freight shall be delivered from a single source to each demand point. In reality, this may not be necessarily true since each demand point is mostly composed of a number of customers depending on the industries involved. Also, there can be multiple freight terminals in the FAZ. In this section, the original ring-sweep algorithm is modified to resolve these issues and to be applied to the current freight distribution/collection modeling in the intra-regional level.
Figure 6.2 describes more detail about the development of our model. The entire freight delivery region (i.e., FAZ) is represented by an arbitrary shape. Each freight delivery region is composed of a set of mutually disjointed census tracts. In this figure, there are six census tracts in total and the centroid of each census tract is marked by a black dot. The number near each census tract centroid indicates total number of employees in industries considered in the census tract. The truck terminal (or depot) is represented by the black circle and located on the left side of the freight delivery region. Since a fleet of trucks need to make a round trip from the terminal daily to satisfy the freight demand across the region, the entire freight delivery region is partitioned into several disjointed delivery zones such that the freight demand in one delivery zone is covered by one truck shipment. In Figure 6.2, two partitioned freight delivery zones are shown as examples in blue lines. The truck movements are illustrated by red lines (including both the line-haul and the local travels). The objective is to minimize the total transportation cost which can be represented by the total travel distance.

Figure 6.2 Application of the ring-sweep algorithm to the regional freight delivery problem
Additionally, the following assumptions are made. First, freight demand (i.e., employees) in each census tract is concentrated at the centroid of the census tract and will be assigned to the nearest terminal for service by delivery vehicles (in case there are multiple terminals in the freight delivery region). Second, freight is delivered by identical short-haul trucks with constant low speed (i.e., roadway congestion is not considered). Third, distance along the local roadway network for freight delivery can be approximated by the Euclidean metric. Finally, the delivery zones can be generated as shown in Figure 6.2.

To formulate the model, let $d_i$ be the distance from the terminal to the centroid of the census tract $i$ and $E_{ij}$ be the number of employees in an industry type $j$ in the census tract $i$. Also, we let $I$ be the total number of census tracts in a given freight delivery region and $J$ be the total number of industry types considered in this study. Additionally, the truck capacity is represented by $C$ (tons) and the total daily freight demand in a freight delivery region is represented by $D$ (tons per day). Then, the total line-haul distance ($L_1$) can be formulated as in Newell and Daganzo (1986a) as follows:

$$L_1 = \frac{2D \sum_{i=1}^{I} \sum_{j=1}^{J} E_{ij} d_i}{C \sum_{i=1}^{I} \sum_{j=1}^{J} E_{ij}}.$$  (6.1)

To construct the total local travel distance, let $N_i$ be the total number of demand points in each census tract $i$. To calculate $N_i$, the average number of employees per firm (establishment) in an industry type $j$ is defined as $a_j$, which represents how many employees are served on
average by one truck visit and may vary across industries. As such, \( N_i \) can be obtained using the following equation for each census tract \( i \):

\[
N_i = \sum_{j=1}^{I} \frac{E_{ij}}{a_j}
\]  

(6.2)

The total number of demand points in a given freight delivery region is represented by \( N \), sum of \( N_i \) across all census tracts from 1 to \( I \) (i.e., \( N = \sum_{i=1}^{I} N_i \)). Also, area of a freight delivery region is denoted by \( A \) and a uniformly distributed demand point density in a freight delivery region is defined as \( \delta = \frac{N}{A} \). Then, the total local travel distance (\( L_2 \)) can be formulated as follows (Newell and Daganzo, 1986a):

\[
L_2 = \frac{0.57N}{\sqrt{\delta}}
\]  

(6.3)

Finally, the total cost to serve spatially distributed freight demand within a freight delivery region can be obtained by summing up the total line-haul distance (\( L_1 \)) and the total local travel distance (\( L_2 \)). This modeling approach is very general and suitable for cases in which the number of vehicle routes is expected to be much larger than the number of visits per tour (\( C_r \)), i.e., \( N \gg C_r^2 \) (Daganzo, 2005); our regional freight delivery problem fully satisfies this condition. The model described above focuses on “distribution” of freight demand using freight “attraction” data in a “destination” FAZ. It shall be obvious that the “collection” of items by a
fleet of vehicles also can be addressed in a similar fashion using freight “production” data in each “origin” FAZ.

6.3 Case Study

A case study is conducted to estimate future regional freight delivery activities and the related emissions from 2010 to 2050 in 30 FAZs which cover 22 major metropolitan areas in the U.S. Intra-regional freight deliveries from truck and railroad terminals are modeled separately considering different industry types. Both distribution and collection of freight are included to obtain the total cost. For conciseness of presentation, however, only procedures related to the freight distribution are explained in the following sections.

6.3.1 Regional freight delivery from truck terminals

This section deals with a regional freight delivery problem from truck terminals. Commodities in this study are categorized into two groups considering employment types as shown in Table 6.1 (provided by the urban spatial structure model) since the groups of commodities required by different industries often need to be delivered separately by different trucks. The first group is shown in column (c) which includes commodities for employees in wholesale and retail trade industry. The second group is described in column (d) which includes commodities for employees in the manufacturing industry. As such, the total freight demand will be assigned to two industry groups according to their shares.
Table 6.1 Shares of two industry groups for each commodity type

<table>
<thead>
<tr>
<th>(a) Commodity type</th>
<th>(b) Commodity description</th>
<th>(c) Wholesale and retail trade (%)</th>
<th>(d) Manufacturing (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Agriculture products and fish</td>
<td>25.44</td>
<td>74.56</td>
</tr>
<tr>
<td>2</td>
<td>Grain, alcohol and tobacco products</td>
<td>73.29</td>
<td>26.71</td>
</tr>
<tr>
<td>3</td>
<td>Stones, nonmetallic minerals, and metallic ores</td>
<td>37.13</td>
<td>62.87</td>
</tr>
<tr>
<td>4</td>
<td>Coal and petroleum products</td>
<td>22.02</td>
<td>77.98</td>
</tr>
<tr>
<td>5</td>
<td>Basic chemicals, chemical and pharmaceutical</td>
<td>35.44</td>
<td>64.56</td>
</tr>
<tr>
<td>6</td>
<td>Logs, wood products, and textile and leather</td>
<td>41.19</td>
<td>58.81</td>
</tr>
<tr>
<td>7</td>
<td>Base metal and machinery</td>
<td>46.72</td>
<td>53.28</td>
</tr>
<tr>
<td>8</td>
<td>Electronic, motorized vehicles, and precision</td>
<td>72.00</td>
<td>28.00</td>
</tr>
<tr>
<td>9</td>
<td>Furniture, mixed freight, and miscellaneous</td>
<td>46.05</td>
<td>53.95</td>
</tr>
<tr>
<td>10</td>
<td>Commodity unknown</td>
<td>93.86</td>
<td>6.14</td>
</tr>
</tbody>
</table>

We assume a number of truck terminals are located near the junctions of major highways. The freight demand in each census tract is assigned to the closest truck terminal, and thus equations (6.1)-(6.3) can be rewritten for a specific terminal \( k \) in the form of (6.4)-(6.5) for commodities related to the wholesale and retail trade industry, and (6.6)-(6.7) for commodities related to the manufacturing industry. Equations (6.4) and (6.6) represent the line-haul distance and equations (6.5) and (6.7) denote local travel distance. Summing (6.4)-(6.7) across all truck terminals \( k \in K \) yields the total freight delivery cost in the FAZ.

\[
L_{j1}^k = \frac{2\alpha_i \left( D_W + D_R \right) \sum_{i=1}^{I} \sum_{j=1}^{2} E_{ij} d_{ki}}{C \sum_{i=1}^{I} \sum_{j=1}^{2} E_{ij}} , \quad (6.4)
\]

\[
L_{j2}^k = \frac{0.57 N_f^k}{\sqrt{\delta_f^k}} , \quad \text{where } N_f^k = \frac{\alpha_i}{\alpha_i} \sum_{i=1}^{I} \sum_{j=1}^{2} E_{ij} \text{ and } \delta_f^k = \frac{N_f^k}{A_k} , \quad (6.5)
\]
\[ L_{p1}^k = \frac{2\alpha_2 D M \sum_{i=1}^{l_k} E_{i3} d_{ki}}{C \sum_{i=1}^{l_k} E_{i3}}, \quad (6.6) \]

\[ L_{p2}^k = \frac{0.57 N_p^k}{\sqrt{\delta_p^k}}, \text{ where } N_p^k = \frac{\alpha_2}{a_2} \sum_{i=1}^{l_k} E_{i3} \text{ and } \delta_p^k = \frac{N_p^k}{A_k}. \quad (6.7) \]

In an FAZ, parameters \( \alpha_1 \) and \( \alpha_2 \) respectively represent percentage of employees in wholesale and retail trade industry and manufacturing industry served by truck terminals. The total daily freight demand of wholesale trade, retail trade, and manufacturing industries in the FAZ are respectively denoted by \( D_w \), \( D_r \), and \( D_m \). Recall that \( l_k \) is the total number of census tracts assigned to the truck terminal \( k \) and \( d_{ki} \) is the distance from the truck terminal \( k \) to the centroid of the census tract \( i \). Subscript \( j \) has been defined to describe different industries and \( j = 1, 2, \) and 3 respectively represents wholesale trade, retail trade, and manufacturing industries. The sum of total area assigned to the truck terminal \( k \) is represented by \( A_k \). The average number of employees per firm (establishment) in the wholesale and retail trade industry is \( a_1 = 13.94 \), and that in the manufacturing industry is \( a_2 = 40.19 \) (U.S. Census Bureau, 2012). Finally, the total freight delivery cost within the FAZ across all industries from all truck terminals \( (T) \) is as follows:

\[ G_T = \sum_{k=1}^{K} \left( L_{f1}^k + L_{f2}^k + L_{p1}^k + L_{p2}^k \right). \quad (6.8) \]

6.3.2 Regional freight delivery from railroad terminals

A significant share of inter-regional freight shipment is also carried by railroads in the U.S. Commodities in this section will be combined into two groups considering ways for the shipment
to reach customers: (i) direct shipment from the railroad terminals without trucks, and (ii) short-haul truck delivery from the railroad terminals. The groups are defined differently in freight shipment destination zone and origin zone. In the destination zones, commodity types 1 (i.e., agriculture products and fish), 2 (i.e., grain, alcohol and tobacco products), 3 (i.e., stones, nonmetallic minerals, and metallic ores), 4 (i.e., coal and petroleum products), 5 (i.e., basic chemicals, chemical and pharmaceutical products), 6 (i.e., logs, wood products, and textile and leather), and 7 (i.e., base metal and machinery) will be shipped directly by rail, while commodity types 8 (i.e., electronic, motorized vehicles, and precision instruments), 9 (i.e., furniture, mixed freight, and miscellaneous manufactured products), and 10 (i.e., commodity unknown) will be delivered by trucks. In the origin zones, commodity types 3, 4, 5, and 7 will be shipped directly, while commodity types 1, 2, 6, 8, 9, and 10 will be delivered by trucks. Different methodologies will be applied to these commodity groups to address the regional freight delivery problem from railroad terminals.

