Center for Advanced Multimodal Mobility

Solutions and Education

Project ID: 2019 Project 02

OPTIMIZING TRANSIT EQUITY AND ACCESSIBILITY BY INTEGRATING RELEVANT GTFS DATA PERFORMANCE METRICS

Final Report

by

Wei Fan (ORCID ID: https://orcid.org/0000-0001-9815-710X)
Yang Li (ORCID ID: https://orcid.org/0000-0001-5358-7835)

Wei Fan, Ph.D., P.E.
Director, USDOT CAMMSE University Transportation Center
Professor, Department of Civil and Environmental Engineering
The University of North Carolina at Charlotte
EPIC Building, Room 3261, 9201 University City Blvd, Charlotte, NC 28223
Phone: 1-704-687-1222; Email: wfan7@uncc.edu

for

Center for Advanced Multimodal Mobility Solutions and Education
(CAMMSE @ UNC Charlotte)
The University of North Carolina at Charlotte
9201 University City Blvd
Charlotte, NC 28223

September 2020

ACKNOWLEDGEMENTS

This project was funded by the Center for Advanced Multimodal Mobility Solutions and Education (CAMMSE @ UNC Charlotte), one of the Tier I University Transportation Centers that were selected in this nationwide competition, by the Office of the Assistant Secretary for Research and Technology (OST-R), U.S. Department of Transportation (US DOT), under the FAST Act. The authors are also very grateful for all of the time and effort spent by DOT and industry professionals to provide project information that was critical for the successful completion of this study.

DISCLAIMER

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACS – American Community Survey

ADA – Americans with Disabilities

CATS – Charlotte Area Transit System

FAST – Fixing America's Surface Transportation

GIS – Geographic Information System

GTFS – General Transit Feed Specification

NCDOT – North Carolina Department of Transportation

RUC – Residential Units Covered by Transit Stop/Station

RUT –Residential Units (Total) in Block groups

TGI – Transit Gap Index

TD – Transportation Disadvantaged/Transit Dependent

TDS – Transit Dependent Score

TDS' – Standardized Transit Dependent Score

TOI – Transit Opportunity Index

TSC – Transit Service Coverage Ratio

TSS – Transit Service Score

TSS' – Standardized Transit Service Score

TTP – Total Population

EXECUTIVE SUMMARY

Public transit mode continues to be a crucial part of transportation planning in the United States. Building upon the assessment of public transit equity and accessibility, one important task of transit planning is to optimize relevant metrics and measurements by modifying current service parameters (e.g., route layouts, schedules, and frequencies) or having new investments. In this context, many studies had been done to perform the gap analysis and discuss the potential uses of such analysis for further optimizing the public transit services via modification of current transit systems. However, efforts still need to be made to enhance and enrich relevant research on integrating transit equity into a network design problem. In the meantime, recent development of the General Transit Feed Specification (GTFS), a well formatted transit feeds open data, provides new opportunities for a better understanding of both spatial and temporal characteristics of public transit because such data is easy to handle and is proved efficient in relevant analysis. By taking such advantages, it is necessary to further leverage such data source to contribute to both the state-of-the-art and the state-of-the-practice in this field, and efforts are needed to review the current practices and develop appropriate mathematical optimization models for improving the public transit equity and accessibility by integrating GTFS data relevant performance metrics.

This research develops and recommends practical optimization models for improving transit equity and accessibility for people by integrating performance metrics with using GTFS data. In such sense, two optimization models under two different conditions (with limited budget and without limited budget) for improving the public transit equity of blockgroups suffering transit deficiency are built. A case study in the City of Charlotte is conducted and the associating comprehensive numerical results and analyses based on the proposed models are provided. This research also develops guidelines and recommends best practices for the use of GTFS data as a main data source to better develop mathematical optimization models for improving the public transit equity and accessibility. Summary and conclusions are made, and further research directions are also given.

Chapter 1. Introduction

1.1 Problem Statement

Public transit mode continues to be a crucial part of transportation planning in the United States. Building upon the assessment of public transit equity and accessibility, one important task of transit planning is to optimize relevant metrics and measurements by modifying current service parameters (e.g., route layouts, schedules, and frequencies) or having new investments. For an existing public transit system, balancing between the service redundancies and less accessible regions will likely provide expanded regional coverage of equity through a redesign or redistribution of public transit systems. In this context, many studies had been done to perform the gap analysis and discuss the potential usages of such analysis for further optimizing the public transit services via modification of current transit systems (El-Geneidy et al., 2014; Karner et al., 2016). To establish new services of public transit, planners and managers will need to take budgets and maximization of equity into consideration. Both improvements are particularly important for disadvantaged populations. Thus, there is a strong need for effective and efficient models and solution approaches to analyze and tackle those problems. In the meantime, recent development of the General Transit Feed Specification (GTFS), a well formatted transit feeds open data, provides new opportunities for a better understanding of both spatial and temporal characteristics of public transit because such data is easy to handle and is proved efficient in relevant analysis. Although years of research efforts have been made by using various accessibility measures in network design optimization models, only a limited number of studies have used and explored GTFS data. In order to further leverage such data source to contribute to both the state-of-the-art and the state-of-the-practice, this research will review the current practices in this field and develop an appropriate mathematical optimization model for improving the public transit equity and accessibility by integrating GTFS data relevant performance metrics and measurements for public transportation planning and operation.

Many studies have discussed the capability of GTFS data utilization in handling transit equity related research (Antrim and Barbeau, 2013; Wong, 2013; Schweiger, 2015; Rodnyansky, 2018). However, not much attention has been given to integrate transit equity into a network design problem. A report published by the State of Florida Department of Transportation showed that GTFS data can be utilized for stop location and spacing optimization (Catala et al. 2011). Coffey et al. (2012) leveraged the GTFS data and used an optimization model that seeks to modify existing schedules to improve transit connectivity. Such study can be treated as optimization of the temporal accessibility of public transit for people to improve its efficiency and equity. Another technical report from the Oregon Department of Transportation presented a proof of concept on how to optimize its transit network using GTFS data (Porter et al. 2014). In order to incorporate equity into the transit network design problem, Bertolaccini and Lownes (2015) developed several Transit Network Design Problem (TNDP) models by combining the Transit Opportunity Index (TOI) and GTFS data, showing the ability of GTFS data to solve optimization problems. Furthermore, Chen (2016) proposed a hybrid transit system that integrated the fixed-route service with demand-adaptive service to improve the accessibility (i.e., connectivity). At the same time, several research efforts have proved the capability of GTFS data in benefiting social welfare. For instance, Anderson et al. (2017) presented a multi-criteria suitability analysis framework to help municipal governments determine the optimal locations for the mobility hubs and included GTFS data as one of the major data sources. Also, Zhong et al. (2017) utilized GTFS data to develop a model for optimizing the locations of affordable housing to maximize residents' accessibility to public transit, subjected to land acquisition and construction budget constraints. In addition, Zygo (2017) combined TOI and GTFS data to optimize the transit network to improve the access to medical facilities for seniors.

1.2 Objectives

The main objective of this research project is to develop optimization models for improving transit equity and accessibility for people by integrating performance metrics with using GTFS data. The objectives of this project are to: (1) conduct a comprehensive review of the state-of-the-art and state-of-the-practice on public transit equity optimization, especially those with optimizing the use of performance metrics utilizing GTFS data; (2) develop suitable optimization models to improve public transit equity of blockgroups with transit deficiency; (3) design an efficient solution method to solve the optimization models; and (4) analyze the results and provide recommendations on future research directions.

1.3 Expected Contributions

To accomplish these objectives, several tasks have been undertaken. A literature review of the public transit equity optimization associated with optimizing the use of performance metrics utilizing GTFS data has been conducted. According to the literature and the research results from previous studies, optimization models for improving public transit equity has been developed, associating with approaches to solving the developed optimization models. Based on the results, primary recommendations on the best-practice/policy to improve the public transit equity will be provided. All products will be integrated into current practices for better investments to optimize the public transit equity and accessibility, help provide method and identify further opportunities to maximize the equity under future uncertain public transit demand/ridership forecasting.

1.4 Report Overview

The remainder of this report is organized as follows: Chapter 2 present a comprehensive review of the state-of-the-art and state-of-the-practice literature on public transit equity, optimization problems of public transit accessibility and equity, the general transit feed specification (GTFS) and also the use of GTFS in public transit equity optimization problems. Chapter 3 gives a brief introduction of the transit gap index (*TGI*). Chapter 4 provides a detailed explanation of and formulation for the developed optimization models for improving public transit equity of blockgroups with transit deficiency. Chapter 5 presents detailed data descriptions of all datasets used and those associated with the methodology developed in this research. Chapter 6 presents a real-world case study as an example. Comprehensive analyses and detailed numerical results based on the data in the City of Charlotte are provided. Finally, Chapter 7 concludes this report with a summary and a discussion of the directions for future research.

Chapter 2. Literature Review

2.1 Introduction

This chapter provides a comprehensive review of the current state-of-the-art and state-of-the-practice on public transit equity and accessibility optimization problem, and also the use of GTFS data related measurements/metrics in public transit equity optimization. This should give a clear picture of public transit equity optimization problem, the methodologies for tackling the problem, and potential and available usage of GTFS in such problem.

The following sections are organized as follows. Section 2.2 presents a general view of the research background. Section 2.3 gives brief descriptions of previous studies that mainly used GTFS data and GTFS relevant performance metrics to optimize public transit equity in public transit network optimization problem. Finally, section 2.4 concludes this chapter with a summary.

2.2 Research Background

Not much attention has been given to integrate transit equity into a network design problem. Along this line, though there have been rapid developments of and in the use of GTFS data, few studies have included GTFS data in the analysis of building the optimization model for the improvement of public transit accessibility. A report published by the State of Florida Department of Transportation showed that GTFS data can be utilized for stop location and spacing optimization (Catala et al., 2011). Another technical report from the Oregon Department of Transportation presented a proof of concept on how to optimize its transit network using GTFS data (Porter et al., 2014). In order to incorporate equity into the transit network design problem, Bertolaccini and Lownes (2015) developed several Transit Network Design Problem (TNDP) models, by combining the Transit Opportunity Index (TOI) and GTFS data. Anderson et al. (2017) presented a multi-criteria suitability analysis framework to help municipal governments determine the optimal locations for the mobility hubs and included GTFS data as one of the major data sources. Zhong et al. (2017) utilized GTFS data to develop a model for optimizing the locations of affordable housing to maximize residents' accessibility to public transit, subjected to land acquisition and construction budget constraints. Zygo (2017) also combined TOI and GTFS data to optimize the transit network to improve the access to medical facilities for seniors. However, efforts are still needed to review the current practices and develop appropriate mathematical optimization models for improving the public transit equity and accessibility by integrating GTFS data relevant performance metrics.

2.3 Previous Studies Using GTFS Data for the Optimization of Public Transit Equity

Public transit equity optimization problem itself indeed is an important topic within the research area of transportation. The major focus of this research project is to take the immediate opportunity of the available GTFS data and use it as a basic data source for optimizing public transit equity. And with the rapid development of GTFS over past few years and its relative convenient and powerful nature in network analysis, research efforts have shown the efficiency of using GTFS data for improving accessibility and equity of public transit by developing and using a variety of methodologies, measures and indicators. However, such progress is relatively slow and the amount of the related research and studies are still very limited. This section presents and lists some most relevant previous studies within this topic.

In 2011, a research project conducted by the National Center for Transit Research examined opportunities of using GTFS data for service planning and operational activity. Though it was not directly related to public transit equity optimization, it was found that by combining location and time elements, the GTFS data provide new opportunities for better supporting the optimization of transit service with considering equity and accessibility based on time and location (Catala et al., 2011).

In Bertolaccini's work (2015), the author tried to incorporate equity into the stop sequencing or stop grouping components of the transit network design problem. And three models and nine possible inequity minimizing objective function formulations were explored, drawing from horizontal, vertical, and intermodal equity perspectives. The Sioux Falls, SD and Willimantic, CT networks were used to test the single route model and to develop the GA. These experiments narrowed the list of possible equity objective functions from nine to six. Extensive testing was conducted on the GA, on both its algorithmic structure and its input parameters, to validate its quality and efficiency. After testing the single route model and developing the GA, the model was expanded into a multiple route model which includes the stop grouping component of the TNDP. This model also considered route transfers, walking connections to transit stops, demand zones, multiple paths between demand zones, and idle time. Last, this model was applied to a subset of the University of Connecticut's (USA) shuttle bus system and solved using an expanded and updated version of the GA applied to the single route model. The GTFS data was used as the only data source for accomplishing the project.

Zhong et al. (2017) developed a new optimization model for locating affordable housing units in order to maximize the accessibility of disadvantage population (i.e., the low-income workers) to appropriate jobs by public transit, with incorporating the GTFS data. Transit accessible housing allows disadvantaged populations to reduce their reliance on automobiles, which can lead to savings on transportation-related expenditures. The housing location model developed in the study maximizes transit accessibility while reducing the clustering of affordable housing units in space. Accessibility is maximized using a high-resolution space-time metric of public transit accessibility, originally developed for service equity analysis. The second objective disperses subsidized housing projects across space using a new minimax dispersion model based on spatial interaction principles. The multi-objective model trades off accessibility maximization and affordable housing dispersion, subject to upper and lower bounds on the land acquisition and construction budget. The model was tested using data for Tempe, AZ including actual data for vacant parcels, travel times by light rail and bus, and the location of low-wage jobs. This model or similar variants could provide insightful spatial decision support to affordable-housing providers or tax-credit administrators, facilitating the design of flexible strategies that address multiple social goals.

In the same year, a final report from Transportation Research Center for Livable Communities (TRCLC), Michigan (Oh et al., 2017) put forth a series of time-sensitive, general transit feed system (GTFS)-enhanced employment accessibility models that account for multiple transportation modes, categories of functional limitation and design characteristics of existing public transit infrastructure, to employ fine-scale performance measures for people with disability. Then the study extended the measurements from the previous models by incorporating multimodal accessibility estimates into an optimization model designed to prioritize investments in transit stops that most improve employment access for all, including people with limited mobility. The optimization model used cost estimates for bus stop retrofits and multimodal accessibility estimates to develop a multi-phase implementation strategy toward ADA compliance. The optimized retrofit strategy was compared with a random or ad hoc strategy in order to evaluate respective gains in both employment accessibility and network connectivity.

Considering that both operational efficiency and access equity are critical to the well-being of public transit services, it is necessary to take into account both of them to provide a comprehensive transit service performance assessment and further help achieve the optimization of the public transit service. Wei et al. (2018) proposed a combination of data envelopment analysis (DEA), GIS and spatial optimization model, the maximal covering location problem (MCLP) to allow the exploration of trade-offs between both operational efficiency and access equity, which are two potentially competing goals and enable the performance of transit services to be assessed in a holistic manner. GTFS data is considered as one major input of the model in this study to achieve the goal.

