Estimating the Location of Private Infrastructure for Delivery and Pick-Up Operations in Dense Urban Areas

Jose Luis Machado Leon

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Science in Civil Engineering

University of Washington

2018

Committee:

Anne V. Goodchild, Chair

Edward McCormack

Program Authorized to Offer Degree:
Civil and Environmental Engineering

University of Washington

Abstract

Estimating the Location of Private Infrastructure for
Delivery and Pick-Up Operations in Dense Urban Areas

Jose Luis Machado Leon

Chair of the Supervisory Committee:
Associate Professor Anne V. Goodchild
Civil and Environmental Engineering

The growth of home deliveries, lower inventory levels and just-in-time deliveries drive the fragmentation of freight flows, increased frequency, more delivery addresses and smaller volumes. This leads to trucks inefficiently loaded and consequently more trucks in the road contributing to the growing congestion in cities. According to a study by INRIX and the Texas Transportation Institute, travelers in the U.S. are stuck 42 hours per rush hour commuter in their cars in 2014, that is twice what it was in 1982 and the problem is four times worse than in 1982 for cities of 500,000 people or less [28]. At the same time, a historical lack of integration of the freight transportation system into city planning efforts has left local governments unprepared. Under these circumstances, there is growing need for best practices for freight planning and management in U.S. cities.

There is anecdotal evidence that the lack of areas for trucks to park and load/unload freight is one of the main causes of an inefficient urban freight parking infrastructure that leads to illegal parking and more congestion. The problem of lack of parking for freight

loading/unloading has been studied with a focus on on-street parking. Meanwhile, the contribution of areas out of the public right of way (i.e. private) such as loading bays in buildings has not benefited from research. More importantly, the location and features of private freight parking are often unknown by local governments due to their private character.

This thesis presents the first predictive tool to estimate the presence of private freight loading/unloading infrastructure based on observable characteristics of property parcels and their buildings. The predictive model classifies parcels with and without these infrastructures using random forest, a supervised machine learning algorithm. The model was developed based on a rich geodatabase of private truck load/unload spaces in the City of Seattle and the King County tax parcel database.

The performance of the random forest model was evaluated through cross-validated estimates of the test error. The distribution of the outcome variables is unbalance with over 90% of parcels without private freight infrastructure. To consider the problem of unbalance sample, the optimum model was set to maximize the area under the ROC curve (AUC). The authors investigated the confusion matrix and the model classifier was design to balance the sensitivity and specificity of the model. Model results showed AUC of 81.5%, a true positive rate of 82.1% and a misclassification error of 22.5%.

This research provides an assessment tool that reduces the field work required to develop a quality inventory of private freight loading/unloading infrastructure by targeting the parcel stock and making data collection methods more effective. Local governments can use this research to inform efforts to revise and update delivery operations and regulations of truck parking and loading.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Number | Page

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my advisor, Prof. Anne Goodchild and my fellow researchers at the Supply Chain Transportation and Logistics Center for the opportunity of working on high quality research projects with a team of incredibly hard-working and fun people.

I would specially like to express my very profound gratitude to my parents and my sister for providing me with unfailing support and continuous encouragement throughout my life. This accomplishment would not have been possible without them. Thank you.

I would also like to express my gratitude to my girlfriend Gabriela Giron, thank you for all your love and support.

Chapter 1

INTRODUCTION

Urban Freight Transport has recently received increasing attention from research and is more present in the local policy context [5]. Cities have been pressed in the last decades to pay attention to urban freight negative impacts, mainly due to increasing congestion and pollution from increased urbanization. Urban freight transportation is predominantly road oriented and despite its relatively low share as part of total traffic, it contributes disproportionately to air pollution, noise pollution, congestion, traffic accidents and reduced accessibility of urban areas [16, 27].

Deliveries in urban areas are recently becoming more and more challenging. The growth of home deliveries, lower inventory levels and just-in-time deliveries drive the fragmentation of freight flows with increased frequency, more delivery addresses and smaller volumes. This leads to more presence of vehicles because they are inefficiently loaded. [1, 18].

Beyond its economic significance, urban policy-making has paid relatively little attention to urban freight planning when compared to the existing extensive research into public and private passenger transport. There is a lack of systematic methodology for studying and planning urban freight activities [17]. The author of this thesis has only been able to find four cities in the U.S. that have advanced city freight master plans and half of these cities developed their plans after 2015: Portland (2006) [24], Atlanta (2008) [23], Seattle (2016) [29] and Washington D.C. (2017) [11]. These advanced plans were motivated to better understand freight-related issues and are evidence of the urgent need of cities to cope with the growing negative impacts of urban freight transport.

Freight cannot simply be viewed as logistics networks operating in abstract urban space, but must rather in a dynamic relationship in which urban places, land use, and connections mutually shape the other [10]. The most advanced city freight planning documents in the U.S. are stand-alone documents exclusively dedicated to goods movement, which elevate freight to the same level as the other modal plans. Advanced city freight planning documents evaluate freight needs and propose policies and actions integrated into a multi-modal comprehensive strategy such as a general transportation plan. Additionally, they are comprehensive, that is, consider several freight modes in geographical areas at the city level greater than a neighborhood or an urban industrial district.

There is anecdotal evidence that the lack of delivery areas is one of the main causes of illegal parking and congestion [1]. Moreover, it is stated that illegal parking drives costs of operations through parking citations, unsafe situations and more congestion [31]. Along this line, cities with advanced city freight plans such as Portland and Seattle consider the evaluation and update of on-street and off-street truck loading regulations and operations to address the negative impacts of the freight loading and parking issue. However, current approaches to study this problem are limited because they only focus on on-street parking for loading and unloading of goods [2, 12, 22]. Meanwhile, the contribution of delivery areas out of the public right of way (i.e. private) such as loading bays in buildings has not benefited from research.

More importantly, data about private freight parking supply is not available as there is not a U.S. city that maintains an open database of truck load/unload spaces that also includes private spaces. The information regarding private freight loading / unloading facilities location, design and capacity is an important piece of the urban goods movement system. This information is necessary to understand and evaluate its capacity and performance and, therefore, to make informed decisions regarding urban freight planning, especially in dense and constrained urban areas.

For these reasons, this research aims to answer the following research question: Can cities use data readily available to estimate the location of off-street freight loading/unloading parking in dense urban areas with a reasonable level of accuracy? The objective of this research is to provide a validated model to estimate the location of private freight load/unload infrastructure in downtown Seattle. The model is developed using a random forest approach, a supervised machine learning technique based on decision trees. Additionally, we use observable characteristics of the built environment for the prediction.

The rest of the document is structured as follows: Chapter 2 includes a literature review on sources of information about urban off-street freight parking infrastructure and predictors of their location, Chapter 3 includes data and methods, Chapter 4 describes the results and Chapter 5 includes the conclusions of this work.

Chapter 2

LITERATURE REVIEW

2.1 Sources of private freight parking supply data in urban areas

There are three main ways to collect data about private freight parking in cities: property history research, direct observations and survey of property managers.