For commodities related to direct shipments, we assume trucks are not involved. The rest of freight demand is distributed over the FAZ daily by short-haul trucks. All trucks start their travel from several railroad terminals, each of which is assumed to be located near the junctions of major railroad links. The freight demand in each census tract is assigned to the closest railroad terminal. We differentiate freight demand considering industry types as shown in Table 6.1, and thus the total freight commodities will be combined into two industry groups as well (i.e., wholesale and retail trade industry and manufacturing industry) according to their shares. Equations (6.1)-(6.3) can be rewritten for a specific terminal \( q \) in the form of (6.9)-(6.10) for commodities related to the wholesale and retail trade industry, and (6.11)-(6.12) for commodities related to the manufacturing industry. Equations (6.9) and (6.11) represent the line-haul distance
and equations (6.10) and (6.12) denote local travel distance. Summing (6.9)-(6.12) across all railroad terminals \( q \in Q \) yields the total freight delivery cost in the FAZ.

\[
I_{s1}^q = \frac{2\beta_1 (D_w + D_R) \sum_{i=1}^{I_q} \sum_{j=1}^{2} E_{ij}^q d_{qi}}{C \sum_{i=1}^{I_q} \sum_{j=1}^{2} E_{ij}^q}, \quad (6.9)
\]

\[
I_{s2}^q = \frac{0.57N_q^s}{\sqrt{\delta^q_s}}, \quad \text{where} \quad N_q^s = \frac{\beta_1}{d_1} \sum_{i=1}^{I_q} \sum_{j=1}^{2} E_{ij}^q \quad \text{and} \quad \delta^q_s = \frac{N_q^s}{A_q}, \quad (6.10)
\]

\[
I_{m1}^q = \frac{2\beta_2 D_M \sum_{i=1}^{I_q} E_{i3}^q d_{qi}}{C \sum_{i=1}^{I_q} E_{i3}^q}, \quad (6.11)
\]

\[
I_{m2}^q = \frac{0.57N_m^q}{\sqrt{\delta^q_m}}, \quad \text{where} \quad N_m^q = \frac{\beta_1}{d_2} \sum_{i=1}^{I_q} E_{i3}^q \quad \text{and} \quad \delta^q_m = \frac{N_m^q}{A_q}. \quad \text{(6.12)}
\]

In an FAZ, parameters \( \beta_1 \) and \( \beta_2 \) respectively represent percentage of employees in wholesale and retail trade industry and manufacturing industry served by trucking service from the railroad terminals. Recall that \( I_q \) is the total number of census tracts assigned to the railroad terminal \( q \) and \( d_{qi} \) is the distance from the railroad terminal \( q \) to the centroid of the census tract \( i \). The sum of total area assigned to the railroad terminal \( q \) is denoted by \( A_q \). Other variables are the same as before in the regional freight delivery from the truck terminals. Finally, the total freight delivery cost within the FAZ across all industries from all railroad terminals \( (G_R) \) is as follows:

\[
G_R = \sum_{q=1}^{Q} \left( I_{s1}^q + I_{s2}^q + I_{m1}^q + I_{m2}^q \right) \quad \text{(6.13)}
\]
6.3.3 Data preparation and results

The urban spatial structure model provides forecast of employment distribution (i.e., $E_{ij}$) in 30 FAZs that cover 22 major Metropolitan Statistical Areas (MSAs) in the U.S. from 2010 to 2050 in ten-year increments (Song, 1994; Anas et al., 1998; Lee, 2007; Lee and Gordon, 2011); the number of total population in the selected 22 MSAs are greater than or equal to 2,000,000 in Year 2000. Employees in wholesale trade, retail trade, and manufacturing industries are considered in this study, which cover most of the employees across all business sectors in the U.S. We have three urban form scenarios (generated by the urban spatial structure model) as follows: (i) “business as usual” in which the current urban sprawl continues in the U.S., (ii) “polycentric development” in which the development of Central Business District (CBD) follows the current trend but sub-centers also experience high growth, and (iii) “compact development” in which both CBD and sub-centers follow high-growth. Since different global economic growth and environmental policy scenarios (proposed in Chapter 3) do not affect employment distribution but scale up or down the total employee number in the FAZs, only “high GDP growth with business as usual” in Table 3.1 is considered as a representative macro scenario. In most cases, one FAZ includes one MSA (for 17 MSAs). However, three MSAs at Chicago, Philadelphia, and St. Louis MSAs are each associated with two FAZs; New York MSA is associated with three FAZs; and Washington MSA is associated with four FAZs. The urban spatial structure model also provides (i) truck and railroad terminal locations and distance from each terminal to each census tract centroid (i.e., $d_{ki}$ and $d_{qi}$) to obtain $I_k$ and $I_q$ in 30 FAZs, and (ii) area of each census tract to obtain $A_k$ and $A_q$.

The four-step inter-regional freight demand model provides future truck and rail freight demand for each FAZ, from which $\alpha_1$, $\alpha_2$, $\beta_1$, $\beta_2$, $D_w$, $D_R$, and $D_M$ can be obtained using
information in Table 6.1. Inter-regional freight demand is assumed to follow a high GDP growth with business as usual macro scenario. We consider both freight distribution and collection to estimate the total regional freight delivery cost; freight attraction data in destination FAZs are used to distributing freight, while freight production data in origin FAZs are applied to collecting freight. Freight demand originating from and destined to the same FAZ is excluded due to the lack of data. We assume light and medium trucks are used for freight delivery service with an average speed of 30 mph and their capacity (i.e., $C$) is assumed to be 4 tons (FHWA U.S. DOT, 2007a; Davis et al., 2012).

Numerical results from the proposed models are described in Table 6.2. Columns (a) and (b) list the 22 MSAs and the three urban form scenarios considered in this study. Columns (c) and (d) respectively describe the total regional freight delivery cost in miles and ton-miles. Column (d) also shows percentage differences of the total freight delivery cost from the one associated with scenario 3 (i.e., compact development urban form scenario) for each delivery region. Columns (e)-(h) present four different emission estimations from freight activities related to each urban form scenario, which include CO$_2$, NO$_X$, PM, and VOC. Emission factors are obtained from TRL (1999).
<table>
<thead>
<tr>
<th>MSA</th>
<th>Scenario</th>
<th>2020</th>
<th>2021</th>
<th>2022</th>
<th>2023</th>
<th>2024</th>
<th>2025</th>
</tr>
</thead>
<tbody>
<tr>
<td>Los Angeles</td>
<td>1</td>
<td>1,157,935</td>
<td>1,358,324</td>
<td>1,534,257</td>
<td>1,693,859</td>
<td>1,839,171</td>
<td>1,963,126</td>
</tr>
<tr>
<td>Minneapolis</td>
<td>2</td>
<td>1,219,457</td>
<td>1,396,234</td>
<td>1,552,345</td>
<td>1,693,859</td>
<td>1,839,171</td>
<td>1,963,126</td>
</tr>
<tr>
<td>Washington</td>
<td>3</td>
<td>1,219,457</td>
<td>1,396,234</td>
<td>1,552,345</td>
<td>1,693,859</td>
<td>1,839,171</td>
<td>1,963,126</td>
</tr>
<tr>
<td>New York</td>
<td>4</td>
<td>1,219,457</td>
<td>1,396,234</td>
<td>1,552,345</td>
<td>1,693,859</td>
<td>1,839,171</td>
<td>1,963,126</td>
</tr>
<tr>
<td>Pittsburgh</td>
<td>5</td>
<td>1,219,457</td>
<td>1,396,234</td>
<td>1,552,345</td>
<td>1,693,859</td>
<td>1,839,171</td>
<td>1,963,126</td>
</tr>
<tr>
<td>Phoenix</td>
<td>6</td>
<td>1,219,457</td>
<td>1,396,234</td>
<td>1,552,345</td>
<td>1,693,859</td>
<td>1,839,171</td>
<td>1,963,126</td>
</tr>
<tr>
<td>San Francisco</td>
<td>7</td>
<td>1,219,457</td>
<td>1,396,234</td>
<td>1,552,345</td>
<td>1,693,859</td>
<td>1,839,171</td>
<td>1,963,126</td>
</tr>
<tr>
<td>Seattle</td>
<td>8</td>
<td>1,219,457</td>
<td>1,396,234</td>
<td>1,552,345</td>
<td>1,693,859</td>
<td>1,839,171</td>
<td>1,963,126</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>9</td>
<td>1,219,457</td>
<td>1,396,234</td>
<td>1,552,345</td>
<td>1,693,859</td>
<td>1,839,171</td>
<td>1,963,126</td>
</tr>
<tr>
<td>St. Louis</td>
<td>10</td>
<td>1,219,457</td>
<td>1,396,234</td>
<td>1,552,345</td>
<td>1,693,859</td>
<td>1,839,171</td>
<td>1,963,126</td>
</tr>
<tr>
<td>New York</td>
<td>11</td>
<td>1,219,457</td>
<td>1,396,234</td>
<td>1,552,345</td>
<td>1,693,859</td>
<td>1,839,171</td>
<td>1,963,126</td>
</tr>
<tr>
<td>Washington</td>
<td>12</td>
<td>1,219,457</td>
<td>1,396,234</td>
<td>1,552,345</td>
<td>1,693,859</td>
<td>1,839,171</td>
<td>1,963,126</td>
</tr>
</tbody>
</table>

Table 6.2 Intra-regional freight shipment cost and the related emission estimation
Table 6.2 (continued)