2.4 Summary

A comprehensive review and synthesis of the current and historical researches related to public transit network optimization problem, general transit feed specification (GTFS), and also the use of GTFS in public transit equity optimization have been discussed and presented in the preceding sections. This is intended to provide a solid reference and assistance in formulating public transit equity optimization problem and developing effective improvement strategies for future tasks.

Chapter 3. A Brief Introduction to Transit Gap Index (*TGI*)

3.1 Introduction

This chapter gives brief introduction to the transit gap index (*TGI*) and its associated components, which were developed in the previous study by Fan and Li (2019). Based on this developed evaluation method, the next chapter develops the optimization models for improving the public transit equity for blockgroups.

3.2 Modelling the Transit Gap Index (*TGI*)

A "gap analysis" between transit supply and demand is generally the most common form used to evaluate the equity/accessibility of a public transit service system. Such analysis is categorized as spatial analysis, where ArcGIS is deployed to undertake the task. On the other hand, when considering the availability of features in the GTFS data, the definition of equity in this study is designed around those components available through GTFS.

3.2.1 Transit Supply

3.2.1.1 Transit Service Coverage

In this study, the transit service coverage is defined as a ratio as follows:

$$TSC_j = \frac{RUC_j}{RUT_j} \tag{1}$$

where TSC_j is the transit service coverage ratio of blockgroup j, RUC_j is the number of residential units (non-overlapping) covered by all stops within a 0.5-mile walking catchment area in blockgroup j and RUT_j is the total number of residential units in blockgroup j. In most of the previous studies, ¼ miles (or 400 meters), or equivalent five-minute walking distance is considered as "acceptable walking distance" (O'Sullivan et al., 1996; Jiang et al., 2012; Daniels et al., 2013; Zhao et al., 2013; El-Geneidy et al., 2014). O'Sullivan et al. (1996) and Daniels et al. (2013) also pointed out that the distance would vary based on the type of the transit service. For example, people will be willing to walk even further when they take a light rail instead of a bus. Moreover, one recent research study (Durand et al., 2016) has shown that individuals seem to be willing to walk further to reach transit stops/stations than what "rule of thumb" guidelines indicated (¼ miles, or 400 meters). It further exhibited that with other factors being the same, there is a 50% chance that people will walk to a stop from a distance that is two miles away from the transit stop, and this probability will increase to 80% for a one-mile distance. Thus, in order not to underestimate the transit service coverage, a 0.5-mile walking distance has been applied in this study.

Adopted from the work of Bejleri et al. (2018), the calculation of service coverage at the stop level considers the actual residential units within the blockgroup. Such ratio calculation here shows an improvement with the consideration of the actual spatial coverages of the residential units instead of simply measuring it as a ratio of the service area to the total area of the blockgroup.

3.2.1.2 Per Capita Maximum Daily Available Seats

Per capita maximum daily available seats for specific blockgroup can be computed as:

$$D_j = \frac{\sum_i \frac{\sum_l F_l \times C_l \times RUC_{lij}}{RUC_i}}{P_j} \tag{2}$$

where D_j is the per capita maximum daily available seats for blockgroup j, F_l denotes the frequency of route l, C_l represents the typical capacity per bus of route l, RUC_{lij} means the number of residential units covered by stop i along route l within the 0.5-mile walking catchment area in blockgroup j, RUC_i is the total number of residential units covered by stop i within the 0.5-mile walking catchment area, and P_j denotes the total population in blockgroup j.

Per Capita Maximum Daily Available Seats estimates the level of service provided by the transit service for the total population within one blockgroup area other than the people who have access to the service. This concept is adopted from (Mamun et al., 2013) and a modification has been made here with the usage of “residential units” instead of simply allocating the capacity to each blockgroup in the original form. This parameter presents an average daily basic level of service for specific blockgroup served by all relevant public transit services.

3.2.1.3 Transit Service Score

Finally, by combining the transit service coverage ratio with per capita maximum daily available seats, the transit service score can be computed as follows:

$$TSS_j = TSC_j \times D_j \tag{3}$$

where TSS_j is the transit service score for blockgroup j. In a sense, the transit service score covers the spatial and temporal (daily basis) characteristics for the public transit service (supply).

3.2.2 Transit Demand

The formulation developed to compute the transit dependent populations at the census block group level is adopted from and modified based on studies conducted by U.S. Department of Transportation (Steiss 2006), and Capital Area Transit Authority in Lansing, Michigan (CATA 2011). This method has also been used in Jiao (2013, 2015). Even though transit dependent populations are normally referred to as the people who are too young, too old, too poor, or who are physically handicapped and unable to drive (Grengs 2001), the internal overlapping characteristics of census data among these topics will unavoidably result in the potential for “double-counts” when computing transportation demand by simply adding each criterion together. Therefore, the following formulation has been developed and used in this study, which is shown as follows:

Household drivers = (population age 16 and over) – (people living in group quarters)
Transit-dependent household population = (household drivers) – (vehicles available)

TD_j **= Transit-dependent population** = (transit-dependent household population) + (population age 10–15) + (non-institutionalized population living in group quarters) (4)

A group quarters is a place where people live or stay, in a group living arrangement manner, which is owned or managed by an entity or organization providing housing and/or services for the residents (U.S. Census Bureau, 2010). There are two types of group quarters, institutional (e.g., correctional facilities, nursing homes, or mental hospitals) and non-institutional (e.g., dormitories, military barracks, group homes, missions, or shelters). And institutionalized populations living in group quarters will move/travel together in a group manner, whereas non-institutionalized populations living in group quarters are not.

Such calculation shifts the focus from why individuals may not drive (due to different reasons that are related to age, income, and mobility, etc.) to the determination of where there are limited vehicles available for the whole population to use (Jiao, 2013; Jiao, 2015), which can effectively eliminate the overlapping among each topic (which are relevant to age, income, and mobility, etc.). Negative values might be obtained and will be adjusted to zero whenever necessary. The reasoning for this is that no blockgroup should have a negative number of people who are transit dependent.

After obtaining the total transit dependent population for each blockgroup, a transit dependent score (TDS_j) can be achieved by using the following equation:

$$TDS_j = \frac{TD_j}{TTP_j} \qquad (5)$$

where TTP_j is the total population of blockgroup j.

2.1.3 Transit Gap Index

Finally, the transit service gap index could be obtained by comparing the differences between supply and demand in a standardized manner. The values from both the transit supply and demand will be standardized in a scale of 0 to 1 based on the equation as shown below:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \qquad (6)$$

where X' can be referred to as the transit supply (TSS_j) and demand (TDS_j). The transit gap index can then be calculated by subtracting TDS_j' from TSS_j' for blockgroup j:

$$TGI_j = TSS_j' - TDS_j' \qquad (7)$$

3.3 Summary

The brief introduction of *TGI* and its associated components is introduced in this chapter, along with the discussions about meanings of each component. Procedures developed for conducting the analysis are also provided.

Chapter 4. Optimization Models for Improving Transit Equity

4.1 Introduction

This chapter will mainly focus on developing optimization models for improving the public transit equity of blockgroups with transit deficiency based on the results of the previous study of Fan and Li (2019).

The following sections are organized as follows. Section 4.2 provides a general overview of the basic idea of developing optimization models for improving the public transit equity of blockgroups with transit deficiency. The brief definitions of transit deficiency are introduced in Section 4.3. Section 4.4 presents two optimization models for improving the public transit equity of blockgroups with transit deficiency, where one comes with limited budget and the other is not. Finally, Section 3.5 concludes this chapter with a summary.

4.2 Overview

In this project, the objective is to optimize the transit equity by mitigating the transit deficiency based on the study of Fan and Li (2019). *TGI* and its associated components will be utilized in the model developments. The basic idea is to improve the level of transit service in those blockgroups that suffer transit deficiency. And two conditions are considered while developing the optimization models, which are: 1) maximizing the level of transit service with constraint of limited budget that will be invested into the new constructions of public transit stops; 2) minimizing the total cost for constructing new public transit stops with constraint that requires certain improvements of transit service for a certain amount of blockgroups.

4.3 Transit Deficiency

In previous study (Fan and Li, 2019), *TSS'* and *TDS'* are classified into seven categories (i.e., Very Low, Low, Medium-Low, Medium, Medium-High, High and Very High). Therefore, blockgroups with transit deficiency can be defined as those blockgroups with “Very Low”/ “Low” *TSS'* and “High”/ “Very High” *TDS'*, which means there are not enough transit services provided to meet the transit demands.

The case study in this project is the City of Charlotte. There are 28 blockgroups that suffer transit deficiency in the study area and the basic information on those blockgroups, including the total population, *TDS'*, *TSS'* and *TGI* will be presented in Chapter 5.

4.4 Models for Improving the Transit Equity

This section will present two optimization models for improving the public transit equity and accessibility for blockgroups with transit deficiency, where one comes with limited budget and the other is not but requires certain improvements of transit service for a certain amount of blockgroups.

4.4.1 Model with Limited Budget

$$\text{Minimize} \sum_{i \in I} TGI_i'^2 = \sum_{i \in I} (TSS_i'' - TDS_i')^2 = \sum_{i \in I} \{(D_i + \frac{a_i x_i}{P_i \times TSS'_{\max}}) \times [1 - z_i + z_i \times TSC_i] - TDS_i'\}^2 \qquad (1)$$

Subject to:

$$\sum_{i \in I} c_i x_i \leq B, \forall i \in I \qquad (2)$$

$$TSS_i'' \leq TDS_i', \forall i \in I \qquad (3)$$

$$0 \leq x_i \leq s_{\max}, \forall i \in I \qquad (4)$$

$$z_i \in \{0,1\}, \forall i \in I \qquad (5)$$

$B -$ total budget

$I = \{1, 2, ..., 28\}$

$P_i -$ population of blockgroup *i*;

$TSS'_{\max} = 100$;

$s_{\max} -$ the maximal no. of stops that can be added to one blockgroup;

$c_i -$ cost for constructing new stop in blockgroup *i*;

$a_i -$ average capacity of stop for blockgroup *i*;

$x_i -$ decision variable, the no. of stops constructed in blockgroup *i*;

$z_i -$ indicator, if $x_i = 0$, then $z_i = 1$, otherwise $z_i = 0$.

The formulation, parameters and variables associated with brief explanations of the model are introduced above. The objective (1) of this model is to minimize the summation of the square of new *TGI*s, which belong to blockgroups with transit deficiency. The original *TGI*s of such blockgroups are negative so that the objective of this model can also be viewed as the maximization of *TSS'*s by adding a certain amount of stops with particular capacity (a_i) of transit services. Constraint (5) states that z_i is an indicator variable, which equals to "1" when blockgroup *i* is not chosen to have new stop(s) added and equals to "0" when there will be new stop(s) adding to blockgoup *i*. Furthermore, it is assumed that if blockgroup *i* is chosen for new stop(s) construction, its transit service coverage (TSC_i) could be increased to "1" due to the reason that location(s) of stop(s) could be well examined to cover almost all residential units. Thus, the expression, "$1 - z_i + z_i \times TSC_i$", simply implies that if blockgroup *i* is chosen, then its TSC_i will be set to "1", otherwise the TSC_i will remains the same.

Constraint (2) denotes the constraint of limited budget, which explains that the total cost of new stop constructions cannot exceed the total budget invested. Constraint (3) shows that the level of transit service of one blockgroup should not be larger than its transit demand to avoid transit service redundancy. Constraint (4) is boundary condition of decision variable x_i for each blockgroup, including the non-negativity and the maximal number of new stops that can be added to one single blockgroup.

4.4.2 Model without Budget

$$\text{Minimize} \sum_{i \in I} c_i x_i \tag{1}$$

Subject to:

$$\sum_{i \in I} y_i \geq N, \forall i \in I \tag{2}$$

$$TSS_i'' = (D_i + \frac{a_i x_i}{P_i \times TSS'_{\max}}) \times [1 - z_i + z_i \times TSC_i] \leq TDS_i', \forall i \in I \tag{3}$$

$$0 \leq x_i \leq s_{\max}, \forall i \in I \tag{4}$$

$$z_i \in \{0,1\}, \forall i \in I \tag{5}$$

$$y_i \in \{0,1\}, \forall i \in I \tag{6}$$

$I = \{1, 2, ..., 28\}$;

N – no. of blockgroups set to meet the goal of improvement on TGI;

P_i – population of blockgroup i;

$TSS'_{\max} = 100$;

$s_{\max}$ – the maximal no. of stops that can be added to one blockgroup;

c_i – cost for constructing new stop in blockgroup i;

a_i – average capacity of stop for blockgroup i;

x_i – decision variable, the no. of stops constructed in blockgroup i.

y_i – indicator, if $\frac{\Delta TGI_i}{|TGI_i|} = \frac{TSS_i'' - TSS_i'}{|TGI_i|} \geq L$, then $y_i = 1$, otherwise $y_i = 0$;

L – goal of improvement on TGI in the study area (e.g., 0.1 means 10% improvement);

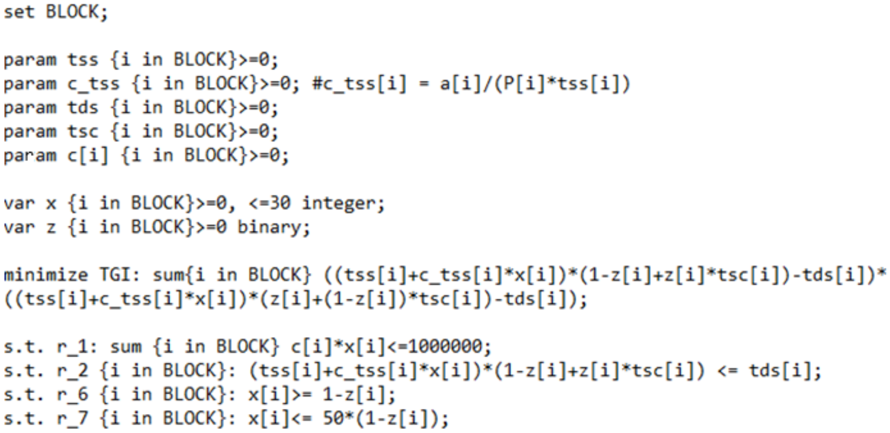
z_i – indicator, if $x_i = 0$, then $z_i = 1$, otherwise $z_i = 0$.

The formulation, parameters and variables associated with brief explanations of the second model are introduced as above. Unlike the first model, there will not be a certain limited budget defined in this model. Instead, the objective (1) of this model is to minimize the total cost of new stop constructions in the study area.

In the model, constraint (6) states that y_i is an indicator variable, which equals to "1" when the improvement of TGI_i meets the goal (i.e., $\frac{\Delta TGI_i}{|TGI_i|} = \frac{TSS_i'' - TSS_i'}{|TGI_i|} \geq L$, certain improvement to the transit service level, such as 10%, 20% and etc.), and equals to "0" otherwise. Therefore, constraint (2) of this model is to require that a certain amount (N) of blockgroups should have their transit service levels meet the goal of certain improvement. Constraints (3), (4) and (5) have the same meanings with the ones in the first model, correspondingly.