There are different ways to conduct property history research such as investigating building permits, parcel and property tax data, city directories, fire insurance maps and land use codes. Parcel refers to parcels of real property or tax parcels that are used for locating property within a region. These data sources can be fragmented and the data warehoused in different formats. In the case of a data driven city such as Seattle, all this information must be obtained from up to six different databases. Additionally, property records are available online or in the format of microfilm copies depending on the date they were issued [3].

Land use data is one of the data types most relevant to the presence of private parking supply and it is commonly available on-line in GIS databases together with property parcel data. Cities such as New York, Washington D.C., Seattle, Los Angeles, Chicago, Dallas, Boston and Atlanta require buildings of certain size and use to have freight parking facilities [20, 22]. For instance, Seattle considers three categories of loading demand based on land-use: high, medium, and low demand, and has a different set of thresholds and requirements for loading berths depending on the demand category (Figures 2.1 and 2.1). Seattle's land use code also considers width and space for measurability [25]. Width requirements are segmented according to demand and the largest weekly delivery truck.

Different to history property research, three studies were based on direct observations in New York City [22], Toronto [21] and Seattle [13], respectively, and one study used surveys of property manager in New York City [20].

Morris [20] surveyed property managers to document the presence of private freight parking in a sample of 82 commercial office buildings in New York Central Business District (CBD). The study presents the buildings categorized into Premier Class A buildings and a second group Class B buildings that are a step below based on Building Owners and Managers Association (BOMA)' s standards. According to Morris [20] , BOMA ranks buildings based on a combination of factors such as rent per square foot, building finishes, system standards and efficiency, building amenities, location/accessibility and market perception. This study shows the prevalence of private freight parking spaces in higher quality buildings (see Table 1).

Additionally, most of the 59 Premier Class A buildings in Morris' study were newer, bigger and had more rentable floors than Class B buildings. Most Class A buildings were constructed between 1950 and 1985 whereas most Class B buildings dated between 1910 and 1929. Class A buildings had in average 730,000 sq ft while Class B only averaged 200,000 sq ft. Last, 87% of Class A buildings had 16 or more rentable floors while only 57% of Class B had that number of floors.

Figure 2.1: Seattle code' s loading bay requirements for buildings [25]

Table A for Section 23.54.035

Type of Use	Square Feet of Aggregate Gross Floor Area	Required Number of Loading Berths
Low Demand	40,000 to 60,000	1
	60,001 to 160,000	2
	160,001 to 264,000	3
	264,001 to 388,000	4
	388,001 to 520,000	5
	520,001 to 652,000	6
	652,001 to 784,000	7
	784,001 to 920,000	8
	For each additional 140,000	1 additional berth
Medium Demand	10,000 to 60,000	1
	60,001 to 160,000	2
	160,001 to 264,000	3
	264,001 to 388,000	4
	388,001 to 520,000	5
	520,001 to 652,000	6
	652,001 to 784,000	7
	784,001 to 920,000	8
	For each additional 140,000	1 additional berth
High Demand	5,000 to 16,000	1
	16,001 to 40,000	2
	40,001 to 64,000	3
	64,001 to 96,000	4
	96,001 to 128,000	5
	128,001 to 160,000	6
	160,001 to 196,000	7
	For each additional 36,000	1 additional berth

Figure 2.2: Seattle code' s categories of land uses based on level of demand [25]

Table for Section 23.54.035 A

Low Demand	Medium Demand	High Demand
Animal services	Agricultural uses	Airport, land-based
Business incubator	Airport, water-based	
	Assisted living facilities	
Business support services	Automotive parts or accessory sales	Cargo terminals
Car wash	Eating and drinking establishments	Commercial laundries
Custom and craft work	Heavy commercial services except commercial laundries and construction services	Construction services
Entertainment uses	Institute for advanced study	Food proccessing for human consumption
Gas station	Mini-warehouse	High-impact uses
Helistop and heliport	Mortuary services	Hospitals
Institutions, except hospitals and institutes for advanced study	Passenger terminal	Manufacturing
Lodging	Personal and household retail sales and services	Outdoor storage
Marine retail sales, services	Recycling collection stations	Recycling center (separate facilities)
Medical services	Research and development laboratory	Sale of heating fuel
Offices	Sales, service and rental of equipment	Sales, service and rental of commercial equipment and construction materials
Personal transportation services	Transit vehicle base	Salvage yard
Sales and rental of motorized vehicles	Utilities	Warehouse
Towing services	Vehicular repair, major and minor	Wholesale showroom

Table 2.1: Breakdown of freight accommodation and building class from Morris' data [20]

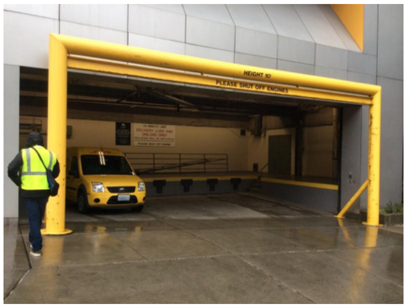
Freight Accommodation	**Premier Class A buildings (%)**	**Class B buildings (%)**
Operating freight docks	80	9
Separate freight doors	12	74
Neither a dock nor a separate freight door	8	17

Lubinsky et al. [22] used direct observations to document types of off-street freight infrastructure commonly found in the city. This study identifies up to eight different categories of these infrastructures for which they use the umbrella terms loading berths and loading docks interchangeably. The study points out the lack of comprehensive review of off-street freight parking design standards and the negative impact of missing data regarding this infrastructure. However, the study did not result in a geodatabase that gathered location or attributes of this infrastructure.

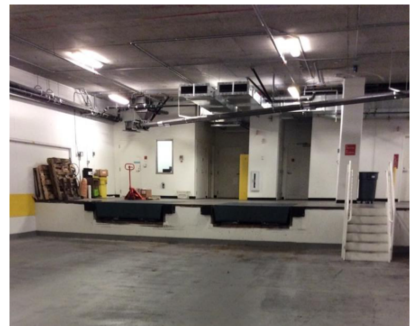
Nourinejad et al. [21] uses location data of off-street freight parking facilities to evaluate truck parking policy impacts in urban areas. The infrastructure data is obtained from an inventory of truck parking supply in a 5x4 city block area of Toronto CBD. The information about off-street parking supply is limited to the location and type of the facilities. The inventory includes approximately 37 loading bays in addition to other off-street public freight parking options such as surface parking, public parking and alleys in the area, however, there is not further information available about the facilities.

Goodchild et al. [13] developed a data collection method to map private freight parking in urban areas through direct observations. This work includes a Seattle case study that resulted in an inventory of 246 facilities in four centric neighborhoods of the city. This work shows the largest database of these infrastructures, which took approximately 210 person-hours just for data collection tasks in field by hourly staff. The inventory includes the geolocation of these facilities and detailed features such as width and height of entrances and exits, number of parking spaces and the presence of an elevated platform to facilitate loading/unloading operations from trucks (i.e. dock). Goodchild et al. [13] built a typology of private freight parking facilities that includes the following definitions:

- **Loading bay (Figure 2.1)**. An enclosed space inside the building with an entrance/exit point (e.g. roll up doors, garage doors) that act as a continuation of the upper parts of the building. This space is partially or completely dedicated to unloading and loading activities. It has entrances and exits greater than 8 feet x 8 feet for commercial vehicles. Loading bays can have loading docks and truck parking spaces with or without access to a loading dock.