<table>
<thead>
<tr>
<th>(a)</th>
<th>Scenario</th>
<th>2020</th>
<th>2025</th>
<th>2030</th>
<th>2040</th>
<th>2050</th>
<th>2020</th>
<th>2025</th>
<th>2030</th>
<th>2040</th>
<th>2050</th>
<th>2020</th>
<th>2025</th>
<th>2030</th>
<th>2040</th>
<th>2050</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Los Angeles</td>
<td>1435</td>
<td>203.5</td>
<td>209.6</td>
<td>346.4</td>
<td>433.0</td>
<td>1408</td>
<td>1997</td>
<td>2664</td>
<td>3400</td>
<td>4250</td>
<td>45</td>
<td>6.3</td>
<td>8.4</td>
<td>10.8</td>
<td>13.5</td>
<td></td>
</tr>
<tr>
<td>Phoenix</td>
<td>1359</td>
<td>162.8</td>
<td>208.4</td>
<td>261.3</td>
<td>322.1</td>
<td>1334</td>
<td>1598</td>
<td>2045</td>
<td>2565</td>
<td>3161</td>
<td>4.2</td>
<td>5.1</td>
<td>6.5</td>
<td>8.1</td>
<td>10.0</td>
<td></td>
</tr>
<tr>
<td>Miami</td>
<td>1317</td>
<td>140.7</td>
<td>183.4</td>
<td>225.3</td>
<td>274.4</td>
<td>1312</td>
<td>1443</td>
<td>1800</td>
<td>2212</td>
<td>2693</td>
<td>4.2</td>
<td>4.6</td>
<td>5.7</td>
<td>7.0</td>
<td>8.5</td>
<td></td>
</tr>
<tr>
<td>Boston</td>
<td>427</td>
<td>514</td>
<td>604</td>
<td>704</td>
<td>815</td>
<td>1.4</td>
<td>1.6</td>
<td>1.9</td>
<td>2.2</td>
<td>2.6</td>
<td>4.6</td>
<td>5.7</td>
<td>7.0</td>
<td>8.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cleveland</td>
<td>418</td>
<td>482</td>
<td>563</td>
<td>653</td>
<td>753</td>
<td>1.3</td>
<td>1.5</td>
<td>1.8</td>
<td>2.1</td>
<td>2.4</td>
<td>3.8</td>
<td>4.4</td>
<td>5.6</td>
<td>7.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detroit</td>
<td>412</td>
<td>489</td>
<td>566</td>
<td>670</td>
<td>760</td>
<td>1.3</td>
<td>1.5</td>
<td>1.8</td>
<td>2.1</td>
<td>2.4</td>
<td>3.8</td>
<td>4.4</td>
<td>5.6</td>
<td>7.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Los Angeles</td>
<td>1398</td>
<td>560</td>
<td>657</td>
<td>768</td>
<td>876</td>
<td>4.3</td>
<td>5.2</td>
<td>6.1</td>
<td>7.3</td>
<td>8.4</td>
<td>4.2</td>
<td>4.6</td>
<td>5.7</td>
<td>7.0</td>
<td>8.5</td>
<td></td>
</tr>
<tr>
<td>Denver</td>
<td>467</td>
<td>529</td>
<td>611</td>
<td>704</td>
<td>806</td>
<td>1.4</td>
<td>1.6</td>
<td>1.9</td>
<td>2.2</td>
<td>2.6</td>
<td>4.6</td>
<td>5.7</td>
<td>7.0</td>
<td>8.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phoenix</td>
<td>433</td>
<td>507</td>
<td>580</td>
<td>663</td>
<td>753</td>
<td>1.3</td>
<td>1.5</td>
<td>1.8</td>
<td>2.1</td>
<td>2.4</td>
<td>3.8</td>
<td>4.4</td>
<td>5.6</td>
<td>7.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miami</td>
<td>1395</td>
<td>1155</td>
<td>1391</td>
<td>1444</td>
<td>1795</td>
<td>2.1</td>
<td>2.8</td>
<td>3.6</td>
<td>4.7</td>
<td>5.9</td>
<td>3.2</td>
<td>4.0</td>
<td>5.1</td>
<td>6.4</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td>Los Angeles</td>
<td>1398</td>
<td>560</td>
<td>657</td>
<td>768</td>
<td>876</td>
<td>4.3</td>
<td>5.2</td>
<td>6.1</td>
<td>7.3</td>
<td>8.4</td>
<td>4.2</td>
<td>4.6</td>
<td>5.7</td>
<td>7.0</td>
<td>8.5</td>
<td></td>
</tr>
<tr>
<td>Denver</td>
<td>467</td>
<td>529</td>
<td>611</td>
<td>704</td>
<td>806</td>
<td>1.4</td>
<td>1.6</td>
<td>1.9</td>
<td>2.2</td>
<td>2.6</td>
<td>4.6</td>
<td>5.7</td>
<td>7.0</td>
<td>8.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phoenix</td>
<td>433</td>
<td>507</td>
<td>580</td>
<td>663</td>
<td>753</td>
<td>1.3</td>
<td>1.5</td>
<td>1.8</td>
<td>2.1</td>
<td>2.4</td>
<td>3.8</td>
<td>4.4</td>
<td>5.6</td>
<td>7.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miami</td>
<td>1395</td>
<td>1155</td>
<td>1391</td>
<td>1444</td>
<td>1795</td>
<td>2.1</td>
<td>2.8</td>
<td>3.6</td>
<td>4.7</td>
<td>5.9</td>
<td>3.2</td>
<td>4.0</td>
<td>5.1</td>
<td>6.4</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td>Los Angeles</td>
<td>1398</td>
<td>560</td>
<td>657</td>
<td>768</td>
<td>876</td>
<td>4.3</td>
<td>5.2</td>
<td>6.1</td>
<td>7.3</td>
<td>8.4</td>
<td>4.2</td>
<td>4.6</td>
<td>5.7</td>
<td>7.0</td>
<td>8.5</td>
<td></td>
</tr>
<tr>
<td>Denver</td>
<td>467</td>
<td>529</td>
<td>611</td>
<td>704</td>
<td>806</td>
<td>1.4</td>
<td>1.6</td>
<td>1.9</td>
<td>2.2</td>
<td>2.6</td>
<td>4.6</td>
<td>5.7</td>
<td>7.0</td>
<td>8.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phoenix</td>
<td>433</td>
<td>507</td>
<td>580</td>
<td>663</td>
<td>753</td>
<td>1.3</td>
<td>1.5</td>
<td>1.8</td>
<td>2.1</td>
<td>2.4</td>
<td>3.8</td>
<td>4.4</td>
<td>5.6</td>
<td>7.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miami</td>
<td>1395</td>
<td>1155</td>
<td>1391</td>
<td>1444</td>
<td>1795</td>
<td>2.1</td>
<td>2.8</td>
<td>3.6</td>
<td>4.7</td>
<td>5.9</td>
<td>3.2</td>
<td>4.0</td>
<td>5.1</td>
<td>6.4</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td>Los Angeles</td>
<td>1398</td>
<td>560</td>
<td>657</td>
<td>768</td>
<td>876</td>
<td>4.3</td>
<td>5.2</td>
<td>6.1</td>
<td>7.3</td>
<td>8.4</td>
<td>4.2</td>
<td>4.6</td>
<td>5.7</td>
<td>7.0</td>
<td>8.5</td>
<td></td>
</tr>
<tr>
<td>Denver</td>
<td>467</td>
<td>529</td>
<td>611</td>
<td>704</td>
<td>806</td>
<td>1.4</td>
<td>1.6</td>
<td>1.9</td>
<td>2.2</td>
<td>2.6</td>
<td>4.6</td>
<td>5.7</td>
<td>7.0</td>
<td>8.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phoenix</td>
<td>433</td>
<td>507</td>
<td>580</td>
<td>663</td>
<td>753</td>
<td>1.3</td>
<td>1.5</td>
<td>1.8</td>
<td>2.1</td>
<td>2.4</td>
<td>3.8</td>
<td>4.4</td>
<td>5.6</td>
<td>7.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miami</td>
<td>1395</td>
<td>1155</td>
<td>1391</td>
<td>1444</td>
<td>1795</td>
<td>2.1</td>
<td>2.8</td>
<td>3.6</td>
<td>4.7</td>
<td>5.9</td>
<td>3.2</td>
<td>4.0</td>
<td>5.1</td>
<td>6.4</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td>Los Angeles</td>
<td>1398</td>
<td>560</td>
<td>657</td>
<td>768</td>
<td>876</td>
<td>4.3</td>
<td>5.2</td>
<td>6.1</td>
<td>7.3</td>
<td>8.4</td>
<td>4.2</td>
<td>4.6</td>
<td>5.7</td>
<td>7.0</td>
<td>8.5</td>
<td></td>
</tr>
</tbody>
</table>
| Denver | 467 | 529 | 611 | 704 | 806 | 1.4 | 1.6 | 1.9 | 2.2 | 2.6 | 4.6 | 5.7 | 7.0 | 8.5 | 118
In most cases, scenario 1 (i.e., business as usual) shows the largest and scenario 3 (i.e., compact development) shows the least total freight delivery cost in miles and ton-miles. The results demonstrate significant advantage of compact and polycentric urban forms which are known to lead to high-density and sustainable urban development by combining residential and commercial zones. Note that the percentage differences of the total freight delivery cost in column (d) grow significantly larger over the years for the FAZs that include Atlanta, Dallas, Denver, Houston, Minneapolis, Phoenix, Portland, Seattle, Tampa, and Washington MSAs, which is caused by two factors: (i) increase in the number of employees far from the truck or railroad terminals, which causes rapid increase in the total long-haul distance in scenario 1, and (ii) faster growth of estimated freight demand than increase in the number of employees in the FAZs. Since the amount of emissions generated from trucks are proportional to the freight delivery activities, the largest and the least amount of air pollutants and greenhouse gases are observed in scenario 1 and scenario 3, respectively.

6.4 Conclusion

In this chapter, the methodology for the intra-regional freight distribution and collection is investigated. A large number of spatially distributed freight demand and supply points in a freight delivery region need to be served, thus this problem is addressed by the large-scale VRP. The ring-sweep algorithm from Newell and Daganzo (1986a) is adopted and modified since each demand/supply point in delivery regions is composed of a number of customers depending on the industries involved. A set of formulae are constructed to estimate large-scale freight delivery efficiency. The total travel distance of a fleet of trucks within an FAZ is obtained as sum of total line-haul distance and total local travel distance. Since it is an asymptotic approximation method
and the number of demand/supply points in our setting is significantly large, output is expected to be quite accurate.

A case study is provided to forecast daily regional freight delivery cost using employment distribution data in 30 FAZs which include 22 major MSAs in the U.S. from 2010 to 2050. Also, future truck and rail freight demand estimation for each FAZ is obtained from the previous four-step inter-regional freight demand forecasting model. We consider employees in wholesale trade, retail trade, and manufacturing industries. Three urban form scenarios are constructed including business as usual, polycentric development, and compact development. Intra-regional freight deliveries from truck and railroad terminals are modeled separately considering different industry types, in which light and medium trucks are assumed to be used to satisfy freight demand. The numerical results are found to effectively estimate future regional freight delivery cost and the related various emissions for each urban form scenario.
CHAPTER 7
URBAN FREIGHT TRUCK ROUTING UNDER STOCHASTIC CONGESTION AND EMISSION CONSIDERATIONS

7.1 Introduction

Traffic congestion in large urban areas is responsible for a significant portion of air pollution and the related human health problems (Copeland, 2011). According to Sjodin et al. (1998), CO and NO\textsubscript{X} emissions from vehicles increase substantially in congested traffic, causing adverse effect on air quality and human health. In addition, greenhouse gas such as CO\textsubscript{2} emission from vehicles also increases greatly when roadway is severely congested, which plays a significant role in global climate change (Barth and Boriboonsomsin, 2008). As such, vehicle emissions in congested urban areas and their environmental impacts should be carefully addressed. It is also well known that trucking as a dominant mode for freight delivery is a major source of air pollutant as well as greenhouse gas emissions (ICF Consulting, 200). Therefore, improvements in fleet operations and efficiencies from the trucking service sector in congested urban roadway networks and the following reduction in vehicle emissions could result in huge benefit with respect to urban air quality, human exposure, and eventually global climate change.