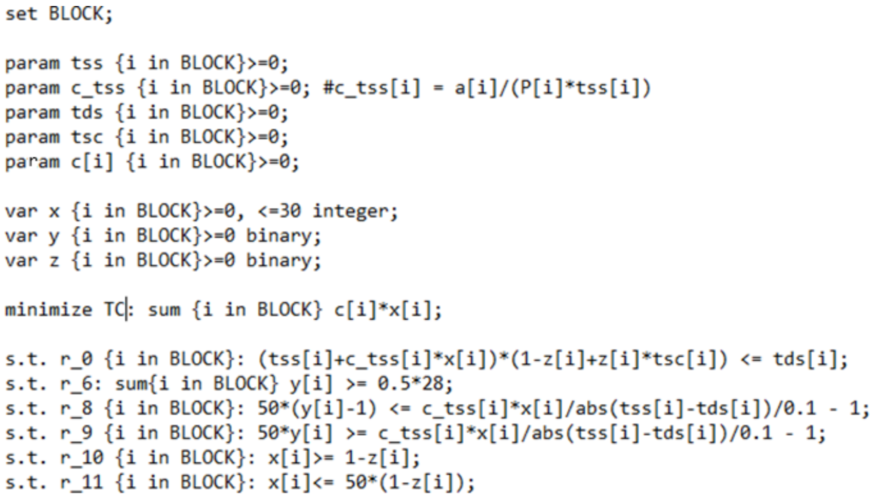
The proposed models can be classified as mixed integer nonlinear program (MINLP) models, since the decision variables in both models are discrete and both models contain nonlinear constraint(s). Both models are programmed using the AMPL modeling language (Fourer et al., 2003) and, due to the models' nonlinear nature, they will be run on server BARON from NEOS (https://neos-server.org/neos/solvers/minco:BARON/AMPL.html), which used solver Baron 19.3.22. Detailed numerical results will be provided in Chapter 6.

Figure 4.1 and 4.2 are the codes of both optimization models in AMPL, and Figure 4.3 displays partial of the data inputs used in AMPL. As can be seen in the coding files, constraints for indicator variables in both models have been rephrased by applying the big-M method. The "50" in both files represents the M, since it is sufficient big for the case study that is conducted in this project.



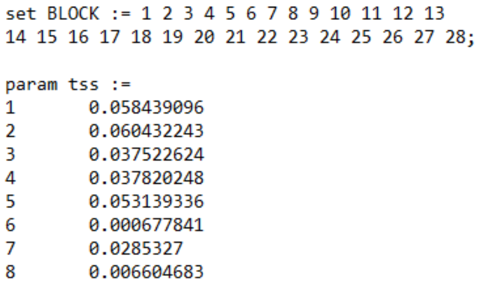
```
set BLOCK;

param tss {i in BLOCK}>=0;
param c_tss {i in BLOCK}>=0; #c_tss[i] = a[i]/(P[i]*tss[i])
param tds {i in BLOCK}>=0;
param tsc {i in BLOCK}>=0;
param c[i] {i in BLOCK}>=0;

var x {i in BLOCK}>=0, <=30 integer;
var z {i in BLOCK}>=0 binary;

minimize TGI: sum{i in BLOCK} ((tss[i]+c_tss[i]*x[i])*(1-z[i]+z[i]*tsc[i])-tds[i])*
((tss[i]+c_tss[i]*x[i])*(z[i]+(1-z[i])*tsc[i])-tds[i]);

s.t. r_1: sum {i in BLOCK} c[i]*x[i]<=1000000;
s.t. r_2 {i in BLOCK}: (tss[i]+c_tss[i]*x[i])*(1-z[i]+z[i]*tsc[i]) <= tds[i];
s.t. r_6 {i in BLOCK}: x[i]>= 1-z[i];
s.t. r_7 {i in BLOCK}: x[i]<= 50*(1-z[i]);
```

Figure 4.1 AMPL Code for Model with Limited Budget



```
set BLOCK;

param tss {i in BLOCK}>=0;
param c_tss {i in BLOCK}>=0; #c_tss[i] = a[i]/(P[i]*tss[i])
param tds {i in BLOCK}>=0;
param tsc {i in BLOCK}>=0;
param c[i] {i in BLOCK}>=0;

var x {i in BLOCK}>=0, <=30 integer;
var y {i in BLOCK}>=0 binary;
var z {i in BLOCK}>=0 binary;

minimize TC: sum {i in BLOCK} c[i]*x[i];

s.t. r_0 {i in BLOCK}: (tss[i]+c_tss[i]*x[i])*(1-z[i]+z[i]*tsc[i]) <= tds[i];
s.t. r_6: sum{i in BLOCK} y[i] >= 0.5*28;
s.t. r_8 {i in BLOCK}: 50*(y[i]-1) <= c_tss[i]*x[i]/abs(tss[i]-tds[i])/0.1 - 1;
s.t. r_9 {i in BLOCK}: 50*y[i] >= c_tss[i]*x[i]/abs(tss[i]-tds[i])/0.1 - 1;
s.t. r_10 {i in BLOCK}: x[i]>= 1-z[i];
s.t. r_11 {i in BLOCK}: x[i]<= 50*(1-z[i]);
```

Figure 4.2 AMPL Code for Model without Budget



```
set BLOCK := 1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20 21 22 23 24 25 26 27 28;

param tss :=
1       0.058439096
2       0.060432243
3       0.037522624
4       0.037820248
5       0.053139336
6       0.000677841
7       0.0285327
8       0.006604683
```

Figure 4.3 Partial Data Inputs in AMPL

4.5 Summary

The idea of developing optimization models to improve the public transit equity of blockgroups with transit deficiency is presented in this chapter. Both transit deficiencies developed from a previous study of Fan and Li (2019) are briefly introduced. Two optimization models under two conditions (with limited budget and without limited budget) for improving the public transit equity of blockgroups with transit deficiency are also provided with the presentation of detailed information on both formulations.

Chapter 5. Case Study

5.1 Introduction

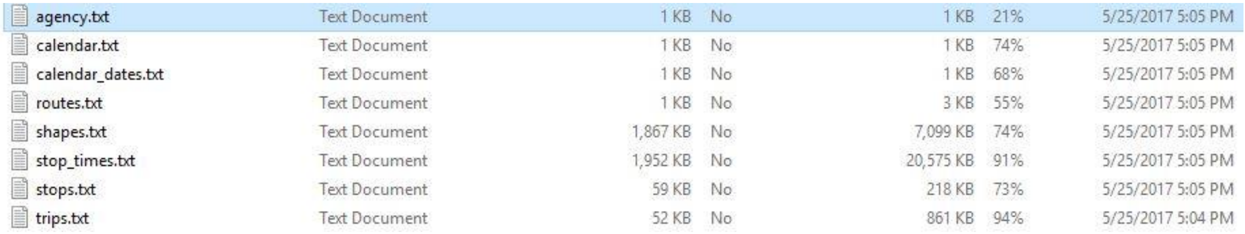
The chapter presents the data obtained and used in this study. Most of the data are used to conduct the transit gap analysis, whose results are utilized as the inputs of the optimization models developed in this project.

The following sections are organized as follows. Section 5.2 describes GTFS data of the Charlotte Area Transit System (CATS). Section 5.3 presents the demographic data in the City of Charlotte. Section 5.4 shows the transportation data of the City of Charlotte. Section 5.5 lists the other regional data that are used in this case study. Section 5.6 displays the information on blockgroups with transit deficiency. Finally, section 5.7 concludes this chapter with a summary.

5.2 GTFS Data

As discussed in Chapters 1 and 2, GTFS as a standard transit feeds data format has been demonstrated to be extremely useful, due to its contents associated with spatial and temporal characteristics.

This project uses the GTFS data of CATS that are obtained from TRANSITLAND (https://github.com/transitland/gtfs-archives-not-hosted-elsewhere/blob/master/charlotte-cats.zip). The data include all the required files of a standard GTFS data as shown in Figure 5.1 below:



agency.txt	Text Document	1 KB	No	1 KB	21%	5/25/2017 5:05 PM
calendar.txt	Text Document	1 KB	No	1 KB	74%	5/25/2017 5:05 PM
calendar_dates.txt	Text Document	1 KB	No	1 KB	68%	5/25/2017 5:05 PM
routes.txt	Text Document	1 KB	No	3 KB	55%	5/25/2017 5:05 PM
shapes.txt	Text Document	1,867 KB	No	7,099 KB	74%	5/25/2017 5:05 PM
stop_times.txt	Text Document	1,952 KB	No	20,575 KB	91%	5/25/2017 5:05 PM
stops.txt	Text Document	59 KB	No	218 KB	73%	5/25/2017 5:05 PM
trips.txt	Text Document	52 KB	No	861 KB	94%	5/25/2017 5:04 PM

Figure 5.1 GTFS Data of CATS

This version of GTFS data was updated on May 25th, 2017. Table 5.1 shows the general information about CATS based on the obtained GTFS data, and the typical capacity per bus is 40 seats in CATS.

Table 5.1 General Characteristics of CATS based on GTFS Data

Number of Routes	Number of Stops	Number of Trips	Typical Capacity/bus (seats)
75	3,307	10,047	40

The “shapes.txt” and “stops.txt” files are integrated into ArcGIS to create the shapefiles of the public transit system (both routes and stops/stations) in the City of Charlotte. This has been mentioned in Figure 4.4 in section 4.3, which is the output of “Display GTFS” tool in ArcGIS for the CATS.

The “stops.txt”, “stop_times.txt”, “trips.txt” and “routes.txt” files in the CATS GTFS data are used to determine the stop-route pairs and matrix, which have already been discussed in section 4.4. There are 4,678 unique stop-route pairs in total and an example of the stop-route pair of “Route 590” is shown in Table 5.2 below:

Table 5.2 Example of Stop-Route Pairs

Stop ID	Route ID
23520	590
45710	590
45711	590
45815	590
46439	590
52240	590

Associated with blockgroups shapefile, the capacities for each stop within specific blockgroup can be obtained by conducting spatial analysis using ArcGIS, and the following provides an example of the blockgroup with specified stop capacity:

Table 5.3 Example of Blockgroup with Specified Stop Capacity

Stop ID	Blockgroup ID	Capacity
11430	371190017011	933
11440	371190017011	933
17540	371190014002	2042
17550	371190014002	2033
17580	371190014002	2447
17620	371190014002	2379

5.3 Demographic Data

As discussed in section 3.4, in order to calculate the transit dependent (TD) population, several necessary demographic data are obtained from US Census Bureau database and most of the data are available at the blockgroup level.

The first dataset of the demographic profile is the “total population, sex by age, 2012-2016 American Community Survey (ACS) 5-Year Estimates” in Mecklenburg County, North Carolina. This dataset has very fine resolutions on age groups. Particularly, it contains the age groups below and above 10 years old, which are the major components when calculating the TD populations as shown in section 3.4. Figure 5.2 displays the spatial distribution of the total populations within each census blockgroup in the City of Charlotte. The total population of Charlotte is 842,629. By exploring the dataset, three blockgroups in the City of Charlotte are found to have no residential population and therefore are excluded from further analyses.

The next dataset is the “total population, household type (including living alone) by relationship, 2012-2016 American Community Survey (ACS) 5-Year Estimates” in Mecklenburg County, North Carolina. According to US Census Bureau (2010), “A group quarters is a place where people live or stay, in a group living arrangement that is owned or managed by an entity or organization providing housing and/or services for the residents.” Thus, group quarter is not a typical household-type living arrangement. Statistics are used to exclude the population living in the group quarters as illustrated in section 3.4. Figure 5.3 shows the spatial distribution of population living in the group quarters within each census blockgroup in the City of Charlotte. The total number of people living in the group quarter is 12,840.

The last demographic profile dataset is the “aggregate number of vehicles available by tenure, Occupied housing units, 2012-2016 American Community Survey (ACS) 5-Year Estimates”. Again, as mentioned in section 3.4, excluding the vehicle numbers from the total population is a very crucial part of determining the potential maximum TD population. With simple calculations, this dataset can provide the vehicle numbers of each blockgroups. Figure 5.4 displays the spatial distribution of vehicles within each census blockgroup in the City of Charlotte. The total number of vehicles is 502,276.