(a)

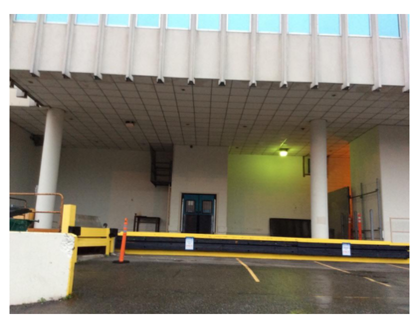


(b)

Figure 2.3: Examples of loading bays in downtown Seattle. (a) Vehicle door of loading bay, (b) detail of loading dock inside loading bay

- **Exterior loading dock (Figure 2.1)**. A loading dock that is located outside of building exterior wall. Exterior loading docks can be completely open to the sky or partially or completely covered by a canopy or upper part of the building. Additionally, exterior loading docks can also include inside loading platforms, where trucks dock the cargo compartment to a dock door.

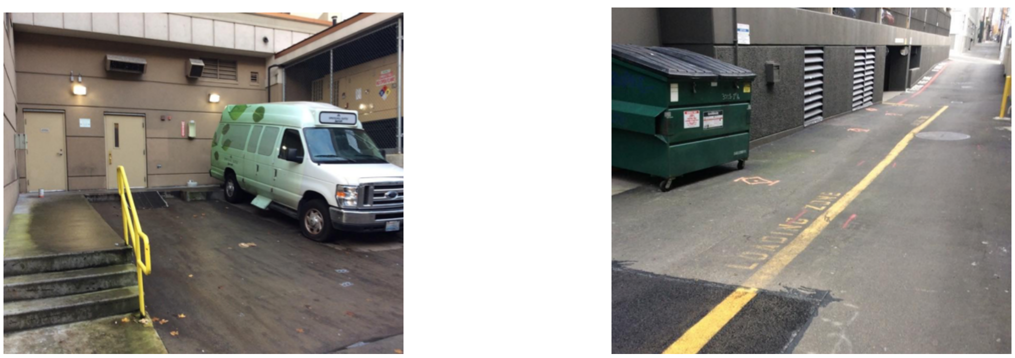
(a)



(b)

Figure 2.4: Examples Of exterior loading docks in downtown Seattle. (a) Loading dock with platform inside building, (b) loading dock with platform outside the building

- **Exterior loading area (Figure 2.1)**. Space for loading and unloading located outside of building exterior walls and without a loading dock. Exterior loading zones can be completely open to the sky, or partially or completely covered by a canopy or upper part of the building.



(a) (b)

Figure 2.5: Examples of exterior loading areas in downtown Seattle. (a) Loading area accessed from street, (b) loading area in alleyway

To summarize, there is a handful of studies that collect data about private freight parking infrastructure through direct observations or surveys of building managers. Property history research is an alternative data collection method but there are not examples of private freight parking databases built through this method.

The existing databases include facilities in dense urban areas such as CBDs and the study areas are different in size, from an area of 20 city blocks in Toronto to an area of 522 city blocks in Seattle. The comparison between studies is difficult as these infrastructures receive different names without an explicit definition including terms like loading bays, loading berths, freight docks and freight doors. The first typology of these facilities is built by Goodchild et al. [13] who coin three definitions based on observable characteristics of the infrastructures: loading bay, exterior loading area and exterior loading dock.

The current approaches to build detailed inventories of these infrastructures are limited. The cost and time required by surveys through direct observations can result in barriers for local governments aiming to build city wide inventories. Inspecting 522 city blocks to document 246 facilities can take up to 210 person-hours just of field data collection tasks; additional costs may come from building the survey form, buying measurement instruments, data quality control and workforce management [13]. The property history research approach may be impractical at a city-wide scale because it is potentially time consuming due to the fragmentation of data sources. The use of partial data such as land use may also make this method to be inaccurate compared to surveys of building managers and direct observations. The author will further explain the inaccuracies related to the use of land use data in Section 2.2.

There are not applications of predictive methods to identifying the location of private freight parking across a city's entire building stock. This approach has helped to overcome similar data collection barriers in analogous urban policy challenges. For instance, the application of building energy conservation measures requires costly and time consuming audits to determine building eligibility and implementation costs of these measures.

Marasco et al. [19] explored ways to utilize available data to identifying and predicting building retrofit opportunities to improve building energy conservation. Predictions can increase the impact of audits and provide alternative rapid assessment tools. In a similar manner, there is an opportunity to use predictive approaches for the identification of potential private freight parking location.

2.2 Predictive factors of the location of private freight parking

Land use data could be a partial predictor of the presence of private freight parking. In fact, there is anecdotal evidence supporting this hypothesis; Pivo et al. [26] found that truck drivers in Seattle relate off-street truck parking spaces to specific types of buildings and land uses such as urban office building ground-floor or underground facility, urban grocery or commercial store, and suburban grocery or commercial store.

On the other hand, land use data may not fully explain the variance in the location of these facilities because:

- A significant fraction of the built environment may be made of buildings constructed before zoning laws (i.e. pre-zoning). For instance, 85% of buildings in New York are pre-zoning [22].

- The absence of strong regulations or enforcement regarding loading space requirements in buildings could have led to fail at ensuring the development of these spaces.

- Owners and managers of commercial real estate have not highly prioritized freight deliveries and do not see private freight parking as a marketing tool [20]

- Some existing buildings could have been retrofitted to build or remove private freight parking.

Morris [20] looked into CBD's commercial buildings of different quality levels and their freight loading/unloading facilities (Table 1). This work considers the following list of building features without statistically testing their relationship with the presence of freight parking facilities:

- Size

- Year built

- Number of floors

- Rent per square foot
- Building finishes
- System standards and efficiency
- Building amenities
- Location/accessibility
- Market perception

To conclude, the results of this literature review indicate that building and parcel data should be used as alternative data sources for the prediction of the location of private parking for loading and unloading because they contain detailed change information. They are available in digital format and can be updated on a regular basis as more local government agencies utilize GIS for creating and maintaining parcel maps.

Chapter 3
DATA AND METHODS

3.1 Data selection and preparation

These data sources used in this work are the following:

- **Private freight infrastructure (PFI) survey in Seattle:** Goodchild et al. [13] built this dataset that includes loading bays and docks in four neighborhoods of Seattle: Uptown, South Lake Union, Downtown, International District and Pioneer Square.

- **King County Parcel dataset:** includes data about the property parcel, buildings and accessory structures and parcels' history of changes [7].