Roadway congestion in urban areas is stochastic since roadway traffic can be affected by many uncertain factors. For instance, unexpected traffic accident or adverse weather condition can cause congestion even during off-peak hours. With the advancement of information

---

4 This chapter has been adapted from “Hwang, T. and Ouyang, Y. (2014c). Urban freight truck routing under stochastic congestion and emission considerations. Working Paper, University of Illinois at Urbana-Champaign, Urbana, Illinois.”
technology, a truck driver is often able to receive real time information about the congestion pattern of roadway networks, so that he/she might be able to avoid heavy congestion by dynamically choosing an alternative path that is more likely to yield a minimum expected cost. As such, the optimal routing decision is solved by the shortest path problem in a stochastic network. Previously, such studies considered travel delay as the only travel cost component and focused on minimizing the expected total travel time. To the best of the authors’ knowledge, the environmental externalities caused by the vehicle movements in a stochastic network have not been extensively considered. Obviously, the amount of vehicle emissions per mile depends on vehicle speed. For example, the minimum CO_2 emission rate occurs when the vehicle speed is around 45 to 50 mph, while a very high or a very low vehicle speed would lead to a much larger amount of CO_2 emission (Barth and Boriboonsomsin, 2008). Thus, maintaining a moderate travel speed throughout a journey would help reduce vehicle emissions although it usually increases total traveling time.

This chapter explicitly considers the environmental cost caused by truck activities in a stochastic shortest path problem setting. Vehicle emissions including CO_2, VOC, NO_X, and PM are introduced as one of the cost components to capture the environmental impacts from a truck delivery system. The capability of delivering at the scheduled time is an important performance metric in the trucking service industry, which receives higher priority. To ensure delivery punctuality, a penalty for a late or early truck arrival at the delivery destination is included in the total cost. Thus, we define the “total” cost as the sum of cost components associated with the freight travel time, emissions, and a penalty for late or early arrival. Due to the emission and penalty cost functions, simply finding the minimum expected travel time solution using the
classical shortest path approaches does not necessarily guarantee the minimum expected total cost solution in our problem.

In this study, we focus on urban transportation networks which can be simplified to graphs composed of node sets (e.g., major intersections) and directed link sets (e.g., urban freeways and arterials) in a time period whose length is comparable of that needed for making a local delivery (e.g., morning rush hour). The traffic congestion state on each network link, represented by the prevailing speed, follows a known probability distribution. Given the origin and destination pair of a freight delivery, a truck driver needs to decide at each network node (i.e., intersection), upon observing the latest realization of congestion pattern on the current roadway, which network link he/she will use for the next part of travel so as to minimize the expected total cost. We develop a mathematical formulation of this problem and design a solution approach based on stochastic dynamic programming. A deterministic shortest path heuristic is also provided to obtain a quick feasible solution even for large-scale networks. A set of test case studies with different network sizes are provided to analyze and compare the performance of the suggested algorithms in existing U.S. urban networks. As such, the work presented in this chapter can be used to find the environmental-friendly shortest path to the destination in a stochastic setting. Eventually, this effort will be useful to help transportation planners and policy makers in both public and private sectors reduce adverse social impacts caused by freight truck emissions in metropolitan areas.

The remainder of this chapter is organized as follows: Section 7.2 formulates the mathematical model to estimate the minimum expected total cost of truck delivery in a transportation network with stochastic congestion. Section 7.3 provides solution methods including a dynamic programming approach and a deterministic shortest path heuristic. Section
7.4 presents the results from a set of numerical examples. Conclusions are presented in Section 7.5.

7.2 Model Formulation

An urban roadway network is represented by a graph $D(V, A)$ where $V$ is a node set and $A$ is a directed link set. A truck needs to deliver freight from its origin $g \in V$ to its destination $s \in V$ through the given network. As soon as a truck driver leaves the origin, he needs to choose his first travel link to minimize the expected total travel cost. During the trip, the truck driver needs to decide the next travel link whenever he arrives at a node, until the travel ends at the destination. We assume congestion on a link is stochastic and the truck speed on the link is uniquely determined by the current congestion state on this link, thus the truck speed on each link is also stochastic. The truck speed on each link follows a certain probability distribution that is fixed throughout the period of routing study (e.g., morning rush hour), but not necessarily identical across the links. Since only major arterial roads and freeways are considered to represent the urban network links, we assume queue formed on a certain link does not spill over into immediate downstream links and congestion states are independent across the links.

Let $d_{ij}$ and $U_{ij}$ respectively denote the length of a link $(i, j) \in A$ (miles) and the stochastic truck speed (mph) on this link. Let $W(\cdot)$ be the emission rate (grams/veh-mile) which is a function of the truck speed $U_{ij}$. In this study, emission rate equations for CO$_2$, VOC, NO$_X$, and PM are obtained from TRL (1999) and a combined truck emission cost equation for link $(i, j) \in A$ is formulated in the case study section; see equation (7.5). We have one binary decision variable $x_{ij}$ to represent whether the delivery vehicle passes link $(i, j) \in A$. Then, our
model can be formulated into the following mathematical problem using the inputs and decision variables described above:

Minimize

\[
E \left[ \alpha \sum_{(i,j) \in A} d_{ij} U_{ij} + \beta \sum_{(i,j) \in A} d_{ij} x_j W(U_{ij}) + P \left( \sum_{(i,j) \in A} \frac{d_{ij}}{U_{ij}} x_{ij} \right) \right],
\]

subject to

\[
\sum_{\{j \mid (i,j) \in A\}} x_{ij} - \sum_{\{j \mid (j,i) \in A\}} x_{ji} = \begin{cases} 
1, & \text{if } i = g, \\
-1, & \text{if } i = s, \\
0, & \text{otherwise}, 
\end{cases}
\]

\( x_{ij} \in \{0,1\}, \quad \forall (i, j) \in A. \)  

The objective function (7.1) minimizes the expected total travel cost of a freight truck along its path from its origin to its destination. The total travel cost is composed of three components including the total delivery time, emissions, and a penalty for late or early arrival at the destination. The parameters \( \alpha \) and \( \beta \) are used to convert travel time (hours) and the amount of total emissions (grams) into monetary value. The penalty cost for late or early truck arrival at the destination is represented by \( P(\cdot) \) which is a function of the total travel time \( t \). We assume an exogenously scheduled travel time \( h \) is given, which represents shippers’ preference on the total delivery time. In this problem, we assign a high penalty if \( t > h \), and a low penalty if \( t < h \), and zero penalty otherwise. The consideration of penalty cost for early arrival is general. In fact, such a penalty cost can be eliminated by setting it to be zero. Constraints (7.2) ensure flow conservations at all network nodes including origin and destination. Finally, constraints (7.3) define the binary decision variables associated with the link set \( A \); that is, \( x_{ij} = 1 \) if link \((i, j)\) is chosen for the delivery, or 0 otherwise.
7.3 Solution Methods

To solve the proposed truck routing problem, we first develop an algorithm based on dynamic programming. Then, a deterministic shortest path heuristic is presented to obtain a feasible solution in a short computation time for large-scale roadway networks.

7.3.1 Dynamic programming approach

We define stages of the system such that stage \( i \) represents node \( i \in V \) in a given network. Also, states of the system can be described such that state \( m \) at stage \( i \) represents truck arrival time \( m \) at node \( i \). For each state \( m \) at each stage \( i \), we have a finite set of decisions on the next link to move onto \( \{(i, j) | (i, j) \in A\} \).

In the rest of this section, an algorithm for obtaining the minimum expected total cost from the origin is described in a dynamic programming framework. Observing structure of the solution algorithm, we need to define countable, explicit states at each stage before the iterative procedures are applied. Thus, a discretization technique is presented to partition a continuous state space and selectively choose the arrival times as states at all stages of the problem.

7.3.1.1 Application of dynamic programming

Let \( T_i \) denote the set of possible arrival times at node \( i \) and obviously, \( T_g = \{0\} \) at origin node \( g \).

Recall that \( U_{ij} \) is a positive stochastic truck speed on a link \((i, j) \in A\) and let \( u_{ij} \) be its realization. The probability density function of the truck speed \( U_{ij} \) on a link \((i, j) \in A\) is described by \( f_{U_{ij}}(u_{ij}) \). Lastly, the minimum expected cost-to-go value of the freight truck from node \( i \) to the destination node \( s \) can be represented as \( Q_i(\delta) \), a function of truck arrival time (i.e., the state)
\( \delta \in T_i \) at node \( i \in V \). Using the variables described above, the algorithm to obtain the minimum expected total cost of the freight truck from its origin node \( g \) to its destination node \( s \) can be written into a recursive Bellman equation with backward induction, as follows (Powell, 2011):

**Step 0. Initialization.** For each node \( i \in V \) in a given network, let

\[
Q_i(\delta) = \begin{cases} 
  P(\delta), & \forall \delta \in T_i, \text{if } i = s, \\
  \infty, & \forall \delta \in T_i, \text{otherwise.}
\end{cases}
\]

Set the candidate list \( C = \{s\} \).

**Step 1.** Select node \( j \) from the set \( C \), which is the first element in the set.

**Step 2.** If \( j \neq g \), conduct the following analysis for all nodes \( i \) where link \((i, j)\in A\) exists:

**Step 2a.** 
\[
Q_i(\delta) = \int_0^\infty f_{u_j}(u_j) \left\{ \frac{d_{ij}}{u_{ij}} + \beta W(u_j) + Q_j \left( \delta + \frac{d_{ij}}{u_{ij}} \right) \right\} du_j, \forall \delta \in T_i.
\]

**Step 2b.** If \( Q_i(\delta) < Q_i(\delta) \) for any \( \delta \in T_i \),

(i) Update \( Q_i(\delta) \leftarrow Q_i(\delta) \).

(ii) If \( i \not\in C \), include node \( i \) to the set \( C \) as the last element in the set.

**Step 3.** Eliminate node \( j \) from the set \( C \). If \( C = \{\} \), terminate the algorithm and \( Q_g(0) \) is the minimum expected total cost of the truck; otherwise, go to Step 1.

In this recursive framework, Step 0 describes the first step of the iteration, in which the minimum expected cost-to-go value at the destination is simply the penalty cost related to the arrival time (i.e., state). We assign significantly large numbers to all other nodes except destination as the initial minimum expected cost-to-go values. In the rest of the iterative steps, the minimum expected cost-to-go value at each node is updated sequentially from the destination.
The optimal solution to the problem is the minimum expected cost-to-go value at the origin at the end of the iteration.