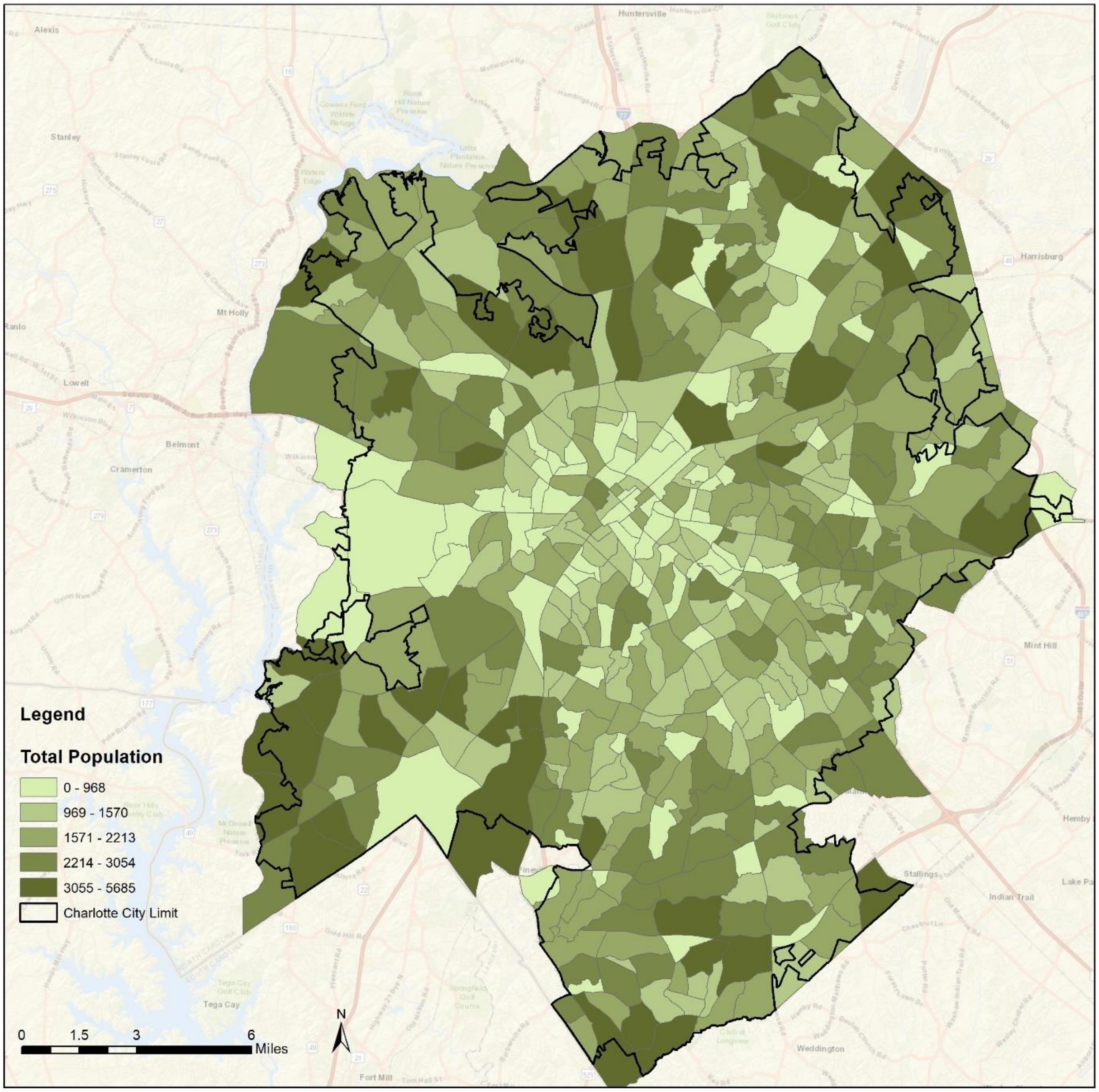


Figure 5.2 Distribution of Total Population in the City of Charlotte

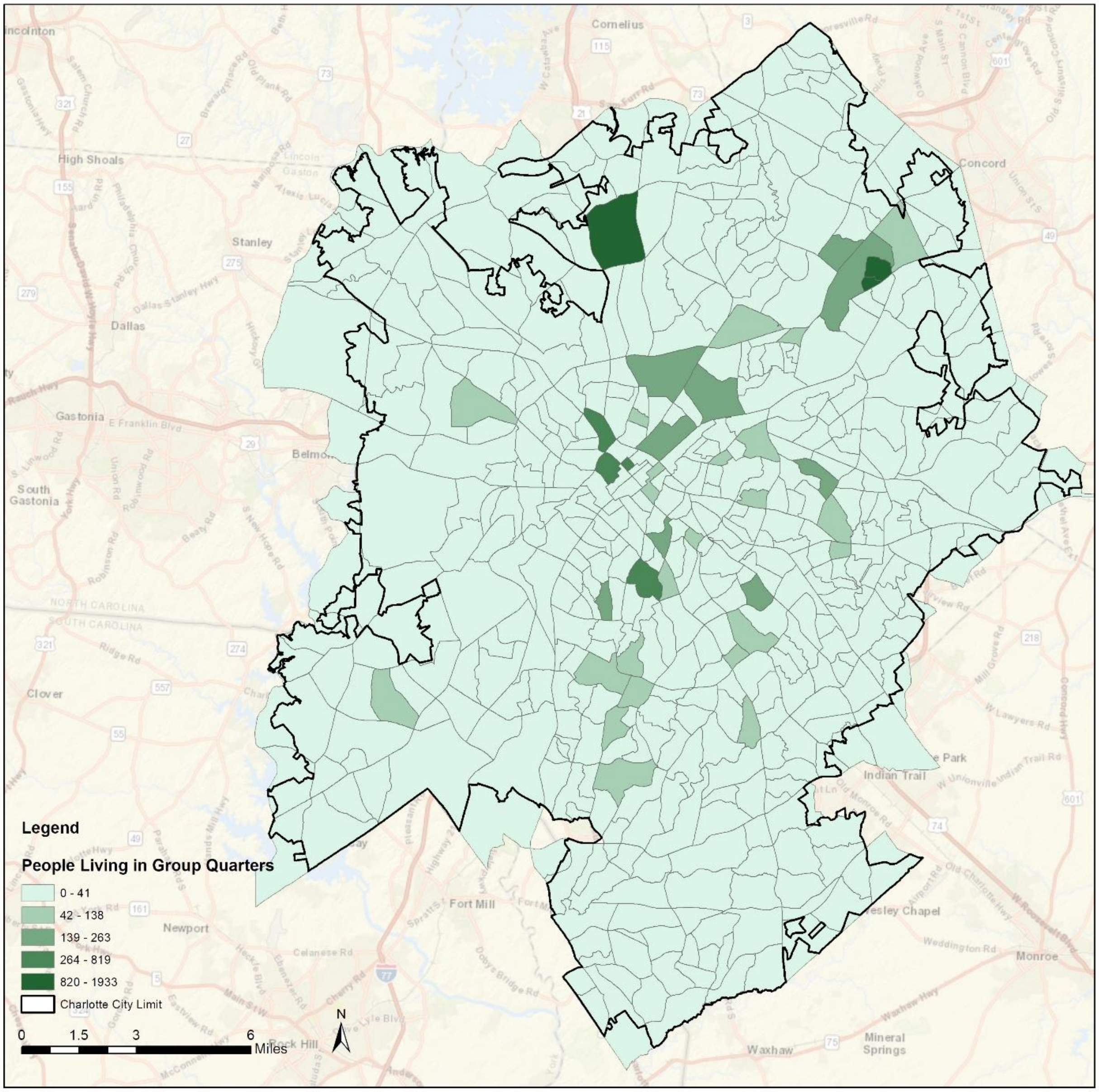


Figure 5.3 Distribution of People Living in Group Quarters in the City of Charlotte

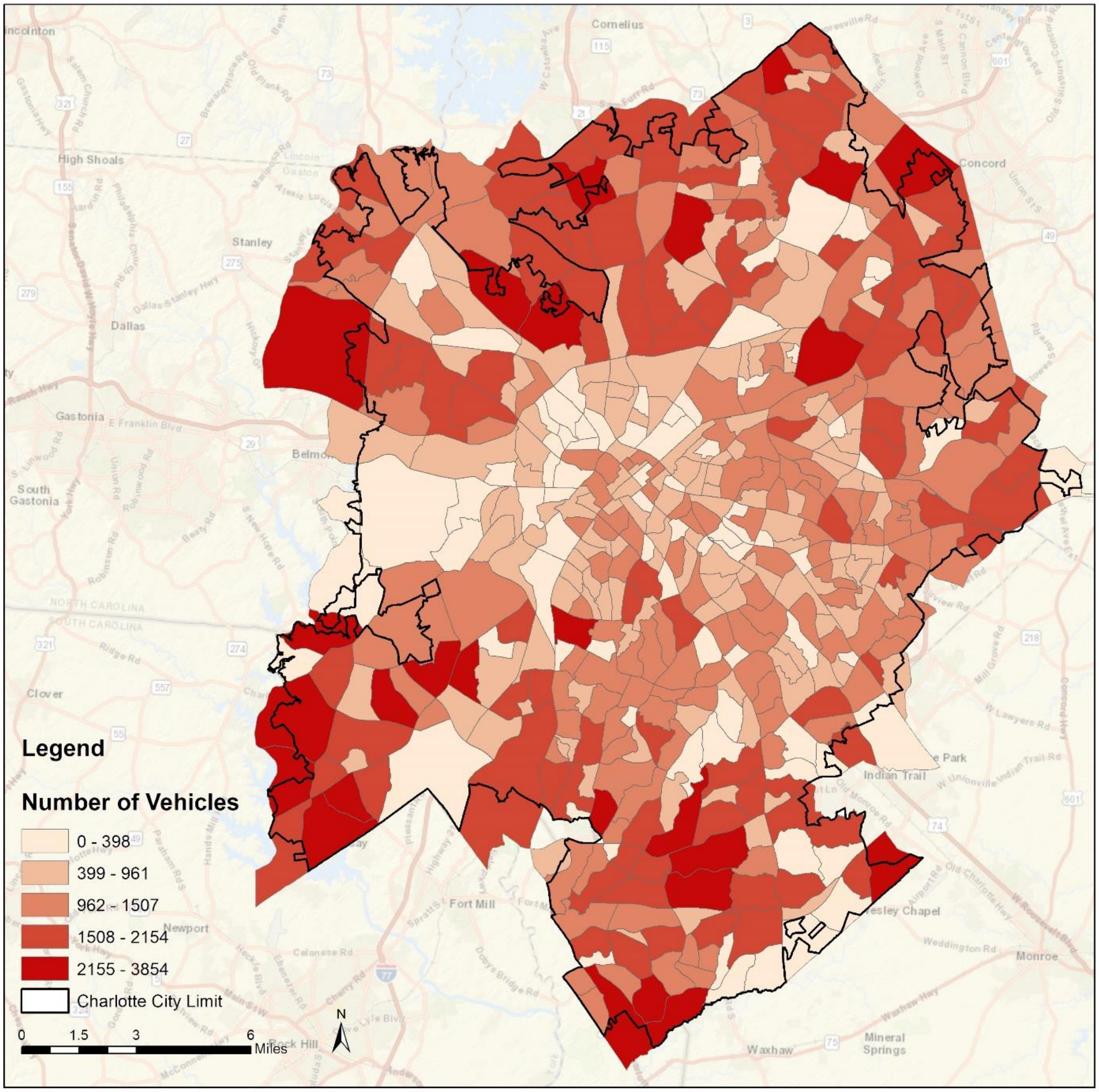


Figure 5.4 Distribution of Vehicles in the City of Charlotte

Table 5.3 gives a summary of the demographic datasets used in this research, which shows the number of people in each category of interest with respective percentage compared to the total population.

Table 5.4 Summary of the Demographic Datasets

Number of People Living in the Quarter Group	Number of Vehicles	Number of People Over 10 Years Old	Number of People Under 10 Years Old	Total Population
12,840	502,276	722,305	120,324	842,629
1.52%	59.61%	85.72%	14.28%	100.00%

5.4 Transportation Data

Despite the public transit route system, the roadway system in the City of Charlotte is also required to implement the methodology in this study. The primary purpose of the use of roadway system is to determine the 0.5-mile walking catchment area for each public transit stop/station. The North Carolina Statewide System and Non-System Road system, an ArcGIS shapefile acquired from "GIS Data Layers-Connect NCDOT" (https://connect.ncdot.gov/resources/gis/pages/gis-data-layers.aspx), is used in this study. Figure 5.5 is the dataset input to the ArcGIS by showing the roadways and their associating roadway classes in the region of the City of Charlotte.

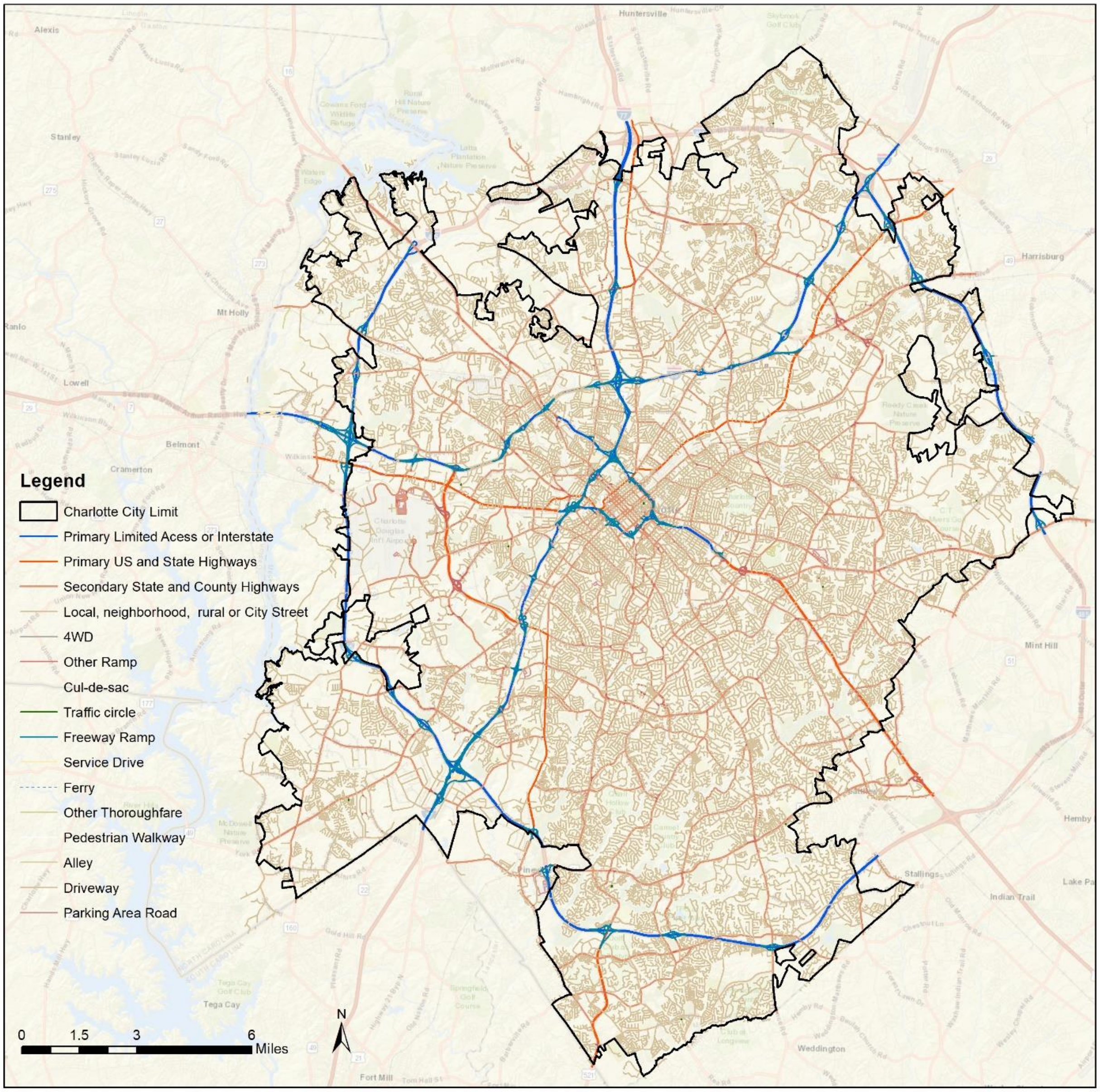


Figure 5.5 North Carolina Statewide System and Non-System Road System in the City of Charlotte

5.5 Other Regional Data

This section lists the other regional datasets in the City of Charlotte. The first one is the shapefile of “Charlotte City Council Districts” and it is obtained from the “City of Charlotte Open Data Portal” (http://clt-charlotte.opendata.arcgis.com/datasets/dc81ea7a87a440f282776f79fa7e1485_0). It contains the boundaries and contact information about Charlotte’s City Council Districts. The second one is the “Parcel Boundaries and Standard Fields, Integrated Cadastral Data Exchange, Mecklenburg County, North Carolina” dataset, which contains the parcel data. Such data can be found in “NC OneMap GeoPortal” (http://data.nconemap.gov/geoportal/catalog/main/home.page).