The author selected a subset of the parcel dataset that matched the extension of the PFI dataset. Furthermore, some processing of the data was required to set the parcel as the unit of analysis and have values of the variables in the dataset at the parcel level. This was necessary because the relationship between the parcel datasets is a one-to-many, that is, each parcel can have one or more private freight infrastructures, accessories and changes in their history.

The author used the geolocation of private freight infrastructures in the PFI dataset to create a categorical variable (INF_YN) indicating if the parcel had these facilities with two values of the variable: 1 – Yes, 2 – No. Figure 3.1 shows the result of this transformation for each of the 2282 parcels in the PFI survey area.

King County parcel dataset include and extensive list of parcel features. Based on the literature review and a preliminary analysis ten variables were preselected and grouped in the following categories (see Table 3.1):

- **Land use:** refers to the best use of the parcel based on appraisal judgment, zoning and history of the parcel. It is used as a proxy of the requirements of freight parking space that come from city's land use codes and the functionality of the property.

- **Size:** includes variables related to the size of the parcel lot and its buildings.

- **Accessory features:** describe the presence of structures in the parcel that in some cases are related to parking infrastructure.

- **History of changes:** describe the history of physical and legal changes made to the parcel. This variable helps to account for possible retrofits to buildings that resulted in freight parking infrastructures being built or removed.

Figure 3.1: Parcels with and without private freight infrastructure in the PFI database study area

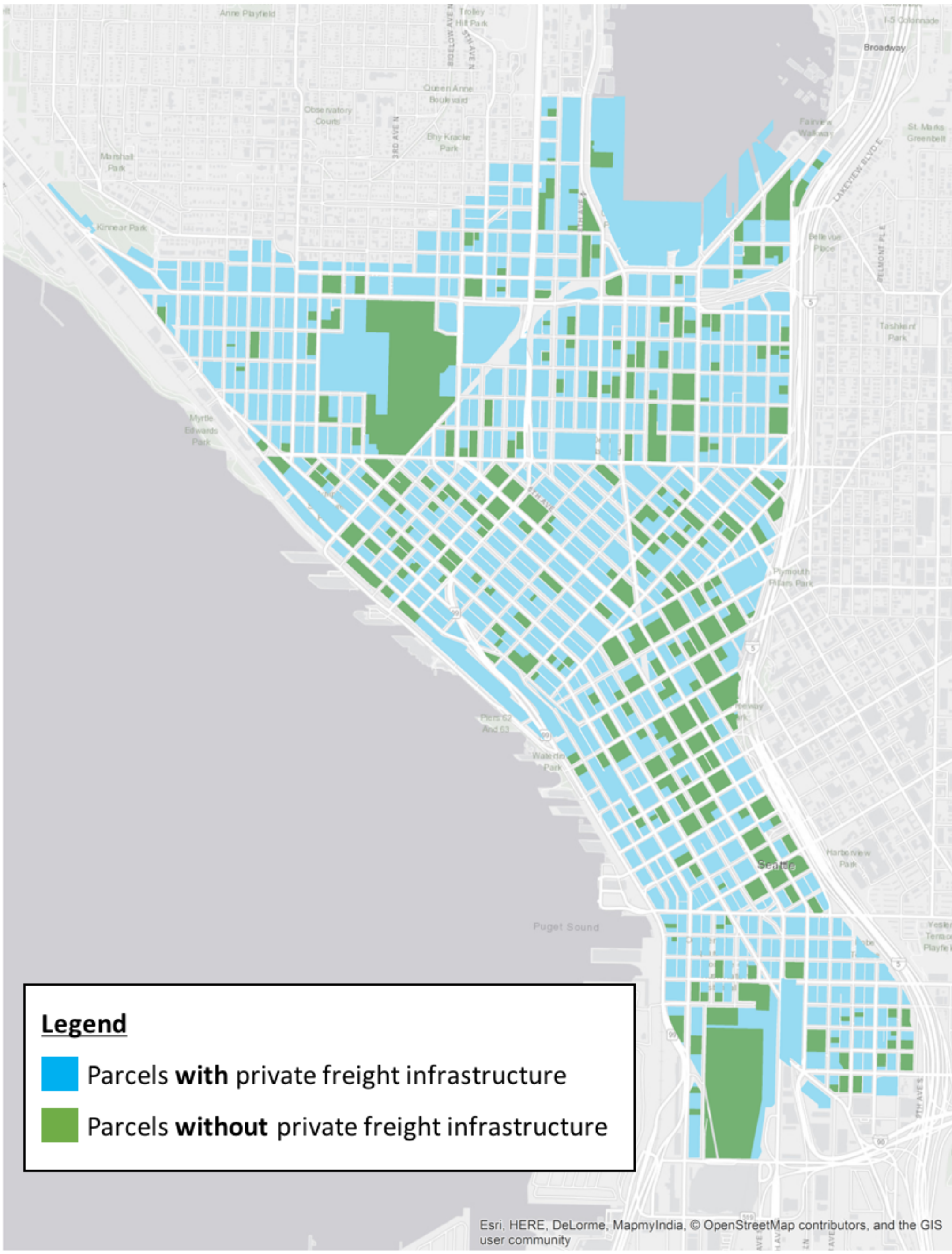


Table 3.1: Variables included in the analysis

Category	**Variable**
Land use	Best use as if vacant is retail or wholesale
	Best use as if vacant is temporary lodging
	Best use as if vacant is manufacturing
Size	Mean gross area of commercial buildings in parcel (sqft)
	Square footage of lot (sqft)
	Mean number of stories of commercial buildings in parcel
History of changes	Count of changes in history of the parcel
Accessory features	There is an accessory structure related to covered or open parking
	There is an accessory structure related to other types different to covered or open parking

3.2 Descriptive analysis

Table 3.2 describes a cross-classification between land uses of the parcel and the presence of freight parking infrastructures. The categories that showed the highest proportion of parcels with off-street infrastructure were Cultural (42.86%), Amusement/Entertainment (10%) and Temporary lodging (17.65%).

According to Seattle's code (Figure 2.1), high demand uses include warehousing/retail, manufacturing and commercial services and have higher requirements of loading space. Parcels in the study area with these uses showed private freight infrastructure in 11.90% of commercial services, 10.71% of manufacturing and 4.35% of retail and wholesale parcels.

The author assumes that can improve the accuracy of estimates by developing models specific to similar land uses. This approach has been previously used in the building energy performance literature that includes models specific to uses such as commercial and residential [30, 19]. For these reason, this thesis is focused on building a predictive model specific to parcels with a predominant commercial use including commercial services (1227 parcels), warehousing/retail (276 parcels) and manufacturing (28 parcels). To achieve this, the parcels with these three uses were selected to calibrate the model resulting in a sample of 1531 parcels of which 161 (10.52%) have private freight parking infrastructure.