7.3.1.2 Discretization of the state space

The proposed approach based on a dynamic programming has a backward recursion structure and requires explicit computation of the total travel time from the origin to each node. The arrival time at each node can in theory be any continuous value spanning from zero to infinity. Thus, an adaptive discretization approach is used to partition the continuous state space into discrete grids. In this way, we can define countable and explicit states at all stages of the given problem before the backward recursive procedures are conducted.

![Diagram](image)

**Figure 7.1 Expansion and selection of the possible states**

Figure 7.1 describes how the states are selectively chosen in the next stage. In this figure, \( x \)-axis represents a number of stages, nodes in a given network, and \( y \)-axis represents the states, arrival times at each stage. We observe that as the stage proceeds from \( i \) to \( j \) where \((i, j) \in A\),
the number of possible arrival times at a stage \( j \) rapidly increases. Thus, we discretize the state space into mutually disjointed uniform grids with a certain size \( \gamma \) and selectively consolidate all arrival times in each grid as one state (i.e., represented by black dots in Figure 7.1). The possible arrival times at a stage \( k \) where \((j, k) \in A\) can be obtained based on the selected arrival times (i.e., black dots) in stage \( j \). The detailed procedures are described as follows:

**Step 0. Initialization.** Let \( T_i = \{0\} \) if \( i = g \); \( T_i = \{\} \), otherwise. Set the candidate list \( S = \{g\} \). Obtain in-degree of a node \( i \), \( \text{indeg}(i) \), \( \forall i \in V \setminus \{g\} \); \( \text{indeg}(i) = 0 \), otherwise.

**Step 1.** Select node \( i \) from the set \( S \), which is the first element in the set.

**Step 2.** For all nodes \( j \) where link \((i, j) \in A\) exists, if \( \text{indeg}(j) > 0 \), conduct the following analysis:

**Step 2a.** Calculate all possible states (i.e., arrival times) at node \( j \) based on set \( T_i \).

**Step 2b.** Incorporate the result obtained in Step 2a into the current state set at node \( j \), \( T_j \), and selectively choose some elements among the total elements in \( T_j \).

**Step 2c.** If \( j \notin S \), include node \( j \) to the set \( S \) as the last element in the set.

**Step 2d.** Update \( \text{indeg}(j) \leftarrow \text{indeg}(j) - 1 \).

**Step 3.** Eliminate node \( i \) from the set \( S \). If \( S = \{\} \) or \( \text{indeg}(i) = 0, \forall i \in V \), terminate the algorithm and set \( T_i \) contains selectively chosen states at node \( i \), \( \forall i \in V \); otherwise, go to Step 1.

### 7.3.2 Deterministic shortest path heuristic

In many real roadway networks, truck drivers need to select the next travel link in real time (e.g., within several seconds). Thus, this section proposes a heuristic to find not only a feasible solution in a very short computation time even for very large networks, but also an upper bound
to the optimum solution. In this algorithm, the shortest path from the origin to the destination is obtained using the expected link cost which includes only link travel time and the related emissions as cost components. Once the truck reaches the destination, penalty based on the arrival time at the destination is incorporated to generate the expected total cost. The detailed algorithm is described as follows:

Step 0. Initialization. Set $\lambda_i = 1$, $\forall i \in V \setminus \{g\}$ to represent candidate nodes which have been scanned; $\lambda_i = 0$, otherwise to represent a current node. Set the total travel time to a node $i$, $\tau_i = \infty$, $\forall i \in V \setminus \{g\}$; $\tau_i = 0$, otherwise. Set the expected total cost to reach a node $i$, $\mu_i = \infty$, $\forall i \in V \setminus \{g\}$; $\mu_i = 0$, otherwise. Set an index $I = 0$.

Step 1. For each node $i' \in V$ where $\lambda_{i'} = 0$, conduct the following analysis for each link $(i', j) \in A$:

Step 1a. Calculate expected cost for traveling link $(i', j)$, $\pi_{(i', j)}$, as follows:

If node $j \neq s$, $\pi_{(i', j)} = \int_{0}^{\infty} f_{U_{i'}}(u_{i'}) \left( \alpha \frac{d_{ij}}{u_{ij}} + \beta W(u_{i'}) \right) du_{ij}$;

otherwise, $\pi_{(i', j)} = \int_{0}^{\infty} f_{U_{i'}}(u_{i'}) \left( \alpha \frac{d_{ij}}{u_{ij}} + \beta W(u_{i'}) + P \left( \tau_i + \frac{d_{ij}}{u_{ij}} \right) \right) du_{ij}$.

Step 1b. If $\mu_j > \mu_i + \pi_{(i', j)}$, update $\mu_j \leftarrow \mu_i + \pi_{(i', j)}$,

$$\tau_j \leftarrow \tau_i + \int_{0}^{\infty} f_{U_{i'}}(u_{i'}) \left( \frac{d_{ij}}{u_{ij}} \right) du_{ij}, \text{ and } I \leftarrow I + 1, \text{ and let } \lambda_j = 2.$$  

Step 2. For all nodes $i \in V$, if there are nodes with $\lambda_i = 0$, let $\lambda_i = 1$; if there are nodes with $\lambda_i = 2$, let $\lambda_i = 0$.

Step 3. If $I \neq 0$, let $I = 0$ and go to Step 1; otherwise, terminate the algorithm and $\mu_s$ represents the expected total cost.
7.4 Case Studies

The proposed model and solution algorithms are coded on a personal computer with a 3.4 GHz CPU and 8 GB memory. Numerical experiments are first conducted on two simple and small imaginary networks: a 5-node and 7-link network from Powell (2011), as shown in Figure 7.2(a), and a 15-node and 25-link network from Papadimitriou and Steiglitz (1998), as shown in Figure 7.2(b). The focus with these small networks is to investigate the performance of the proposed solution algorithms. Then, the proposed approach is applied to more complex and large-scale urban transportation networks obtained from Bar-Gera (2009): a 24-node and 76-link Sioux Falls network, as shown in Figure 7.2(c), and a 416-node and 914-link Anaheim network, as shown in Figure 7.2(d).

In Figures 7.2(a)-7.2(c), the numbers in circles represent node indices and those near links denote link indices. Due to the complexity of the network, Figure 7.2(d) only shows selected node indices. Origin-destination pairs with relatively long shipment distances are randomly selected and represented by grey circles. Let $t$ be the total travel time and $h$ be the scheduled travel time (hours). We assume that the penalty cost function $P(t)$ is piece-wise linear, as follows:

\[
P(t) = \begin{cases} 
100(t - h), & \text{if } t \geq h, \\
-10(t - h), & \text{otherwise}. 
\end{cases}
\]  

(7.4)
Figure 7.2 Four test networks
(c) 24-node and 76-link Sioux Falls network (Bar-Gera, 2009)

(d) 416-node and 914-link Anaheim network (Bar-Gera, 2009)

Figure 7.2 (continued)
Let $W_1(U_{ij})$, $W_2(U_{ij})$, $W_3(U_{ij})$, and $W_4(U_{ij})$ be the freight truck emission rate functions (grams/veh-mile) for CO$_2$, VOC, NO$_X$, and PM respectively on a link $(i, j) \in A$ (TRL, 1999). Thus, the combined truck emission cost for the link $(i, j)$ shown in equation (7.1) can be represented as follows:

$$\beta W(U_{ij}) = \beta_1 W_1(U_{ij}) + \beta_2 W_2(U_{ij}) + \beta_3 W_3(U_{ij}) + \beta_4 W_4(U_{ij})$$

$$= k_1 + k_2 U_{ij} + \frac{k_3}{U_{ij}} + \frac{k_4}{U_{ij}^2} + \frac{k_5}{U_{ij}^3}$$

(7.5)

where $\beta_1=280$ ($$/\text{tonCO}_2$), $\beta_2=200$ ($$/\text{tonVOC}$), $\beta_3=200$ ($$/\text{tonNO}_X$), and $\beta_4=300$ ($$/\text{tonPM}_{10}$) are parameters that convert the weight of various emissions into the monetary value (Muller and Mendelsohn, 2007; Winebrake et al., 2008). Note that the unit price of PM$_{10}$ is used to estimate the total PM emission cost (due to the data availability). Coefficients in the combined emission cost function (7.5) are $k_1=0.7121$, $k_2=-0.0128$, $k_3=0.0848$, $k_4=6.2065$, and $k_5=2.1976 \times 10^{-6}$. The function hence has a parabolic shape which is minimized when the truck speed is around 44 mph. Also, in equation (7.1) we choose $\alpha=20$ ($$/\text{hr}$) (Bai et al., 2011).

The probability density function of the truck speed on each link is assumed to follow a randomly generated log-normal distribution whose mean and standard deviation are uniformly and independently drawn from [20, 60] mph and [10, 15] mph, respectively. States at each stage are obtained based on the states at the previous stage, using five representative vehicle speeds including minimum (truck speed at the 0.1th percentile), maximum (truck speed at the 99.9th percentile), mean speed, average between the minimum and mean speed, and average between the maximum and mean speed. Let $h$ be the average of possible arrival times at the destination.
For each test network, initial grid size ($\gamma < 1$) is arbitrarily selected and decreased with a small step size (0.005) until convergence.

Table 7.1 summarizes computational results of the case studies. Column (a) presents four test networks and column (b) specifies the two proposed algorithms applied in each test example. While applying dynamic programming approach, the grid size ($\gamma$) at which the objective cost converged is also included. Column (c) represents the minimum expected total costs associated with different algorithms and column (d) shows the percentage differences between the heuristic solutions and the dynamic programming solutions. Column (e) shows computation times associated with these two algorithms.

<table>
<thead>
<tr>
<th>(a) Network</th>
<th>(b) Algorithm</th>
<th>(c) Min. expected total cost ($)</th>
<th>(d) = (Heuristic – DP)/DP (%)</th>
<th>(e) Solution time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-node and 7-link network</td>
<td>Shortest path heuristic</td>
<td>33.33</td>
<td>2.54</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>Dynamic programming</td>
<td>32.51</td>
<td></td>
<td>0.218</td>
</tr>
<tr>
<td></td>
<td>($\gamma = 0.025$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15-node and 25-link network</td>
<td>Shortest path heuristic</td>
<td>20.49</td>
<td>2.61</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>Dynamic programming</td>
<td>19.97</td>
<td></td>
<td>0.725</td>
</tr>
<tr>
<td></td>
<td>($\gamma = 0.030$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24-node and 76-link</td>
<td>Shortest path heuristic</td>
<td>49.55</td>
<td>2.82</td>
<td>0.011</td>
</tr>
<tr>
<td>Sioux Falls network</td>
<td>Dynamic programming</td>
<td>48.19</td>
<td></td>
<td>160.052</td>
</tr>
<tr>
<td></td>
<td>($\gamma = 0.050$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>416-node and 914-link</td>
<td>Shortest path heuristic</td>
<td>132.60</td>
<td>21.02</td>
<td>0.071</td>
</tr>
<tr>
<td>Anaheim network</td>
<td>Dynamic programming</td>
<td>109.57</td>
<td></td>
<td>8,741.145</td>
</tr>
<tr>
<td></td>
<td>($\gamma = 0.040$)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Column (c) of Table 7.1 shows that the minimum expected total cost obtained from the dynamic programming approach is always lower than that from the deterministic shortest path heuristic. The percentage difference in column (d) gradually increases with the network size; the
difference is notable in the largest network tested (i.e., the Anaheim network example). Column (e) shows that the solution times of the deterministic shortest path heuristic are much shorter than those of the dynamic programming approach.