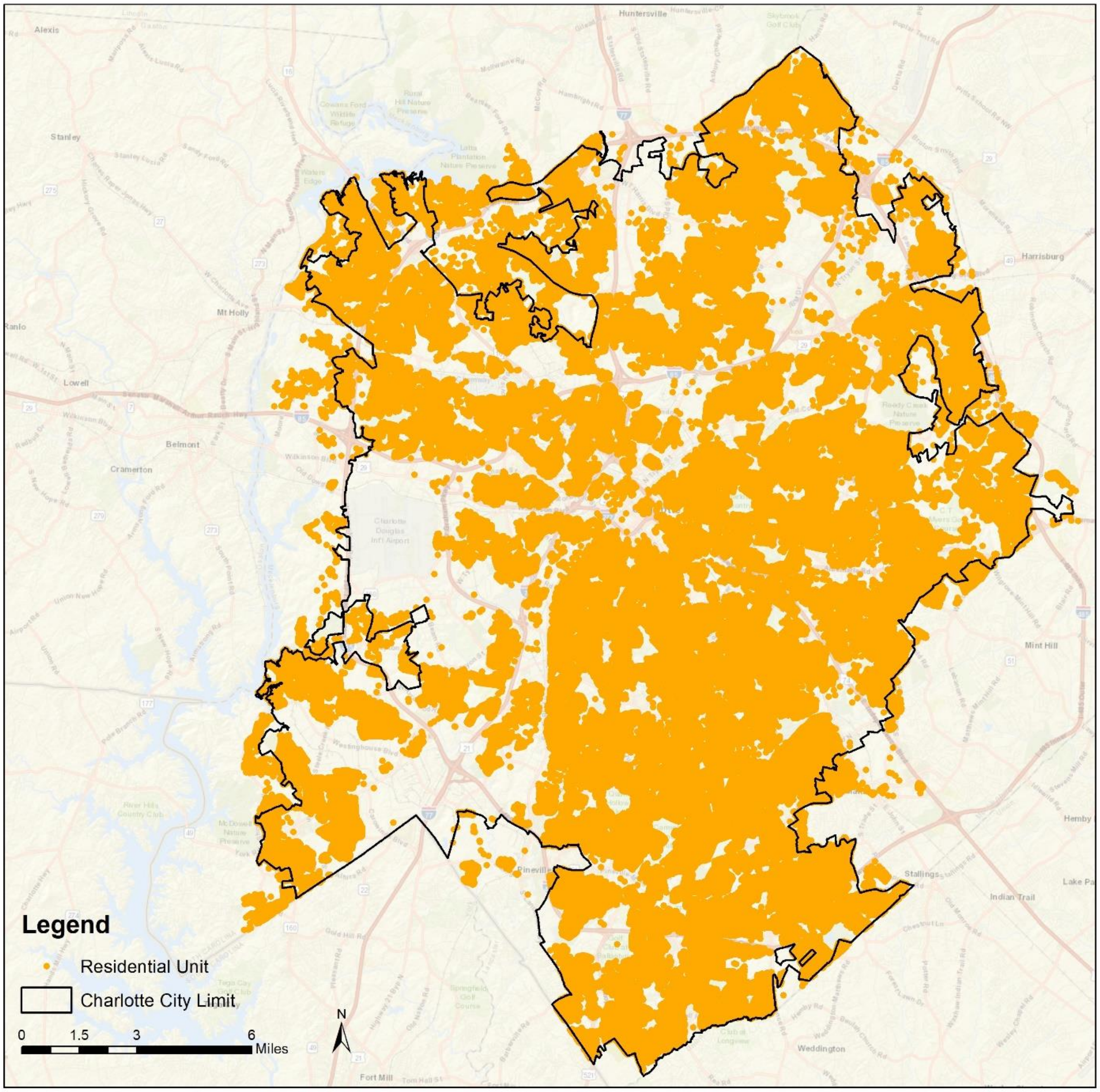


Figure 5.6 Distribution of Residential Units in the City of Charlotte

With some filtrations, there are 260,531 residential units in the City of Charlotte. Figure 5.6 shows both data in ArcGIS, and the parcel data displayed in the figure has already been filtered with only residential buildings being left.

Due to the reason that GTFS data of CATS do not include any transit route frequency information in the “frequencies.txt” file, as a supplement, information about the bus capacities, routes and schedules is collected on the website of CATS (http://charlottenc.gov/cats/Pages/default.aspx). Furthermore, since the version of the only available GTFS data of CATS is a little behind the current CATS, coordination between GTFS data and current CATS has to be made as follows: 1) non-existed routes and stops/stations in current CATS are removed from GTFS; and 2) routes with unmatched names from GTFS data are adjusted to the actually existing routes of CATS. A total of 68 out of 75 routes are kept and 3074 stops/stations are left.

5.6 Blockgroups with Transit Deficiency

Based on the results from (Fan and Li, 2019), there are 31 areas (blockgroups) with public transit service deficiency (“High” and “Very High” *TDS’* with “Low” and “Very Low” *TSS’*) and redundancy (“Low” and “Very Low” *TDS’* with “High” and “Very High” *TSS’*). Figure 5.7 presents the distribution of the blockgroups for both *TSS’* and *TDS’* in each category combination area, as well as the corresponding Jenks natural breaks of *TSS’* and *TDS’*. And Figure 5.8 shows the areas with both public transit deficiency and redundancy spatially.

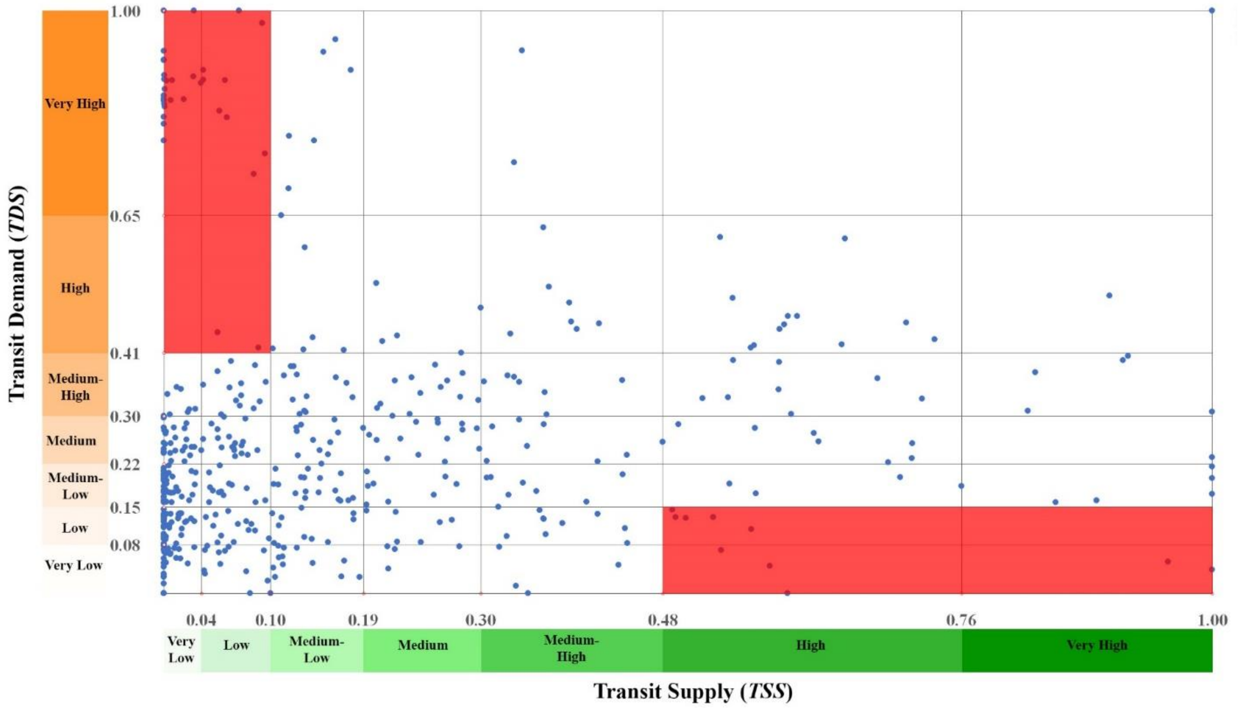


Figure 5.7 Scatter Plot of Transit Supply (*TSS*) and Transit Demand (*TDS*)

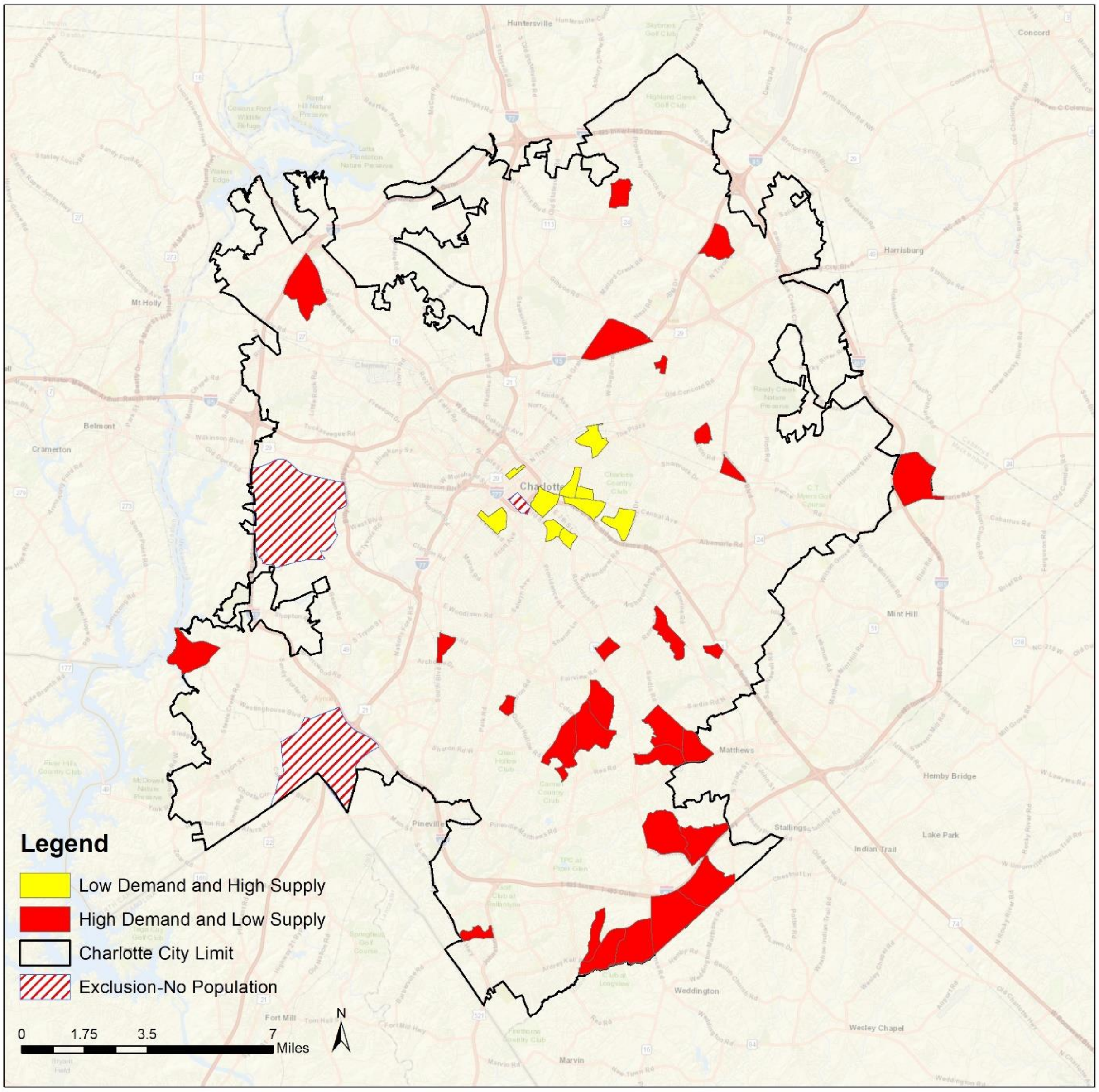


Figure 5.8 Spatial Distribution of Transit Supply Redundancies and Deficiencies

Table 5.5 displays the basic information of blockgroups with transit deficiency. The information includes the total population, *TDS'*, *TSS'* and *TGI* will be presented in Chapter 5Such data is the output of transit gap analysis (i.e., analysis of *TGI*) and also can be seen as the input of the optimization models in this project.

Table 5.5 General Characteristics of CATS based on GTFS Data

Blogkgroup ID	Total Population	*TSS'*	*TDS'*	*TGI*
371190031023	1371	0.06	0.88	-0.82
371190015071	1712	0.06	0.82	-0.76

Blogkgroup ID	**Total Population**	***TSS'***	***TDS'***	***TGI***
371190015083	884	0.04	0.88	-0.84
371190019153	968	0.04	0.90	-0.86
371190020024	1207	0.05	0.83	-0.77
371190020031	1299	0.00	0.84	-0.84
371190029041	1578	0.03	0.89	-0.86
371190030072	2267	0.01	0.85	-0.84
371190030073	1824	0.01	0.88	-0.87
371190030112	510	0.09	0.98	-0.88
371190030152	1685	0.00	0.88	-0.88
371190030153	1244	0.04	0.88	-0.84
371190030162	898	0.02	0.85	-0.83
371190053082	1103	0.09	0.72	-0.63
371190055133	846	0.00	0.88	-0.88
371190055233	1667	0.00	0.89	-0.89
371190055246	1339	0.10	0.75	-0.66
371190056212	707	0.00	0.78	-0.78
371190058231	992	0.00	0.92	-0.92
371190058232	1578	0.00	0.85	-0.85
371190058373	859	0.00	0.82	-0.82
371190058451	1234	0.00	0.81	-0.81
371190058461	1938	0.00	0.93	-0.93
371190058462	1973	0.00	0.91	-0.91
371190058471	1371	0.00	0.87	-0.86
371190058482	1923	0.00	0.85	-0.85
371190059072	1277	0.00	0.85	-0.85
371190060101	2298	0.00	0.84	-0.83

5.7 Summary

This chapter presents the detail information about all the data that are needed to conduct the case study in the City of Charlotte to implement the methodology that has been developed in this research. Meanwhile, the ways of handling and utilizing each dataset are also provided.

Chapter 6. Numerical Results

6.1 Introduction

As described in Chapter 4, two models aiming at improving the public transit equity are developed by considering two different conditions (i.e., with limited budget and without budget constraint). This chapter focuses on the numerical results of the both developed models. Numerical results of the case study in the City of Charlotte are analyzed and presented in detail.

The remainder of this chapter is organized as follows: Section 6.2 provides the parameter setup with supplemental data used in the case study. Section 6.3 and Section 6.4 give detailed results and their analysis for both models. Finally, a summary concludes this chapter in Section 6.5.