Table 3.2: Cross-classification of parcel use type and presence of freight infrastructure

Best Use as if Vacant	**Private Freight Parking Infrastructure**	
	No	**Yes**
Cultural	57.10%	42.90%
	(8)	(6)
Amusement/Entretainment	90.00%	10.00%
	(9)	(1)
Temporary lodging	82.30%	17.70%
	(42)	(9)
Mixed use	84.20%	15.80%
	(48)	(9)
Commercial service	88.10%	11.90%
	(1081)	(146)
Manufacturing	89.30%	10.70%
	(25)	(3)
Regional land use	90.30%	9.70%
	(28)	(3)
Retail/Wholesale	95.60%	4.40%

Table 3.2 continued from previous page

Best Use as if Vacant	**Private Freight Parking Infrastructure**	
	No	**Yes**
	(264)	(12)
Multi-family dwelling	96.00%	4.00%
	(432)	(18)
Single family	100.00%	0.00%
	(71)	(0)
Duplex	100.00%	0.00%
	(2)	(0)
Triplex	100.00%	0.00%
	(2)	(0)
Group residence	100.00%	0.00%
	(4)	(0)
Mobile home	100%	0.00%
	(1)	(0)
Educational service	100.00%	0.00%
	(3)	(0)
Park/Recreation	100.00%	0.00%
	(14)	(0)

Table 3.2 continued from previous page

Best Use as if Vacant	Private Freight Parking Infrastructure	
	No	**Yes**
Grand total	90.80%	9.20%
	(2034)	(207)

Descriptive statistics of the subset of 1531 commercial parcels is shown in Table 3.3, including all the independent variables preselected as predictive factors in the random forests model.

The variables related to size the parcel and its commercial buildings indicate a positive relationship between size and presence of private freight parking infrastructures. The stock of commercial buildings in the group of parcels with these infrastructures averaged a mean square footage six times bigger than the building stock in parcels lacking private freight parking. Moreover, 75% of the parcels with these infrastructures had commercial buildings that averaged 12 stories high or less while the equivalent number of stories for parcels without private freight parking was only three. In a similar manner, parcels with private freight parking were approximately three times bigger in average terms than parcels without these facilities.

There were eight parcels with private freight parking that did not have commercial buildings. There are two possible explanations for this, the first being that other types of buildings such as residential structures may also be equipped with these infrastructures. It is also possible that the King County parcel dataset is not up to date and newly constructed commercial buildings with private freight parking have not been entered in the database yet. For the scope of this thesis, the cases of parcels with private freight parking and no commercial buildings will be considered in the model as if the corresponding commercial building gross square footage and number of stories were zero. This supports the idea that these facilities may not only be found in commercial buildings.

The information related to accessory structures is insufficient to derive meaningful conclusions. Less than 60 parcels had accessory structures of any type including covered and open parking and other types of structures. Furthermore, only eight parcels with private freight parking showed accessory structures related to parking. Due to the limited sample size, these variables were ruled out of the final model.

Information about changes in the history of the parcel does not show determinative trends. Parcels with private freight parking tend to have slightly more changes in their history with 46.85 changes in average, approximately six more changes than parcels without these infrastructures. 75% of parcels with private freight parking had 52 changes or less whereas in the case of parcels without private freight parking this number was 45 changes. Table 3.4 shows the different types of changes and their frequencies for the 1531 commercial parcels considered in the analysis. The most frequent type of change was char update that relates to updates of characteristics of the parcel.

Table 3.3: Statistics of predictors for three groups: all sample of parcels, parcels with PFI and without PFI

Variable	**Statistic**	**All**	**Parcels without PFI**	**Parcels with PFI**
INF_YN (presence of PFIs)	Sample Size	1,531	1,370	161
Square footage of lot	Min.	0	0	1,885.00
	1st Quartile	6,480.00	6,480.00	14,110.00
	Median	10,440.00	8,734.00	25,920.00
	3rd Quartile	17,470.00	14,400.00	42,840.00
	Max.	381,900.00	381,900.00	177,700.00
	Mean	16,610.00	14,600.00	33,710.00
	Sandard Dev.	22,719.96	21,098.41	28,258.13
Mean gross area of commercial buildings	Min.	0	0	0
	1st Quartile	0	0	39,200.00
	Median	9,520.00	7,021.00	157,100.00
	3rd Quartile	51,360.00	33,540.00	380,300.00
	Max.	1,952,000.00	1,545,000.00	1,952,000.00
	Mean	71,480.00	46,920.00	280,500.00
	Sandard Dev.	179,463.14	126,826.70	348,182.90
	Min.	0	0	0

Table 3.3 continued from previous page

Variable	Statistic	All	Parcels without PFI	Parcels with PFI
Mean number of stories of commercial buildings in parcel	1st Quartile	0	0	3
	Median	2	1	6
	3rd Quartile	4	3	12
	Max.	76	46	76
	Mean	3.62	2.81	10.51
	Sandard Dev.	6.89	5.16	13.14
Count of changes in history	Min.	9	9	13
	1st Quartile	34	34	40
	Median	39	39	45
	3rd Quartile	46	45	52
	Max.	104	94	104
	Mean	40.97	40.28	46.85
	Sandard Dev.	10.13	9.69	11.78
Accessory Parking	Count	57	49	8
Accessory Other	Count	50	47	3

The age of the commercial building stock was not considered as a predictor in the analysis because the use of this variable would result in a significant reduction of the sample size.

Table 3.4: Types of changes in parcel's history and their frequency

Types of change	**Frequency**
Char Update	57,653
Transfer	139
New Parcel	4
Levy Code	2,632
Board Order	1,218
Seg Merge	524
Legal Change	443
Kill	11
Unkill	32

Age of commercial buildings is information not applicable for parcels without these types of buildings, that is a total of 450 commercial parcels and 30% of the sample. Under the assumption that the input of missing values is not an applicable approach in this case, 450 parcels would be excluded from the analysis if commercial building age is considered.

Table 3.5 shows statistics of the age distribution of commercial buildings. These results are informative but will not be further considered in the analysis with random forests for the reason stated above. The parcels with private freight infrastructure tend to have newer commercial buildings with 50% of these parcels having buildings built on 1978 or previously and 75% having buildings from 2001 or earlier years. On the other hand, 50% of the parcels without private freight parking had buildings that dated from 1932 or earlier and 75% had buildings from 1970 or before that year.

Table 3.5: Age distribution of commercial buildings stock

Variable	**Statistic**	**All**	**Parcels without PFI**	**Parcels with PFI**
Oldest year built of commercial buildings in parcel	Sample size	1,081	928	153
	Min.	1900	1900	1900
	1st Quartile	1917	1915	1929
	Median	1945	1932	1978
	3rd Quartile	1979	1970	2001
	Max.	2016	2016	2016
	Parcels without commercial buildings	450	442	8

3.3 Classification trees and random forests

Roughly speaking, there are two steps in the process of building a classification tree (James et al., 2013):

- A. The predictor space, that is a set of possible values for X1, X2, . . ., Xp, is divided into J distinct and non–overlapping regions, R1,R2, . . . , RJ .

- B. For every observation that falls into the region Rj , the same prediction is made, which is simply the most commonly occurring class of training observations in Rj.