To investigate how emission cost consideration will affect the truck driver’s route decision and eventually influence the expected total cost, two scenarios are designed as follows: (i) a conventional approach in which the emission cost is ignored in the truck driver’s routing decision process (but we do include the emission cost into the total cost evaluation based on the truck driver’s routing decision), and (ii) the proposed dynamic programming approach in which the emission cost is incorporated into the truck driver’s route decision process. Columns (c)-(f) of Table 7.2 show the minimum expected total cost and the monetary values of its itemized components: total travel time, emissions, and penalty. Note that the cost differences between two scenarios in columns (c)-(f) are presented in every third and fourth rows of each network example.

Compared to the solutions obtained from the conventional approach, the proposed approach results in a smaller minimum expected total cost for all study examples, as shown in column (c). In the first, second, and last examples, decrease in the total cost has been mainly due to a large reduction in the emission cost. The third example shows a different case in which the proposed approach has caused a large cost reduction in the total travel time, which is offset by significant cost increase in the penalty. Note that savings from the proposed modeling approach are larger in the last two large-scale urban network examples, in which each truck can save more than $4.5 for that trip. If the proposed environmental friendly routing policy can be implemented for all freight trucks operated in the study area, a large reduction in adverse social cost can be expected. This will improve urban air quality and contribute to enhanced public social welfare.
Table 7.2 Comparisons between two different scenarios

<table>
<thead>
<tr>
<th>(a) Network</th>
<th>(b) Scenario</th>
<th>(c) Min. expected total cost ($)</th>
<th>(d) Travel time ($)</th>
<th>(e) Emissions ($)</th>
<th>(f) Penalty ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-node and 7-link</td>
<td>Conventional approach*</td>
<td>35.45</td>
<td>13.88</td>
<td>13.76</td>
<td>7.80</td>
</tr>
<tr>
<td>network</td>
<td>Proposed approach**</td>
<td>32.51</td>
<td>14.39</td>
<td>10.81</td>
<td>7.30</td>
</tr>
<tr>
<td></td>
<td>Cost difference ***</td>
<td>2.94</td>
<td>-0.51</td>
<td>2.95</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.29%</td>
<td>-3.67%</td>
<td>21.42%</td>
<td>6.41%</td>
</tr>
<tr>
<td>15-node and 25-link</td>
<td>Conventional approach</td>
<td>21.22</td>
<td>6.35</td>
<td>5.90</td>
<td>8.96</td>
</tr>
<tr>
<td>network</td>
<td>Proposed approach</td>
<td>19.97</td>
<td>6.73</td>
<td>4.46</td>
<td>8.78</td>
</tr>
<tr>
<td></td>
<td>Cost difference</td>
<td>1.25</td>
<td>-0.38</td>
<td>1.44</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.87%</td>
<td>-5.97%</td>
<td>24.38%</td>
<td>2.08%</td>
</tr>
<tr>
<td>24-node and 76-link</td>
<td>Conventional approach</td>
<td>52.75</td>
<td>24.57</td>
<td>15.57</td>
<td>12.61</td>
</tr>
<tr>
<td>Sioux Falls network</td>
<td>Proposed approach</td>
<td>48.19</td>
<td>14.52</td>
<td>9.22</td>
<td>24.45</td>
</tr>
<tr>
<td></td>
<td>Cost difference</td>
<td>4.56</td>
<td>10.05</td>
<td>6.36</td>
<td>-11.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.64%</td>
<td>40.90%</td>
<td>40.82%</td>
<td>-93.93%</td>
</tr>
<tr>
<td>416-node and 914-link</td>
<td>Conventional approach</td>
<td>114.38</td>
<td>67.00</td>
<td>47.00</td>
<td>0.39</td>
</tr>
<tr>
<td>Anaheim network</td>
<td>Proposed approach</td>
<td>109.57</td>
<td>67.29</td>
<td>41.81</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>Cost difference</td>
<td>4.81</td>
<td>-0.29</td>
<td>5.19</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.21%</td>
<td>-0.44%</td>
<td>11.04%</td>
<td>-21.76%</td>
</tr>
</tbody>
</table>

*Conventional approach: Ignoring emission cost in the truck driver’s routing decision process;  
**Proposed approach: Considering emission cost in the truck driver’s routing decision process;  
***Cost difference = Cost from the conventional approach - Cost from the proposed approach.

7.5 Conclusion

Freight truck shipments contribute to the largest portion of the total freight delivery in the U.S. metropolitan areas. At the same time, truck fleets have been pointed out as a dominant source of air pollutant and greenhouse gas emissions in various studies and truck emissions significantly affected air quality and global climate change. Thus, it is necessary for the trucking industry to consider the environmental impacts of vehicle delivery activities so as to mitigate the freight truck emissions. This study takes into account various factors including CO₂, VOC, NOₓ, and PM emissions incurred by freight truck movements, as well as penalty cost for late or early
arrivals, as some of the cost components. We then formulate the routing problem into a stochastic shortest path problem.

We focus on urban transportation networks in which traffic congestion on each link represented by the prevailing speed follows a known independent probability distribution. A truck driver needs to decide the next traveling link at each network node so as to minimize the expected total cost which includes total delivery time, emissions, and a penalty for late or early arrival. This problem is formulated into a mathematical model and two solution algorithms including a dynamic programming approach and a deterministic shortest path heuristic are provided. Suggested algorithms are tested on both small, imaginary networks and complex, large scale urban transportation networks. Although solutions from the dynamic programming approach are always lower than those from the deterministic shortest path heuristic, the latter requires much shorter solution times. The results from the proposed approach are compared with those from the conventional approach (in which emission cost is ignored in truck driver’s routing decision process) to investigate the trade-off among different cost components. It was shown that the minimum expected total cost can be reduced by applying the proposed methodology mainly due to a large reduction in the emission cost. The suggested methodology will be useful to transportation planners and policy makers in both public and private sectors, who aim to improve urban air quality and contribute to enhanced public social welfare.
CHAPTER 8
CONCLUSIONS AND FUTURE RESEARCH PLAN

8.1 Conclusions

This Ph.D. research is a part of the collaborative efforts that aim to develop an integrated modeling framework to estimate future emissions from freight transportation systems and its environmental impacts at global, regional, and urban levels based on future economic growth and climate policy scenarios, projections of urban spatial structure, and vehicle emission characteristics. The freight demand has been continuously increasing since the world wide industrialization began and its growth rate has become more rapid in recent several decades. On the other hand, sharp increase of the size of the freight transportation market caused serious air pollution and greenhouse gas emission problems; emissions from various freight shipment modes have been threatening human life. Therefore, the relationship between the freight shipment activities and the environmental problems motivated this dissertation work to investigate the freight transportation systems. Our study focuses on constructing the freight demand models to forecast freight movements between and within the U.S. geographical regions (i.e., FAZs) and by modes (i.e., truck and rail) for future years. The estimated freight shipment activities at the national and regional levels will be useful to predict global, regional, and urban air quality and its impacts on human health and social welfare.

The freight demand models presented in this dissertation are divided into two groups: one for inter-regional freight movements and one for intra-regional freight flow. For the inter-regional freight transportation analysis, the four-step freight demand forecasting framework is adopted, which is composed of trip generation, trip distribution, modal split, and traffic
assignment. The intra-regional freight transportation analysis uses various network modeling and optimization methodologies to construct logistics systems that capture large-scale freight delivery in each freight analysis zone and individual truck routing on stochastic road networks.

In the beginning of this dissertation work, freight trip generation and trip distribution, as part of the four-step inter-regional freight demand model, are investigated. The objective is to forecast future freight demand for all commodity types that will be produced in and attracted to each FAZ and distribute the estimated freight shipment demand among all FAZ O/D pairs. Given future I/O economic value growth factors, the total production and attraction for all commodity types and all FAZs are generated. Then, the RAS algorithm is applied to distribute the future freight demand on all FAZ O/D pairs assuming future freight demand distribution is proportional to the base-year freight demand distribution. The suggested methodology is applied to the case study problem and the future freight demand and its distribution among all O/D pairs can be obtained in a short time.

The connections among freight shipment demand mode choice, crude oil price, and the air quality and climate impacts have been thoroughly investigated in the next chapter. The macroscopic binomial logit market share model is adopted to represent freight transportation mode choice decision between truck and rail, the two dominating freight shipment modes in the U.S., for 10 typical commodity types. As the explanatory variables, freight transportation activities and shipment characteristics such as freight value per unit weight and average shipping distance for each mode as well as crude oil price are included for each commodity type in the suggested model. To estimate the coefficients of the model, not only FAF² and FAF³ databases but also two years of CFS data are combined together to generate the complete four-year dataset. Model validation shows that the suggested model is reliable in forecasting the modal share. As a
result, it was shown that generally trucks tend to be chosen to handle higher value products, for shorter distance delivery. During the study period from 1993 to 2007, it was also shown that the growth in freight tonnage for railroad is much larger than that for trucks probably due to the oil price increase.

As the last step of the inter-regional freight flow modeling framework, truck and rail freight shipment assignment is conducted while network congestion effect is taken into consideration. Since not only “freight shipment mode choices” but also “route choices” can significantly affect regional and urban air quality and eventually human health, it is important to investigate how to determine the freight transportation demand on each link of each modal network between all FAZ O/D pairs. User equilibrium is adopted as the route choice rule. A traditional convex combinations algorithm is applied to solve traffic routing equilibria problem for the truck network. Link cost function is modified to consider traffic volume that already exists on the U.S. highway network. A customized network assignment model is proposed for rail freight shipment demand, where single and double track lines are represented by an equivalent directed graph. A railroad-specific link cost function adjusted for single and double tracks is developed to capture traffic delay, and an adapted convex combinations algorithm is developed to find the shipment routing equilibrium. Our models are applied to an empirical case study for the U.S. highway and rail networks using national freight shipment demand in Year 2007. The algorithms converged within a short time and the optimal freight flow patterns can be found for both truck and rail. The results include the total freight ton-miles traveled on each link and on the whole network, which help build the basis for transportation emission estimations on the national scale.
The intra-regional freight distribution and collection problem is addressed by the large-scale VRP since a large number of demand and supply points are spatially distributed in freight delivery regions. The ring-sweep algorithm (Newell and Daganzo, 1986a) is adopted with modifications. We consider that each demand point in our model represents a number of customers from all industries of our interest. Then, closed-form formulae are constructed to estimate the asymptotic total travel distance of a fleet of trucks (including line-haul distance and local travel distance) which serve all demand in a delivery region. A case study is conducted to forecast daily regional freight delivery cost from both truck and railroad terminals for 30 FAZs which include 22 major MSAs in the U.S. Employees in wholesale trade, retail trade, and manufacturing industries are considered and freight collection as well as distribution cost is included in the total cost. The numerical results are found to effectively estimate future regional freight delivery cost and the related emissions under three urban form scenarios.