6.2 Parameter Setup with Supplemental Data

Other than the data presented in Chapter 5, in order to conduct the case study under several different designed scenarios, there are some other supplemental data used for parameter setup, such as the budget (B), potential stops' capacities (a_i), maximal number of stops (s_{max}) and construction cost (c_i) for one stop.

6.2.1 Budget Information (B)

According to the City of Charlotte City Council (2018), it is easy to obtain the annual budget of Charlotte Area Transit System (CATS) for "Transit Facilities (Transportation and Planning)" from 2015 to 2018. Table 6.1 shows this basic information.

Table 6.1 Annually Budget of CATS "Transit Facilities (Transportation and Planning)"

FY 2015	FY 2016	FY 2017	FY 2018
$5,411,637	$5,925,558	$5,737,159	$8,337,458

6.2.2 Potential Stops' Capacities (a_i)

In the case study, the potential stops' capacities (a_i) added to each blockgroup with transit deficiency can be determined as the average capacity of the stops within the blockgroup or the capacity of the nearest stops if there are no stops within the blockgroup, which can be viewed as the extensions of existing transit routes. Table 6.2 displays this information.

Table 6.2 Potential Stops' Capacities (a_i) for Each Blockgroup

Blogkgroup ID	Potential Stops' Capacities
371190015071	2960
371190015083	3000
371190019153	3600
371190020024	3240
371190020031	4360
371190029041	320
371190030072	320

Blogkgroup ID	**Potential Stops' Capacities**
371190030073	320
371190030112	2120
371190030152	2080
371190030153	2080
371190030162	2080
371190031023	2920
371190053082	3480
371190055133	3160
371190055233	3680
371190055246	3160
371190056212	400
371190058231	400
371190058232	400
371190058373	1920
371190058451	480
371190058461	480
371190058462	400
371190058471	400
371190058482	400
371190059072	2160
371190060101	240

6.2.3 Maximal Number of Stops (s_{max}) and Construction Cost (c_i) for One Stop

The maximal number of the stops that can be added to each blockgroup with transit deficiency is set to be 40. This case study uses $12,000 as the construction cost (c_i) for one stop, according to Angie Schmitt (2018).

6.3 Results of Model with Limited Budget

According to the budget information introduced in the Section 6.3, the maximal budget that can be invested into new stop construction should not exceed $8,337,458. But in order to further conduct the sensitivity analysis of the budget, there are 12 scenarios designed for the model with limited budget constraint, ranging from $1,000,000 to $12,000,000. Table 6.3 and 6.4 exhibit both number of new stops installations and the improvements of *TGI* for each blockgroup under 12 scenarios with different budget constraints. Figures 6.1 to 6.4 show both results in a combo chart manner and provide a clear picture of the tendencies of both results with the changes of the budget value in the limited budget constraint.

Table 6.3 Number of New Stops Installations under Different Budgets Constraints

Blockgroup	B =$1M	B =$2M	B =$3M	B =$4M	B =$5M	B =$6M	B =$7M	B =$8M	B =$9M	B =$10M	B =$11M	B =$12M
371190031023 (1)	2	12	18	23	29	31	33	34	35	35	36	38
371190015071 (2)	0	3	12	21	29	33	35	37	38	38	40	40
371190015083 (3)	11	14	17	19	21	22	23	23	23	23	24	24
371190019153 (4)	11	14	16	18	20	21	21	22	22	22	22	23
371190020024 (5)	6	12	16	19	23	24	25	26	26	27	27	28
371190020031 (6)	10	14	17	19	21	22	23	23	23	24	24	25
371190029041 (7)	0	0	0	0	0	0	0	0	1	36	40	40
371190030072 (8)	0	0	0	0	0	0	0	0	0	0	28	40
371190030073 (9)	0	0	0	0	0	0	0	0	0	0	40	40
371190030112 (10)	12	14	16	17	19	19	20	20	20	20	21	21
371190030152 (11)	0	0	9	26	40	40	40	40	40	40	40	40
371190030153 (12)	0	7	16	26	35	38	40	40	40	40	40	40
371190030162 (13)	5	13	18	23	28	30	31	32	33	33	34	35
371190053082 (14)	4	8	11	13	16	17	18	18	18	18	19	20
371190055133 (15)	12	15	17	19	20	21	22	22	22	23	23	23
371190055233 (16)	6	15	21	26	31	33	35	36	37	37	38	40
371190055246 (17)	0	6	11	15	20	22	23	24	25	25	26	27
371190056212 (18)	0	0	0	0	1	34	40	40	40	40	40	40
371190058231 (19)	0	0	0	0	0	24	40	40	40	40	40	40
371190058232 (20)	0	0	0	0	0	0	0	7	40	40	40	40
371190058373 (21)	4	12	18	23	28	30	32	32	33	33	34	36
371190058451 (22)	0	0	0	0	0	0	40	40	40	40	40	40
371190058461 (23)	0	0	0	0	0	0	0	30	40	40	40	40
371190058462 (24)	0	0	0	0	0	0	0	0	29	40	40	40
371190058471 (25)	0	0	0	0	0	0	2	40	40	40	40	40
371190058482 (26)	0	0	0	0	0	0	0	0	5	39	40	40
371190059072 (27)	0	7	17	26	35	39	40	40	40	40	40	40
371190060101 (28)	0	0	0	0	0	0	0	0	0	0	0	40

Table 6.4 Improvements of *TGI* for Each Blockgroup under Different Budgets Constraints

Blockgroup	B =$1M	B =$2M	B =$3M	B =$4M	B =$5M	B =$6M	B =$7M	B =$8M	B =$9M	B =$10M	B =$11M	B =$12M
371190031023 (1)	5%	31%	47%	60%	75%	80%	86%	88%	91%	91%	93%	98%
371190015071 (2)	0%	7%	27%	48%	66%	76%	80%	85%	87%	87%	92%	92%
371190015083 (3)	44%	56%	68%	76%	84%	88%	93%	93%	93%	93%	97%	97%
371190019153 (4)	48%	61%	69%	78%	86%	91%	91%	95%	95%	95%	95%	99%
371190020024 (5)	21%	42%	55%	66%	80%	83%	87%	90%	90%	94%	94%	97%
371190020031 (6)	40%	56%	68%	76%	84%	88%	92%	92%	92%	96%	96%	100%
371190029041 (7)	0%	0%	0%	0%	0%	0%	0%	0%	0%	9%	9%	9%
371190030072 (8)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	5%	7%
371190030073 (9)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	8%	8%
371190030112 (10)	56%	66%	75%	80%	89%	89%	94%	94%	94%	94%	99%	99%
371190030152 (11)	0%	0%	13%	37%	56%	56%	56%	56%	56%	56%	56%	56%
371190030153 (12)	0%	14%	32%	52%	70%	76%	80%	80%	80%	80%	80%	80%
371190030162 (13)	14%	36%	50%	64%	78%	84%	87%	89%	92%	92%	95%	98%
371190053082 (14)	20%	40%	55%	65%	80%	85%	90%	90%	90%	90%	95%	100%
371190055133 (15)	51%	64%	72%	81%	85%	89%	93%	93%	93%	97%	97%	97%
371190055233 (16)	15%	37%	52%	65%	77%	82%	87%	89%	92%	92%	94%	99%
371190055246 (17)	0%	22%	39%	54%	72%	79%	83%	86%	90%	90%	93%	97%
371190056212 (18)	0%	0%	0%	0%	1%	25%	29%	29%	29%	29%	29%	29%
371190058231 (19)	0%	0%	0%	0%	0%	11%	18%	18%	18%	18%	18%	18%
371190058232 (20)	0%	0%	0%	0%	0%	0%	0%	2%	12%	12%	12%	12%
371190058373 (21)	11%	33%	49%	63%	77%	82%	88%	88%	90%	90%	93%	98%
371190058451 (22)	0%	0%	0%	0%	0%	0%	19%	19%	19%	19%	19%	19%
371190058461 (23)	0%	0%	0%	0%	0%	0%	0%	8%	11%	11%	11%	11%
371190058462 (24)	0%	0%	0%	0%	0%	0%	0%	0%	6%	9%	9%	9%
371190058471 (25)	0%	0%	0%	0%	0%	0%	1%	14%	14%	14%	14%	14%
371190058482 (26)	0%	0%	0%	0%	0%	0%	0%	0%	1%	10%	10%	10%
371190059072 (27)	0%	14%	34%	52%	70%	78%	80%	80%	80%	80%	80%	80%
371190060101 (28)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	5%

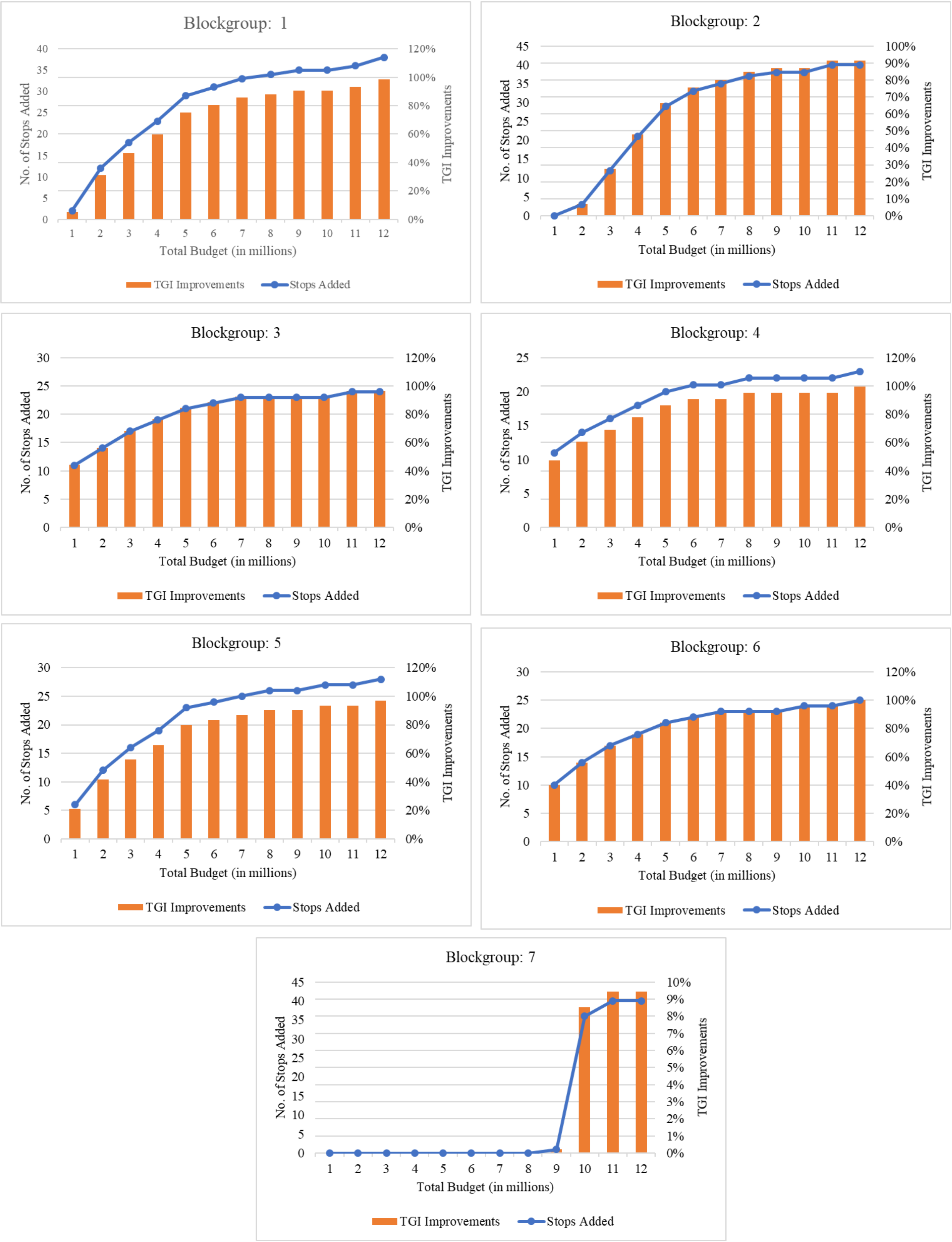


Figure 6.1 Combo Charts of Number of New Stops Installations and Improvements of *TGI* for Each Blockgroup under Different Budget Constraints

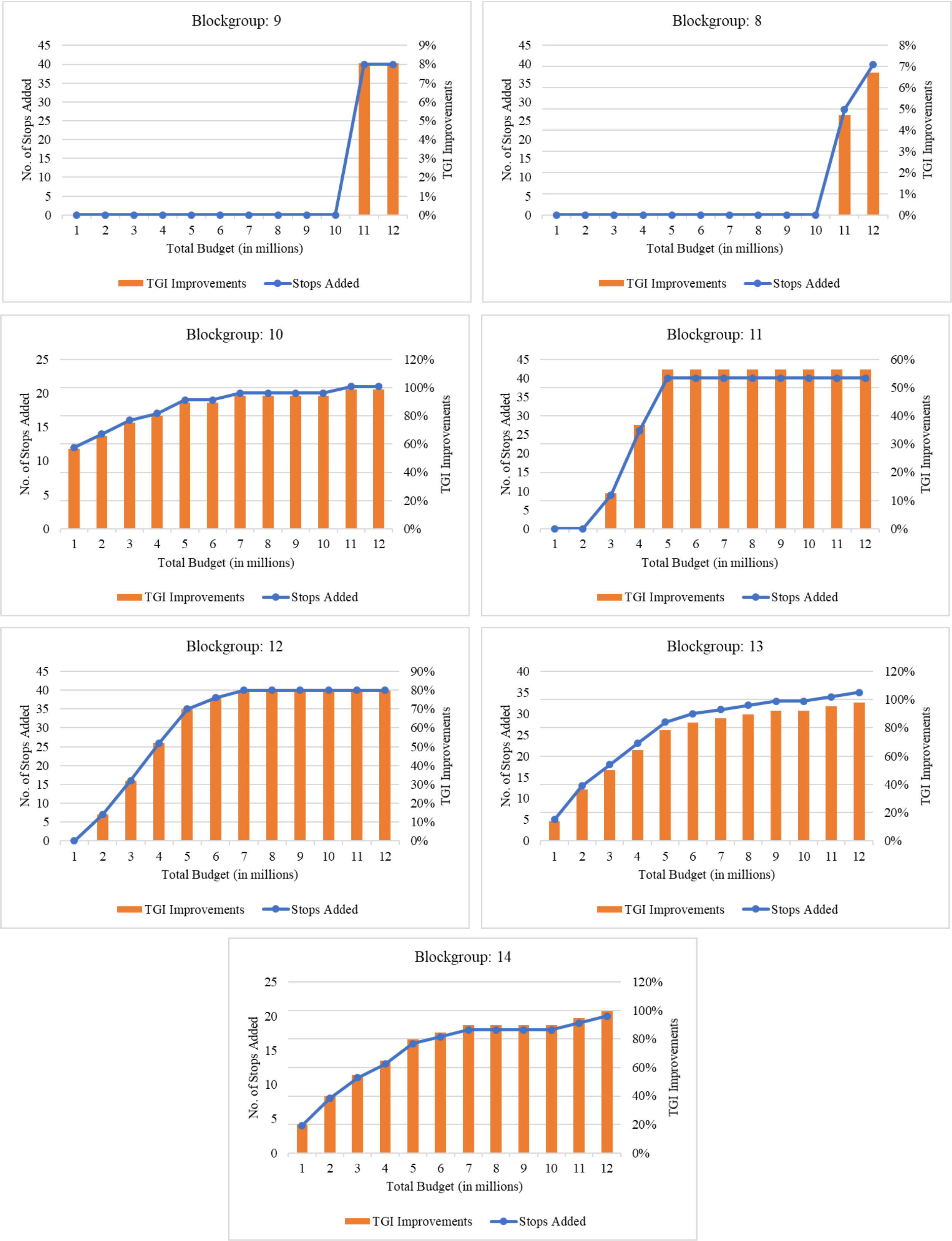


Figure 6.2 Combo Charts of Number of New Stops Installations and Improvements of *TGI* for Each Blockgroup under Different Budget Constraints