Bootstrapped aggregating, also called as bagging, is a machine learning technique to reduce the variance between models calibrated with different samples. Bagging is also used to avoid overfitting. Though the mathematical theory and proof is beyond the scope of this master thesis, bagging trees are the result of averaging the prediction of multiple decision trees calibrated with different bootstrapped training sets. Bootstrapped samples or training sets are obtained by uniformly sampling form the original training set with replacement. The author of this thesis refers the reader to James et al. [14] for a more thorough explanation of the machine learning techniques used in this research.

Bagged trees are not effective in the case of having a strong predictor with moderately strong predictors because all bagged trees tend to use the strong predictor and be similar to each other. This is problematic because averaging many highly-correlated quantities does not effectively reduce variance as averaging many uncorrelated quantities.

Random forests provide an improvement over bagged trees by decorrelateing the trees. Bagged trees are also built with random forest but each time a split in a tree is considered, a random sample of m predictors is chosen as split candidates from the full set of p predictors. The split can use only one of those m predictors. This process allows to make the average of the resulting trees less variable and hence more reliable.

One disadvantage of the collection of bagged classification trees is that these are much more difficult to interpret than a single tree. However, we can obtain an overall summary of the importance of each predictor by adding up the total amount that the Gini index is decreased by splits over the given predictor, averaged over all B trees.

3.4 Cross-validation method and unbalanced data

There are two parameters that can be tuned in the calibration of random forests:

- **m predictors:** number of predictors used to build decision trees on bootstrapped training samples

- **B trees:** number of trees built on bootstrapped training samples.

The number of trees B is not a critical parameter with bagging; using a very large value of B will not lead to overfitting. In practice, we use a value of B sufficiently large that the error has settled down, which was set to 500 trees. Regarding the tuning of number of parameters m, the author chose its optimum value as the one that maximizes the area under the Receiver Operating Characteristic (ROC) curve, also called AUC, through 10-fold cross-validation.

The distribution of the outcome variables is unbalance with over 90% of parcels without private freight infrastructure. The investigation of this aspect is important because even if the model shows high accuracy (e.g. over 90% of cases correctly classified) this could only be a manifestation of the distribution of the outcome. In other words, a model that always predicts INF_YN' = 0 would have a 90% accuracy and could be wrongly considered as a valid model.

To consider the problem of unbalance sample, the author will investigate the confusion matrix to evaluate the True Positive Rate (TPR) and True Negative Rate (TNR). TPR, also called Sensitivity is the proportion of positive outcomes correctly classified as such, in other words, it indicates the proportion of parcels with private freight infrastructure that are correctly classified by the model. True negative rate (TNP) or Specificity relates to the proportion of negatives correctly classified as such.

Chapter 4
RESULTS AND DISCUSSION

Table 4.1 shows the 10-fold cross-validation with random forests. Model parameters are estimated for each of the 10 folds and averaged to obtain the final value. The results include the following information:

- **Area under the ROC curve (AUC):** To obtain the optimum value for the m parameter, The author tested four possible values: p, p/2, p/3 and p/4 where p is the number of predictors of the model (i.e. 9 predictors). The optimum value of m maximized the average AUC between the 10 folds of the cross-validation process. Therefore, optimum m was p/2 = 4 predictors and resulted in a AUC of 0.815.

- **Optimum probability cutoff point (Opt Cutoff).** Indicates the probability threshold of the classifier that maximizes the TPR and TNR. Our results indicate that a cutoff threshold of 9.4% would result in a TPR of 82.1% and a TNR of 72.9%. That is, the classifier would result in a 22.5% misclassification rate if probability estimates of 9.4% or greater are labeled as parcels with private freight parking.

- **TPR and TNC with a cutoff point of 50%.**TPR and TNR are also estimated for a traditional cutoff threshold of 50% and show that the improvements resulting of estimating the optimum cutoff point are large. TPR would only be 20.6%, that is, the model would miss 79.4% of the commercial parcels with private freight infrastructure.

Table 4.1: Cross-validation results with random forests

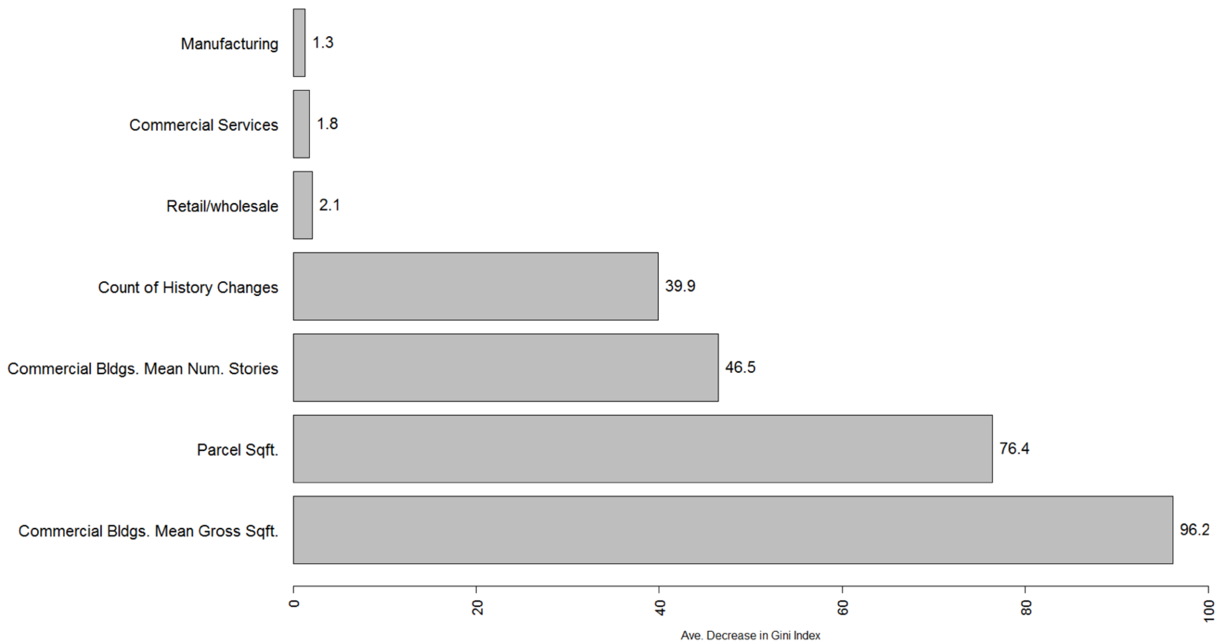
m =	**AUC**				**TPR (cutoff 50%)**	**TNP (cutoff 50%)**	**Opt Cutoff**	**TPR Opt**	**TNR Opt**
	p	**p/2**	**p/3**	**p/4**	**p/2**	**p/2**	**p/2**	**p/2**	**p/2**
Fold 1	0.815	0.829	0.853	0.691	0.188	0.979	0.102	0.812	0.75
Fold 2	0.882	0.807	0.755	0.846	0.111	0.927	0.092	0.833	0.758
Fold 3	0.677	0.853	0.797	0.809	0.25	0.992	0.09	0.85	0.791
Fold 4	0.787	0.925	0.774	0.824	0.1	0.961	0.144	1	0.827
Fold 5	0.781	0.82	0.874	0.847	0.278	0.941	0.09	0.889	0.695
Fold 6	0.836	0.729	0.86	0.752	0.267	0.959	0.092	0.667	0.69
Fold 7	0.676	0.763	0.907	0.812	0.111	0.979	0.07	0.722	0.704
Fold 8	0.772	0.832	0.749	0.862	0.308	0.979	0.032	0.923	0.607
Fold 9	0.838	0.799	0.81	0.8	0.333	0.926	0.146	0.733	0.785
Fold 10	0.843	0.79	0.738	0.857	0.111	0.992	0.08	0.778	0.68
Mean	0.791	0.815	0.812	0.81	0.206	0.963	0.094	0.821	0.729