Finally, we investigate a microscopic urban freight truck routing problem on a stochastic transportation network. Since freight trucking contributes to the largest share of urban air pollutant and greenhouse gas emissions, freight trucks need to consider their environmental impacts in delivery planning. Therefore, this study takes into account the various emissions including CO₂, VOC, NOₓ, and PM incurred by the freight truck delivery activities in the setting of a stochastic shortest path problem. The penalty for late or early truck arrival at the delivery destination is also introduced to ensure delivery punctuality. The urban transportation networks in this study can be simplified to directed graphs and random congestion state on each network link follows an independent probability distribution. Our model finds the best truck routing on a given network so as to minimize the expected total travel cost. This formulated problem is solved by two solution algorithms including a dynamic programming approach and a deterministic
shortest path heuristic. Numerical examples show that the proposed algorithms perform very well even for the large-size U.S. urban networks.

8.2 Future Research Plan

Many future research opportunities have been identified throughout this dissertation work. This section introduces some of them in the following four areas.

First, in the trip generation and trip distribution procedures, once the newer version of the Freight Analysis Framework database becomes available, we can use a more recent base-year (rather than Year 2007) to improve forecast accuracy.

Second, in the modal split procedure, we would like to update the models once additional freight demand data become available, which will be useful to estimate precise environmental impacts of freight transportation systems.

Third, in theory, capacity expansion in some of the U.S. rail network links will affect modal split. It will be interesting to study how infrastructure investment in the rail network will affect future rail freight demand (i.e., against other modes) in a competitive freight shipment market. If railroads can enhance their level of service (e.g., by infrastructure investment), it may attract more freight shippers to choose rail shipment service. Although it is hard to develop and incorporate this feed-back structure into the current four-step inter-regional freight demand forecasting framework, preliminary result can be shown easily. If some of the rail network links connecting a given O/D pair are expanded, travel time on those routes will decrease. Then, we can adjust the railroad shipping “distance” in the modal split procedure in accordance to the decreased travel time (i.e., in the modal split step, “distance” has been implicitly used as a proxy of travel time assuming speeds are constant for both truck and rail). Then, modal split prediction
can be updated using the new railroad shipping “distance”. In this way, we can estimate change in rail shipment share induced by rail network capacity expansion.

Fourth, time-dependent stochastic congestion state on each link can be applied in the freight delivery routing problem in Chapter 7. Then, the link travel time and the following emissions will be affected by stochastic truck speed on the link as well as truck arrival time at the link starting node. In addition, we can include local and collector roads in the urban transportation networks, in which truck speed on downstream links may be correlated to that on upstream links. Besides, environmental impacts from the transportation activities can be further applied to other stochastic network optimization problems such as stochastic traveling salesman problem.
REFERENCES


   *Papers in Regional Science*, 81(2), 223-246.


APPENDIX A
SOFTWARE MANUAL FOR THE FOUR-STEP INTER-REGIONAL FREIGHT DEMAND MODEL

A.1 Introduction
This dissertation work has developed the freight demand model to forecast freight movements between the U.S. geographical regions, Freight Analysis Zones (FAZs), in a national-wide point of view via two major shipment modes, truck and rail. In this inter-regional freight shipment analysis, the well-known four-step freight commodity transportation demand forecasting model is adopted, which consists of trip generation, trip distribution, modal split, and traffic assignment steps. Based on the freight demand forecasting model suggested in this dissertation, the integrated decision-support software has been developed using a Visual Basic Application (VBA) in the Microsoft Excel 2010. This document is a user manual for the program. We hope that the development and dissemination of such decision support tool will help decision makers and analysts in the freight industries as well as in the government agencies assess atmospheric impacts of freight shipment activities in various future economic growth and climate policy scenarios.

A.2 Software Functionality
This section briefly explains overview of the software. Then, it describes how to use the software in detail with a number of screenshots to help users understand better. To start the software, users can simply double-click the provided Microsoft Excel file, “FourStepSoftware.xlsm”.

164
A.2.1 Overview of the software

Figure A.1 describes an overview of the software which is composed of input, main program, and output. Those three main parts are included in one Excel file.

![Diagram](image)

Figure A.1 Overview of the software

Input is made up of eighteen different worksheets in the software, which includes “Attraction_S1”, “Attraction_S2”, “Attraction_S3”, “Attraction_S4”, “Production_S1”, “Production_S2”, “Production_S3”, “Production_S4”, “2007Demand”, “TruckDist”, “RailDist”, “ModalSplit”, “TruckDemand”, “RailDemand”, “TruckNetwork”, “RailNetwork”, “TruckNode”, and “RailNode” sheets. Detailed explanation about each input worksheet will be provided in Section A.4. Using the input data, the main program coded in VBA generates a set of output from each of the four-step procedures, which will be recorded on seven different worksheets including “Trip_Generation”, “Trip_Distribution”, “Modal_Split”, “TruckResult”, “RailResult”, “TruckMap”, and “RailMap”. Detailed explanation about each output worksheet will be provided in Section A.3.

A.2.2 User interface
Figure A.2 illustrates user interface of the software in “Model” sheet, which also shows default setting of the program. Whenever users start the software or click the “Reset” button (denoted by “1” in Figure A.2) on the user interface, every setting in check boxes and drop-down menus that has been changed by users will be restored to the default states as shown in Figure A.2. Also, any results recorded or saved in any output worksheets will be deleted whenever the software is opened or the “Reset” button is clicked.

![Four-step Inter-regional Freight Commodity Transportation Demand Forecasting Model](image)

**Figure A.2 User interface in “Model” sheet**

After users open the program file, they first need to choose “future year” and “economic growth and climate policy scenario” from the two drop-down menus denoted by “2” and “3” in Figure A.2 respectively. In case of future year, users can select any year from 2010 to 2050 in five-year increments. Also, users can select one among four different economic growth and climate policy scenarios which include “1.High_Business_as_usual”, “2.High_Climate_policy”, “3.Low_Business_as_usual”, and “4.Low_Climate_policy”. Therefore, total 36 combinations of
the final results can be obtained based on different future years and different economic growth
and climate policy scenarios.

Users have options whether to execute the entire four-step freight demand model at one
go or to run the individual step one at a time. This function can be selected by checking or
unchecking the check box denoted by “4” in Figure A.2. Once users check the check box,
“Entire Model” button denoted by “5” in Figure A.2 will be activated, and four buttons named
with each step denoted by “6” in Figure A.2 will be deactivated. If users uncheck the check box,
the “Entire Model” button will be deactivated, and only “Trip Generation” button will be
activated among the four individual step buttons. Lastly, if users check the check box denoted by
“7” in Figure A.2, all the input worksheets will be appeared. The default setting is to hide all of
them since they are not necessary information for users in executing the software.

A.2.2.1 Entire four-step model procedure
If users select to run the “Entire Model”, entire four-step model will be executed without breaks.
After every step is completed automatically from the trip generation to the network assignment,
results will be recorded in all the seven output worksheets and users will see the pop-up message
box as shown in Figure A.3.

Figure A.3 Pop-up window after completing the entire four-step model procedure
Figure A.3 shows computation time of the program. After users click the “OK” button to remove the message box, they will encounter the user interface as shown in Figure A.4.

Figure A.4 User interface after completing the entire four-step model procedure

In this stage, users can see results in each individual output worksheet associated with each of the four steps. If users want to save the results, they need to copy and paste the results on another Excel file or create copies of the output worksheets in the software with different worksheet names. Note that the names of all the original output worksheets should remain the same. Otherwise, users will encounter an error and cannot run the model.

A.2.2.2 Individual step of the four-step model
If users want to check the result of each step before proceeding to the next step and execute each step one by one, they can choose to run each step individually by clearing the check box denoted by “4” in Figure A.2. Then, the “Entire Model” button will become deactivated, and only “Trip
Generation” button among the four individual steps will be activated to be able to be clicked. After the trip generation step is completed, the message box as shown in Figure A.5 will appear to notify users to continue to the next step, trip distribution.

![Figure A.5 Pop-up window after completing the trip generation step](image)

If users click the “OK” button in Figure A.5, this message box will disappear and users can see only “Trip Distribution” button is activated in the user interface. The result from the trip generation step will be recorded in an output worksheet named “Trip_Generation”. Also, from hence, users cannot change the selected scenario from the drop-down menus and cannot choose to run the entire model. If users want to change any of these options, they need to click the “Reset” button and start the analysis from the beginning. Once users click the “Trip Distribution” button, the program will run the second step of the entire procedure, and the message box in Figure A.6 will appear to let users know the trip distribution step is completed after a while.

![Figure A.6 Pop-up window after completing the trip distribution step](image)
The message box in Figure A.6 will disappear when users click the “OK” button, and only “Modal Split” button will be activated in the user interface. The result from the trip distribution step will be recorded in an output worksheet named “Trip Distribution”. Once users click the “Modal Split” button, the program will run the third step of the entire procedure and the message box in Figure A.7 will appear to let users know that the modal split step is completed after a while.

Figure A.7 Pop-up window after completing the modal split step

If users click the “OK” button in Figure A.7, this message box will disappear and users will notify that only “Network Assignment” button is activated in the user interface. The result from the modal split step will be recorded in an output worksheet named “Modal_Split”. Once users click the “Network Assignment” button, the program will execute the last step and the message box in Figure A.8 will appear after the analysis is complete. Note that the running time of the last step may be much longer than that of the previous steps.

Figure A.8 Pop-up window after completing the network assignment step
The message box in Figure A.8 will disappear when users click the “OK” button, and users will see the user interface as shown in Figure A.9. Results from the network assignment step will be recorded in four different output worksheets named “TruckResult”, “RailResult”, “TruckMap”, and “RailMap”.