Figure 6.3 Combo Charts of Number of New Stops Installations and Improvements of TGI for Each Blockgroup under Different Budget Constraints

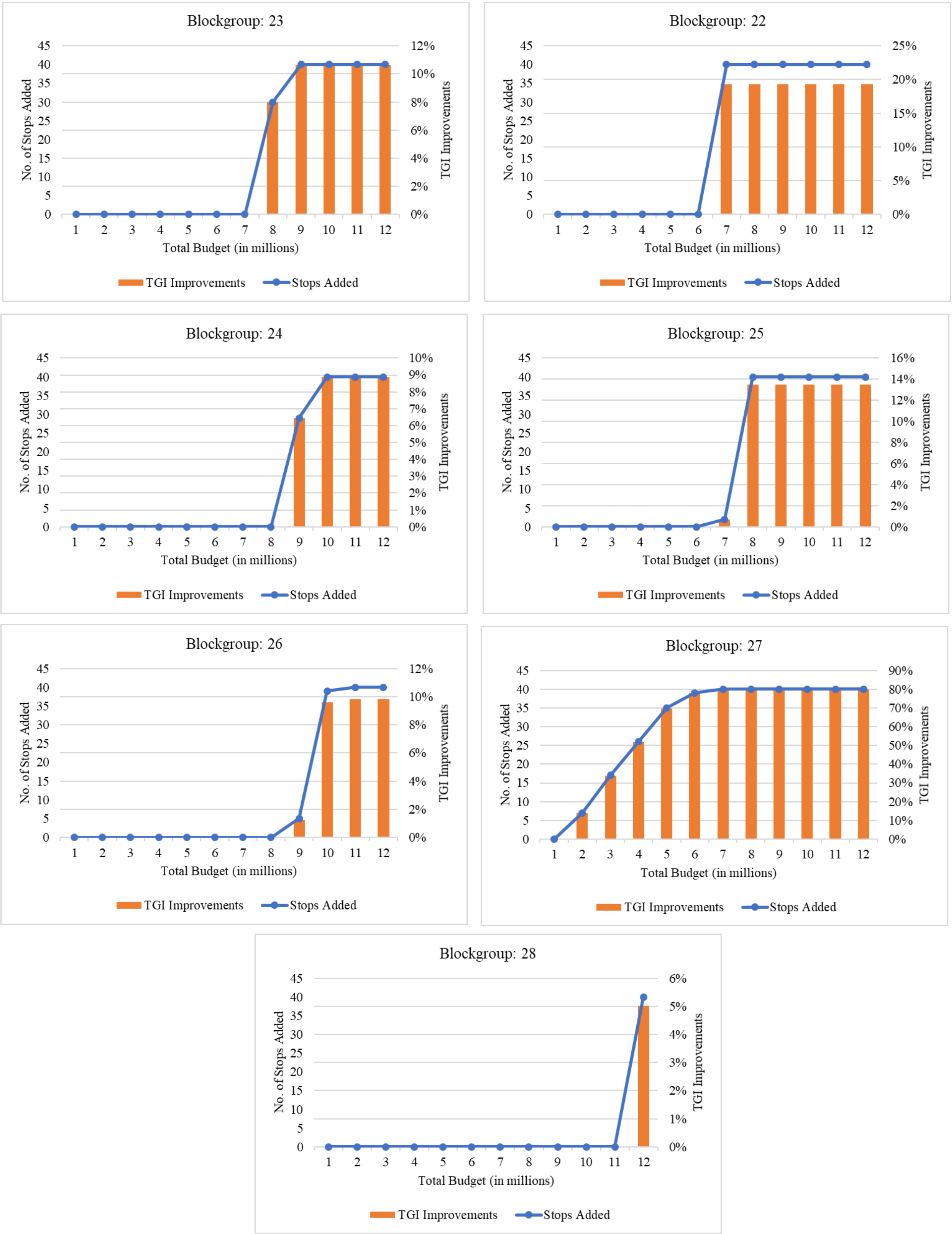


Figure 6.4 Combo Charts of Number of New Stops Installations and Improvements of TGI for Each Blockgroup under Different Budget Constraints

Figure 6.5 displays the changes of model objectives (*TGI*^{2}) with changes of the limited budget constraints. And for the case study in this research, model shows no improvement of the objective function when the budget exceeds $12,000,000. One major reason is the maximal number of stops that can be added to each blockgroup, which limits the further improvements. Another reason is that some of the blockgroups have relatively small potential stops' capacity. This can also be seen in Table 6.3 and 6.4 that though some of the blockgroups have the maximal number of stops added, the improvements are far below 100% (i.e., blockgroups 7, 8, 9, 18, 19, 20, 22, 23, 24, 25, 26, 28). Associating with the fact that most of those blockgroups reside at the edge of the study area and do not even have transit service (i.e., no transit routes pass by), the result of such sensitivity analysis of the limited budget constraint is not hard to understand and also makes sense.

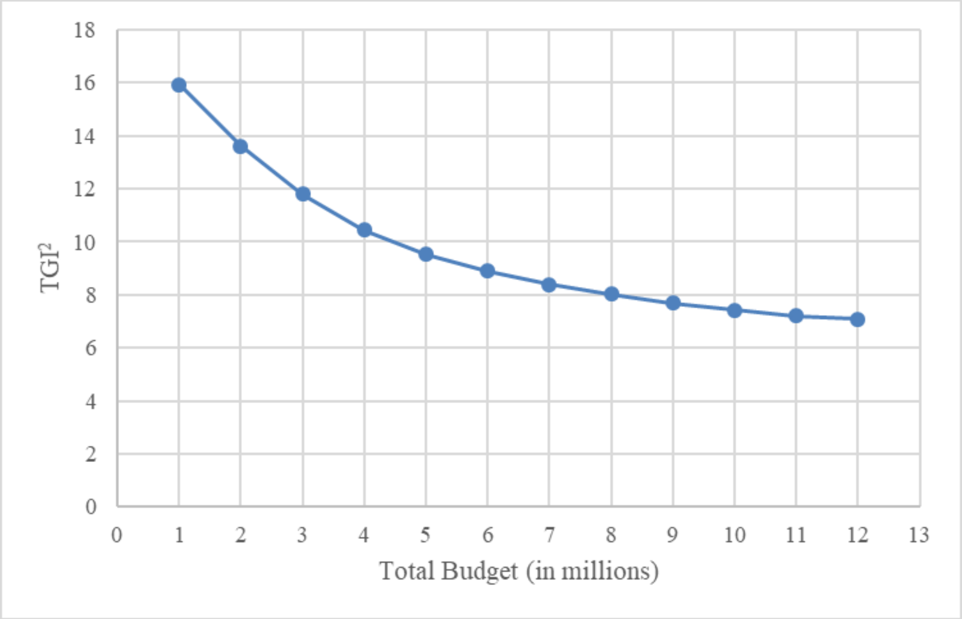


Figure 6.5 Changes of Objectives (*TGI*^{2}) with Changes of the Limited Budget Constraints

6.4 Results of Model without Budget

As introduced in Chapter 4, unlike the previous model with limited budget constraint, this model does not have such constraint but tries to minimize the total construction cost by constraining on certain level of improvements to transit services and a certain number of blockgroups that meet such improvements requirement. Three main scenarios are designed to demonstrate the capability of the model for the case study.

For the first scenario, at least half of blockgroups are set to meet certain improvements of transit services, which is equivalent to "N =14". And there are four different requirements of the improvement requirements within this main scenario, forming four sub-scenarios, which are equivalent to "L = 10%", "L = 25%", "L = 50%" and "L = 75%", respectively. And Figure 6.6 presents results of first scenario with half blockgroups meeting certain improvement requirements, including new stop installations of each blockgroup and total construction cost (i.e., "Budget" in the figure) under four sub-scenarios.

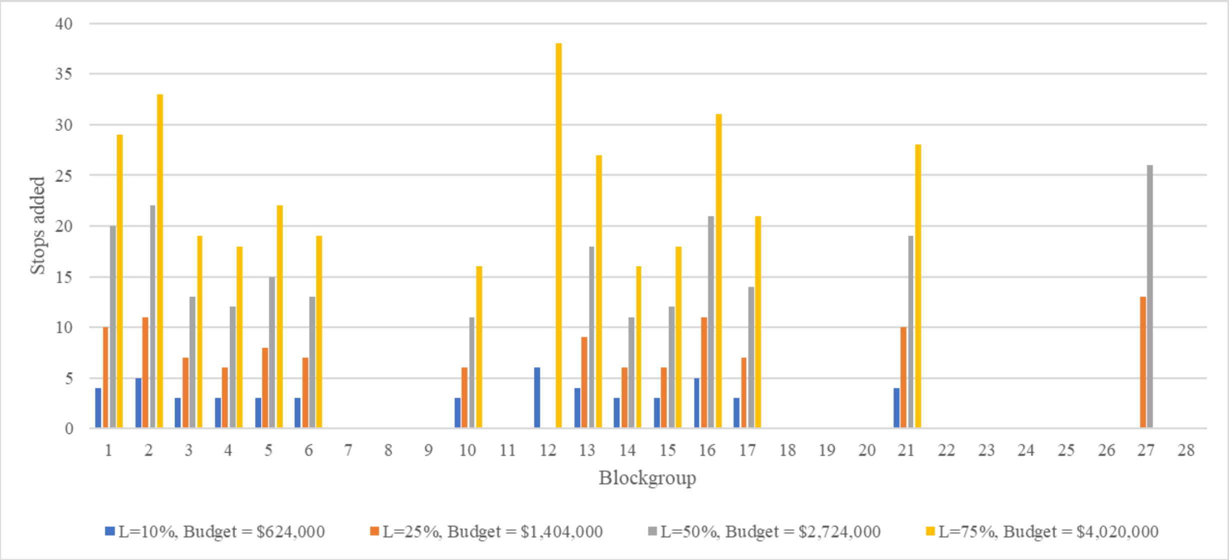


Figure 6.6 Results of First Scenario with Half Blockgroups Meet Certain Improvement Requirements

For the second main scenario, the certain number of blockgroups that meet certain improvements of transit services increases to 75%, which is equivalent to “*N* =21”. Unlike the first scenario, there are two sub-scenarios with “*L* = 5%” and “*L* = 10%”. In this case study, if 75% of the blockgroups are set to meet the criteria, no feasible solutions can be found with higher requirements on the improvement that are designed in the first scenario (i.e., “*L* = 25%”, “*L* = 50%” and “*L* = 75%”). Figure 6.7 displays results of second scenario with 75% of blockgroups meeting certain improvement requirements, including new stop installations of each blockgroup and total construction cost (i.e., “Budget” in the figure) under two sub-scenarios.

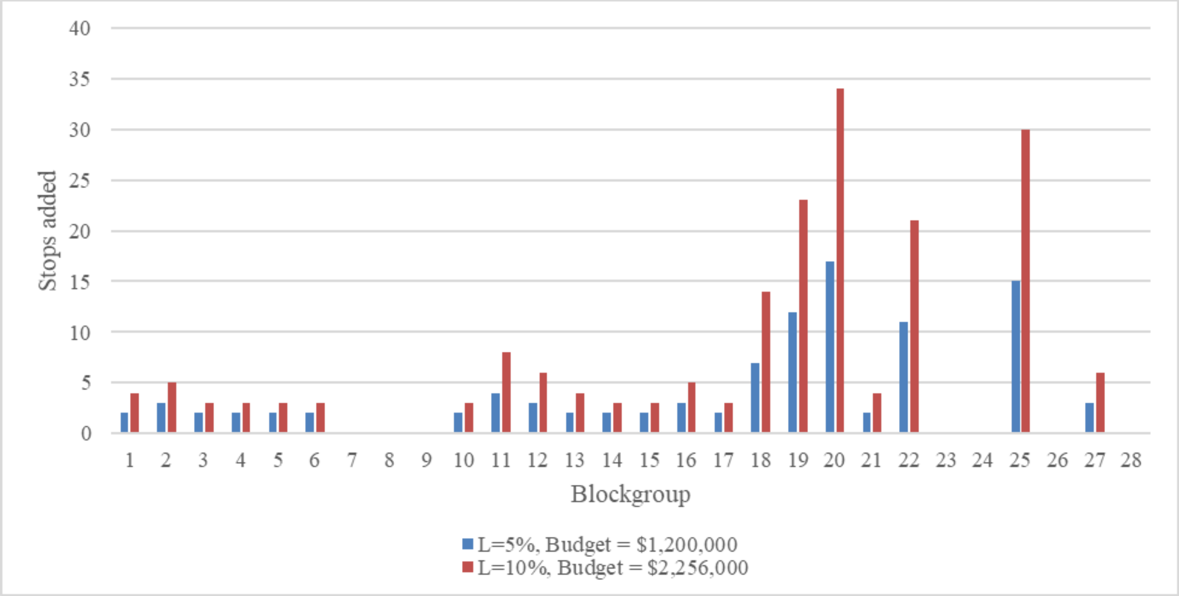


Figure 6.7 Results of Second Scenario with 75% of Blockgroups Meet Certain Improvement Requirements

The third scenario is more comprehensive than the previous two. By expanding the two constraints in the model (i.e., expanding $\sum_{i\in I} y_i \ge N$ and $\frac{TSS_i'' - TSS_i'}{|TGI_i|} \ge L$ into $\sum_{i\in I} y_i^j \ge N_j$ and $\frac{TSS_i'' - TSS_i'}{|TGI_i|} \ge L_j$), we want to have all blockgroups meet the improvement requirement of at least 5% (i.e., N_1 = 28, L_1 = 5%), at least 75% of blockgroups meet the improvement requirement of at least 10% (i.e., N_2 = 21, L_2 = 10%), at least 50% of blockgroups meet the improvement requirement of at least 25% (i.e., N_3 = 14, L_3 = 25%) and at least 25% of blockgroups meet the improvement requirement of at least 50% (i.e., N_4 = 7, L_4 = 50%) at the same time. Figure 6.8 show the results of the third scenario, including the new stop installations for each blockgroup and the total construction cost for all installations, which is $5,688,000.