The results also indicate what variables where more important in the prediction of private freight parking location based on Gini Index values see Figure 4. Variables related to size were the most important predictors having in order of decreasing importance the mean gross square footage of commercial buildings, area of the parcel lot and average number of stories of commercial buildings. This result is in line with Morris [20] results that suggested a relationship between the size of commercial buildings and the presence of freight facilities.

The count of changes in the history of the parcel followed size variables in terms of importance. The interpretation of this result is not straightforward and further research is necessary to derive further conclusions. The history of changes in the parcel could relate to physical changes that upgraded the structures in the parcel to have freight facilities. However, there is little information available to further explore this hypothesis including a detailed description of change history types and if the private freight parking resulted from new construction or a structure retrofit. On the other hand, the amount of changes could be correlated to the number of buildings in the parcel and their age, with older structures having more changes. Morris [20] results suggested that newer structures tend to have more freight facilities. However, the number and age of structures in the parcel was not controlled in the random forest model and the author cannot further explore this hypothesis.

Land use variables had almost no importance indicating that compared to size and change history variables, it is irrelevant if the parcel is best use as for manufacturing, commercial services or retail/wholesale.

Figure 4.1: Predictor importance from random forests results



The model developed with this research should be applicable to locations other than Seattle's CBD. Transferability may hold reasonably well between CBDs with similar land use, density, contemporary economic growth, and building construction patterns, and urban development.

Seattle's CBD is a product of the pre-World War II period, when metropolitan employment was more concentrated. After World War II, cities were built to function around the automobile and suburbanization exploded in late 1940s and 1950s. Based on the percentage of dwelling units constructed before 1940, Seattle ranks the number 25 out of US metropolitan areas over 1 million population in 2010 [8]. Seattle's share of pre-WWII dwelling units in the metropolitan area is 11.1% and in the historical core is 29.9%. Similarly, New Orleans, LA, Portland, OR, and Detroit, MI are other pre-WWII historical core municipalities with equivalent age distribution of dwelling units.

Table 4.2 and Table 4.3 show a comparison of US CBDs in metropolitan areas over 1 million population in 2010 conducted by Cox [9] based on data from the American Community Survey 2006-2010. Seattle CBD's employment in 2010 was 163,830 people and represented 9.7% of the employment in the metropolitan area (see Table 4.2). That places Seattle in the 8th position in terms of relative size of CBD closely followed by New Orleans, LA, (9.6%) and Denver, CO (9.6%) [9]. Additionally, 37% of Seattle's CBD employment were transit commuters, placing Seattle in the 7th position in terms of the share of transit commuting to CBD's jobs (see Table 4.3). This rank also includes Portland, OR, another Pacific North West city like Seattle, in the 10th position with 27% transit share between CBD's commuters.

Seattle is experiencing exceptional growth and is one of the fastest-growing cities in the US. The City is adding nearly 20,000 city population a year since 2010 resulting in 703,352 inhabitants and a 3.1% growth rate between 2014 and 2015 [4, 6]. This places Seattle as the nation's 10th-densest, following Miami, Los Angeles and Long Beach. Seattle has had a strong economy since 1980s, with the rise of Microsoft. Nowadays, the city is a company town with companies such as Amazon, which has more office space than the next 40 biggest employers combined [6]. This growth impacts the urban fabric and Seattle's downtown Skyline, which as of mid-2016, had 16 projects of new construction skyscrapers [15].

To conclude, the selection of candidate CBDs to transfer this model should be based on a comparison of metrics such as CBD's employment share in the metropolitan area, transportation mode share of CBD commuters, age of urban fabric, density and economic and population growth. These metrics represent key factors that influence decision making regarding the urban fabric.

Table 4.2: Rank of U.S. metropolitan areas based on CBD's employment share

Rank	**Metropolitan Area**	**CBD Employment**	**Metropolitan Area Employment**	**CBD Share**
1	New York, NY-NJ-PA	1,981,305	8,983,981	22.10%
2	San Francisco-Oakland, CA	297,420	2,069,673	14.40%
3	Washington, DC-VA-MD-WV	379,215	2,892,018	13.10%
4	Richmond, VA	56,815	498,175	11.40%
5	Chicago, IL-IN-WI	500,450	4,407,655	11.40%
6	Boston, MA-NH	242,900	2,279,803	10.70%
7	Hartford, CT	62,520	589,357	10.60%
8	Seattle, WA	163,830	1,690,490	9.70%
9	New Orleans. LA	49,250	510,454	9.60%
10	Denver, CO	119,565	1,252,889	9.50%
11	Louisville, KY-IN	54,245	591,742	9.20%
12	Cleveland, OH	85,235	958,330	8.90%

Table 4.3: Rank of U.S. metropolitan areas based on transit mode share between CBD commuters

Rank	Metropolitan Area	CBD Transit Commuters	CBD Employment	Transit Market Share
1	New York, NY-NJ-PA	1,517,749	1,981,305	76.60%
2	Chicago, IL-IN-WI	287,245	500,450	57.40%
3	Boston, MA-NH	126,735	242,900	52.20%
4	San Francisco-Oakland, CA	150,724	297,420	50.70%
5	Washington, DC-VA-MD-WV	178,500	379,215	47.10%
6	Philadelphia, PA-NJ-DE-MD	105,869	239,625	44.20%
7	Seattle, WA	60,604	163,830	37.00%
8	Pittsburgh, PA	29,920	92,010	32.50%
9	Minneapolis-St. Paul, MN-WI	31,320	99,315	31.50%
10	Portland, OR-WA	22,970	85,195	27.00%
11	Los Angeles, CA	30,709	136,585	22.50%
12	Denver, CO	23,660	119,565	19.80%

Chapter 5

CONCLUSIONS

Data collection of private freight parking can be costly and time consuming at a city-wide scale, resulting in investments that are not feasible for some local governments. These barriers can be effectively overcome by applying data-driven empirical models that use currently available information.

In combination with data collection through direct observations, the model developed in this thesis an easily accessible low-cost option. For instance, the PFI survey in Seattle investigated 2262 parcels in three dense urban neighborhoods of the city taking 210 person-hours just in data collection in field. The model developed could have correctly ruled out 73% of commercial parcels without freight parking infrastructure. Considering that 1370 of these parcels were commercial parcels without these facilities, data collectors could have avoided visiting over 1,000 parcels at the cost of missing 17.9% of the parcels with infrastructures, that is, 28 parcels.