![Four-step Inter-regional Freight Commodity Transportation Demand Forecasting Model](image)

Figure A.9 User interface after completing all four steps individually

In Figure A.9, all drop-down menus and buttons are deactivated except “Reset” button. If users want to re-run the software with different settings, they need to click “Reset” button to make the user interface default setting. Note that once users click the “Reset” button, all the results obtained previously will be deleted. Thus, users need to copy and paste the results in a separate Excel file or create copies of all output worksheets in the software with different sheet names. Note that the names of all original output worksheets should remain the same. Otherwise, users will encounter an error and cannot run the model.
A.2.2.3 Aborting the program

Whenever users want to abort the program while it is running, they can simply force the program
to close in Windows Task Manager. However, to provide the opportunities for users to abort the
program smoothly, message box as shown in Figure A.10 will pop-up periodically while the
program is running. If users want to continue the program running, they can either click the “No”
button in Figure A.10 or just leave the message box without clicking anything since it will
disappear automatically in 5 seconds and the program will run again. If users want to abort the
program, they can click the “Yes” button in Figure A.10.

![Figure A.10 Message box to abort the program](image)

If users selected to execute the entire model, every intermediate result will be eliminated
and the settings changed by users will return to the default states after aborting the program. If
users selected to run the model by individual step, the software will go back to the default states
after aborting the program if the program is in either trip generation or trip distribution steps. If
the program is in the modal split step, all results will be deleted except the ones in
“Trip_Generation” and “Trip_Distribution” output worksheets after aborting the program. If the
program is in the network assignment step, the software will delete all results except the ones in
“Trip_Generation”, “Trip_Distribution”, and “Modal_Split” output worksheets after aborting the
program.
A.3 Output Worksheets

Results from different steps will be recorded in different output worksheets. For example, the results from trip generation, trip distribution, and modal split steps will be recorded in the “Trip_Generation”, “Trip_Distribution”, and “Modal_Split” worksheets, respectively. The results from truck freight demand network assignment will be recorded in the “TruckResult” worksheet, and those from the rail freight demand network assignment will be recorded in the “RailResult” worksheet. Also, figures generated in the “TruckMap” and “RailMap” worksheets will visually describe the results in “TruckResult” and “RailResult” worksheets.

A.3.1 Output from the trip generation step

Figure A.1 represents sample result from the trip generation step in the “Trip_Generation” worksheet. It shows two groups of columns: one for freight production and the other for freight attraction in each FAZ. In freight production group, origin (i.e., FAZ), commodity type, amount of freight production in terms of tonnage, and the total freight production across all commodity types in each FAZ are described. Similarly, in freight attraction group, destination (i.e., FAZ), commodity type, amount of freight attraction in terms of tonnage, and the total freight attraction across all commodity types in each FAZ are described.
Figure A.11 Sample result from the trip generation step in the “Trip_Generation” worksheet

<table>
<thead>
<tr>
<th>Origin</th>
<th>Commodity Type</th>
<th>Production (ton)</th>
<th>Total Production (ton)</th>
<th>Destination</th>
<th>Commodity Type</th>
<th>Attraction (ton)</th>
<th>Total Attraction (ton)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>2156693.407</td>
<td>72198437.37</td>
<td>1</td>
<td>1</td>
<td>1812931.702</td>
<td>9917030.52</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3058187.921</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2467921.022</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>14389356.68</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>17931404.62</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>17213927.36</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>15771433.69</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>4587813.458</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>4940353.42</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>4108611.136</td>
<td>1</td>
<td>6</td>
<td>6</td>
<td>7915344.542</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>13812536.48</td>
<td>1</td>
<td>7</td>
<td>7</td>
<td>15566269.73</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>781003.5549</td>
<td>1</td>
<td>8</td>
<td>8</td>
<td>1130305.421</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>10440827.53</td>
<td>1</td>
<td>9</td>
<td>9</td>
<td>10974040.65</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>1097790.402</td>
<td>1</td>
<td>10</td>
<td>10</td>
<td>1098962.348</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>479484.39338</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2984939.006</td>
<td>4120598.35</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>945132.2047</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>184819.5874</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>10616501.8</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>494518.966</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>11621761.15</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>12941054.36</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>3950047.191</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>4942167.536</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>5379811.546</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>2615112.74</td>
<td>2</td>
<td>7</td>
<td>7</td>
<td>4327719.556</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>2903134.653</td>
<td>2</td>
<td>8</td>
<td>8</td>
<td>115476.1615</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>2771484.897</td>
<td>2</td>
<td>9</td>
<td>9</td>
<td>5001341.666</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>2</td>
<td>576222.9567</td>
<td>2</td>
<td>10</td>
<td>10</td>
<td>534811.3122</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>3</td>
<td>14336965.93</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1881763.43</td>
<td>21814954.1</td>
</tr>
<tr>
<td>23</td>
<td>3</td>
<td>3208344.048</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>7014275.031</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>3</td>
<td>3464212.39</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>15750758.25</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>3</td>
<td>3113361.79</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>13518862.55</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>3</td>
<td>4897896.35</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>1389672.18</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>3</td>
<td>7856691.46</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>6228212.36</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>3</td>
<td>3544411.62</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td>18277805.56</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>3</td>
<td>400530.786</td>
<td>3</td>
<td>8</td>
<td>8</td>
<td>530525.88</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>3</td>
<td>1532741.55</td>
<td>3</td>
<td>9</td>
<td>9</td>
<td>19797424.83</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>3</td>
<td>2013572.012</td>
<td>3</td>
<td>10</td>
<td>10</td>
<td>1912832.257</td>
<td></td>
</tr>
</tbody>
</table>

A.3.2 Output from the trip distribution step

Figure A.12 shows sample result from the trip distribution step in the “Trip_Distribution” worksheet. It includes origins and destinations of freight flow, different commodity types for each O/D pair, freight demand in terms of tonnage, and total freight demand across all commodity types for each O/D pair.
Sample result from the modal split step is described in Figure A.13, a screenshot of the “Modal_Split” worksheet. It includes origins, destinations, different commodity types for each O/D pair, the total freight demand in tons for each commodity type and each O/D pair, and shares of truck and rail. Note that the sum of freight demand for truck and rail is the same as the total freight demand.
Figure A.13 Sample result from the modal split step in the “Modal_Split” worksheet

A.3.4 Output from the network assignment step

Sample results from the truck and rail freight demand network assignment step are shown in Figures A.14(a) and A.14(b), screenshots of the “TruckResult” and “RailResult” worksheets, respectively. In case of truck, the result consists of link number, link origin and destination nodes, link distance in miles, total traffic volume on each link which includes background traffic as well as assigned traffic on the link, assigned traffic volume on each link, link cost which represents link travel time in hours, average vehicle speed, and ton-miles on the link. In case of rail, structure of the final result is almost the same as that of the truck freight demand assignment result except the former additionally includes assigned traffic volume for each commodity type. Note that in the “RailResult” worksheet the total traffic volume on two links which connect the same pair of nodes with opposite directions are the same, and the assigned flow on one link becomes the opposite direction traffic flow on another link.
Figure A.14 Sample results from the truck and rail freight demand network assignment step

The user equilibrium results from the truck and rail freight demand network assignment described in the “TruckResult” and “RailResult” worksheets in Figures A.14(a) and A.14(b) are visualized in the “TruckMap” and “RailMap” worksheets, respectively. Figures A.15(a) and A.15(b) illustrate screenshots of the sample results in the “TruckMap” and “RailMap” worksheets. In these figures, sum of the assigned traffic flows on two separate links connecting the same pair of nodes with opposite directions are classified by various line thicknesses and
colors as shown in the legends. Also, blue circles in the figures represent centroids of the 120 FAZs.

(a) “TruckMap” worksheet

(b) “RailMap” worksheet

Figure A.15 Visualizations of the truck and rail freight demand network assignment step
A.4 Input Worksheets

Each step in the four-step inter-regional freight demand forecasting model requires different input worksheets to conduct the analysis. The software has eighteen input worksheets in total. Among them, “Attraction_S1”, “Attraction_S2”, “Attraction_S3”, “Attraction_S4”, “Production_S1”, “Production_S2”, “Production_S3”, “Production_S4”, and “2007Demand” worksheets are provided to complete the trip generation and trip distribution steps; “TruckDist”, “RailDist”, and “ModalSplit” worksheets are input for the modal split step; “TruckDemand”, “RailDemand”, “TruckNetwork”, “RailNetwork”, “TruckNode”, and “RailNode” worksheets are used for the network assignment step. It is recommended for users not to change or update the values provided in all input worksheets, otherwise it may generate unexpected errors or wrong results.

A.4.1 Input for the trip generation and trip distribution steps

Estimates of commodity attraction and production (in terms of monetary values) under different economic growth and environmental regulation scenarios for all FAZs are recorded in the “Attraction_S1”, “Attraction_S2”, “Attraction_S3”, “Attraction_S4”, “Production_S1”, “Production_S2”, “Production_S3”, and “Production_S4” worksheets. “S1”, “S2”, “S3”, and “S4” in worksheet names represent different scenarios. They are composed of ten columns; the first column shows data in Year 2007, a base-year. The input data for future years from 2010 to 2050 in five-year increments are included from the second to the last columns. The “2007Demand” worksheet includes the base-year freight demand distribution among 120 FAZs. It is composed of origin, destination, commodity type, and amount of freight flow in terms of tonnage between all shipment O/D pairs in the U.S.
A.4.2 Input for the modal split step

The “TruckDist” and “RailDist” worksheets include distance in miles between each O/D pair using truck and rail, respectively. They are 120-by-120 matrices in which each row represents origin zone and each column describes destination zone. The “ModalSplit” worksheet contains commodity value per unit weight for all commodity types and crude oil price per barrel. There are four groups of rows, each of which is related to the economic growth and climate policy scenario (i.e., “S1”, “S2”, “S3”, and “S4”) as shown in the worksheet.

A.4.3 Input for the network assignment step

The “TruckDemand” and “RailDemand” worksheets contain truck and rail freight shipment demand and highway and rail network information needed for the freight demand network assignment analysis. They include stopping criteria of iteration in solving the user equilibrium problem, the total link and node number for each modal network, and the total O/D pair number for truck and rail, respectively. In case of truck freight shipment demand, origin, destination, and number of vehicles to be assigned on the highway network for each O/D pair are shown in the “TruckDemand” worksheet. In case of rail freight shipment demand, origin, destination, and number of trains to be assigned on the rail network for both each O/D pair and commodity type are included in the “RailDemand” worksheet. Note that the freight shipment demand for truck and rail are not fixed values but will be changed according to the different future years and scenarios. The “TruckNetwork” worksheet contains detailed highway network information. It includes link number, link origin and destination nodes, link distance in miles, link free flow travel time in hours, link capacity, coefficients of the link cost function, and background traffic volume on the link. The “RailNetwork” worksheet also contains detailed rail network information.
information. It includes link number, link origin and destination nodes, link distance in miles, link free flow travel time in hours, coefficients of the link cost function, and background traffic volume which is initially zero across all links. Lastly, the “TruckNode” and “RailNode” worksheets contain $x$ and $y$ coordinates of all nodes in the highway and rail networks, respectively.