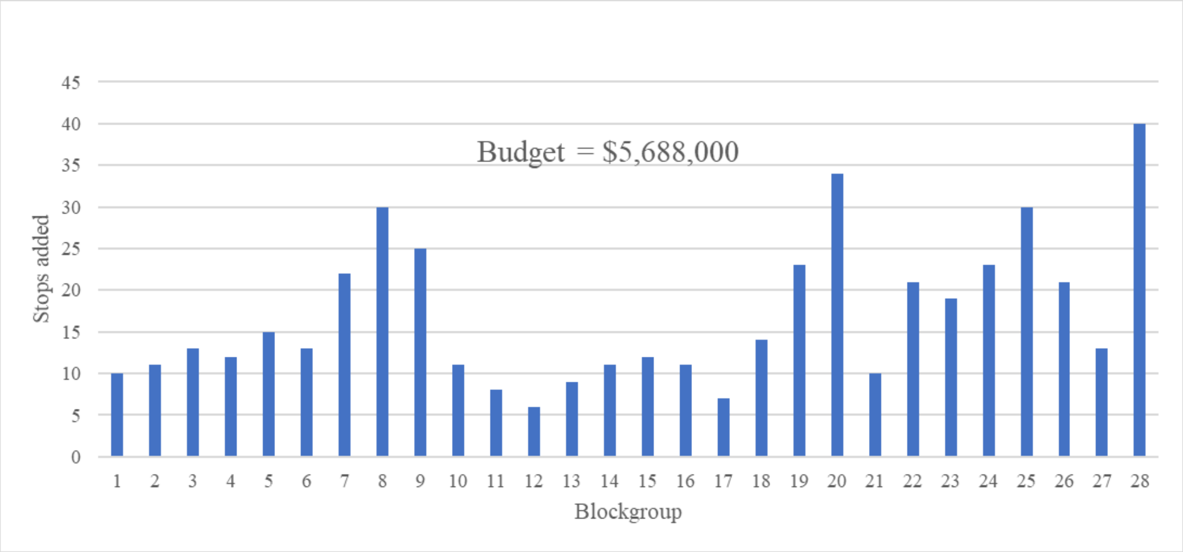


Figure 6.8 Results under the Third Scenario

6.5 Summary

This chapter presents detailed results of two optimization models with/without budget constraints aiming at improving transit equity and accessibility for people by integrating performance metrics by using GTFS data that are developed and applied in the case study of the City of Charlotte. Several scenarios are designed and analyzed to show the capability of models.

Chapter 7. Summary and Conclusions

7.1 Introduction

Public transit mode continues to be a crucial part of transportation planning in the United States. Building upon the assessment of public transit equity and accessibility, one important task of transit planning is to optimize relevant metrics and measurements by modifying current service parameters (e.g., route layouts, schedules, and frequencies) or having new investments. In this context, many studies had been done to perform the gap analysis and discuss the potential use of such analysis for further optimizing the public transit services via modification of current transit systems. However, efforts still need to be made to enhance and enrich relevant research on integrating transit equity into a network design problem.

In the meantime, recent development of the General Transit Feed Specification (GTFS), a well formatted transit feeds open data, provides new opportunities for a better understanding of both spatial and temporal characteristics of public transit because such data is easy to handle and is proved efficient in relevant analysis. By taking such advantages, it is necessary to further leverage such data source to contribute to both the state-of-the-art and the state-of-the-practice in this field, and efforts are needed to review the current practices and develop appropriate mathematical optimization models for improving the public transit equity and accessibility by integrating GTFS data relevant performance metrics.

The primary objective of this research is to develop optimization models for improving transit equity and accessibility for people by integrating performance metrics with using GTFS data. Two optimization models under two different conditions (with limited budget and without limited budget) for improving the public transit equity of blockgroups with transit deficiency are built. A case study in the City of Charlotte and the comprehensive numerical results associated with the proposed models are also conducted and presented.

The rest of this chapter is organized as follows. In section 7.2, major works of this report are reviewed and the capability of both models is discussed with the goal to improve transit equity and accessibility for people by integrating performance metrics with using GTFS data. Section 7.3 presents a brief discussion of the limitations of the current approaches and possible directions for further research are also given.

7.2 Summary and Conclusions

As presented throughout the research, this report has discussed the development of optimization models for improving transit equity and accessibility for people by integrating performance metrics by using GTFS data, which is the *TGI* in our previous research (Fan and Li, 2019). A comprehensive review of the state-of-the-art/practices on public transit equity optimization has been conducted. Especially, those with optimizing the use of performance metrics utilizing GTFS data have been explored. Two optimization models under two different conditions (with limited budget and without limited budget) for improving the public transit equity of blockgroups with transit deficiency are built in this study. A case study in the City of Charlotte is provided along with discussions of detailed numerical results.

Two developed models in this project are developed to optimize the transit equity by mitigating the transit deficiency based on the study of Fan and Li (2019). Results of *TGI* and its associated components have been utilized in the model developments. The main idea is to improve the level of transit services in those blockgroups that suffer transit deficiency. And two different conditions are considered while developing both optimization models, which are: 1) maximizing the level of transit services with constraint of limited budget that will be invested into the new constructions of public transit stops; 2) minimizing the total cost for constructing new public transit stops with constraint that requires certain improvements of transit services for a certain amount of blockgroups. The proposed models can be classified as mixed integer nonlinear program (MINLP) models, since the decision variables in both models are discrete and both models contain nonlinear constraint(s). Both models are programmed using the AMPL modeling language and, due to the models' nonlinear nature, both models have been run on server BARON from NEOS.

Additionally, a case study in which the proposed optimization models are applied has been conducted in the City of Charlotte based on the results from the study of Fan and Li (2019). Sensitivity analysis of different budget constraints has been conducted for the first model with limited budget constraint. 12 scenarios with budget ranging from \$1,000,000 to \$12,000,000 have been introduced to see the changes of both number of new stops installations and the improvements of *TGI* for each blockgroup, along with the objective changes. The second model has no limited budget constraint but tries to minimize the total construction cost by constraining on certain level of improvements to transit services and a certain number of blockgroups that need to meet such improvements requirement. Three main scenarios are designed for this model to demonstrate its capability for the case study.

7.3 Directions for Future Research

In this section, some of the limitations of the proposed optimization models in this research are presented and directions for further research are also discussed.

Transit service deficiency is no doubt an important issue that has negative impact on transit equity and accessibility. Other than such issue, according to the previous study, another key problem would be the transit service redundancy that has been identified in the transit gap analysis. Though this project shows the ability of both models to handle the transit service deficiency, it does not necessarily consider the transit service redundancy. One potential future research direction could focus on integrating the redundancy portion into the model to conduct optimizations for such unbalanced distributions of the transit services.

Following the above statement, other than particular areas, a systematic redesign for all blockgroups might be important. Though not all areas are suffering severe transit deficiency, some areas might still not have sufficient transit services to meet their demands. By considering the redesign of the whole study area, such model would be more realistic and practical for transit planners to further improve overall planning.

In both models, the blockgroups chosen to have new constructions of stops and stations are assumed to have the transit service coverage (TSC_i) increased to "1" due to the reason that location(s) of stop(s) could be well examined to cover almost all residential units. In a more appropriate way to deal with this parameter, one might want to carefully examine the locations of new stop constructions within each blockgroup to see the true increment of this parameter in the future.

References

1. Anderson, K., Blanchard, S. D., Cheah, D., and Levitt, D. (2017). Incorporating Equity and Resiliency in Municipal Transportation Planning: Case Study of Mobility Hubs in Oakland, California. *Transportation Research Record: Journal of the Transportation Research Board*, (2653), 65-74.
2. Angie Schmitt (2018). Why We Need a Bus Shelter at Every Stop. *Streetsblock USA*. Retrieved from https://usa.streetsblog.org/2018/10/01/opinion-we-should-put-a-bus-shelter-at-every-stop-in-america/.
3. Antrim, A., and Barbeau, S. J. (2013). The Many Uses of GTFS Data–Opening the Door to Transit and Multimodal Applications. *Location-Aware Information Systems Laboratory at the University of South Florida*, 4.
4. Bejleri, I., Noh, S., Gu, Z., Steiner, R. L., and Winter, S. M. (2018). Analytical Method to Determine Transportation Service Gaps for Transportation Disadvantaged Populations. *Transportation Research Record*, 0361198118794290.
5. Bertolaccini, Kelly and Lownes, Nicholas E. (2014). Effects of Scale and Boundary Selection in Assessing Equity of Transit Supply Distribution. *Transportation Research Record: Journal of the Transportation Research Board 2350*, 58-64.
6. Catala, M., Dowling, S., and Hayward, D. (2011). *Expanding the Google Transit Feed Specification to Support Operations and Planning* (No. FDOT BDK85# 977-15).
7. Chen, P., & Nie, Y. (2016). A Demand Adaptive Paired-Line Hybrid Transit System. *Transportation. Research. Part B*.
8. City of Charlotte City Council, North Carolina, FY 2018 Proposed Budget. Retrieved from https://charlottenc.gov/newsroom/Documents/PrelimSOP_v9FINALFORWEBSITE.pdf.
9. Coffey, C., Nair, R., Pinelli, F., Pozdnoukhov, A., and Calabrese, F. (2012, November). Missed Connections: Quantifying and Optimizing Multi-Modal Interconnectivity in Cities. In *Proceedings of the 5th ACM SIGSPATIAL International Workshop on Computational Transportation Science* (pp. 26-32). ACM.
10. Daniels, R., and Mulley, C. (2013). Explaining Walking Distance to Public Transport: The Dominance of Public Transport Supply. *Journal of Transport and Land Use*, *6*(2), 5-20.
11. Durand, C. P., Tang, X., Gabriel, K. P., Sener, I. N., Oluyomi, A. O., Knell, G., Sener, I. N., Oluyomi A. O., Knell, G., Porter, A. K., Oelscher, D. M., and Kohl III, H. W. (2016). The Association of Trip Distance with Walking to Reach Public Transit: Data from the California Household Travel Survey. *Journal of Transport & Health*, *3*(2), 154-160.
12. El-Geneidy, A., Grimsrud, M., Wasfi, R., Tétreault, P., and Surprenant-Legault, J. (2014). New Evidence on Walking Distances to Transit Stops: Identifying Redundancies and Gaps Using Variable Service Areas. *Transportation*, *41*(1), 193-210.
13. Fan, W., and Li, Y. (2019). *Using General Transit Feed Specification (GTFS) Data as a Basis for Evaluating and Improving Public Transit Equity* (No. Project ID: 2018 Project 02).
14. Fourer, R. (2003). 2003 Software Survey: Linear Programming. *OR MS TODAY*, *30*(6), 34-43.
15. Grengs, J. (2001). Does Public Transit Counteract the Segregation of Carless Households? *Transportation Research Record* 1753: 3-10.
16. Jiang, Y., Zegras, P. C., and Mehndiratta, S. (2012). Walk the Line: Station Context, Corridor Type and Bus Rapid Transit Walk Access in Jinan, China. *Journal of Transport Geography*, *20*(1), 1-14.

17. Jiao, J., and Dillivan, M. (2013). Transit Deserts: The Gap between Demand and Supply. *Journal of Public Transportation*, *16*(3), 2.
18. Jiao, J., and Nichols, A. (2015). Identifying Transit Deserts in Texas Cities: The Gap between Supply and Demand. *Center for Sustainable Development, School of Architecture, The University of Texas, Austin, TX*.
19. Karner, A. (2016). Planning for Transportation Equity in Small Regions: Towards Meaningful Performance Assessment. *Transport policy*, *52*, 46-54.
20. Mamun, S. A., Lownes, N. E., Osleeb, J. P., and Bertolaccini, K. (2013). A Method to Define Public Transit Opportunity Space. *Journal of Transport Geography*, *28,* 144-154.
21. Oh, J. S., Smith, C. S., Qatra, R., and Al-Akash, M. (2017). *Estimating and Enhancing Public Transit Accessibility for People with Mobility Limitations* (No. TRCLC 15-02). Western Michigan University. Transportation Research Center for Livable Communities.
22. O'Sullivan, S., and Morrall, J. (1996). Walking Distances to and From Light-Rail Transit Stations. *Transportation Research Record: Journal of the Transportation Research Board*, (1538), 19-26.
23. Porter, G. (2014). Transport Services and Their Impact on Poverty and Growth in Rural Sub-Saharan Africa: A Review of Recent Research and Future Research needs. *Transport Reviews*, *34*(1), 25-45.
24. Rodnyansky, S. (2018). Do It Yourself: Obtaining Updated Transit Stop and Route Shapefiles in Urban and Nonurban Areas. *Cityscape, 20*(1), 205-214.
25. Schweiger, C. L. (2015). *Open data: Challenges and Opportunities for Transit Agencies* (No. Project J-7, Topic SA-34).
26. Steiss, T. (2006). Calculating/Analyzing Transit Dependent Populations Using 2000 Census Data and GIS. *Census Transport*. Census Transportation Planning Package 2000 Status Report. U.S. Department of Transportation. Washington, DC.
27. US Census Bureau. (2011). American Community Survey Puerto Rico Community Survey 2009 Subject Definitions.
28. Wei, R., Golub, A., Wang, L. and Cova, T. (2018). Evaluating and Enhancing Public Transit Systems for Operational Efficiency and Access Equity.
29. Wong, James (2013) Leveraging the General Transit Feed Specification for Efficient Transit Analysis. *Transportation Research Record*, *2338*, 11-19.
30. Zhao, J., and Deng, W. (2013). Relationship of Walk Access Distance to Rapid Rail Transit Stations with Personal Characteristics and Station Context. *Journal of Urban Planning and Development*, *139*(4), 311-321.
31. Zhong, Q., Karner, A., Kuby, M., and Golub, A. (2017). A Multiobjective Optimization Model for Locating Affordable Housing Investments While Maximizing Accessibility to Jobs by Public Transportation. *Environment and Planning B: Urban Analytics and City Science*, 2399808317719708.
32. Zygo, Angelika, *Measuring and Improving Seniors' Access to Medical Facilities* (2017). Master's Theses. 1042. http://digitalcommons.uconn.edu/gs_theses/1042, University of Connecticut. Storrs, CT.