The presence of private freight loading/unloading infrastructure can therefore be predicted based on observable characteristics of the built environment with a reasonable level of error. This solution has minimal cost and cities can implement it right away in planning processes because data is readily available.

Future research can add evidence that support the transferability of the results in Seattle for other cities CBDs. The author recommends the selection of candidate CBDs based on a comparison of metrics such as CBDs employment share in the metropolitan area, transportation mode share of CBD commuters, age of urban fabric, density and economic and population growth. Furthermore, models specific to other types of parcels such as residential need to be developed to provide local governments with a complete suite of tools that allow to estimate the location of private freight parking at a city-wide scale.

BIBLIOGRAPHY

[1] Andre Alho, Joao de Abreu e Silva, and Jorge Pinho de Sousa. A state-of-the-art modeling framework to improve congestion by changing the configuration/enforcement of urban logistics loading/unloading bays. *Procedia - Social and Behavioral Sciences*, 111:360 – 369, 2014. Transportation: Can we do more with less resources? 16th Meeting of the Euro Working Group on Transportation Porto 2013.

[2] Andre Romano Alho, Joao de Abreu e Silva, Jorge Pinho de Sousa, and Edgar Blanco. Improving mobility by optimizing the number, location and usage of loading/unloading bays for urban freight vehicles. *Transportation Research Part D: Transport and Environment*, 61:3 – 18, 2018. Innovative Approaches to Improve the Environmental Performance of Supply Chains and Freight Transportation Systems.

[3] Seattle Municipal Archives. Resources for property history research, 2018. Accessed on 04.15.2018.

[4] Gene Balk. Seattle among top 10 most densely populated big cities in the U.S. for first time ever. *The Seattle Times*, 2017.

[5] Erica E.F. Ballantyne, Maria Lindholm, and Anthony Whiteing. A comparative study of urban freight transport planning: addressing stakeholder needs. *Journal of Transport Geography*, 32:93 – 101, 2013.

[6] Scott Beyer. Seattle is becoming America's 7th 'Legacy City'. *www.forbes.com*, 2017.

[7] King County. King County Parcel Assessor: Assessment data download, 2017. Accessed on 08.01.17.

[8] Wendell Cox. American's oldest cities. *NewGeography.com*, 2013.

[9] Wendell Cox. New central business district employment and transit commuting data. *NewGeography.com*, 2014.

[10] Jianqiang Cui, Jago Dodson, and Peter V. Hall. Planning for urban freight transport: An overview. *Transport Reviews*, 35(5):583–598, 2015.

[11] DDOT. MoveDC. The District of Columbia's multimodal long-range transportation plan. freight element. Technical report, District Department of Transportation, October 2014.

[12] Giampaolo Dezi, Giulio Dondi, and Cesare Sangiorgi. Urban freight transport in Bologna: Planning commercial vehicle loading/unloading zones. *Procedia - Social and Behavioral Sciences*, 2(3):5990 – 6001, 2010. The Sixth International Conference on City Logistics.

[13] Anne Goodchild, Barbara Ivanov, Jose Luis Machado-Leon Haena Kim, and Gabriela Giron-Valderrama. The Final 50 Feet of the urban goods delivery system: Research scan and data collection project (final report). Technical report, Supply Chain Transportation Logistics Center and Seattle Department of Transportation, January 2018.

[14] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning: With Applications in R.* Springer Publishing Company, Incorporated, 2014.

[15] Sean Keeley. 16 high-rise projects changing seattle's skyline. *Seattle.curbed.com*, 2016.

[16] MDS Transmodal Limited. DG MOVE European Commission: Study on urban freight transport - final report. Technical Report 210041R4, April 2012.

[17] Maria Lindholm and Snke Behrends. Challenges in urban freight transport planning a review in the Baltic Sea Region. *Journal of Transport Geography*, 22:129 – 136, 2012. Special Section on Rail Transit Systems and High Speed Rail.

[18] Cathy Macharis and Bram Kin. The 4 a's of sustainable city distribution: Innovative solutions and challenges ahead. *International Journal of Sustainable Transportation*, 11(2):59–71, 2017.

[19] Daniel E. Marasco and Constantine E. Kontokosta. Applications of machine learning methods to identifying and predicting building retrofit opportunities. *Energy and Buildings*, 128:431 – 441, 2016.

[20] Anne G Morris. The impact of inadequate off-loading facilities in commercial office buildings. upon freight efficiency and security in urban areas. *European Transport*, 28:85 – 93, 2004.

[21] Mehdi Nourinejad, Adam Wenneman, Khandker Nurul Habib, and Matthew J. Roorda. Truck parking in urban areas: Application of choice modeling within traffic microsimulation. *Transportation Research Part A: Policy and Practice*, 64:54 – 64, 2014.

[22] NYSERDA and NYDOT. Integrated strategies to address emerging freight and delivery challenges in new york city final report. Technical Report 16-33, New York State Energy Research and Development Authority and New York Department of Transportation, January 2017.

[23] City of Atlanta DPCD. Connect Atlanta Plan The City of Atlanta's Comprehensive Transportation Plan. Technical report, City of Atlanta Department of Planning and Community Development, November 2008.

[24] City of Portland. City of Portland Freight Master Plan. Technical report, City of Portland Office of Transportation, May 2006.

[25] City of Seattle. City of Seattle Land Use Code, 2018. Accessed on 04.14.18.

[26] Gary Pivo, Daniel Carlson, Matthew Kitchen, and Don Billen. Learning from truckers: Truck drivers' views on the planning and design of urban and suburban centers. *Journal of Architectural and Planning Research*, 19(1):12–29, 2002.

[27] H.J. Quak and M.B.M. de Koster. Exploring retailers sensitivity to local sustainability policies. *Journal of Operations Management*, 25(6):1103 – 1122, 2007. Supply Chain Management in a Sustainable Environment Special Issue on Frontiers of Empirical Supply Chain Research.

[28] David Schrank, Bill Eisele, Tim Lomax, and Jim Bak. 2015 Urban Mobility Scorecard. Technical report, The Texas A&M Transportation Institute and Inrix, August 2015.

[29] SDOT. City of Seattle Freight Master Plan. Technical report, City of Seattle Department of Transportation, September 2016.

[30] Giovanni Tardioli, Ruth Kerrigan, Mike Oates, James ODonnell, and Donal Finn. Data driven approaches for prediction of building energy consumption at urban level. *Energy Procedia*, 78:3378 – 3383, 2015. 6th International Building Physics Conference, IBPC 2015.

[31] Adam Wenneman, Khandker M. Nurul Habib, and Matthew J. Roorda. Disaggregate analysis of relationships between commercial vehicle parking citations, parking supply, and parking demand. *Transportation Research Record: Journal of the Transportation Research Board*, 2478:28–34, 2015.